

# 5. Ton und Klang

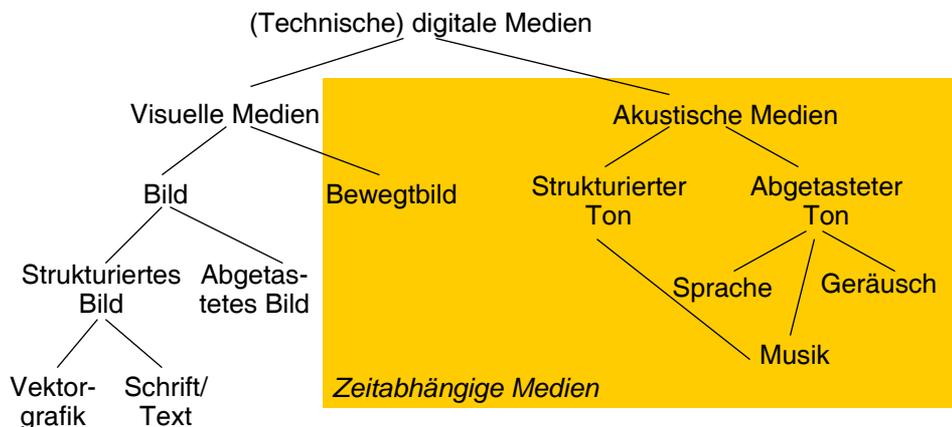
- 5.1 Ton: Physikalische und physiologische Aspekte 
- 5.2 Kompression von Audio-Signalen: MPEG-Audio
- 5.3 Audio-Datenformate: Übersicht
- 5.4 Klangerzeugung und MIDI

Literatur:

Taschenbuch Multimedia, Kapitel 5.1 und 5.2

Hannes Raffaseder: Audiodesign, Fachbuchverlag Leipzig 2002

## Übersicht Medientypen



Weitere Klassifikationskriterien: komprimiert/unkomprimiert, ...

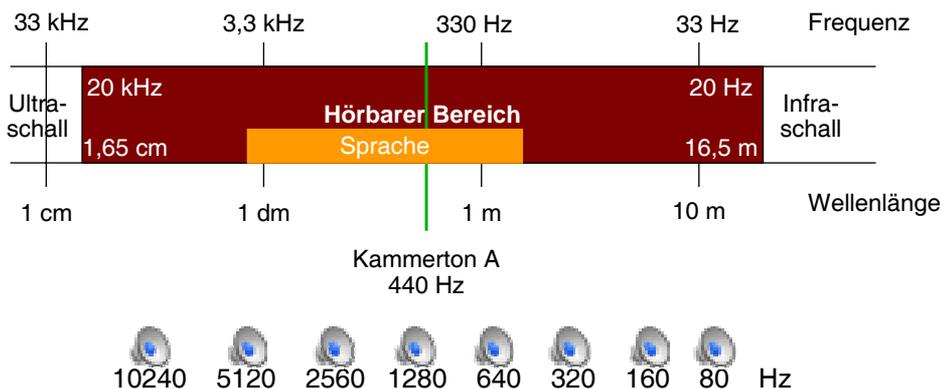
Angelehnt an Eidenberger & Divotkey

# Übermacht der visuellen Eindrücke?

- Tendenz: Visuelle Gestaltung dominiert
  - CD-Cover, Bühnenshow, Interessenlage von Medieninformatik-Studierenden
  - Gibt das Auge den Ton an?
- Warum Gestaltung *akustischer* Medien?
  - Ohr nicht verschließbar, unbewusste Wahrnehmung als Normalfall
    - » Akustische „Szenerie“ bestimmt Grundstimmung, siehe Filmmusik
    - » „Ohrwürmer“
  - Direkte Kopplung der Schallverarbeitung mit Emotionen
    - » Feine „Untertöne“ in der Sprache
  - Verbindender, kollektivierender Effekt
    - » Blick auf Schallereignis, Jugendkultur, Marschmusik
  - Objektivierende Wirkung
    - » Akustische Ereignisse wesentlich schwerer zu verfälschen als optische
  - Aufmerksamkeitseffekt
    - » Akustische Alarme
- Audio-Gestaltung ist wesentlich in Kunst und Ingenieurwesen!

## Ton (*sound*)

- Schallwellen, d.h. schnelle Änderungen des Luftdrucks
  - Longitudinalwellen, keine Schwingungsebenen
  - Ausbreitungsgeschwindigkeit ca. 331 m/s (bei 0°C), ca. 343 m/s (bei 20°C)
- Vom Menschen hörbares Spektrum ca. 20 Hz bis 20 kHz



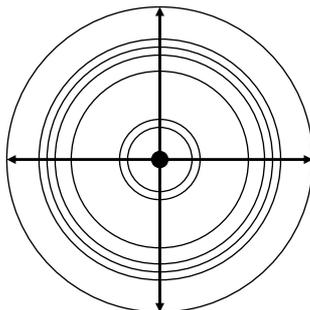
## Longitudinal- und Transversalwellen



- Longitudinalwelle (a):
  - Oszillatoren schwingen in der Ausbreitungsrichtung
- Transversalwelle (b):
  - Oszillatoren schwingen quer zur Ausbreitungsrichtung

## Schallausbreitung

- Schall breitet sich als *Kugelwelle* aus
  - Gleicher Schalldruck und gleiche Schallschnelle zu gegebenem Zeitpunkt auf konzentrischer Kugel um die Quelle
  - Bei größerer Entfernung: Äquivalent zu einer *ebenen Welle*



## Reflexion

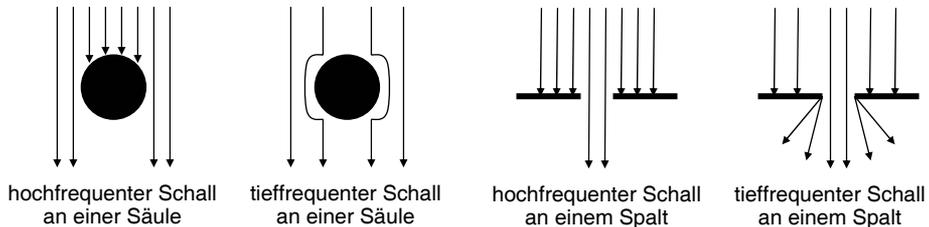
- Für Schallwellen gelten die gleichen Gesetze wie für Lichtwellen
- Aber: Effekte abhängig von Wellenlänge
  - Wellenlänge bei Schallausbreitung in der Luft bei 20°C zwischen ca. 20 m (tiefe Frequenzen) und 2 cm (hohe Frequenzen)
  - Lichtwellen erheblich kürzer (Mikrometer-Bereich)
- *Reflexion:*
  - Reflexionsgesetz (Einfallswinkel gleich Ausfallswinkel) gilt nur, wenn Grenzfläche groß im Vergleich zur Wellenlänge
    - » d.h. in kleineren Räumen keine Reflexion tiefer Frequenzen
  - Rauigkeit der Oberfläche führt zu *diffuser* (zerstreuender) Reflexion, wenn Unebenheiten in der Größenordnung der Wellenlänge
    - » d.h. auch bei zentimetergroßen Unebenheiten wirkt Wand auf tieffrequenten Schall als "glatt"

## Absorption, Brechung

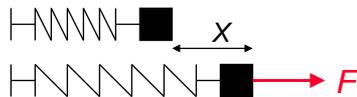
- *Absorption:*
  - Ein Teil der Schallenergie wird nicht reflektiert
    - » Absorptionsgrad abhängig vom Material und der Schallfrequenz
  - *Dissipation:* Umwandlung in Wärmeenergie
  - *Transmission:* Weiterleitung im absorbierenden Medium
  - Aufteilung der absorbierten Energie in Dissipation und Transmission
    - » frequenzabhängig: niedrige Frequenzen höhere Transmission (*Körperschall*)
- *Brechung:*
  - Auch Schall wird gebrochen
  - Z.B. bei Luftschichten verschiedener Temperatur

## Beugung

- Lichtwellen erzeugen immer einen Schatten hinter einem undurchsichtigen Gegenstand
  - Weil Gegenstand groß im Vergleich zur Wellenlänge
- Schall-"Schatten" entsteht erst, wenn Gegenstand groß im Vergleich zur Wellenlänge
- Bei Gegenständen in der Größenordnung der Wellenlänge tritt *Beugung* auf
  - Wellen fließen "um das Hindernis herum"
  - Wellen werden von einem Spalt zerstreut



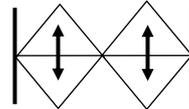
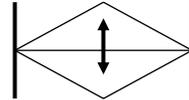
## Schallentstehung: Schwingung



- *Anregungskraft*: Wirkt von außen auf Körper ein
- *Elastizität*: Beschreibt Verformbarkeit des Körpers
  - Beispiel oben: Feder
  - Bewirkt Tendenz eines verformten bzw. ausgelenkten Körpers, in die Grundstellung zurückzukehren
- *Massenträgheit*:
  - Masse kehrt nicht sofort in Ruhestellung zurück, sondern bewegt sich in Gegenrichtung der ursprünglichen Auslenkung weiter
  - Elastizität bewirkt entgegengesetzte Kraft
  - Gesamtergebnis: *Schwingung*
  - Frequenz umso höher, je kleiner die Masse und je geringer die Elastizität

# Schwingungssysteme

- Schwingungssystem besteht aus:
  - Oszillator (z.B. Saite, Membran)
  - Anregung (z.B. zupfen, anstreichen, anschlagen)
  - Resonanzkörper (akustischer Verstärker, bestimmt Klangfarbe)
- Modell eines Schwingungssystems:
  - Gekoppelte einfache Schwingungssysteme (z.B. gekoppelte Federn)
    - » Konzeptionell auch unendlich viele gekoppelte Systeme (z.B. Modell einer Saite)
  - Randbedingungen
    - » z.B. eingespannte Saite: Auslenkung an den Enden immer Null
  - Begrenzte Anzahl von Wellenformen mit jeweils spezifischer Eigenfrequenz (“stehende Wellen”)
    - » Beidseitig eingespannte Saite der Länge  $L$ : Sinuswellen
    - » größtmögliche Wellenlänge  $2 \cdot L$  (Grundfrequenz)
    - » ganzzahlige Vielfache der Grundfrequenz



# Effektivwert und Pegel eines Signals

- Lautstärkenvergleich:
  - Wie bestimmen wir die „mittlere Amplitude“?
  - Arithmetischer Mittelwert falsch wegen negativer Werte
- *Effektivwert*:
  - Quadriert Signalstärkenwerte
  - Bildet Mittelwert über Zeitintervall

$$S_{\text{effektiv}} = \sqrt{\frac{1}{T} \int_T s^2(t) dt}$$

- *Pegel*:
  - Verhältnis zweier Größen
  - Riesige Wertunterschiede in der Schalldruck-Wahrnehmung des Menschen:
    - » Faktor  $32 \times 10^{12}$  (32 Trillionen)
  - > *logarithmische Pegel*

## Dezibel, Absolute Pegel

- Exponentieller Zusammenhang:
  - Verdoppelung eines physikalischen Basismaßes wird als eine Stufe (plus 1) aufgefasst (sog. logarithmisches Gesetz)
  - Beispiele für das logarithmische Gesetz:
    - » Empfindlichkeit des menschlichen Gehörs für Schalldruck
    - » Verlust in Telefonkabeln (als Funktion der Länge)
- Bel (B, benannt nach Graham Bell):
  - Logarithmisches Maß zu einem Referenzwert (0 Bel)
  - Referenzwert traditionell 1 mW bei 600  $\Omega$
- DeziBel (dB, „de-be“, „dee-bee“): 1 dB = 1/10 B
- *Absolute Pegel*: Dezibel-Angabe relativ zu verschiedenen Referenzen:
  - Verlustleistung (m, W)
  - Schalldruck (SPL) relativ zu  $2 \times 10^{-5}$  Pa (Hörschwelle)
  - Bewerteter Schalldruck (A)

## Relative Pegel

Relativer *Leistungs*pegel in dB:

$$L_P = 10 \cdot \log\left(\frac{P_A}{P_E}\right)$$

Verdopplung:

$$L_{P'} = 10 \cdot \log\left(\frac{2 \cdot P_A}{P_E}\right) = 10 \cdot \log(2) + L_P = 3 + L_P$$

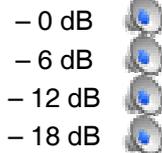
Bezogen auf *Amplituden*pegel (effektive Amplitudenwerte):

$$L_P = 10 \cdot \log\left(\frac{V_A^2}{V_E^2}\right) = 20 \cdot \log\left(\frac{V_A}{V_E}\right)$$

Verdopplung:

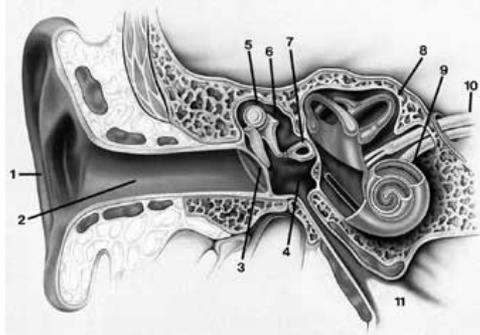
$$L_{P'} = 20 \cdot \log\left(\frac{2 \cdot V_A}{V_E}\right) = 20 \cdot \log(2) + L_P = 6 + L_P$$

220Hz



„Plus 6 dB  
ist doppelte  
Lautstärke“

# Das menschliche Ohr

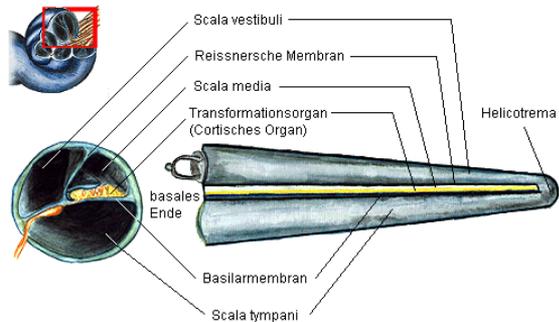


- 1 Ohrmuschel (*pinna*)
- 2 Gehörgang
- 3 Trommelfell (*ear drum*)
- 4 Mittelohr
- 5 Hammer (*malleus*)
- 6 Amboss (*incus*)
- 7 Steigbügel (*stapes*)
- 8 Gleichgewichtsorgan
- 9 Schnecke (*cochlea*)
- 10 Hörnerven

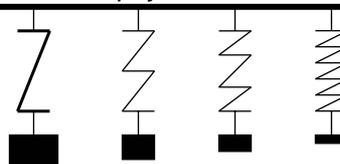
- Aussenohr:
  - HRTF = Head Related Transfer Function
- Mittelohr:
  - Trommelfell, Hammer, Amboss, Steigbügel: Verstärkung der Kraftwirkung
- Innenohr:
  - Schnecke (*cochlea*)
  - Aufgerollte Röhren (Gänge), gefüllt mit Lymphflüssigkeit

# Schnecke

- Aufgerollt: Im abgerollten Zustand ca. 30 mm lang
- Enthält 3 Röhren und die *Basilarmembran*
- Auf der *Basilarmembran*: *Corti'sches Organ*
  - 20000 Haarzellen
  - überträgt Bewegungen der Lymphflüssigkeit auf Haarzellen und damit auf Nervenreize
- Ort der Reizung im *Corti'schen Organ* abhängig von der Frequenz
  - konischer Aufbau und veränderte Elastizität der Membran zur Spitze hin



## Vereinfachtes physikalisches Modell:



## Räumliches Hören

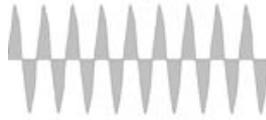
- Stereo-Effekt
  - Zeitliche Verzögerung des selben Schallereignisses in seiner Wahrnehmung durch beide Ohren
  - Verzögerungsmessung liefert Information über Entfernung der Quelle
- Kann ein einseitig tauber Mensch räumlich hören?
  - Eingeschränkt: ja!
  - Reflexion und Beugung an Umwelt und Ohrmuscheln liefern umfangreiche Information
- Frequenzabhängigkeit der Ortung:
  - Niedrige Frequenzen generell schlechter zu orten
  - Konsequenz physikalischer Tatsachen (Wellenlänge:Hindernis)
  - Siehe z.B. „Subwoofer“-Technologie

## Empfindungen: Klang und Geräusch

- Töne sind vom Menschen wahrnehmbare kleine Luftdruckänderungen
  - Warum empfinden wir manchen Ton als "Musik"?
- Primärempfindungen der Tonwahrnehmung:
  - Tonhöhe (Bsp. verschiedene Klaviertasten)
  - Lautstärke (Bsp. Trommelanschlag)
  - Klangfarbe (Bsp. gleicher Ton auf verschiedenen Instrumenten)
- Klang:
  - alle drei Primärempfindungen wahrnehmbar 
- Geräusch:
  - nur Lautstärke wahrnehmbar 

## Periodizität

- Klänge sind, als Signalform betrachtet, *periodisch* (d.h. wiederholen Teilabschnitte)
- Geräusche sind schlechter strukturiert und meist aperiodisch.



Blockflöte 



LKW 



Violine 



Fahrradkette 

## Frequenzspektrum, Oberschwingungen

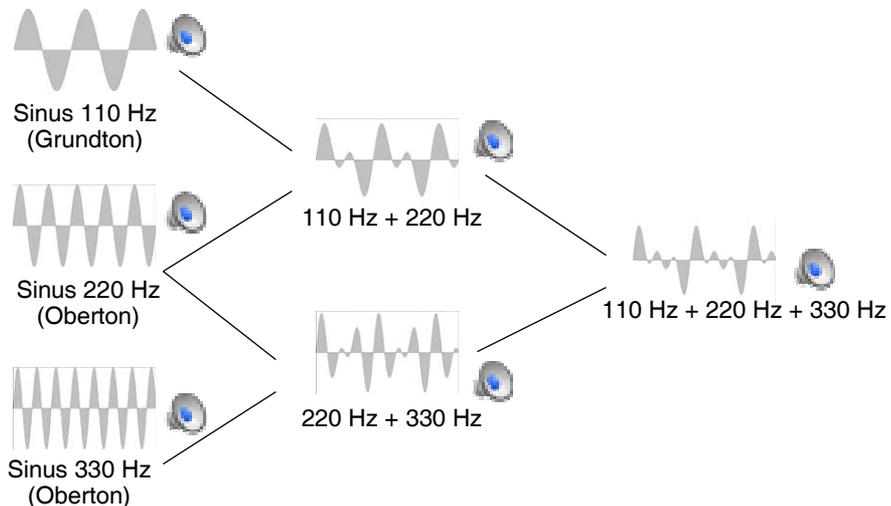
- *Frequenz* (Tonhöhe):
  - Maß für die Häufigkeit, mit der sich positive und negative Spannungen abwechseln, Maß 1 Hertz = 1 Schwingung/s
- Audiosignal:
  - besteht aus Vielzahl von überlagerten Frequenzen (Frequenzspektrum)
  - *Bandbreite*: Differenz zwischen höchster und niedrigster Frequenz
  - Beispiel: Ton eines Musikinstrumentes
- *Grundton*: Wahrgenommene Tonhöhe
  - Größter gemeinsamer Teiler aller am Signal beteiligten Frequenzen
  - Oft: Tiefste enthaltene Teilfrequenz
- *Obertöne*:
  - *reine Obertöne*: ganzzahlige Vielfache der Grundfrequenz
  - zusätzlich: geräuschhafte Tonanteile (z.B. Zupfgeräusch)
  - Obertonspektrum ist für charakteristischen Instrumentklang bestimmend

# Harmonische Schwingungen

$$s(t) = \sum_{n=1}^{\infty} S_n \sin(2\pi \cdot n \cdot f \cdot t + \varphi_n)$$

- *Harmonische Schwingung (harmonischer Ton):*
  - Beschreibbar durch Sinus- und Cosinusfunktionen
- *Harmonisch komplexe Schwingung (Klang):*
  - Zusammengesetzt aus harmonischen Teilschwingungen
  - Grundton und ganzzahlige Vielfache
- *Klänge ohne Grundton:*
  - Zusammengesetzt aus harmonischen Teilschwingungen
  - Keine ganzzahligen Vielfachen einer Grundfrequenz
  - Z.B. Pauken, Gongs

# Überlagerung harmonischer Schwingungen



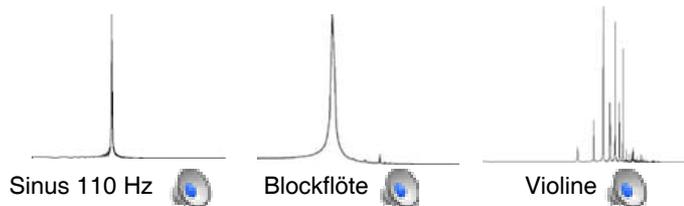
## Interferenz und Schwebung

- **Interferenz:** Überlagerung von Schallwellen exakt gleicher Frequenz
  - konstruktive Interferenz (*in phase*):
    - » Übereinstimmung der Phasenlage
    - » Addition der Amplituden
  - destruktive Interferenz (*out of phase*):
    - » Gegenphasige Lage ( $180^\circ$  verschoben)
    - » Subtraktion der Amplituden – Auslöschung
- **Schwebung:** Überlagerung von Wellen annähernd gleicher Frequenz
  - konstruktive und destruktive Interferenz wechseln sich ab
  - Amplitudenverlauf beschreibt neues Signal mit Frequenz = Differenz der überlagerten Frequenzen

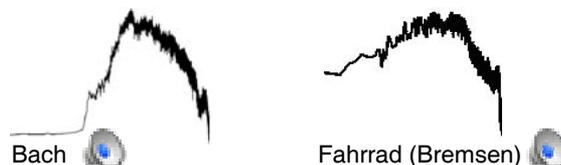
## Frequenzspektren

- Frequenzspektrum von Klängen
  - Anzahl diskreter Spektrallinien (Grund- und Obertöne)
- Frequenzspektrum von Geräuschen
  - kontinuierliches Spektrum diverser Frequenzen

Klänge:

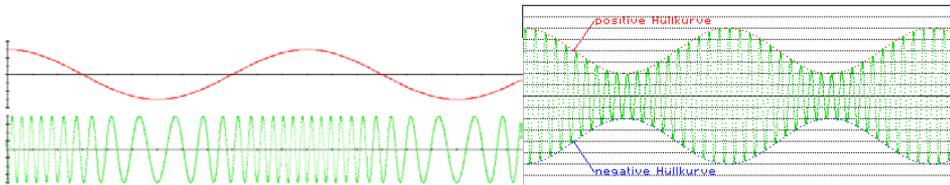


Geräusche:



# Modulation

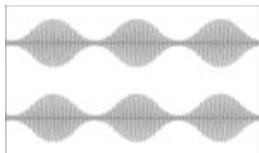
- Modulation:
  - gezielte Überlagerung einer Grundfrequenz (Träger) mit einem Nutzsignal
  - Anwendung in der Übertragungstechnik (z.B. Rundfunk)
- Amplituden-Modulation:
  - Hüllkurve der Trägerfrequenz (Amplitude) in Abhängigkeit vom Nutzsignal verändert
- Frequenz-Modulation;
  - Abweichung von der Basisfrequenz in Abhängigkeit vom Nutzsignal



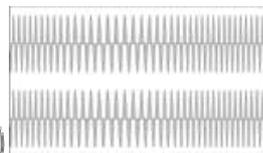
# Zeitlicher Verlauf von Schallsignalen, Transienten

- Mikroskopischer Zeitbereich:  $0,05 \text{ ms} < t \leq 50 \text{ ms}$ 
  - Hörbare Frequenzen (20 Hz bis 20 kHz)
- Zeitbereich der Übergangsklänge (*Transienten*):  $50 \text{ ms} < t \leq 150 \text{ ms}$ 
  - Modulation hörbarer Frequenzen
  - So "schnell" (20 Hz bis 7 Hz), dass noch als Variation des Klangs wahrgenommen
- Makroskopischer Zeitbereich:  $t > 150 \text{ ms}$ 
  - Verlauf eines Klangs (z.B. gespielte Note) über die Zeit (Hüllkurve)
  - Formaler Aufbau eines Musikstücks
  - Tempo, Metrum, Rhythmus
- Beispiele für Transienten:

*Tremolo*  
Amplituden-  
moduliertes  
Sinussignal



*Vibrato*  
Frequenz-  
moduliertes  
Sinussignal



## Informationsgehalt akustischer Szenen

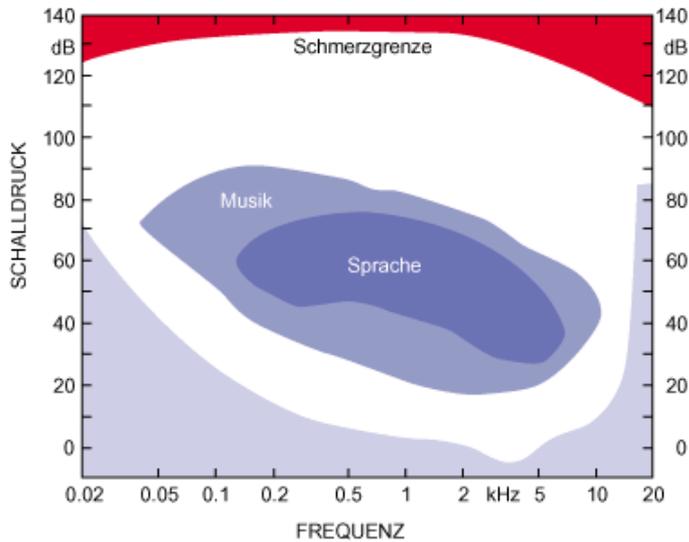
- Umgebung des Menschen ist durch Vielzahl verschiedenartiger Geräusche geprägt
- Menschlicher Gehörsinn (= Hörapparat + Nachverarbeitung)
  - Kann gezielt Aufmerksamkeit auf sehr spezifische Teilsignale richten;  
*Cocktailparty-Effekt*
  - Nimmt Hintergrundinformationen neben selektierter Information weiter wahr (z.B. Alarmsignale, Durchsagen)
  - Kann räumlich weit verstreute Informationsquellen integrieren (wesentlich geringere Verdeckung als in der Optik)
  - Kann sehr genau verschiedene Schallereignisse differenzieren („Mischfarben“ vs. Wahrnehmung gleichzeitiger Töne)
- Digitale Tonverarbeitung muss sich mit der Informationsverarbeitung im Gehirn befassen, nicht nur mit dem physikalischen Hörorgan

## Psychoakustik

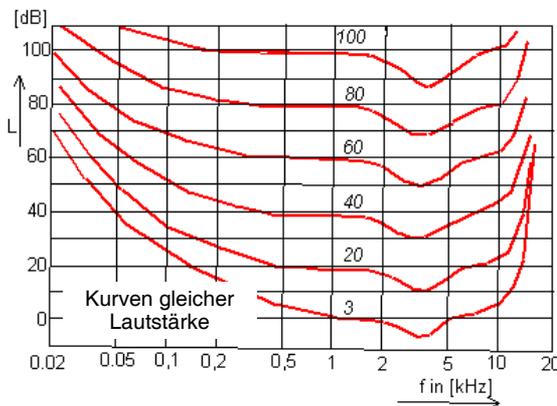
- Die *Psychoakustik* versucht kausale Zusammenhänge zwischen den physikalischen Größen eines Schallsignals und den dadurch ausgelösten Empfindungen zu erfassen.

Mensch wahrnehmbare Eigenschaft	Computer physikalische Kenngröße
Tonhöhe	Grundfrequenz
Lautstärke	Druckamplitude
Klangfarbe	Frequenzspektrum

# Hörfläche



# Lautstärke und Frequenz

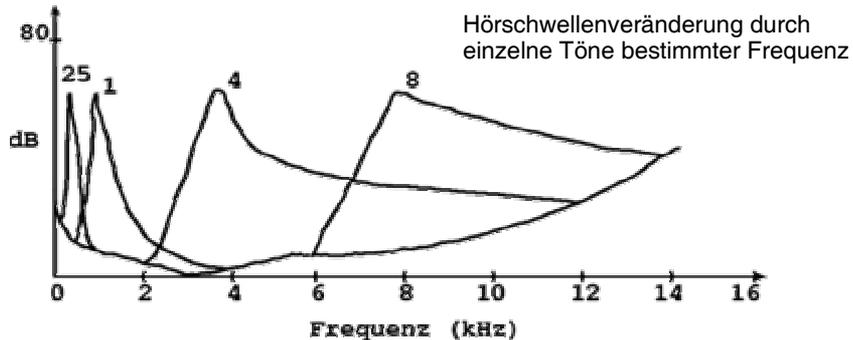


- Die Hörschwelle ist abhängig von der Frequenz:
  - Größte Empfindlichkeit bei ca. 2 – 5 kHz
- Gleich empfundene Lautstärke frequenzabhängig bei verschiedenem Schalldruck
  - *Subjektive* Lautstärke gemessen in *Phon*
  - x Phon: Lautstärke vergleichbar zu 1 kHz Ton bei Schalldruck x dB (SPL)

- Korrektur der Frequenzabhängigkeit durch Gewichtungskurven (A, B, C)
  - meistverwendet A-Kurve (40 Phon), dB(A)
  - für höhere Lautstärken B- und C-Kurven

## Frequenz-Maskierung

- Frequenzwahrnehmung durch die Schnecke:
  - endliche Breite des betroffenen Bereichs
  - dadurch Überlappung benachbarter Frequenzbereiche
- Ein lauter Ton einer bestimmten Frequenz macht leisere Töne ähnlicher Frequenz unhörbar.



## Zeitliche Maskierung

- Die Hörwahrnehmung beruht auf dem Mittelwert eines Zeitintervalls von ca. 30 ms
  - Ein lauter Ton beeinflusst die Wahrnehmung einer frequenzähnlichen Tons auch, wenn der zweite Ton in direkter zeitlicher Nachbarschaft liegt
  - Vorwärtsmaskierung: Nachfolgende Töne kaum wahrnehmbar
  - Rückwärtsmaskierung:
    - » Auch vorausgehende Töne betroffen (in kürzerem Zeitabstand)

