



# Interaktive Medien

## Kapitel Audio

**Prof. Dr.-Ing Timo Gerkmann**  
**Signal Processing, Universität Hamburg**

**Mit Folien von: Prof. Dr. Frank Steinicke**  
Human-Computer Interaction, Universität Hamburg

# Inhalt

- Schall, Frequenz und Lautstärke
- Auditiver Wahrnehmung
- Digitalisierung und Codierung von Audiosignalen
- Praxis: Audiobearbeitung

# Audio

## Motivation

- Ohr nicht verschließbar, unbewusste Wahrnehmung als Normalfall
- direkt Kopplung der Schallverarbeitung mit Emotionen
- akustische Alarme
- ...

# Fokus: Mediendesign

## Auditory Icon

- **Auditory Icon** ist akustische Symbol, welches Objekt oder Vorgang im UI darstellt
- **Auditory Icon** wird meist durch natürliches Geräusch wiedergegeben

# Auditory Icon

## Beispiele

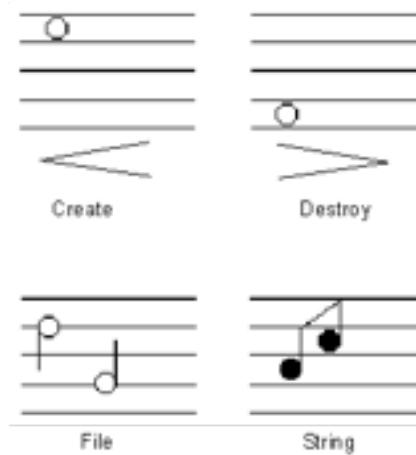


# Auditory Icon Beispiel



# Earcon

- Earcon ist synthetisches Geräusch zur abstrakten Codierung von Information im UI, z.B.



→ Bedeutung von Earcon muss erlernt werden

# Earcon Beispiel





# Interaktive Medien

## Kapitel Audio

Schall, Frequenz und Lautstärke

# Audiosignale

## Physikalische Grundlagen

- **Schall** ist mechanische Bewegung eines physikalischen Mediums, die sich als Druckwelle ausbreitet (**Druckschwankungen**)
  - z.B. Luft, Wasser, ...
- ohne Medium keine Schallausbreitung, z.B. Weltall

# Wellen

## Longitudinal- und Transversal

- **Longitudinalwelle**: Oszillatoren schwingen in der Ausbreitungsrichtung
- **Transversalwelle**: Oszillatoren schwingen quer zur Ausbreitungsrichtung

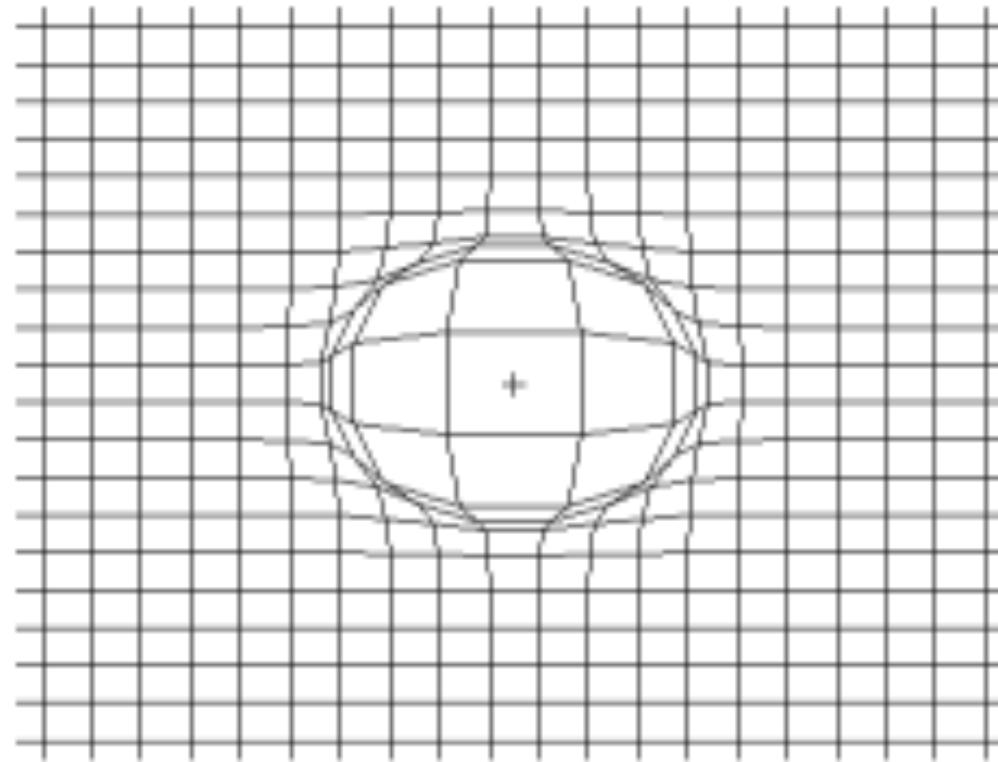
Longitudinalwelle



Transversalwelle

# Longitudinalwelle

## Bespiel

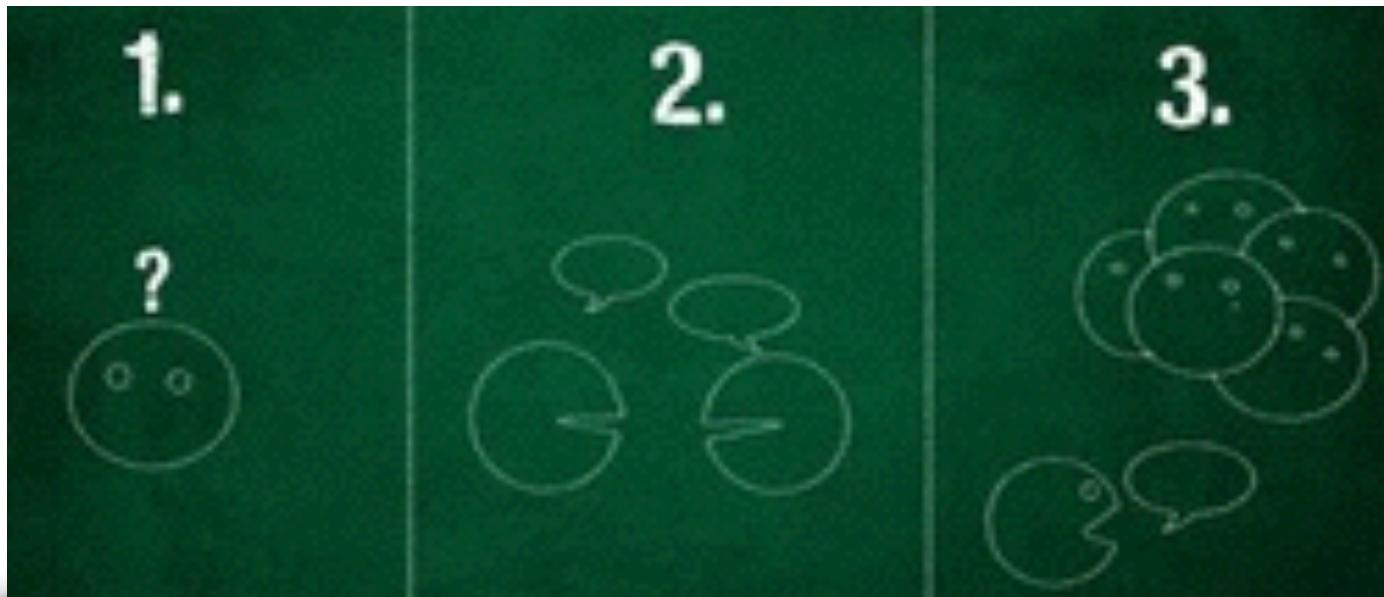


# Audiosignale

## Ausbreitungsgeschwindigkeit

- **Schallwellen** breiten sich im Raum konzentrisch um Entstehungspunkt aus
- **Schallwellen** breite sich in unterschiedlichen Medien mit unterschiedlichen **Ausbreitungsgeschwindigkeiten** aus
  - Beispiele: Luft: 331m/s bei 0°C und 343m/s bei 20°C

# Think! Pair! Share!



Wie weit ist Gewitter bei entfernt, falls zwischen Blitz und Donner 3 Sekunden vergehen?



Wie weit ist Gewitter bei Schallgeschwindigkeit von 340m/s entfernt, falls zwischen Blitz und Donner 3 Sekunden vergehen?

A

ca. 1000m

B

ca. 3000m

C

ca. 2000m

D

ca. 4000m



Wie weit ist Gewitter bei Schallgeschwindigkeit von 340m/s entfernt, falls zwischen Blitz und Donner 3 Sekunden vergehen?

A

ca. 1000m

B

ca. 3000m

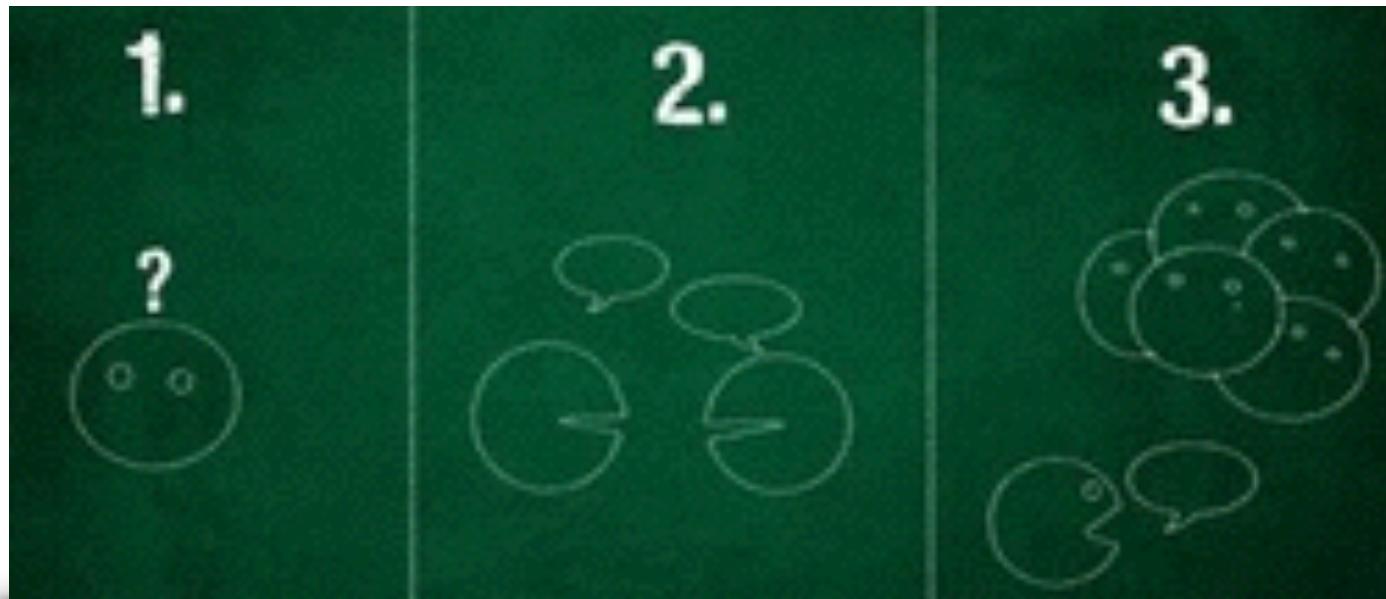
C

ca. 2000m

D

ca. 4000m

# Think! Pair! Share!



Ab welcher Geschwindigkeit in km/h  
durchbricht Flugzeug die Schallmauer?



Ab welcher Geschwindigkeit in km/h durchbricht Flugzeug die Schallmauer?

A

ca. 800km/h

B

ca. 1200km/h

C

ca. 1600km/h

D

ca. 2000km/h



Ab welcher Geschwindigkeit in km/h durchbricht Flugzeug die Schallmauer?

A

ca. 800km/h

B

ca. 1200km/h

C

ca. 1600km/h

D

ca. 2000km/h

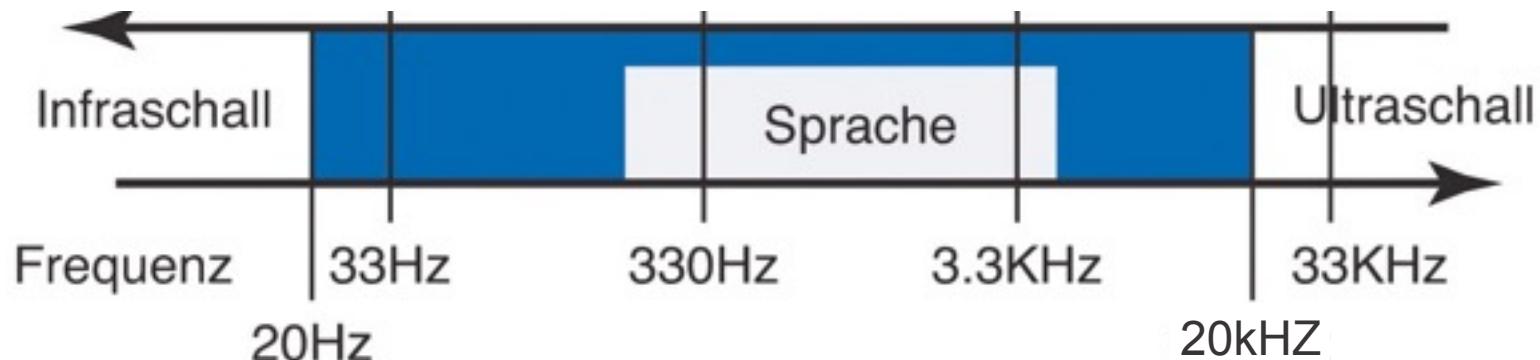
# Ton & Geräusch

- akustischer Reiz entsteht durch (schnelle)  
**Luftdruckschwankungen**
  - **unregelmäßig Luftdruckschwankungen**  
werden als Geräusche wahrgenommen
  - **periodische Luftdruckschwankungen**  
werden als Ton wahrgenommen
    - ▶ **Wellenlänge**, z.B. in mm
    - ▶ **Frequenz**, z.B.  $f$  in Hz

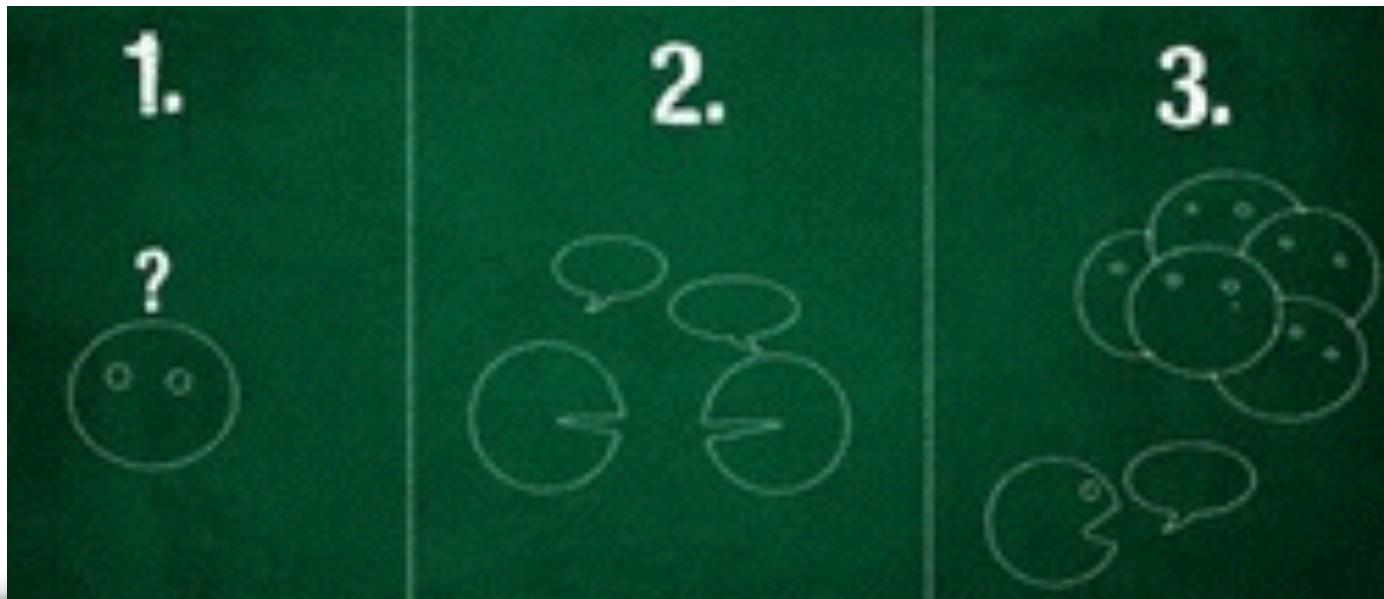
# Frequenzen

## Hörbares Spektrum

- **Hörbares Spektrum** liegt im Idealfall (junger, normal hörender Mensch) im Frequenzbereich von 20Hz bis 20kHz

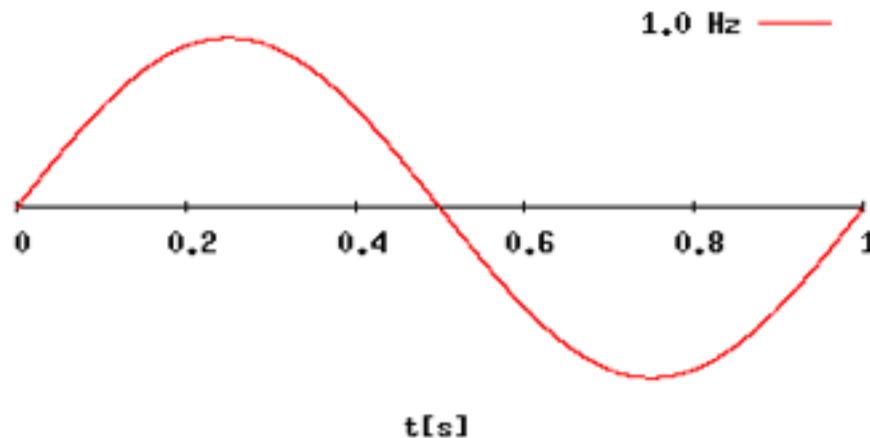


# Think! Pair! Share!



Schallgeschwindigkeit  $c_s = 343\text{m/s}$

Welche Wellenlänge haben somit hörbare  
Audiodrähte?



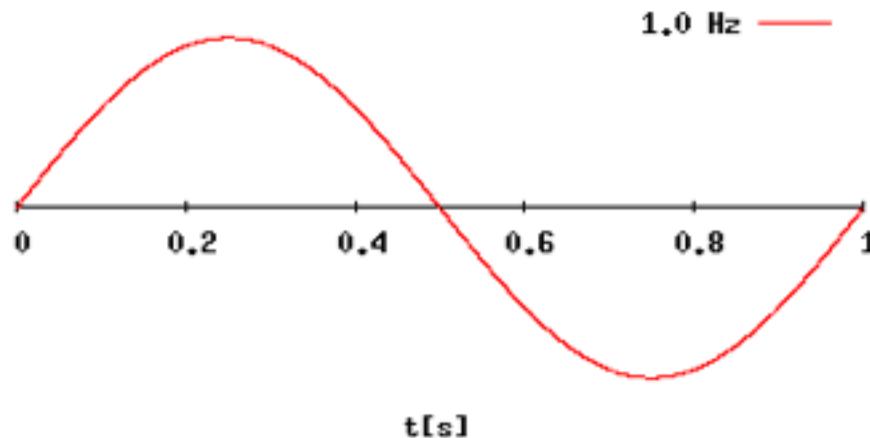
Welche Wellenlänge haben somit hörbare Audiowellen mit einer Schallgeschwindigkeit von  $c_s = 343\text{m/s}$  ?

A ca. 17cm bis 170m

B ca. 0,17cm bis 17m

C ca. 1,7cm bis 17m

D ca. 17cm bis 17m



Welche Wellenlänge haben somit hörbare Audiowellen mit einer Schallgeschwindigkeit von  $c_s = 343\text{m/s}$  ?

A ca. 17cm bis 170m

B ca. 0,17cm bis 17m

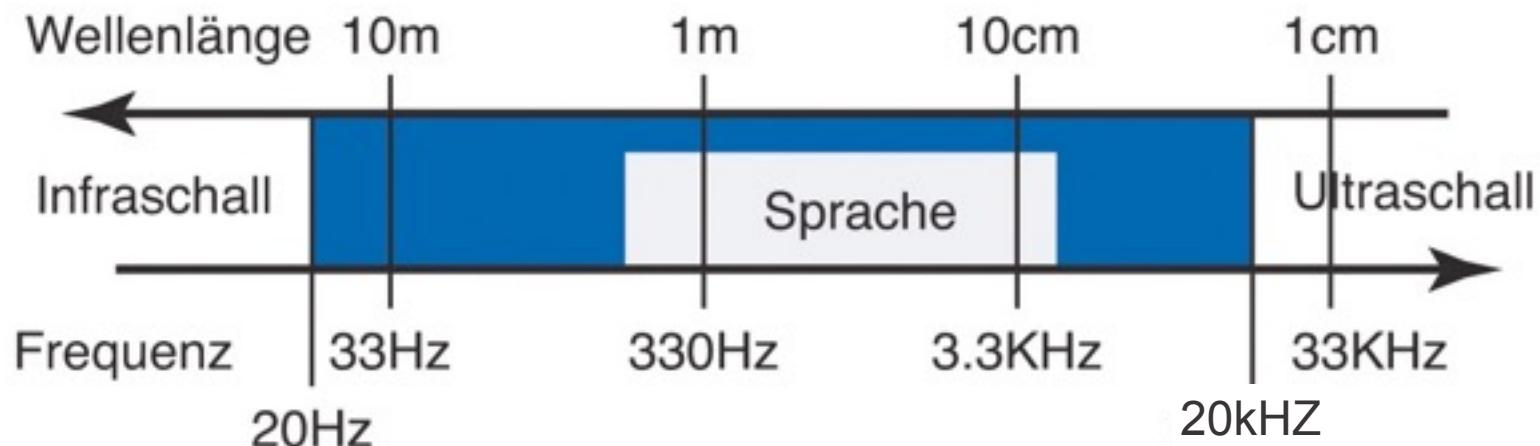
C ca. 1,7cm bis 17m

D ca. 17cm bis 17m

# Frequenzen

## Hörbares Spektrum

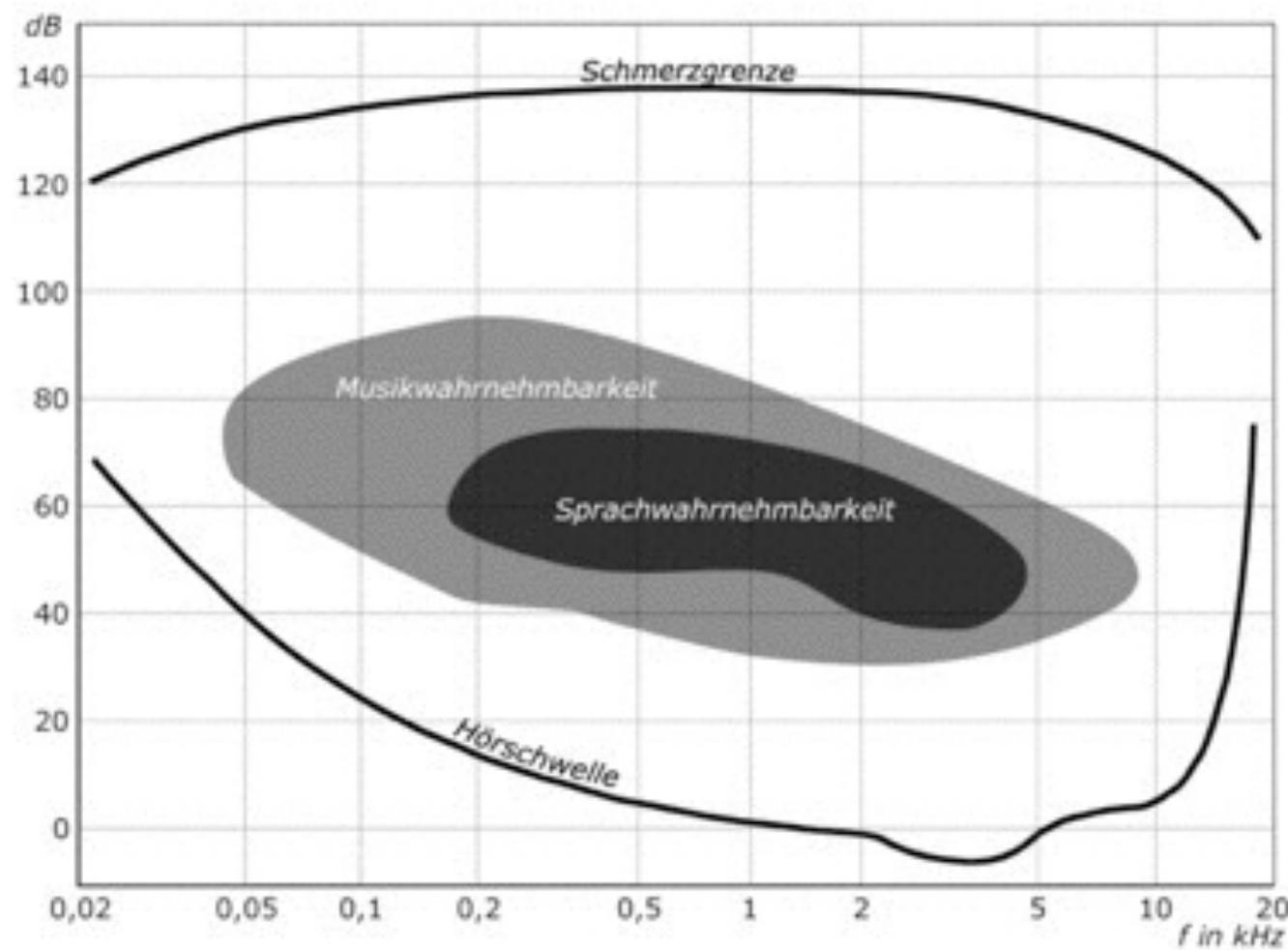
- **Hörbares Spektrum** liegt im Idealfall (junger, normal hörender Mensch) im Frequenzbereich von 20Hz bis 20kHz



# Tonhöhe

- **Tonhöhe / Tonheit** (engl pitch) ist eine psychoakustische Größe. *Tonhöhe* wird aber häufig mit der Grundfrequenz des Schalls gleichgesetzt
  - geringe Grundfrequenz --> niedrige Tonhöhe
  - hohe Grundfrequenz --> hohe Tonhöhe
- Diskriminierungsfähigkeiten des Menschen liegt (bei niedrigen Frequenzen) im Bereich von 1.5Hz

# Lautstärke Wahrnehmung

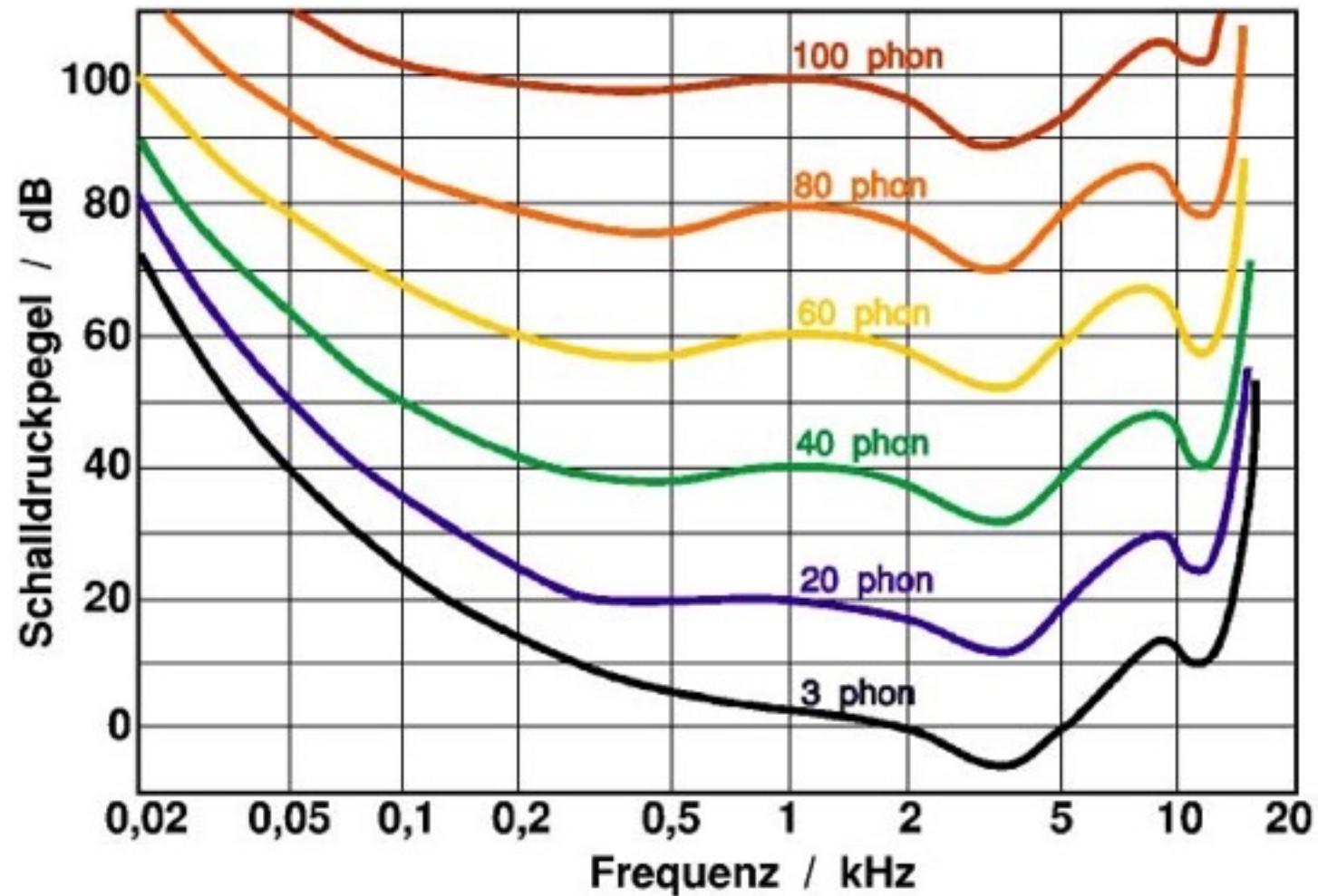


# Lautstärke

- **wahrgenommene Lautstärke** hängt vom Schalldruckpegel, Frequenzspektrum und Zeitverhalten des Schalls ab
- **Schalldruckpegel [dB]** entspricht der physikalisch gemessenen Luftdruckänderung.
- **Lautstärkepegel [phon]** psychoakustisches Vergleichsmaß, z.B.: Hörereignis mit 40phon ist genauso laut wie ein Sinus bei 1000Hz mit Schalldruckpegel 40dB.

# Isophone

## Kurven gleicher Lautstärke





Alexander Graham Bell, 1875

# (Dezi)bel

- **Dezibel** ist logarithmische Maßeinheit zur Kennzeichnung von Pegeln und Maßen
  - **Dezibel (dB)** = 1/10 Bel (B)
- **Dezibel** dient zur Kennzeichnung des Verhältnisses der Energiegröße  $P_2$  und  $P_1$ .
- $P_1 := 2 \cdot 10^{-5}$  Pascal: ursprünglich Schalldruck bei Hörschwelle bei 1kHz)

$$L = \log_{10} \left( \frac{P_2^2}{P_1^2} \right) B = 10 \cdot \log_{10} \left( \frac{P_2^2}{P_1^2} \right) dB = 20 \cdot \log_{10} \left( \frac{P_2}{P_1} \right) dB$$

# Umrechnung dB

## Beispiele

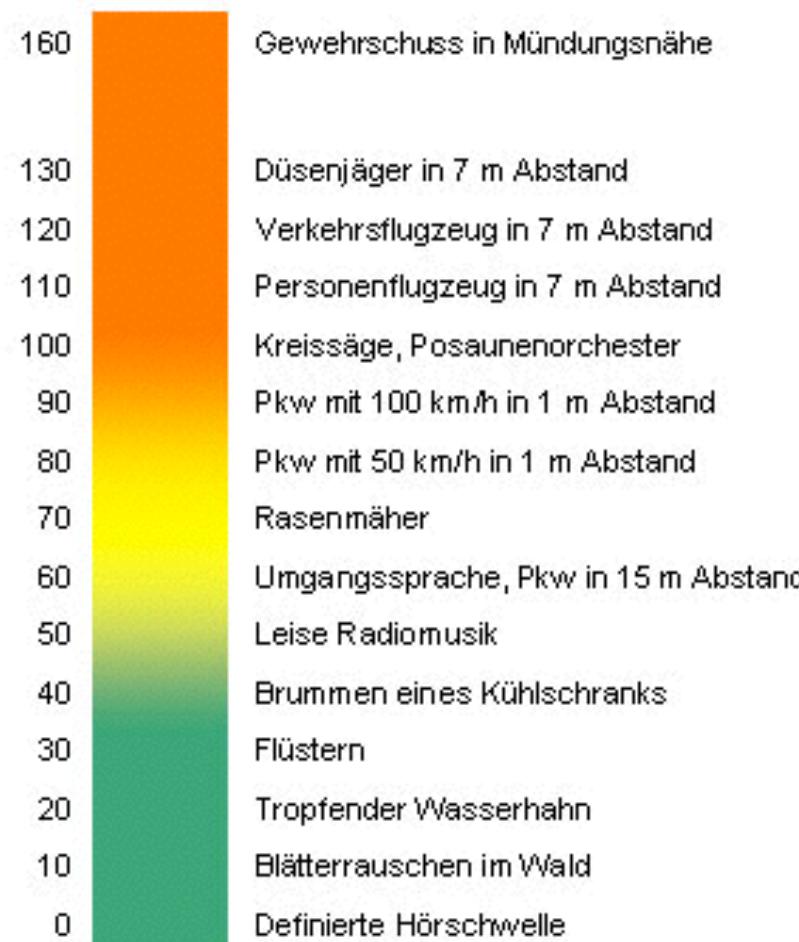
Sprechen →  
Blätter rascheln →

P2 / P1	L
10000	80dB
100	40dB
10	20dB
4	12dB
2	6dB
1,26	2dB
1	0dB
0,79	-2dB
0,5	-6dB
0,25	-12dB
0,1	-20dB
0,01	-40dB
0,0001	-80dB

← Schreien  
← Flüstern  
← Hörschwelle

# Pegelbereiche

Bsp: Lärm in dB



# Lautstärke

Beispiel:  $\pm 10\text{dB}$

Freude, schöner Götterfunken

Allegro maestoso

Ludwig van Beethoven

The musical score consists of two staves of music for piano. The top staff is in common time (indicated by a 'C') and has a key signature of one sharp (F#). The bottom staff is also in common time and has a key signature of one sharp (F#). The music is composed of eighth and sixteenth notes. The tempo is marked as 'Allegro maestoso'. The composer's name, 'Ludwig van Beethoven', is written at the end of the score.

This is a continuation of the musical score from the previous page. It shows two more staves of music for piano, maintaining the same key signature and tempo markings as the first section.



- 10 -

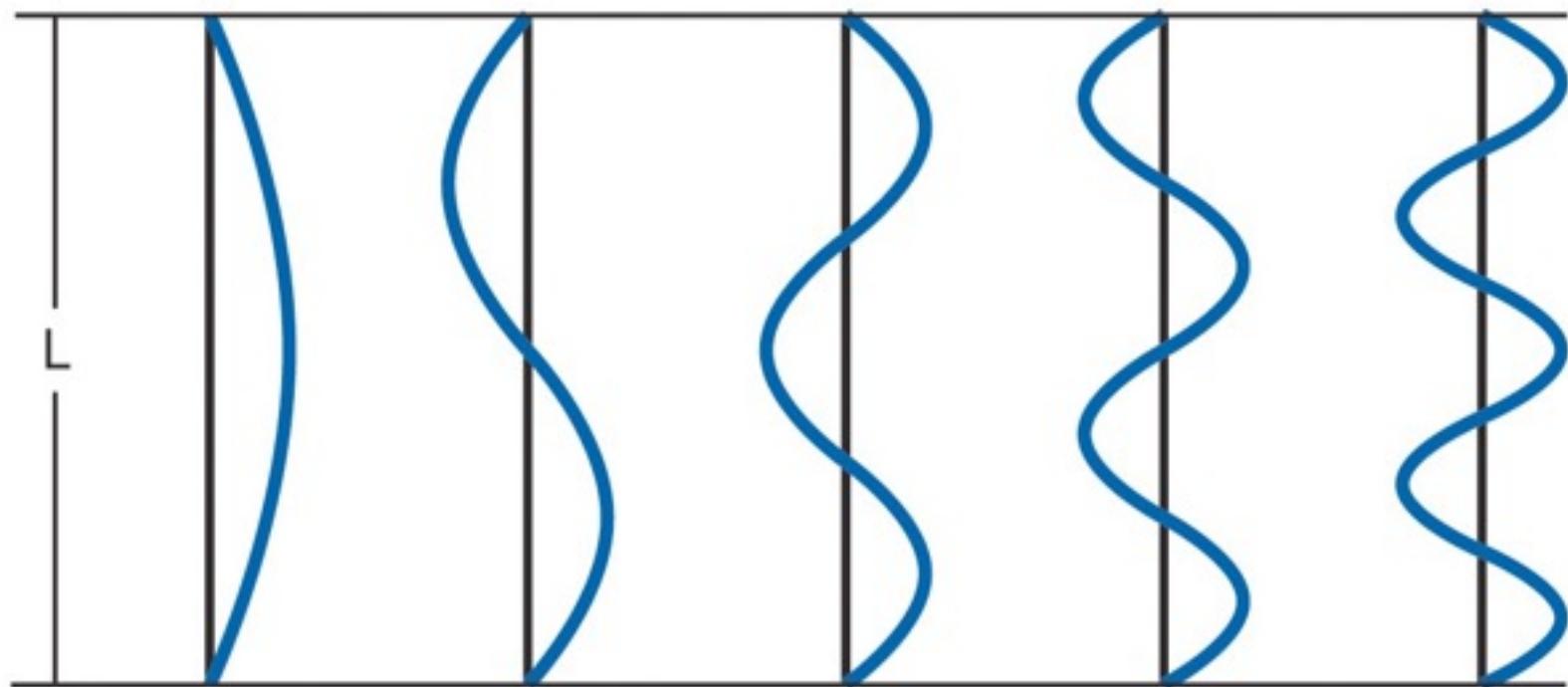
# Töne

## Grundton & Obertöne

- Töne werden oft dadurch erzeugt, dass Gegenstand (Saite, Stimmbänder) in Schwingung versetzt wird
- Schwingung der n-fachen Grundfrequenz wird **n-te Harmonische** oder **(n-1)-te Oberton** genannt

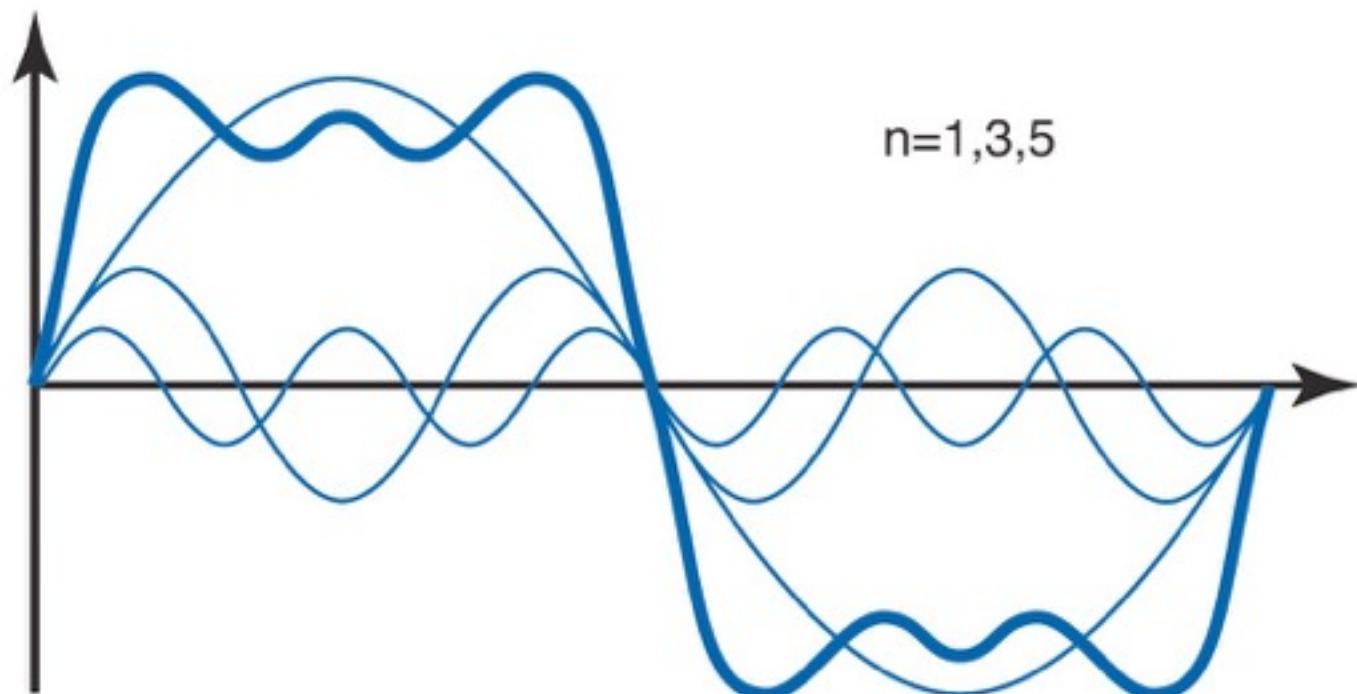
# Harmonische

## Bsp: Grundton & Obertöne



# Addition von Obertönen

## Beispiel



# Klangfarbe

- **Klangfarbe** eines Tons ist spezifisches Gemisch aus **Grundton** und **Obertönen**
  - **Klangfarbe** hängt von Beschaffenheit des Tonerzeugers (z. B. Instrument, Singstimme) ab
  - **Klangfarbe** und **(Ein-)schwingverhalten** lassen uns Töne auditiv unterscheiden

# Fouriertransformation

## Fourier-Reihe

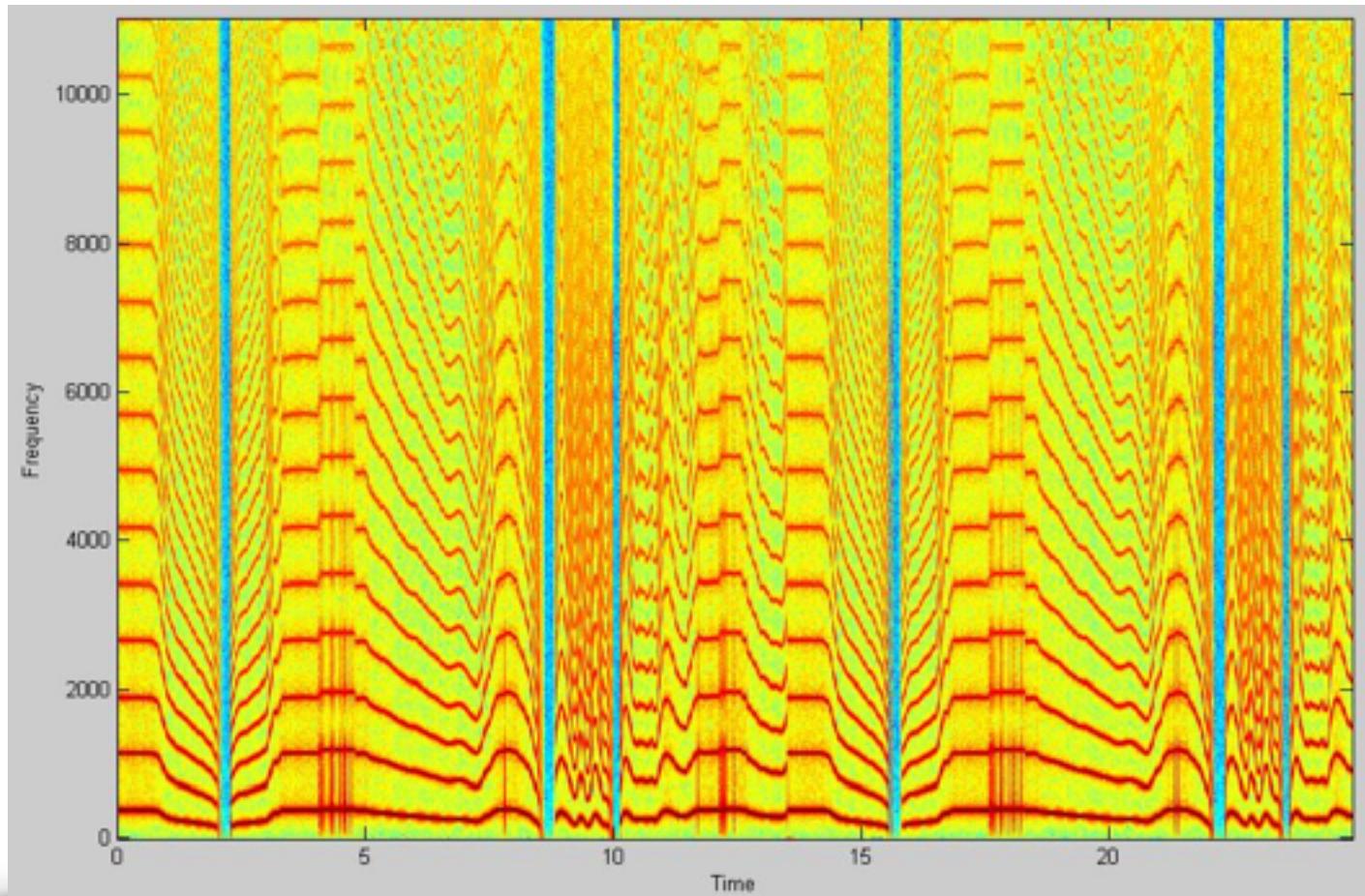
- Jede periodische Schwingung lässt sich durch unendliche Summe von sich überlagerten Cosinus-Schwingungen annähern

$$x(t) = \sum_{k=0}^{\infty} a_k \cdot \cos(k \cdot \omega_0 \cdot t + \theta_k)$$

- Grundfrequenz  $\omega_0$
- Anteil k-ten harmonische Schwingung  $a_k$
- Phasenverschiebung der k-ten harmonischen Schwingung  $\theta_k$

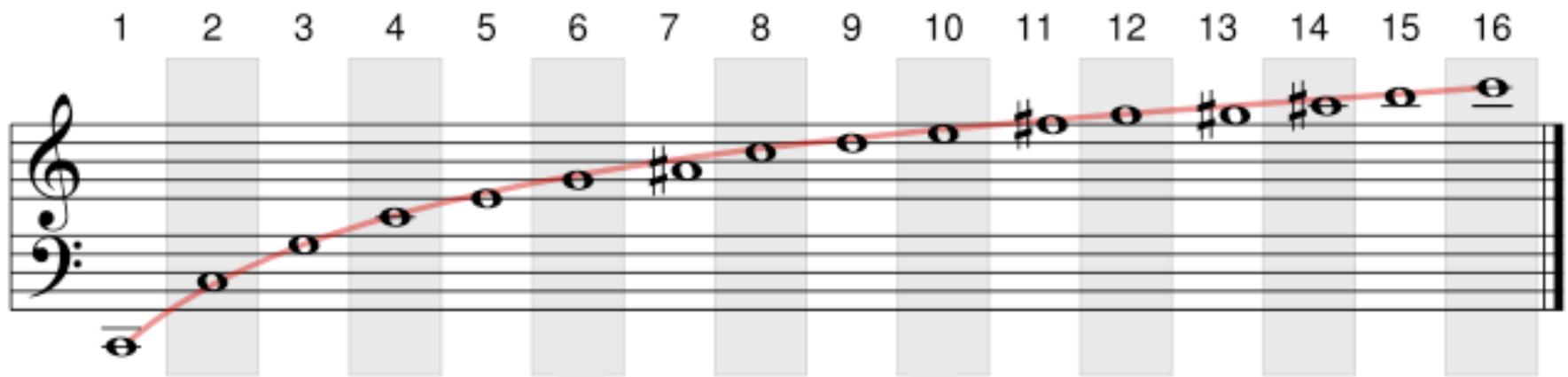
# Frequenzspektrum

## Beispiel: DFT



# Obertonreihe

## Beispiel



# Schallwandler

- am Anfang jeder elektroakustischen Übertragungskette steht **Schallwandler** oder **elektroakustischer Wandler**
- Beispiele
  - Mikrofon
  - einige Musikinstrumente leiten direkt aus mechanischer Schwingung elektrisches Signal ab, E-Gitarre, E-Piano ...

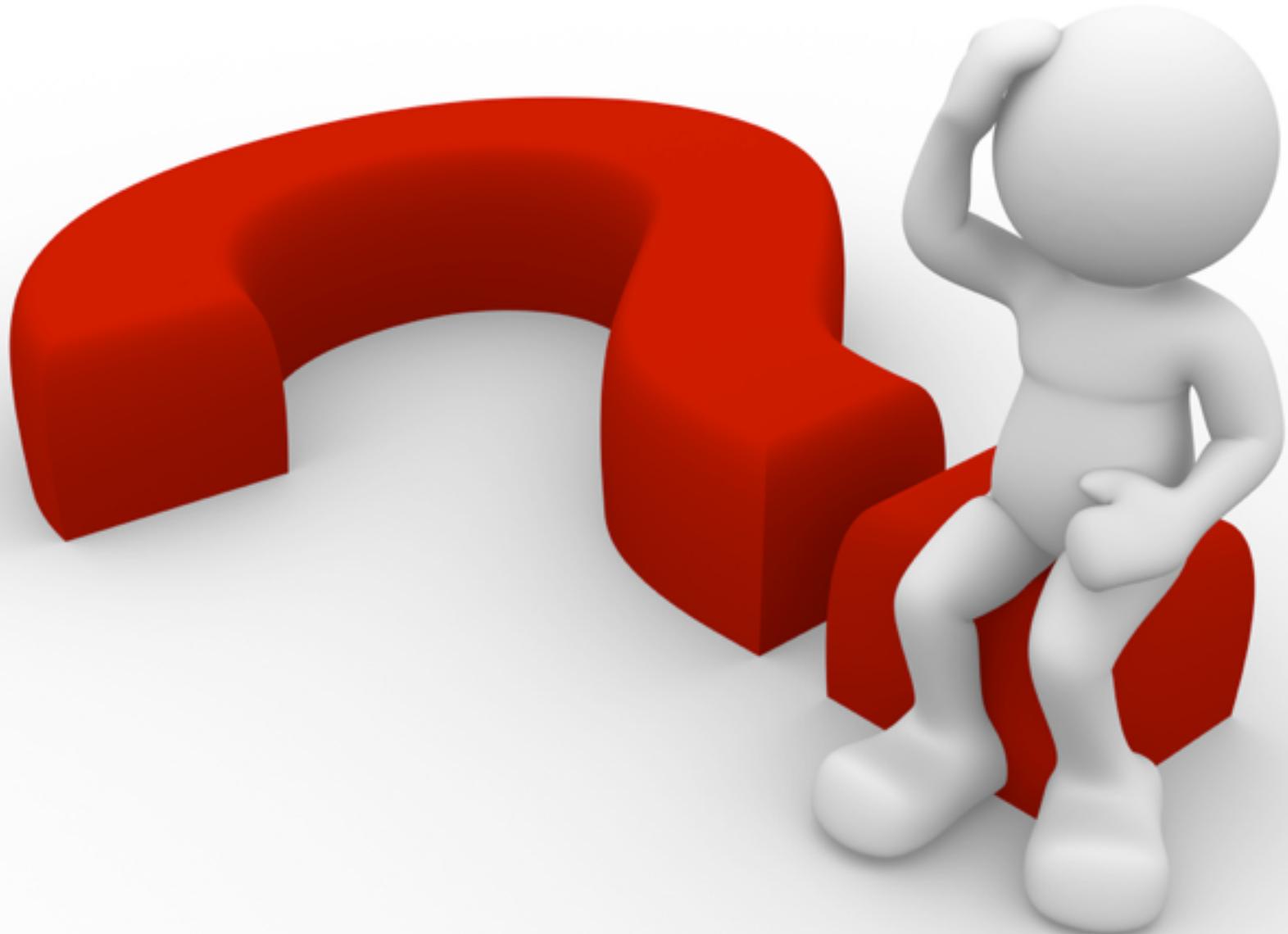
# Fokus Mediendesign

## MIDI

- **Musical Instrument Digital Interface (MIDI)**  
ist Protokoll für Austausch musikalischer Steuerinformationen zwischen elektronischen Instrumenten
- Beispiel:
  1. Byte: MIDI-Kanal 1
  2. Byte: Note C3
  3. Byte: Anschlagstärke 103

# Übertragung

- zur Übertragung von Schallwellen wird analoger (z.B. Magnetband) oder digitaler **Zwischenträger** (z.B. CD) benötigt
- **Zwischenträger** erlaubt Transport ursprünglicher Schallinformation





# Interaktive Medien

## Kapitel Audio

### Psychoakustik

# Psychoakustik

- **Psychoakustik** erfasst Zusammenhänge zwischen physikalischen Größen eines Schallsignals und den dadurch ausgelösten Sinnesempfindungen

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

# Psychophysik

- **Psychophysik** bezieht sich auf Wechselbeziehungen zwischen
  - **subjektivem Erleben** und
  - **quantitativ messbaren Reizen**

# Weber-Fechner-Gesetz

- **Weber-Fechner-Gesetz** beschreibt Zusammenhang zwischen **Reizstärke R** und **Erlebnisintensität E**:

$$E = k \cdot \log R + f$$

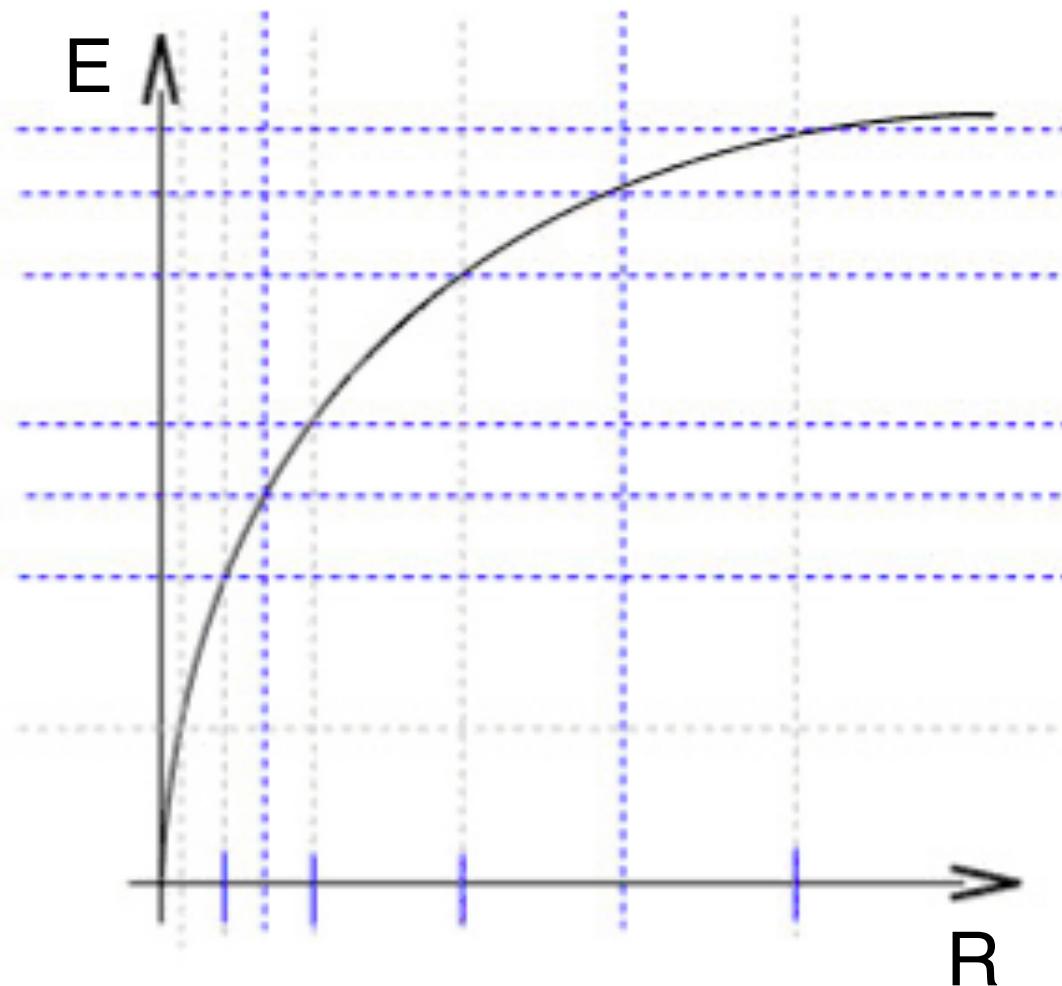
mit Konstanten k und f

# Weber-Fechner-Gesetz

- **Weber-Fechner-Gesetz** besagt, dass subjektiv empfundene Stärke von Sinneseindrücken proportional zum Logarithmus der objektiven Intensität des physikalischen Reizes verhält
- Experimenteller Nachweis durch gerade noch wahrnehmbaren Unterschied (engl. **Just Noticeable Difference**)

# Weber-Fechner-Gesetz

## Beispiel



# Hören

- **Hören** ist Sinneswahrnehmung von Schall, der sich in Form von Wellen in Medium mit **Schallgeschwindigkeit**  $c$  ausbreitet
- **Schwingungen** werden über Umgebungsmedium (Luft, Wasser) oder Untergrund (Vibrationen) übertragen
  - z.B. Feststoffe, Flüssigkeiten oder Gas
- Hören von Geräuschen basiert auf **Vibrationen im Ohr**

# Ohr

## Aufgaben

- **Ohr** ist Sinnesorgan, mit dem Schall aufgenommen wird
- **Ohr** unterteilt sich **Außenohr**, **Mittellohr** und **Innenohr**
- **Ohr** nimmt Änderungen des Luftdrucks wahr, transformiert diese in elektrische Signale, die über Hörnerv an Gehirn weitergeleitet werden

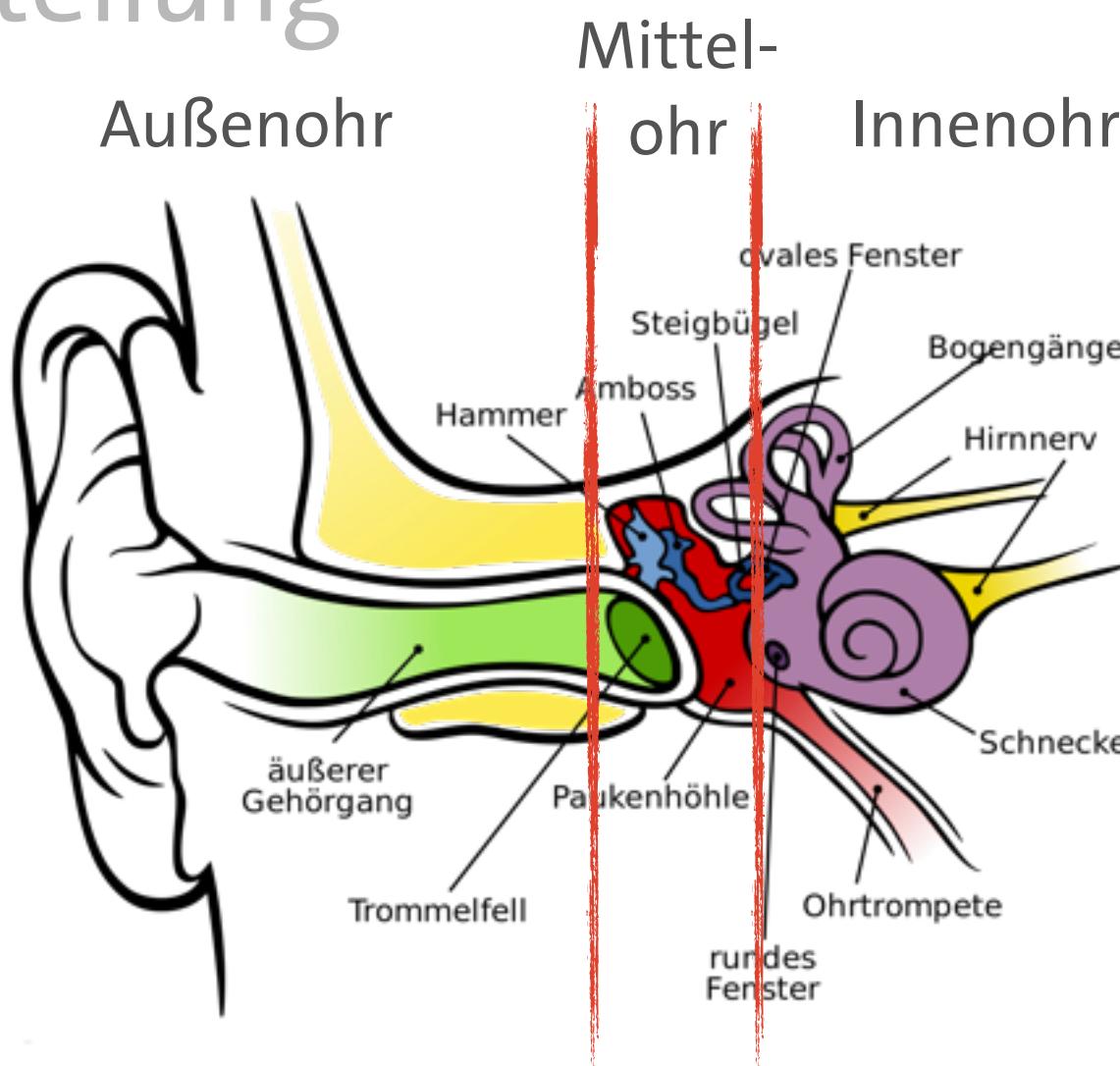
# Ohr

## Unterteilung

1. **Außenohr** beginnt bei Ohrmuschel und geht über Gehörgang zum Trommelfell
2. **Mittelohr** beginnt beim Trommelfell, geht über Gehörknöchelchenkette bis zur Schnecke (**Cochlea**)
3. **Innenohr** ist Schnecke und Hörnerv

# Ohr

## Aufteilung



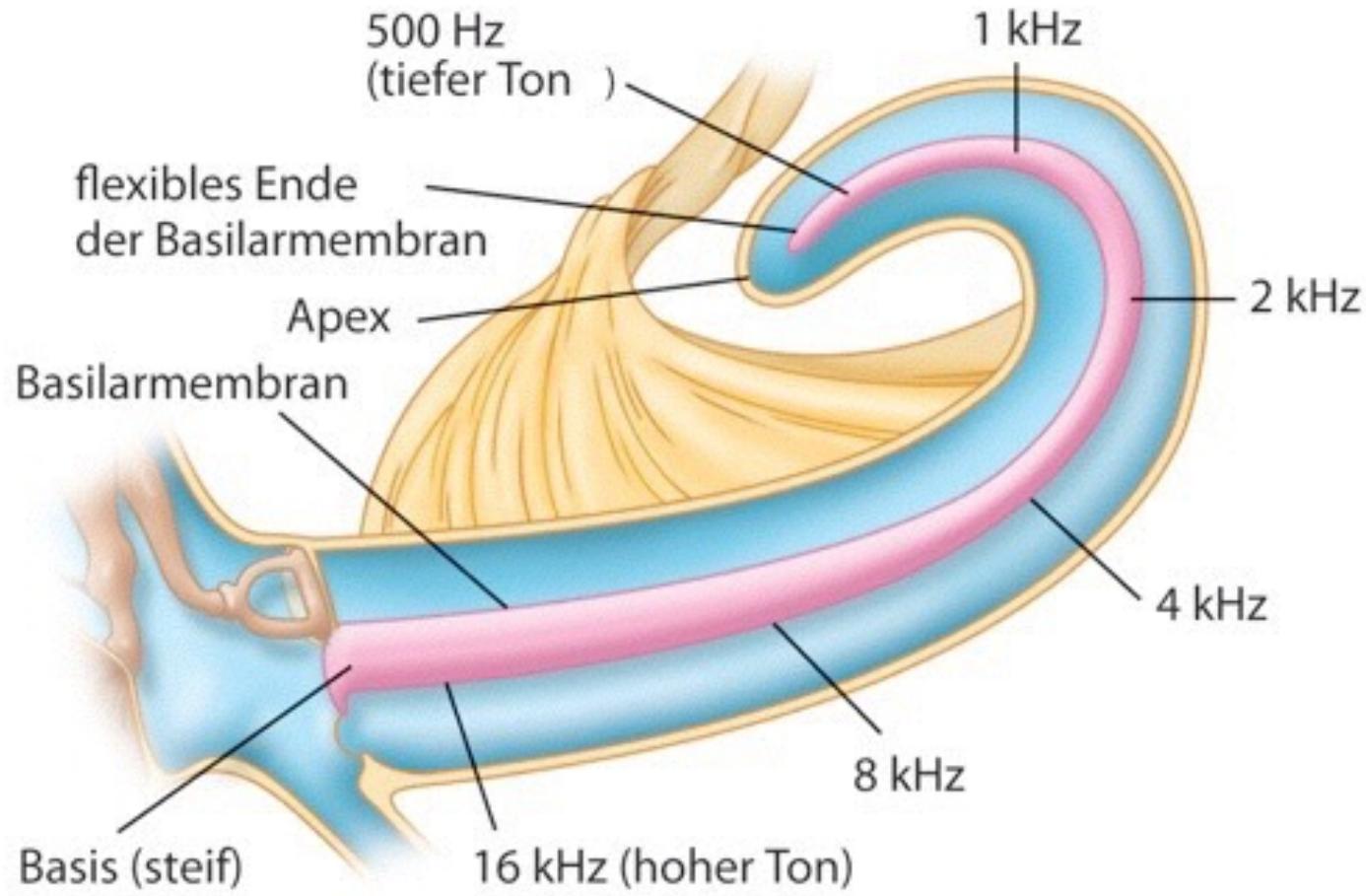
# Außenohr

## Trommelfell

- **Trommelfell** ist dünne (0,1 mm), kegelförmige Hautschicht, die zwischen Gehörgang und Mittelohr liegt
  - **Verdichtung** und **Verdünnung** des Luftdrucks führt dazu, dass Schallwellen im Ohrkanal **Trommelfell** vor und zurück schieben
- Vibration des **Trommelfells**

# Innenohr

## Gehörschnecke





# Zentrales Hören

1. **Transduktion** (Umwandlung von Schallwellen in neuronale Impulse)
2. **Gruppierung** der Klangbilder
3. **Szenenanalyse** und **Extraktion** von auditiven Eigenschaften
4. **Interpretation** der auditiven Umgebung

# Geräuschlokalisierung

- **Lokalisierung** (Bestimmung von Richtung und Distanz) von **Geräuschquellen** wird durch
  - **interaurale Zeitunterschiede**
  - **interaurale Pegelunterschiede**
- **Ursache** für Zeit- und Pegelunterschiede:
  - Ohrenabstand, Kopfabschattung, Form der Ohrmuschel

# Geräuschlokalisierung

## Distanzhinweise

- **Geräuschspektrum:** entfernte Geräusche klingen dumpfer (Beispiel: Gewitter)
- **Lautstärke:** entfernte Geräusche sind leiser
- **Bewegungsparallaxe:** nahe Geräusche bewegen sich schneller am Hörer vorbei als entfernte
- **Reflektionen:** in geschlossenen Räumen erreicht Direktschall Hörer schneller als die Reflexionen; Verhältnis gibt Aufschluss über Distanz

# Geräuschlokalisierung

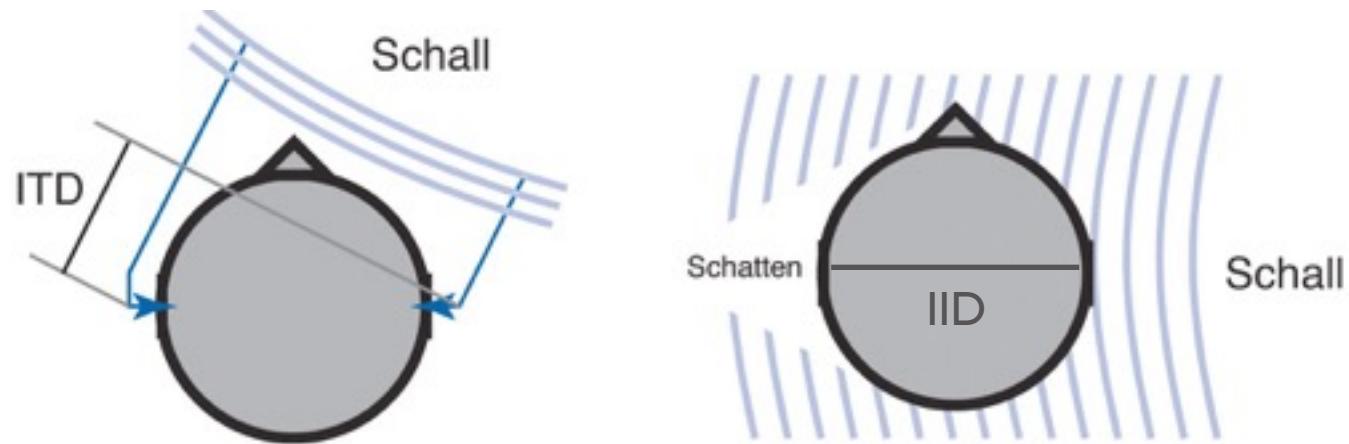
## Bsp: Interaurale Differenzen

- **Interaurale Zeitdifferenz (engl. interaural time difference, ITD)**
  - zeitlicher Unterschied zwischen Hören mit linkem und rechtem Ohr
- **interaurale Intensitätsdifferenz (engl. interaural intensity difference, IID)**
  - Intensitäts-Unterschied zwischen Hören mit linkem und rechtem Ohr

# Interaurale Differenzen

## Head-Related Transfer Functions

- Unterschiede der Signale, welche linkes und rechtes Ohr erreichen, lassen sich durch **Head-related Transfer Functions (HRTFs)** spezifizieren

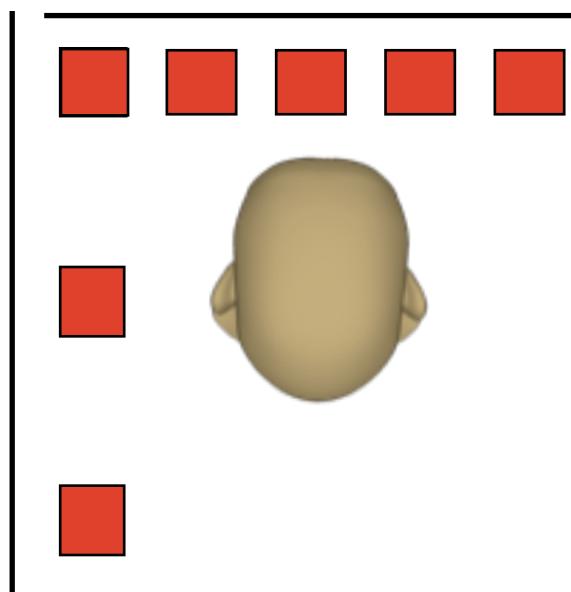


# Geräuschlokalisierung

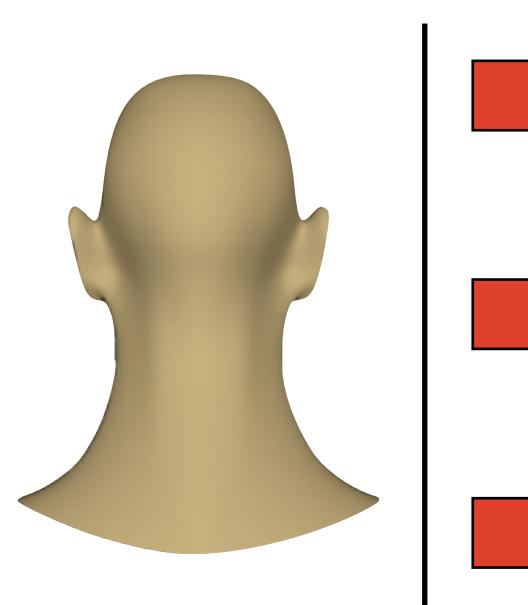
## Beispiel

Horizontal: ca. 5 Quellen

Tiefe: ca. 3 Quellen



Vertikal: ca. 3 Quellen



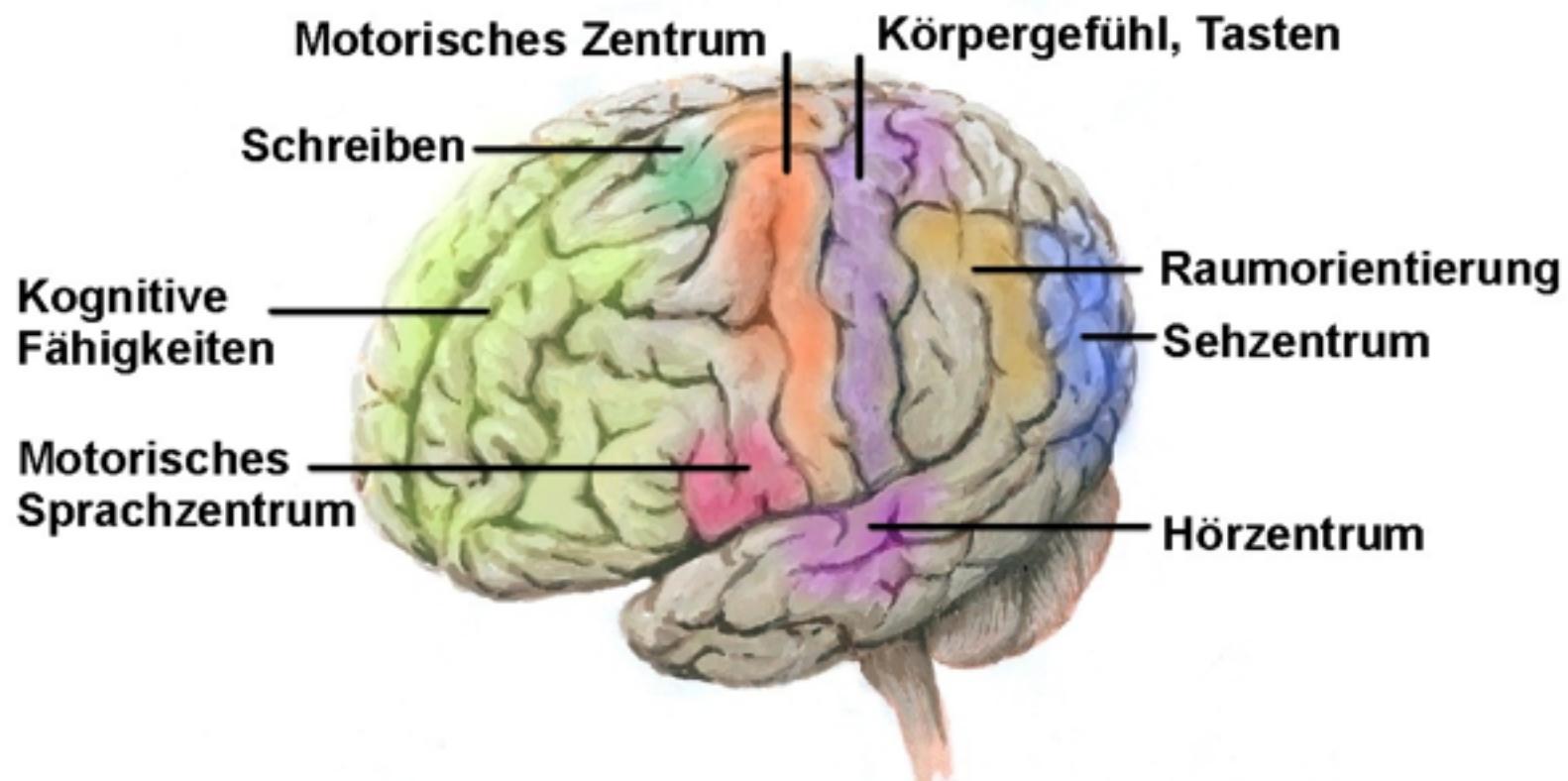
# Hören

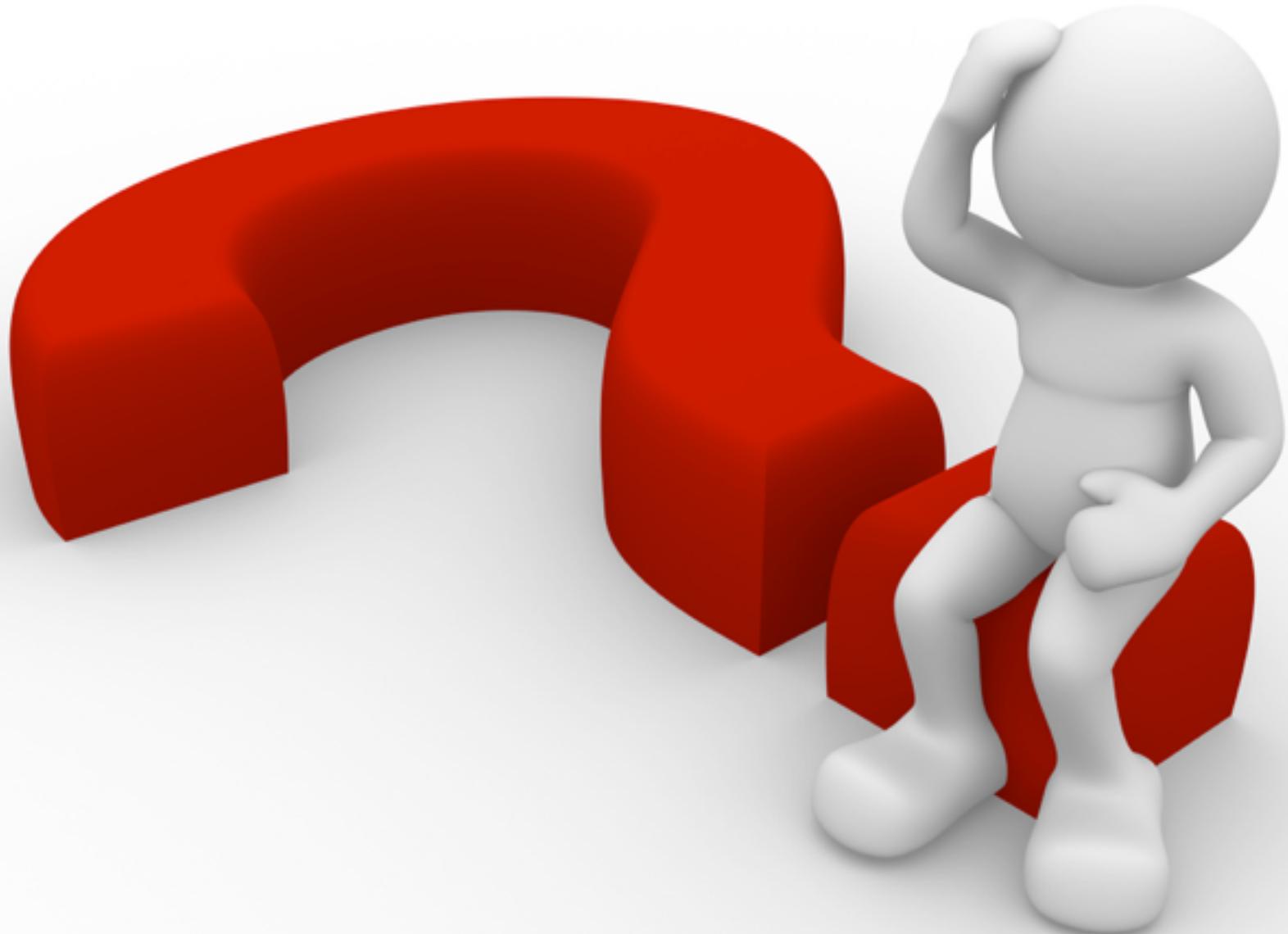
## Untergliederung

- **Periphere Teilstufen:** Außen- und Mittelohr dienen Schallaufnahme und -weiterleitung
- **Zentrale Teilstufen (Zentrales Hören):**
  - Verarbeitung: Vorverarbeitung auditiver Signale in zentraler Hörbahn
  - Wahrnehmung: bewusste Auswertung ankommender Informationen im zentralen auditiven Kortex

# Gehirn

## Auditiver Kortex





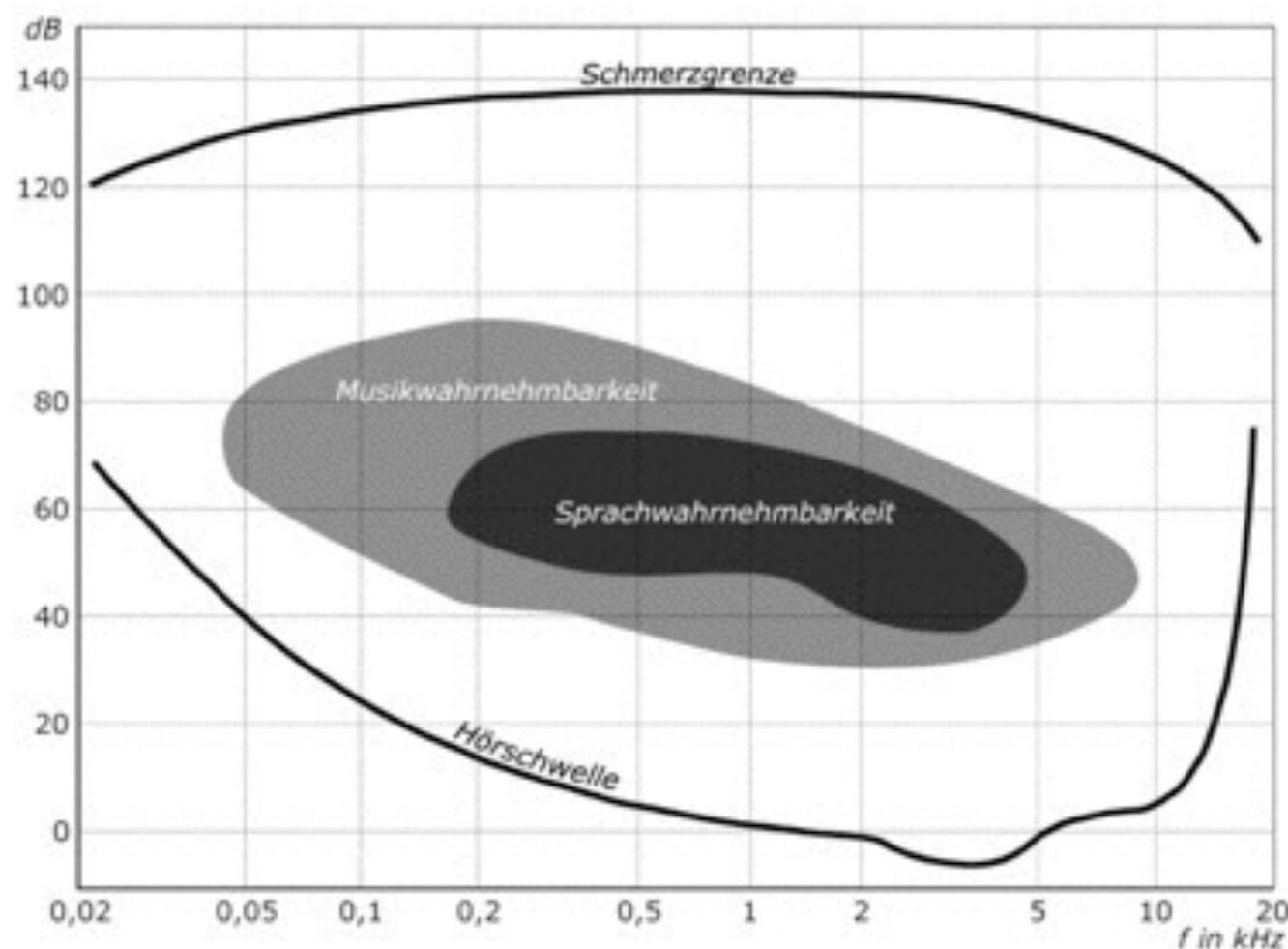


# Interaktive Medien

## Kapitel Audio

### Audiokompression

# Lautstärke Wahrnehmung



# Digitalisierung

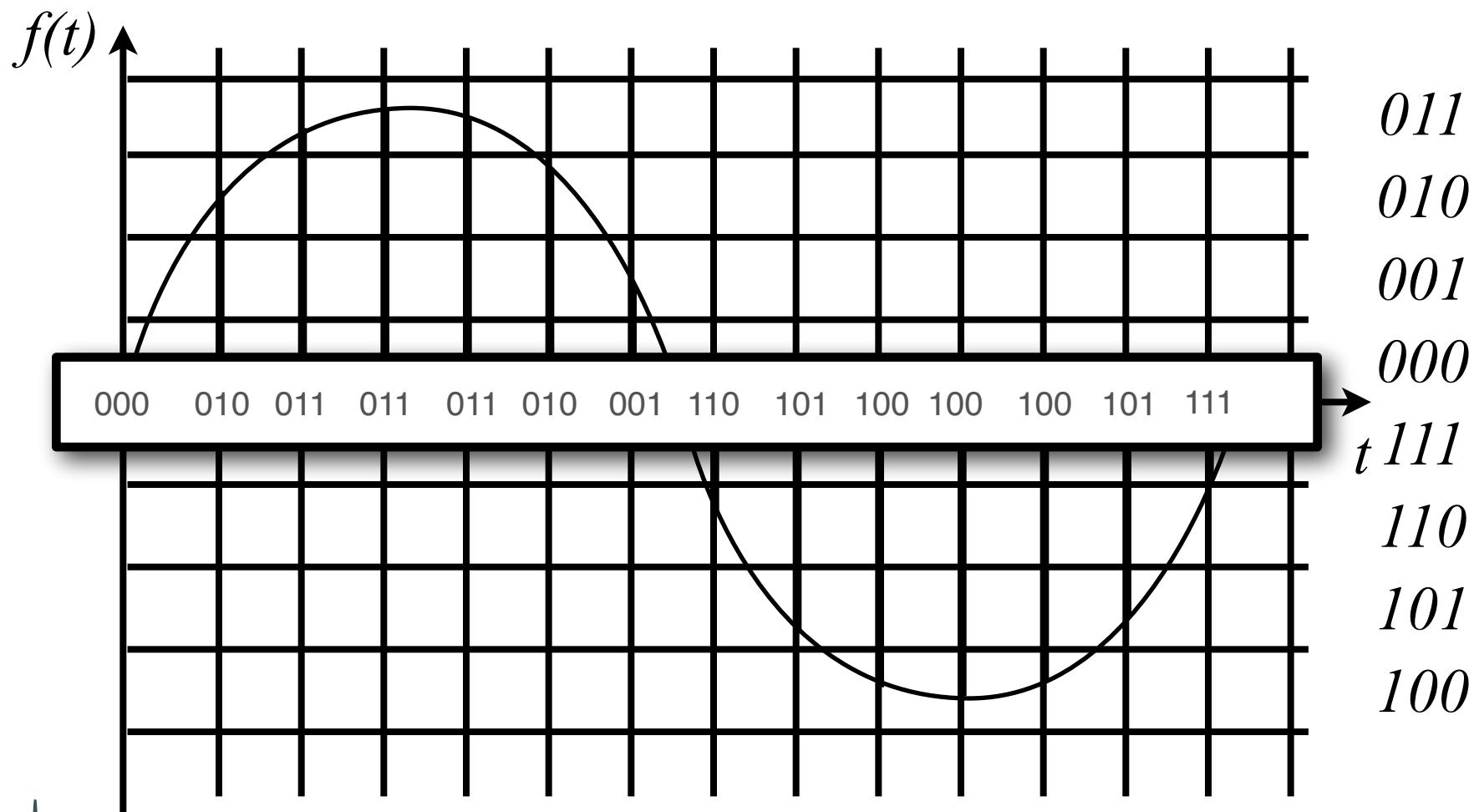
- Beschränkung bei Digitalisierung auf hörbaren Bereich des Menschen
  - ca. 20Hz bis 20kHz
  - ca. 0dB (Hörschwelle) bis 120dB (Schmerzgrenze)
- wegen Abtast-Theorems wird Audio-CD mit 44.100Hz abgetastet werden
  - 8 - 24 Bit Auflösung (bspw. 16 Bit)

# Codierung

- **Pulse Code Modulation (PCM)** ist klassische Vorgehensweise
- **Messwerte** (engl. **Samples**) werden als Bitfolgen hintereinander geschrieben
- Bsp: Standard G.711 für digitale Telefonie (z.B. ISDN oder VoIP)
  - 8 kHz und 8 Bit => Bandbreite eines ISDN-Trägerkanals: 64 kBit/s

# Codierung

## Beispiel: 3-Bit





# Diskussion



Wieviel Minuten Audio (in Stereo) können Sie in etwa auf eine ca. 650MB große Audio-CD (PCM) aufnehmen?

# Diskussion



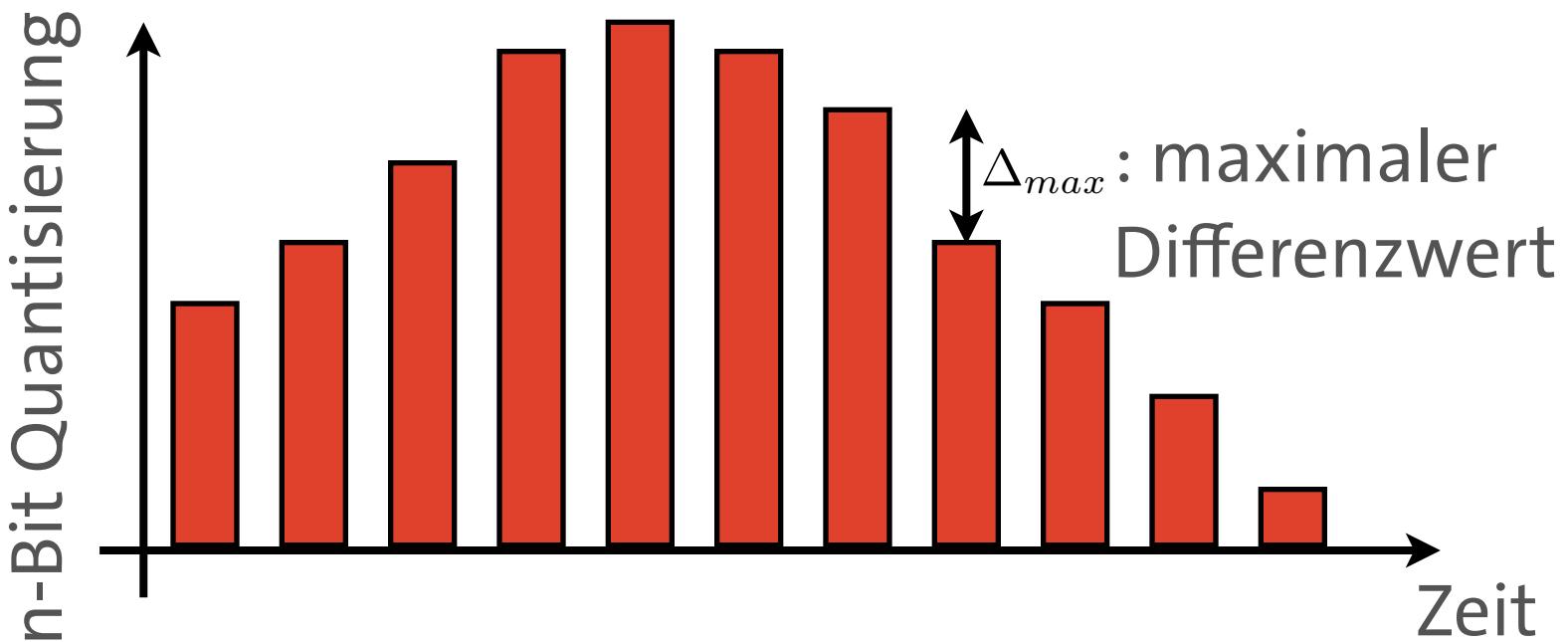
Welche Ideen haben Sie, um die Anzahl der Bits bei PCM zu reduzieren?

# Codierung

- **Differential Pulse Code Modulation (DPCM)**
    - Übertragung der Differenz zwischen zwei Messsignalen statt absoluter Werte
  - **Adaptive Differential Pulse Code Modulation (ADPCM)**
    - Anzahl der für Differenz verwendeten Bits ist variabel
- verlustfreie Codierungen

# Codierung

## DPCM



- Speicherbedarf bei DPCM: notwendige Bits für Startwert + Zeit \* Abtastrate \* notwendige Bits für  $\Delta_{max}$

# Codierung

- **Differential Pulse Code Modulation (DPCM)**
    - Übertragung der Differenz zwischen zwei Messsignalen statt absoluter Werte
  - **Adaptive Differential Pulse Code Modulation (ADPCM)**
    - Anzahl der für Differenz verwendeten Bits ist variabel
- verlustfreie Codierungen

# Audioformate

## (verlustfreie) Beispiele

- WAV: Waveform Audio File
  - entwickelt von Microsoft und IBM
- AIFF: Audio Interchange File Format
  - entwickelt von Apple und Electronic Arts
- CD-DA: Compact Disc Digital Audio
  - entwickelt von Philips und Sony
- FLAC: Free Lossless Audio Codec
  - entwickelt von Xiph.Org Foundation

# Audio-Kompression

- verlustfreie Kompressionsverfahren: begrenzte Reduktion des Speicherbedarfs
- Anwendung **spezieller verlustbehafteter Kompressionsverfahren**
- Beispiele:
  - MP3-Kompression (MPEG-1 Audio Layer III)
  - Advanced Audio Coding (AAC)
  - MPEG-H for spatial audio

# MPEG

- Standardisierungsgremium **Moving Picture Expert Group (MPEG)** arbeitet seit 1988 an Video- und Audiokompression
- MPEG-1 und MPEG-2 beschreiben Formate für digitales Video mit integriertem Audio
- Formate besitzen verschiedene Varianten (**Layer**), die Aufwand und Qualität der Codierung angeben

# MP3-Kompression

- **MPEG-1 Audio Layer III (kurz MP3)** ist Verfahren zur verlustbehafteten Kompression digitaler Audiodaten
- Dateiendung .mp3
- MPEG-Audio-Standards werden auch verwendet bei
  - DAB (Digital Audio Broadcast)
  - DVB (Digital Video Broadcast)

# MPEG-1 Audio Layer

## Beispiele

Verfahren	Kompressionsrate	Anwendungsbereiche
Originaldaten	-	Audio-CD
Layer I (384 kBits/s)	1:4	Digital Compact Cassette (DCC)
Layer II (256...192 kBits/s)	1:6-1:8	Digital Audio Broadcast (DAB), DVD
Layer III (128...112 kBits/s)	1:12	Internet Radio, Satellitenradio



# MPEG-1 Audio Layer III

## Beispiele: Klangqualität

Klangqualität	Modus	Bitrate	Komprimierung
Telefon	Mono	8 kBit/s	1:96
Besser als Kurzwelle	Mono	16 kBit/s	1:48
Besser als Mittelwelle	Mono	32 kBit/s	1:24
UKW	Stereo	56-64kBit/s	1:24
Ähnlich wie CD	Stereo	96 kBit/s	1:16
CD-Qualität	Stereo	128 kBit/s	1:12

# MP3-Kompression

- MP3 basiert auf psychoakustischem (d.h. empirisch ermittelten) Modell der Tonwahrnehmung
- MP3 nutzt **Maskierungseffekte** bezüglich **27 kritischer Bänder**, welche Frequenzbereiche beschreiben, in denen Töne stärker als bestimmter Schwellwert miteinander integrieren

# MPEG-1 Audio Layer III

Beispiel: 128 kBit/s



# MPEG-1 Audio Layer III

## Beispiel: 64 kBit/s



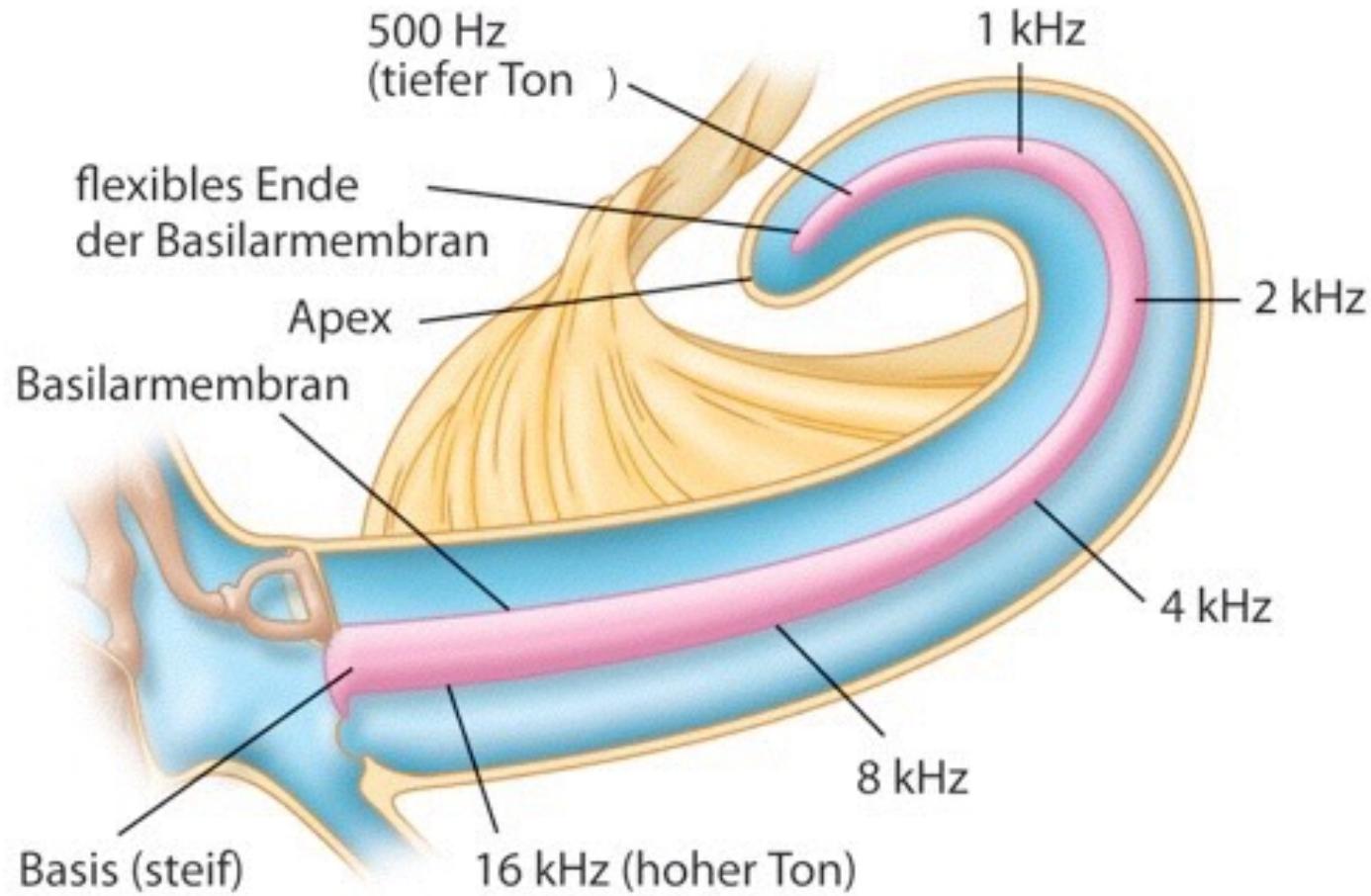
# MPEG-1 Audio Layer III

## Beispiel: 8 kBit/s



# Gehörschnecke

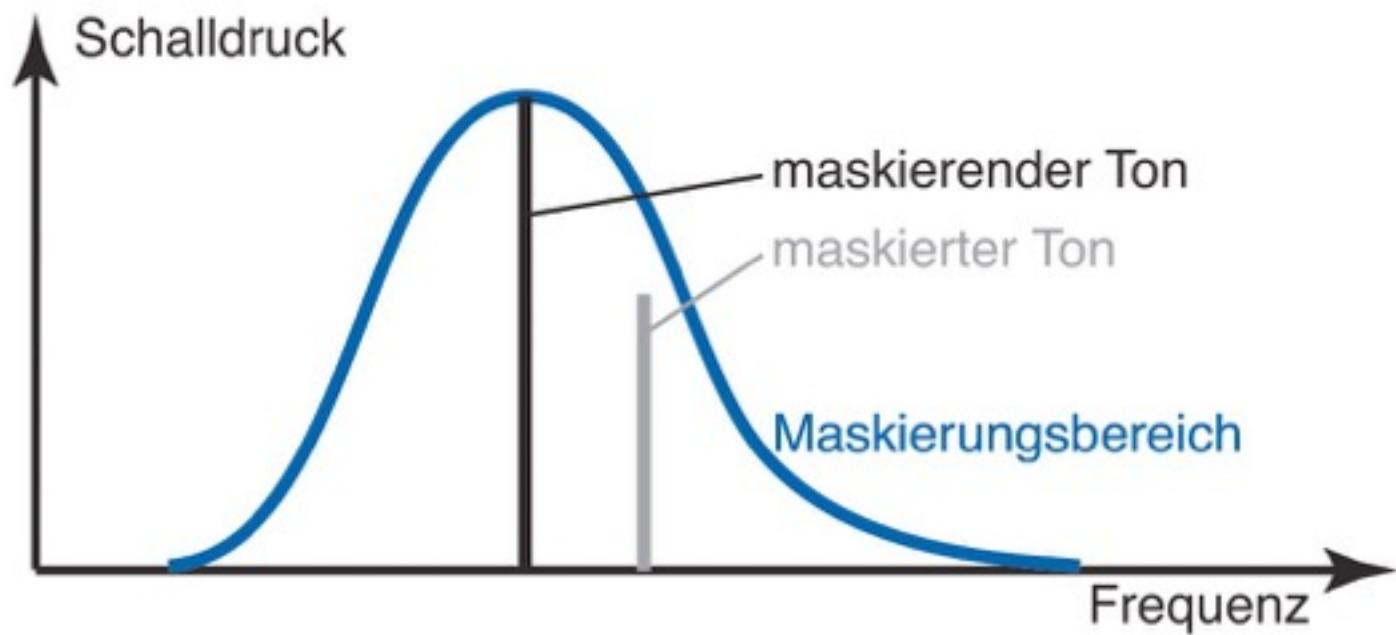
## Sensitivität



# Maskierung

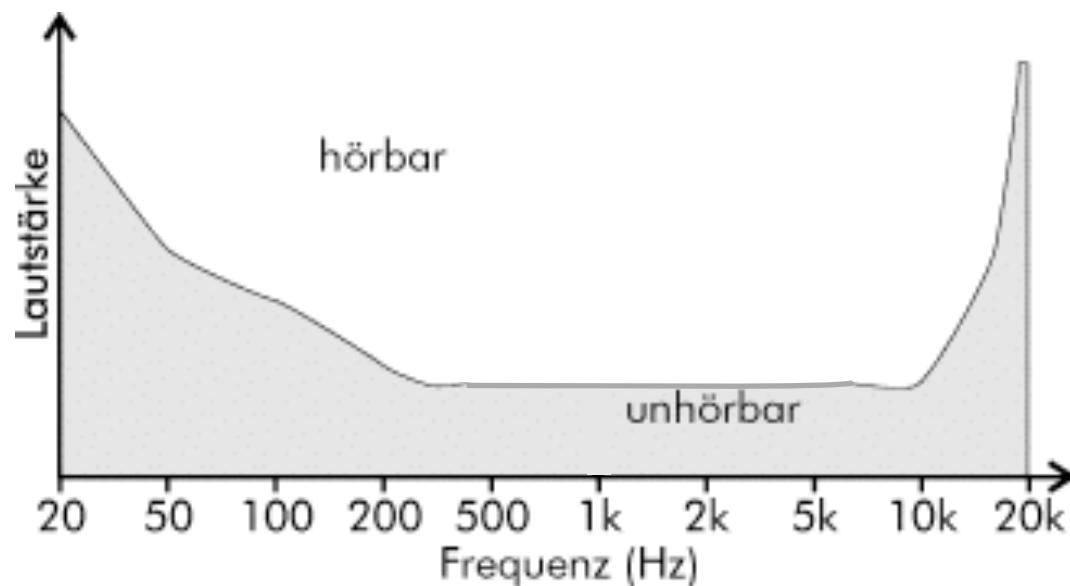
## Simultane Frequenzmaskierung

- **Maskierung** eines Tons durch weiteren Ton tritt auf, falls zweiter Ton ähnlicher Frequenz deutlich lauter ist

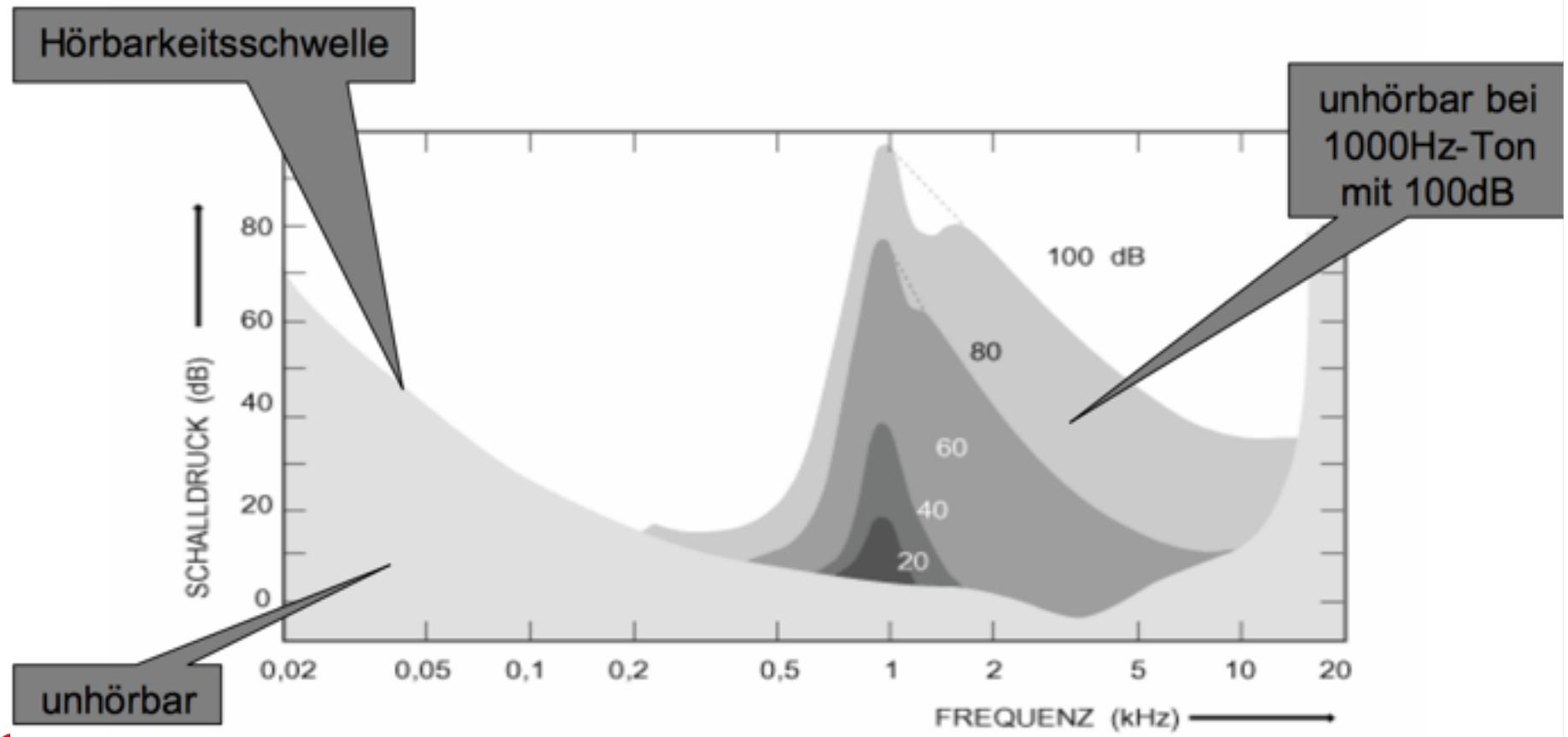


# Modifizierte Hörschwelle

- Präsenz eines Signals modifiziert Hörschwelle, da Gehör in Umgebung dieser Frequenz weniger empfindlich wird



# Modifizierte Hörschwelle



# Maskierungseffekte

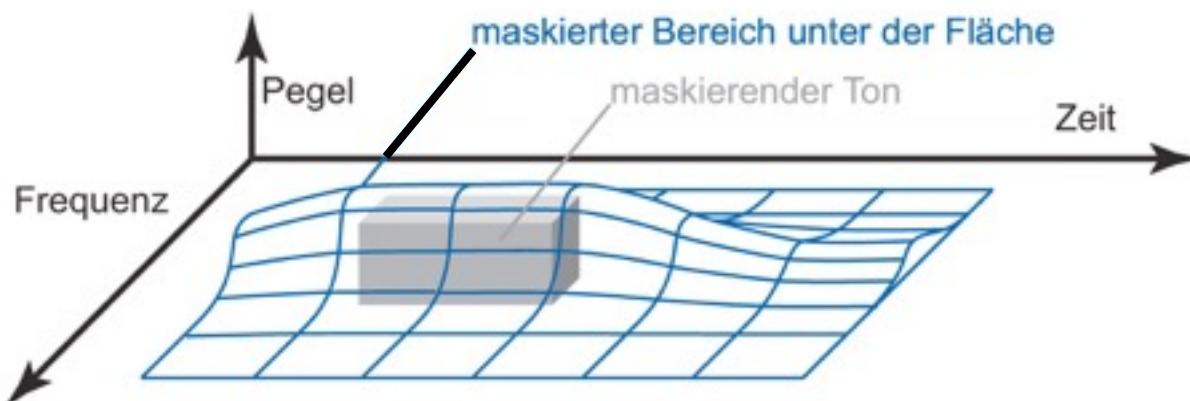
## Vor- & Rückwärtsmaskierung

- Erweiterung von Maskierungseffekten auf Dimension Zeit führt zu weiteren Maskierungseffekten
  - **Vorwärtsmaskierung**
  - **Rückwärtsmaskierung**

# Maskierungseffekte

## Vorwärtsmaskierung

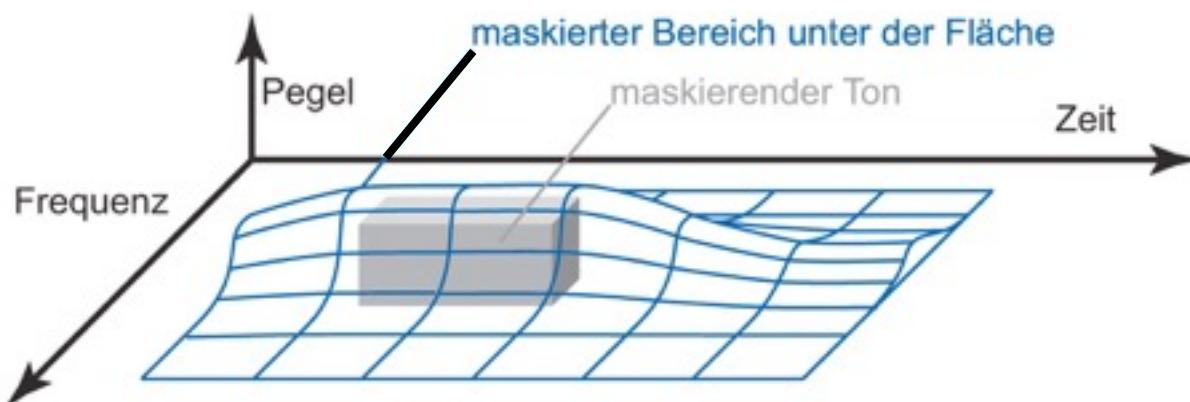
- Tonsignal schwächt Wahrnehmung späterer (bis max. 200 ms) Tonsignale in Umgebung
- Haarzellen benötigen **Erholungszeit** (engl. **Recovery Time**) nach lauterem Signal



# Maskierungseffekte

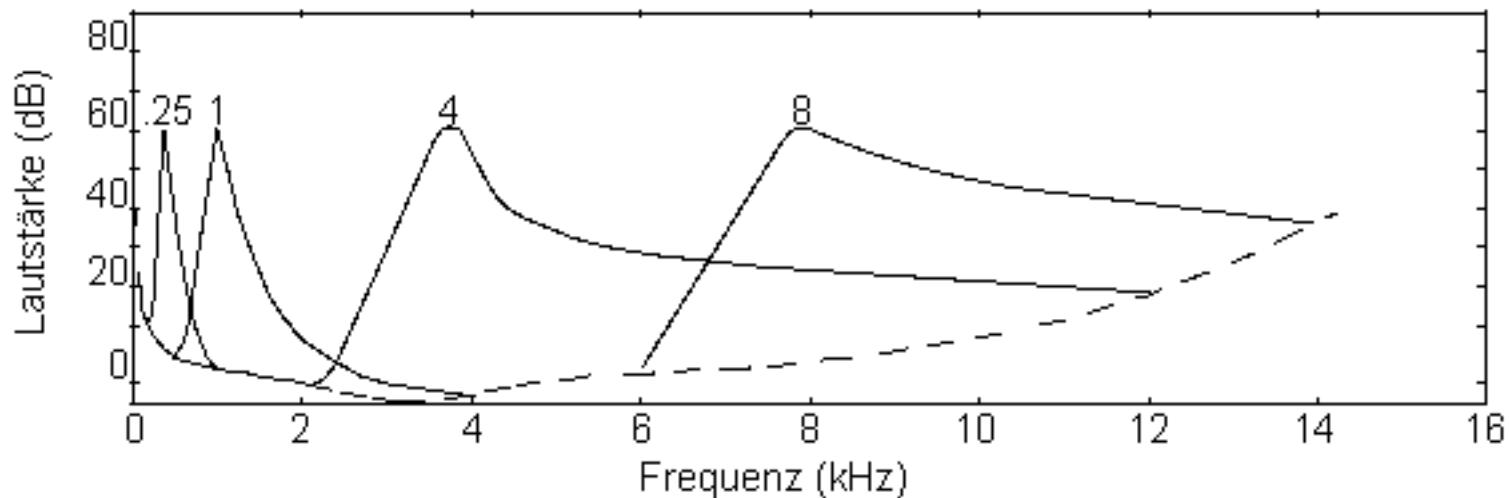
## Rückwärtsmaskierung

- Bei **Rückwärts-Maskierung** maskiert verschleiernder Ton den zuvor (10...20 ms) aufgetretenen Ton
- während dieser Zeit wird ähnlicher, leiserer Ton nicht wahrgenommen



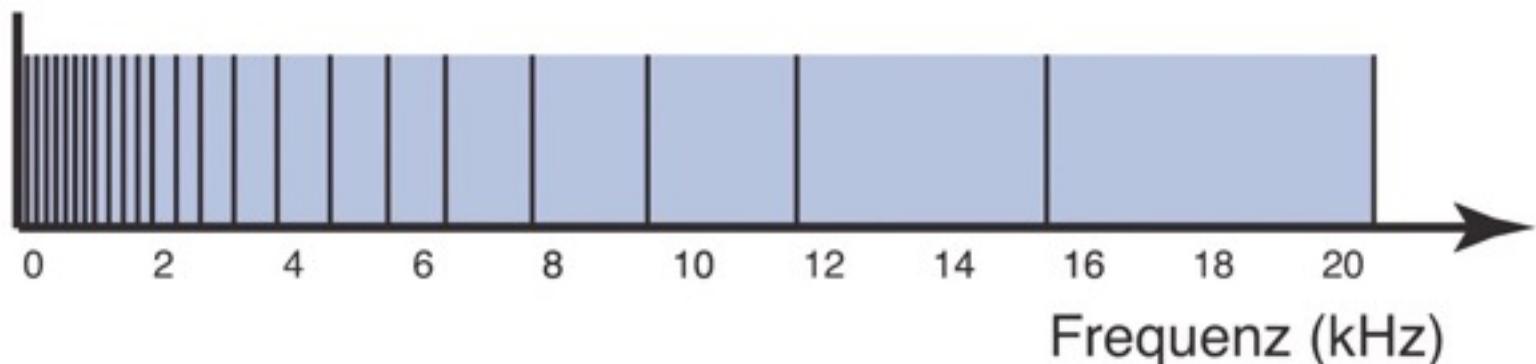
# Kritische Bänder

- **Kritische Bänder** sind Frequenzbereiche, in denen Töne andere Töne stärker als bestimmter Schwellwert beeinflussen (z.B. **Maskierung**)



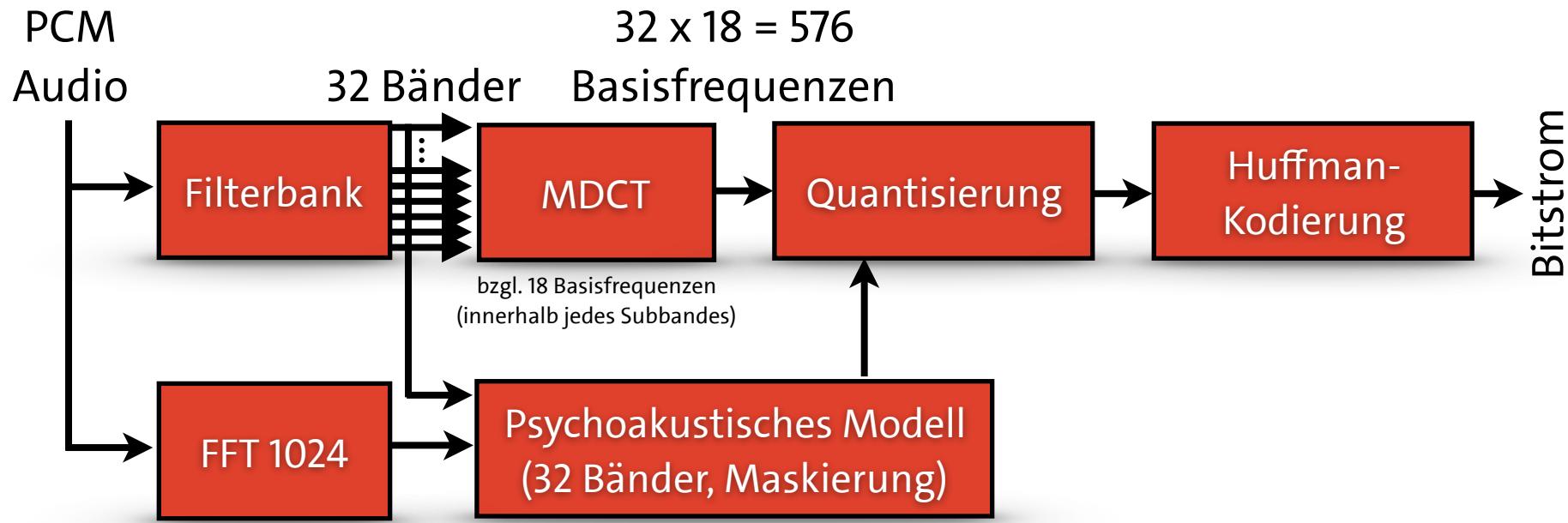
# Kritische Bänder

- **Kritische Bänder** sind Frequenzbereiche, in denen Töne andere Töne stärker als bestimmter Schwellwert beeinflussen (z.B. **Maskierung**)



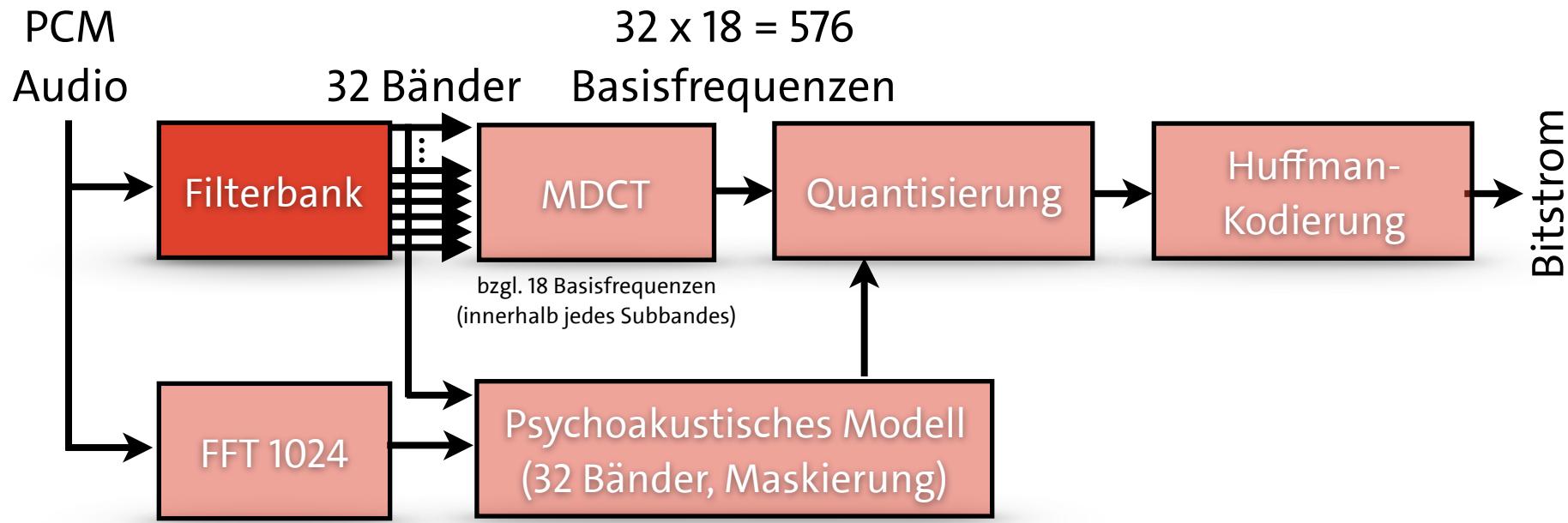
# MP3-Kompression

## Algorithmus



# MP3-Kompression

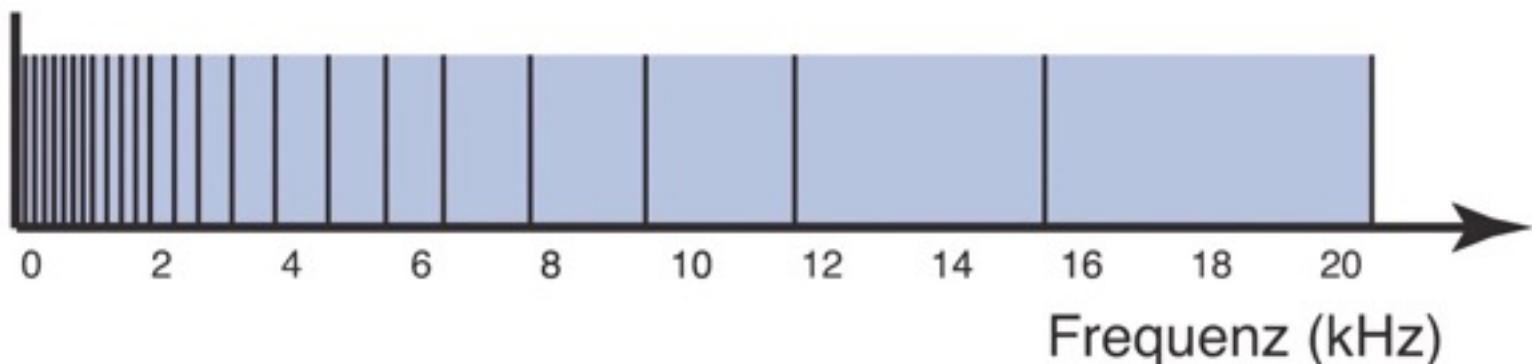
## Algorithmus



# MP3-Kompression

## Filterbank

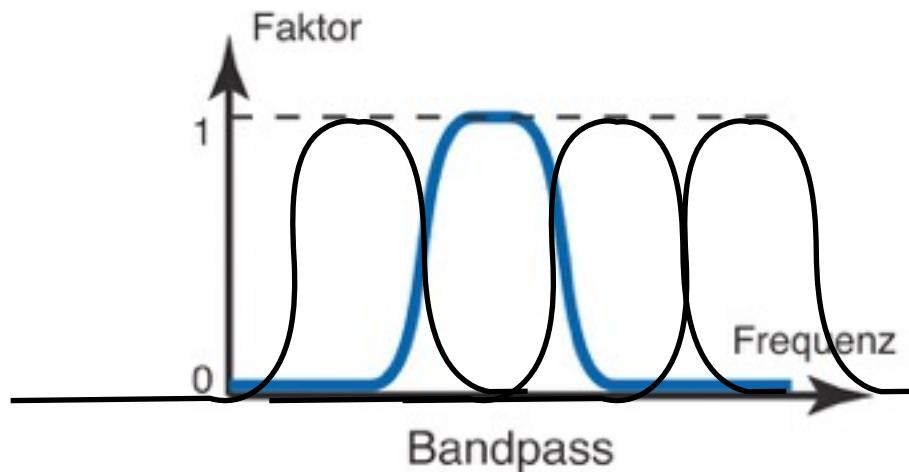
- **Filterbank** unterteilt zu kodierendes Audiosignal in 32 Subbänder, die 27 kritischen Bänder approximieren



# MP3-Kompression

## Bandpass-Filter

- Einordnung in 32 Subbänder durch **Bandpass-Filter**
- bei **Bandpass-Filter** werden Frequenzen zwischen beiden Grenzfrequenzen durchgelassen, andere abgeschwächt



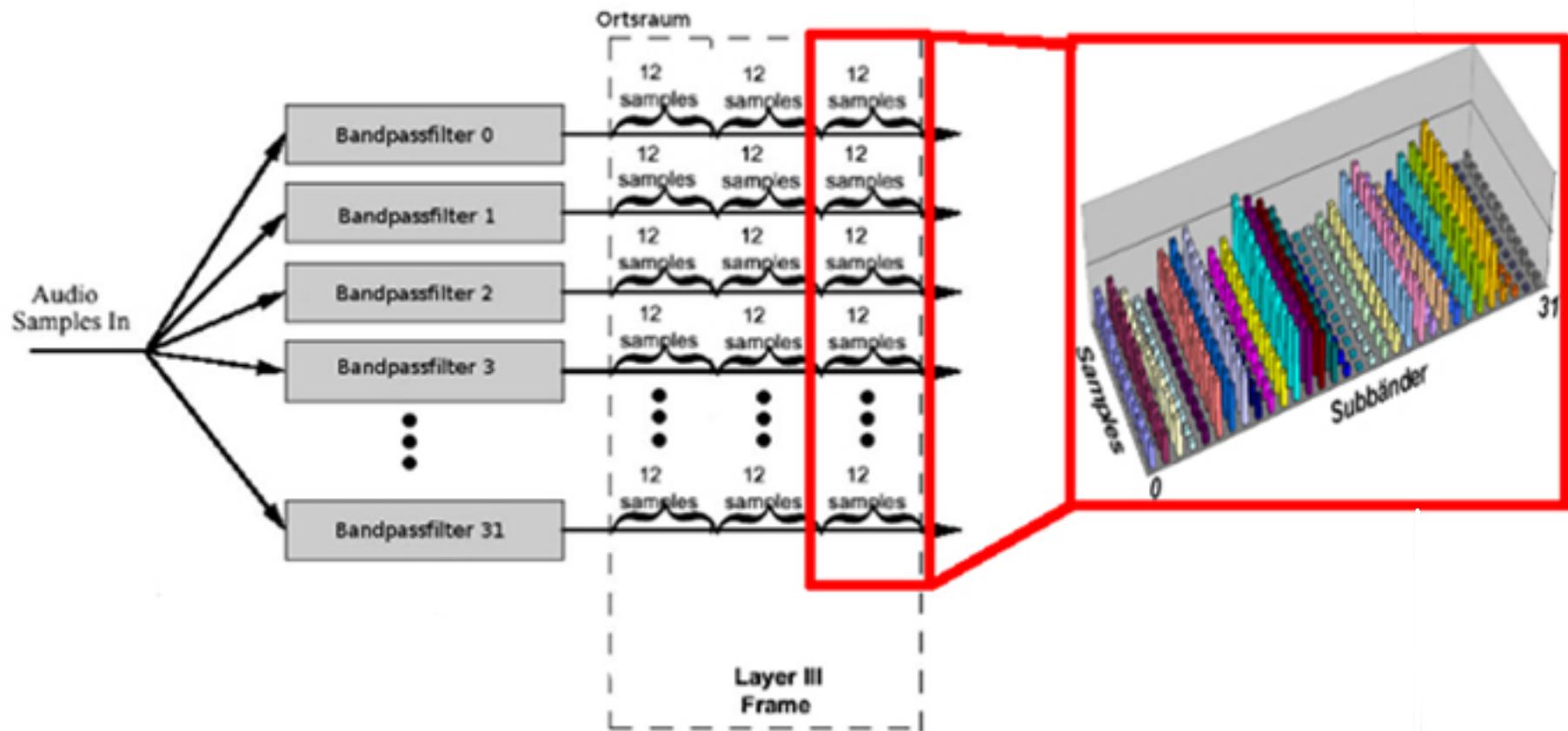
# MP3-Kompression

## Filterbank

- Audiosignal mit Abtastrate 44.100Hz unterteilt in Blöcke von 36 Abtastpunkten (Samples), d.h. insgesamt 1225 Blöcke pro Sekunde
- jeder Block umfasst 36 Samples in allen 32 Subbändern, d.h. insgesamt 1152 Samples pro Block

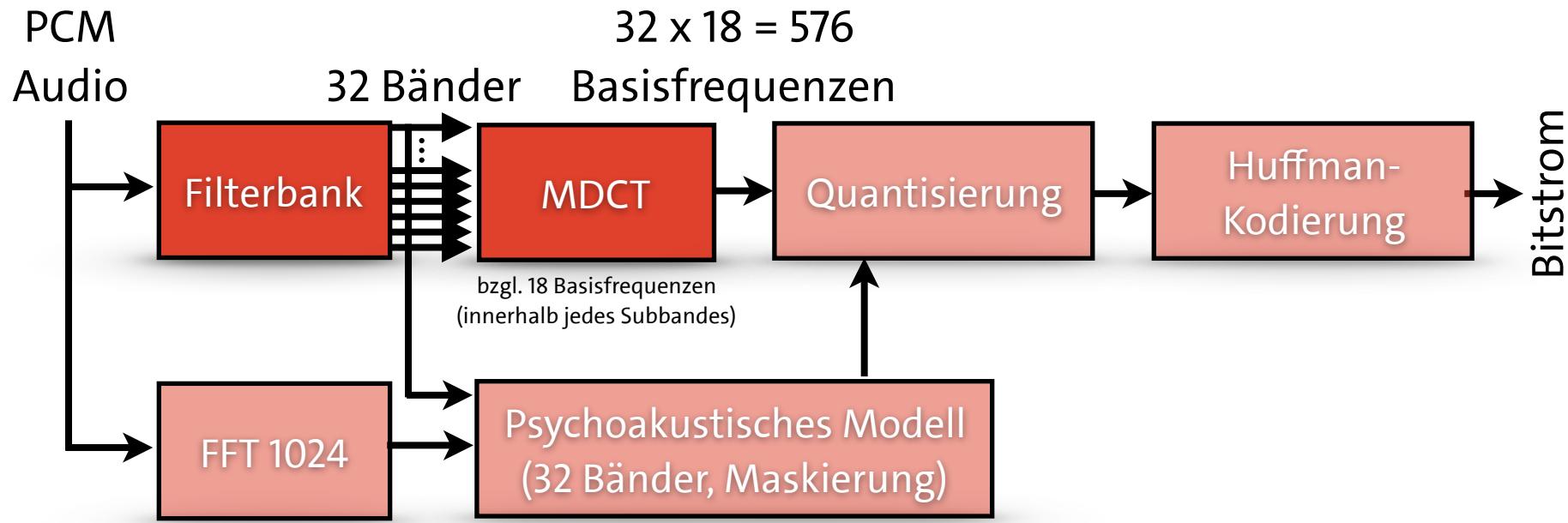
# MP3-Kompression

## Filterbank



# MP3-Kompression

## Algorithmus

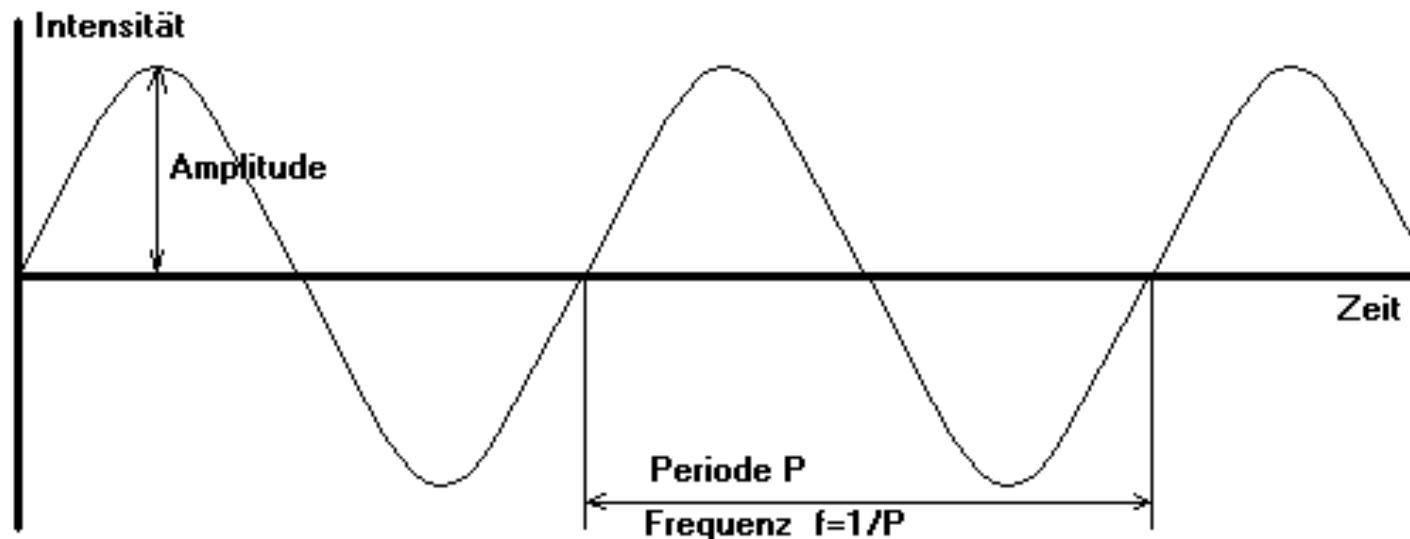


# Frequenzbereich

- Betrachtung der im Audiosignal vorkommenden Frequenzen
- Umcodierung der Signalwerte vom **Zeitbereich in Frequenzbereich**
- Betrachtung für alle Subbänder getrennt

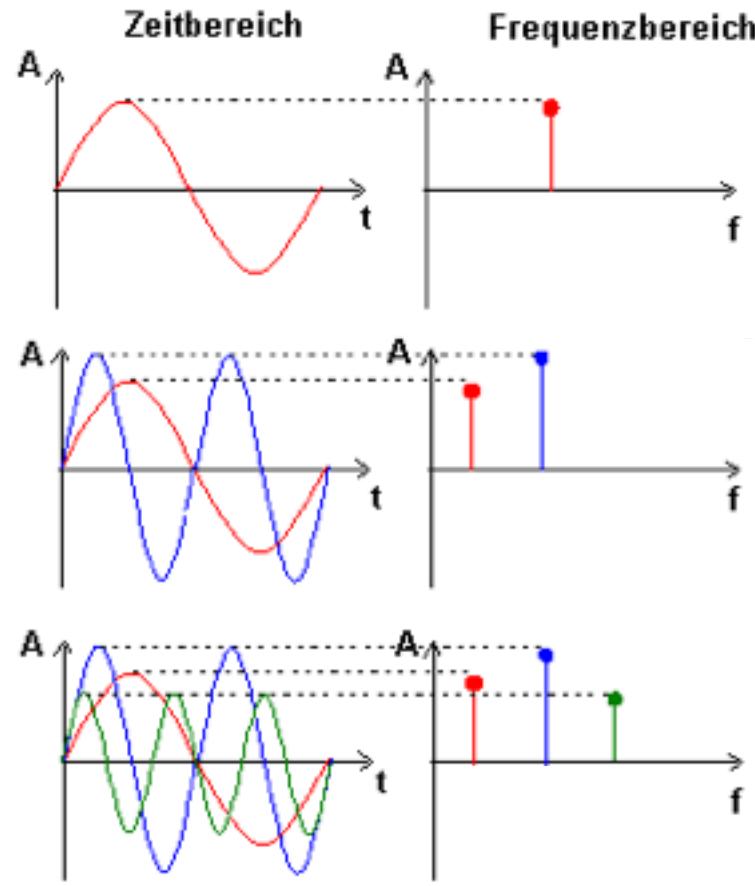
# Frequenzbereich

## Audiosignal



# Frequenzbereich

## Frequenz



# Fouriertransformation

## Fourier-Reihe

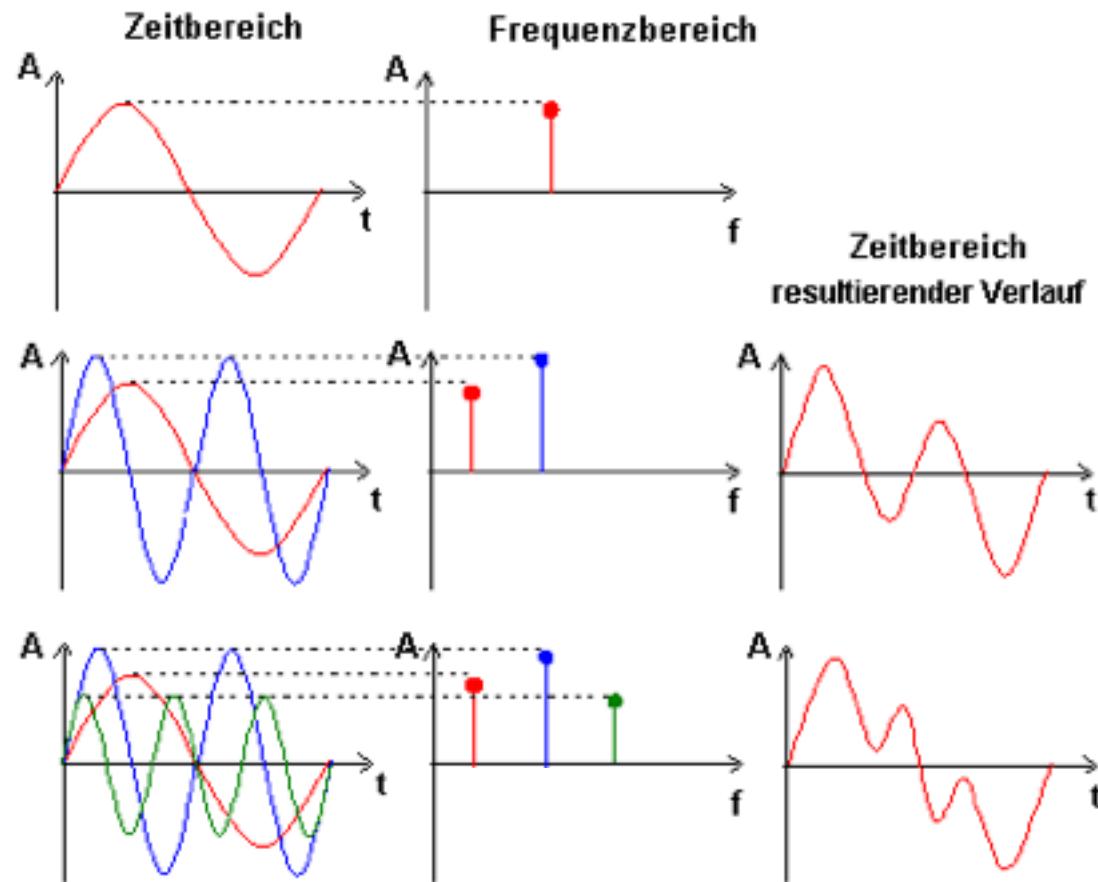
- Jede periodische Schwingung lässt sich durch unendliche Summe von sich überlagerten Kosinus-Schwingungen annähern

$$x(t) = \sum_{k=0}^{\infty} a_k \cdot \cos(k \cdot \omega_0 \cdot t + \theta_k)$$

- Grundfrequenz  $\omega_0$
- Anteil k-ten harmonische Schwingung  $a_k$
- Phasenverschiebung der k-ten harmonischen Schwingung  $\theta_k$

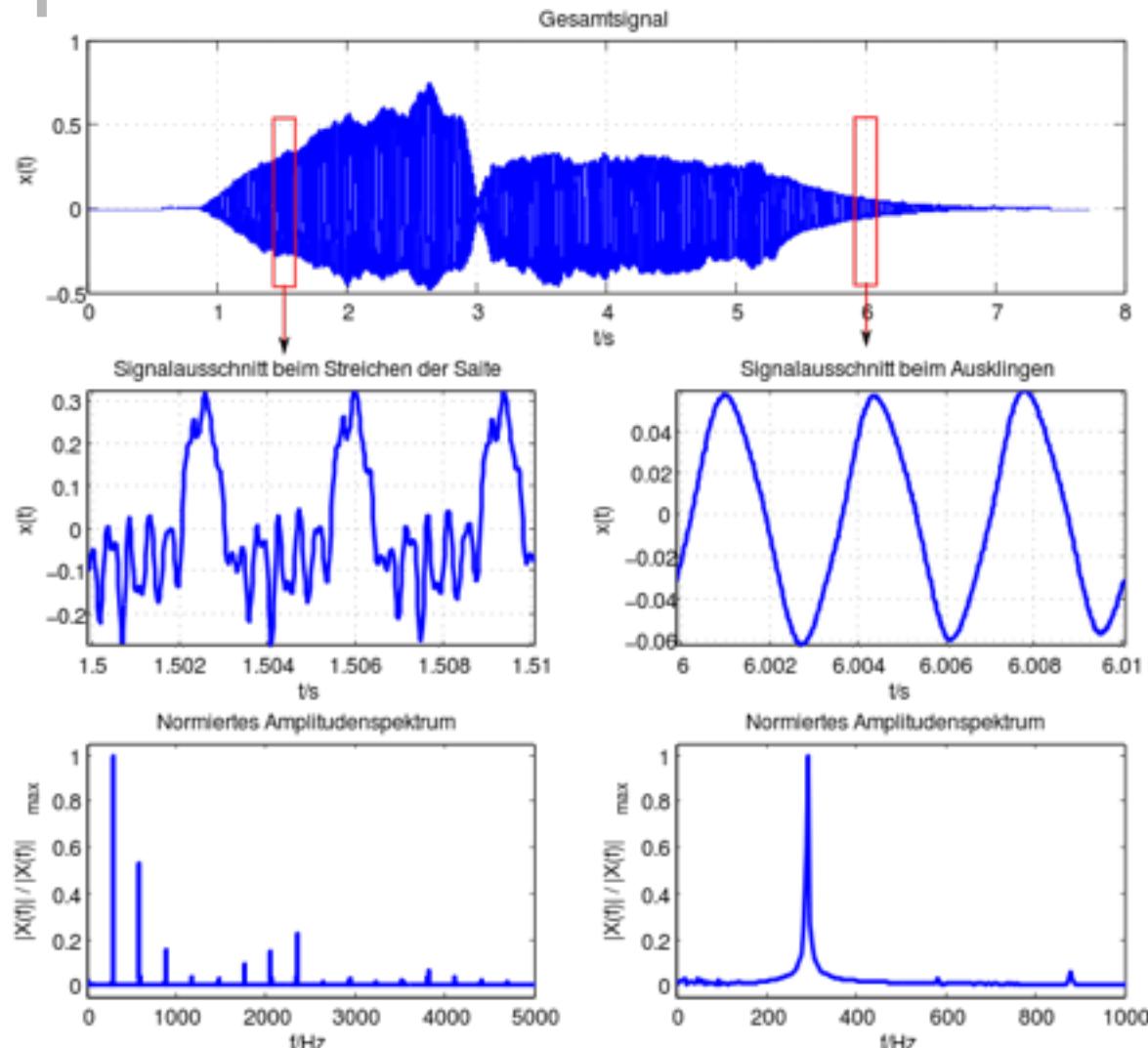
# Frequenzbereich

## Frequenz



# Frequenzspektrum

## Beispiel



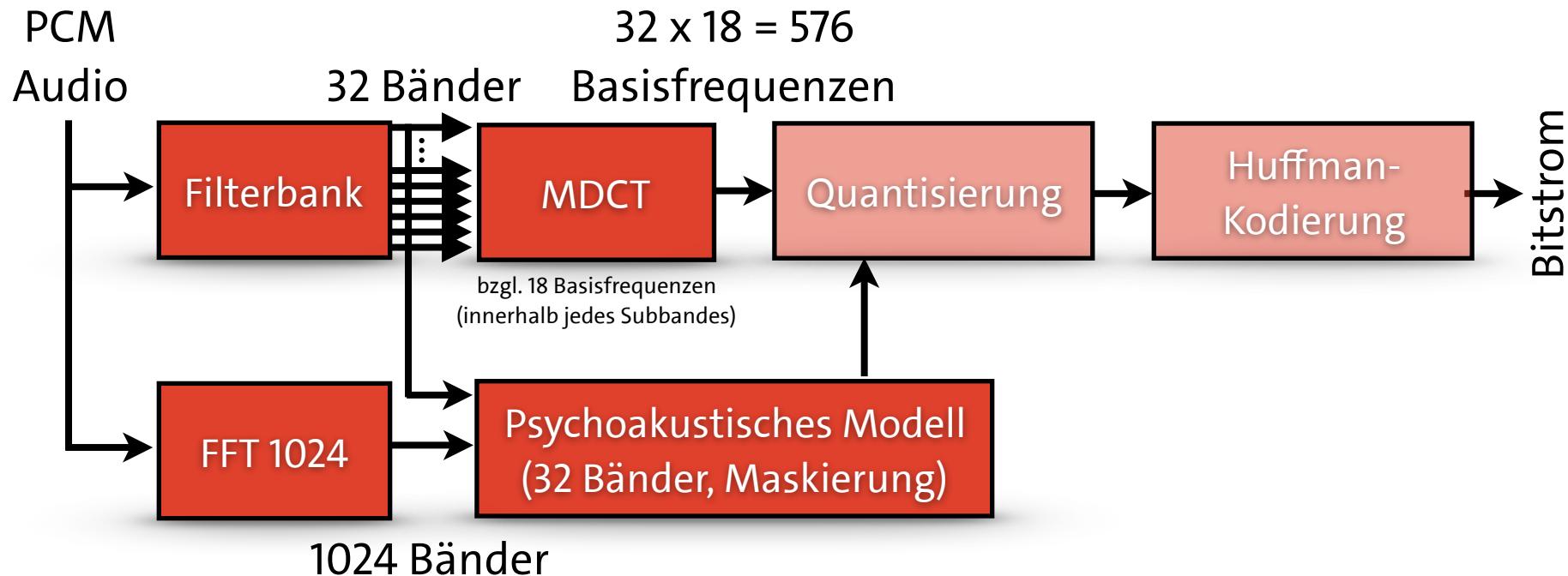
# MP3-Kompression

## MDCT

- 32 Subbänder werden jeweils durch **modifizierte diskrete Kosinus-Transformation (MDCT)** durch 18 **Basisfrequenzen** im jeweiligen Subband dargestellt (vgl. Basismuster bei JPEG)
- $32 \cdot 18 = 576$  **Koeffizienten** der **Basisfrequenzen**

# MP3-Kompression

## Algorithmus



# MP3-Kompression

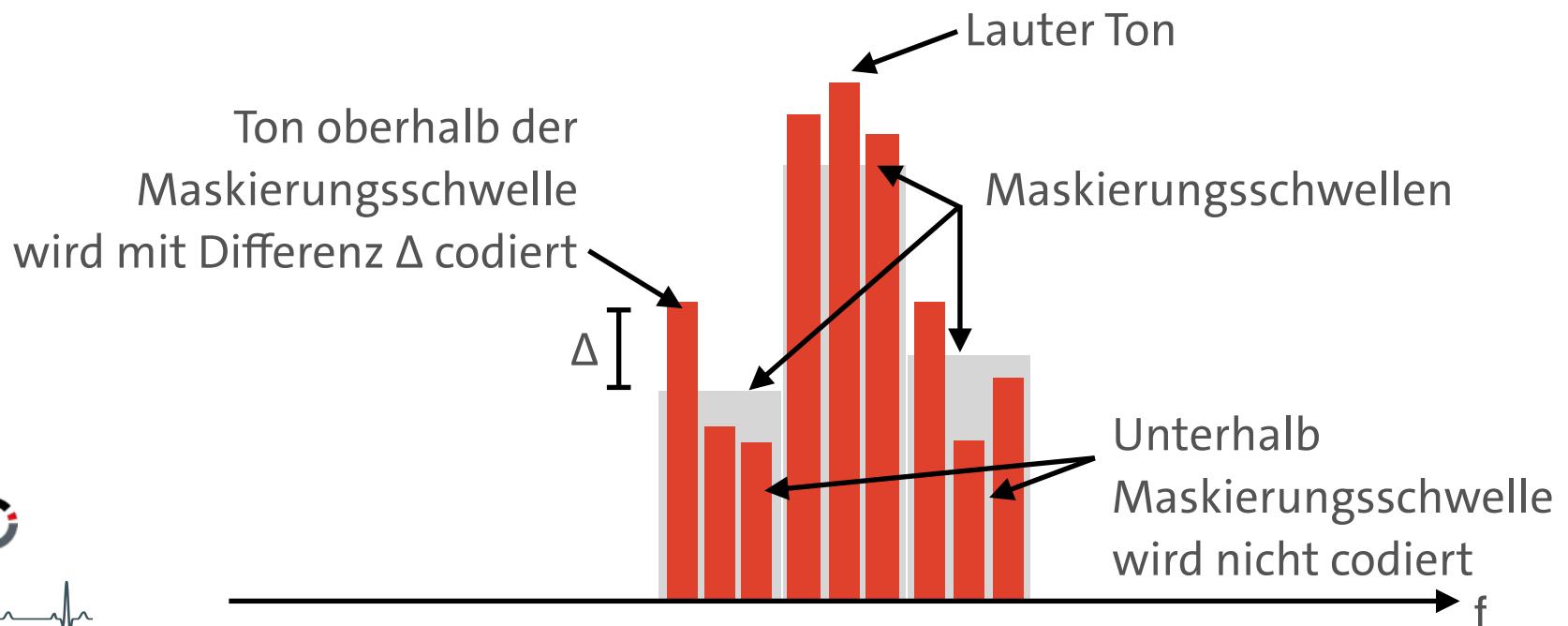
## Maskierung

- **Fast Fourier Transformation (FFT)** zerlegt Frame (12 Samples) in Frequenzraum-darstellung bzgl. 1024 Bänder
- **Psychoakustisches Modell** wird genutzt, um Maskierungsschwellen für Subbänder zu bestimmen
- irrelevante Signalanteile unter Maskierungsschwellen werden mit weniger Bits (oder gar nicht) zu speichern

# MP3-Kompression

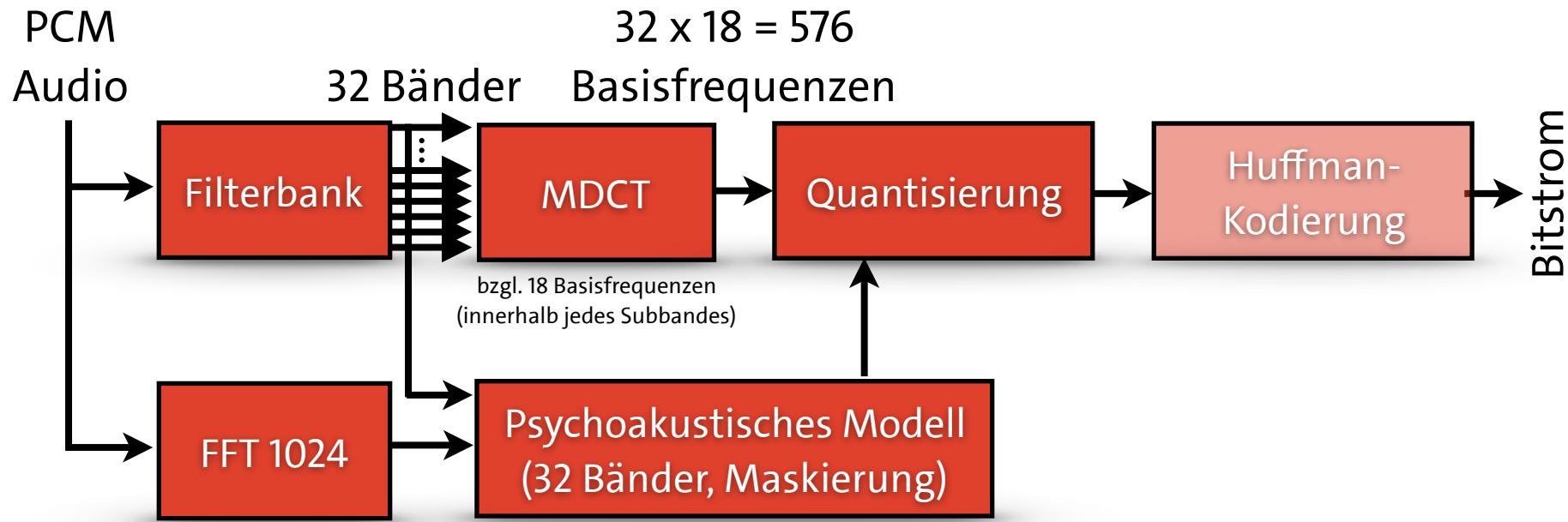
## Quantisierung

- **Maskierungsschwellen** aus psycho-akustischem Modell werden mit tatsächlichen Signalpegel (pro Teilband) verglichen



# MP3-Kompression

## Algorithmus



# Huffman-Kodierung

- nach Quantisierung sind viele Koeffizienten irrelevanter Frequenzanteile durch Nullen repräsentiert
- Koeffizienten werden mit Huffman-Codierung kodiert
- Differenzbildung bei Stereosignal (ähnlich wie bei DPCM bzw. ADPCM)

# MPEG-1 Audio Layer III

## Beispiele: Bitrate

Klangqualität	Modus	Bitrate	Komprimierung
Telefon	Mono	8 kBit/s	1:96
Besser als Kurzwelle	Mono	16 kBit/s	1:48
Besser als Mittelwelle	Mono	32kBit/s	1:24
UKW	Stereo	56-64kBit/	1:24
Ähnlich wie CD	Stereo	96kBit/s	1:16
CD-Qualität	Stereo	128kBit/s	1:12



"I was ready to fine-tune my compression algorithm... somewhere down the corridor, a radio was playing 'Tom's Diner.' I was electrified. I knew it would be nearly impossible to compress this warm a cappella voice."

—Karlheinz Brandenburg

# MPEG-1 Audio Layer III

Beispiel: 128 kBit/s



# MPEG-1 Audio Layer III

## Beispiel: 56 kBit/s



# MPEG-1 Audio Layer III

## Beispiel: 32 kBit/s



# MPEG-1 Audio Layer III

## Beispiel: 16 kBit/s



# MPEG-1 Audio Layer III

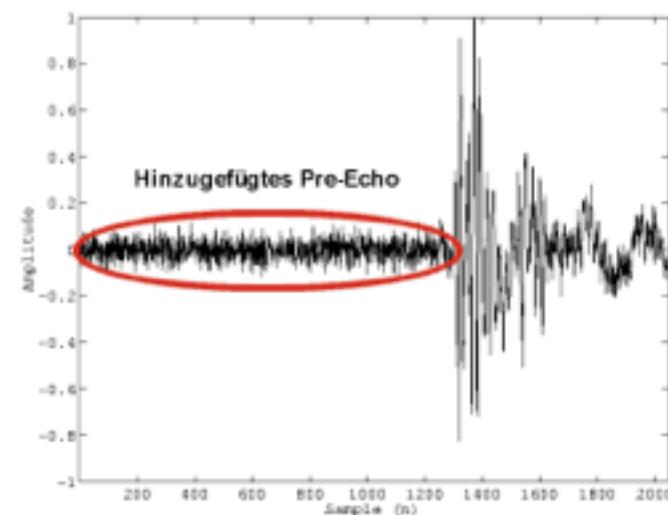
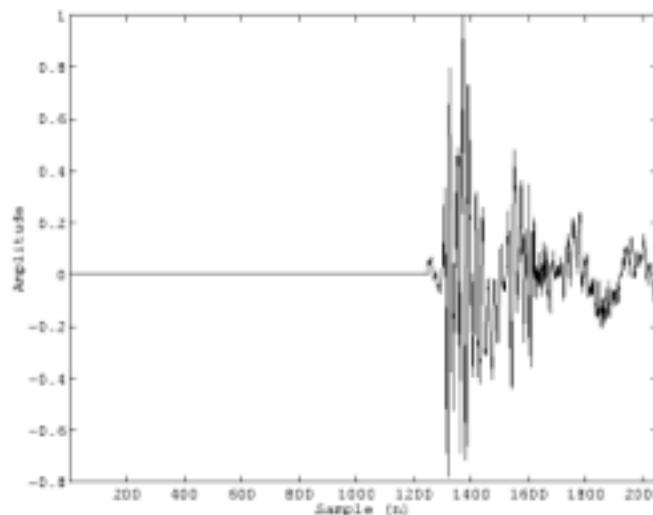
## Beispiel: 8 kBit/s



# MPEG-1 Audio Layer III

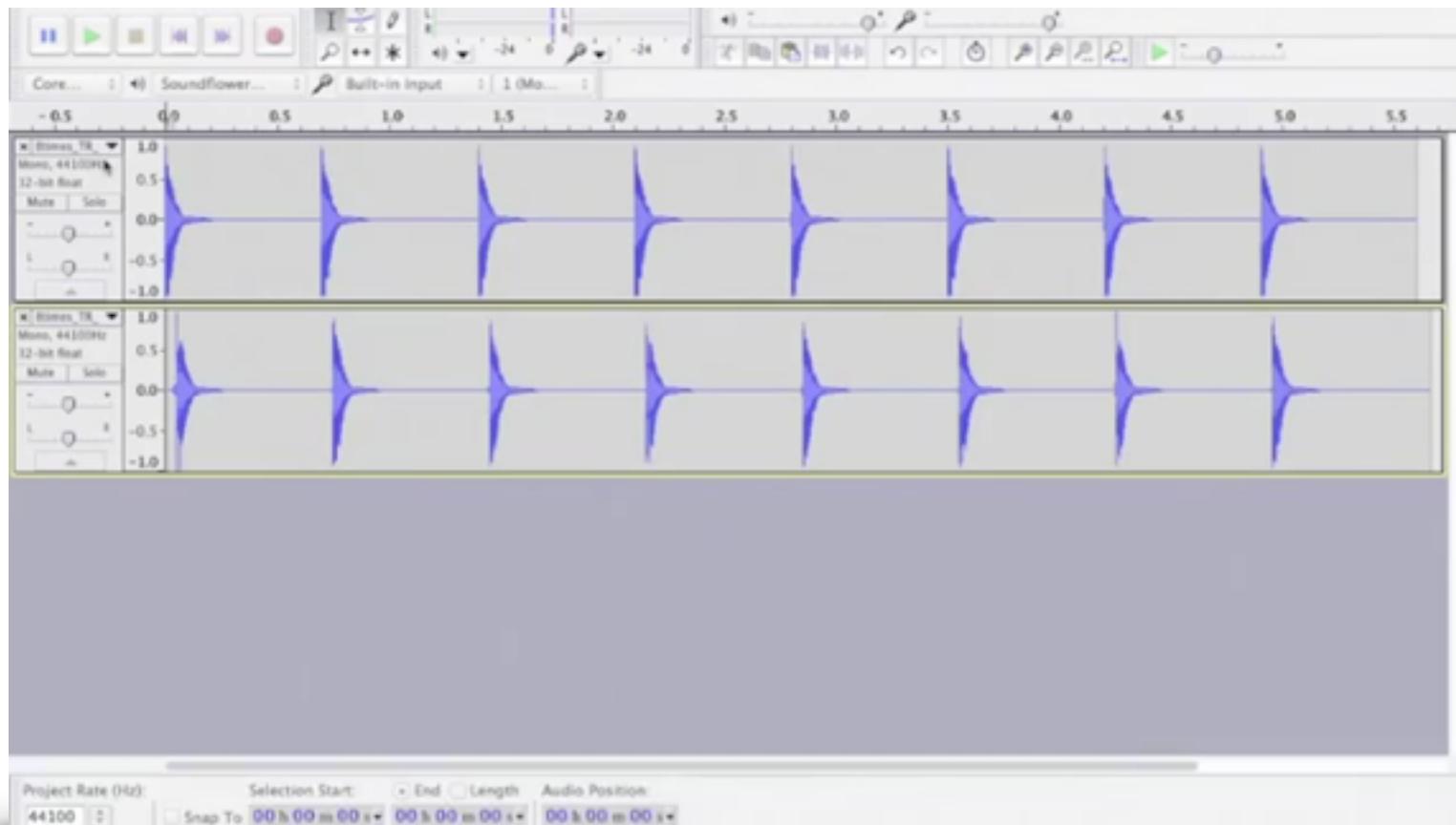
## Problem

- **Frequenzauflösung** ist durch MDCT sehr gut, **Zeitauflosung** allerdings entsprechend schlechter
- **Pre-Echo-Effekte** können auftreten



# MPEG-1 Audio Layer III

## Beispiel: Pre-Echo-Effekte



# Advanced Audio Coding

- **Advanced Audio Coding (AAC)** ist von MPEG-Arbeitsgruppe entwickeltes, verlustbehaftetes Kompressionsverfahren
- Weiterentwicklung von MPEG-2
- AAC wird im **Containerformat** (z.B. Video und Audio) MP4 verwendet

# Advanced Audio Coding

## Eigenschaften

- AAC hat verbesserte Kompressionsrate gegenüber MP3:
- AAC erreicht bereits ab 64 kBit/s akzeptable Stereo-Qualität
  - Bitraten ab 160 kBit/s sind vergleichbar mit verlustfreier CD-Qualität
- AAC findet überall Verwendung, wo auch MP3 verwendet wird

# Advanced Audio Coding Profile

- **Low Complexity (LC)** – für mittlere bis hohe Bitraten
  - Online Musikgeschäfte
- **Low Delay (LD)** – für geringe Verzögerungszeiten (20 ms) bei mittleren bis hohen Bitraten
  - Anwendung im Kommunikationsbereich z.B. Videokonferenzsysteme

# Advanced Audio Coding Profile

- **High Efficiency (HE)** – für niedrige bis sehr niedrige Bitraten
  - **Spektralbandreplikation (SBR)** um hohe Frequenzen synthetisch zu erzeugen
- **Scalable Sample Rate (SSR)** – für Streaming
  - ermöglicht ungestörte Wiedergabe von Streams durch Verringerung der Bitrate, wenn verfügbare Bandbreite plötzlich abfällt

# Spektralbandreplikation

- **Spektralbandreplikation** (SBR, engl. **Spectral Band Replication**) ist patentiertes verlustbehaftetes Verfahren parametrischer Audiokodierung der Firma Coding Technologies
- Anwendung sinnvoll typischerweise bei niedrigen Bitraten (24 kbps bis 64 kbps)

# Spektralbandreplikation

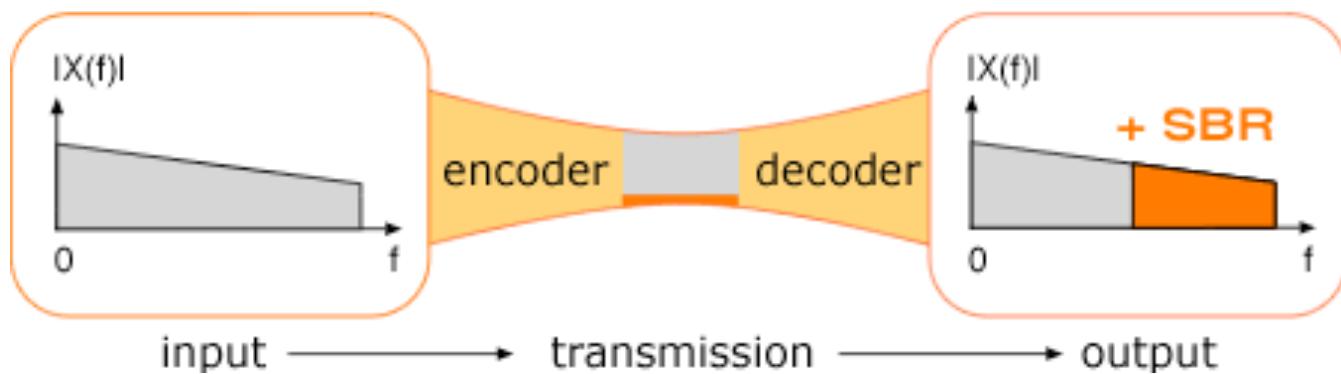
## Generelle Idee

- enge Korrelation zwischen Signalverlauf in hohen und niedrigen Frequenzbändern
  - akustische Wahrnehmung bei hohen Frequenzen weniger genau
- hohe Frequenzanteile aus niedrigeren synthetisiert

# Spektralbandreplikation

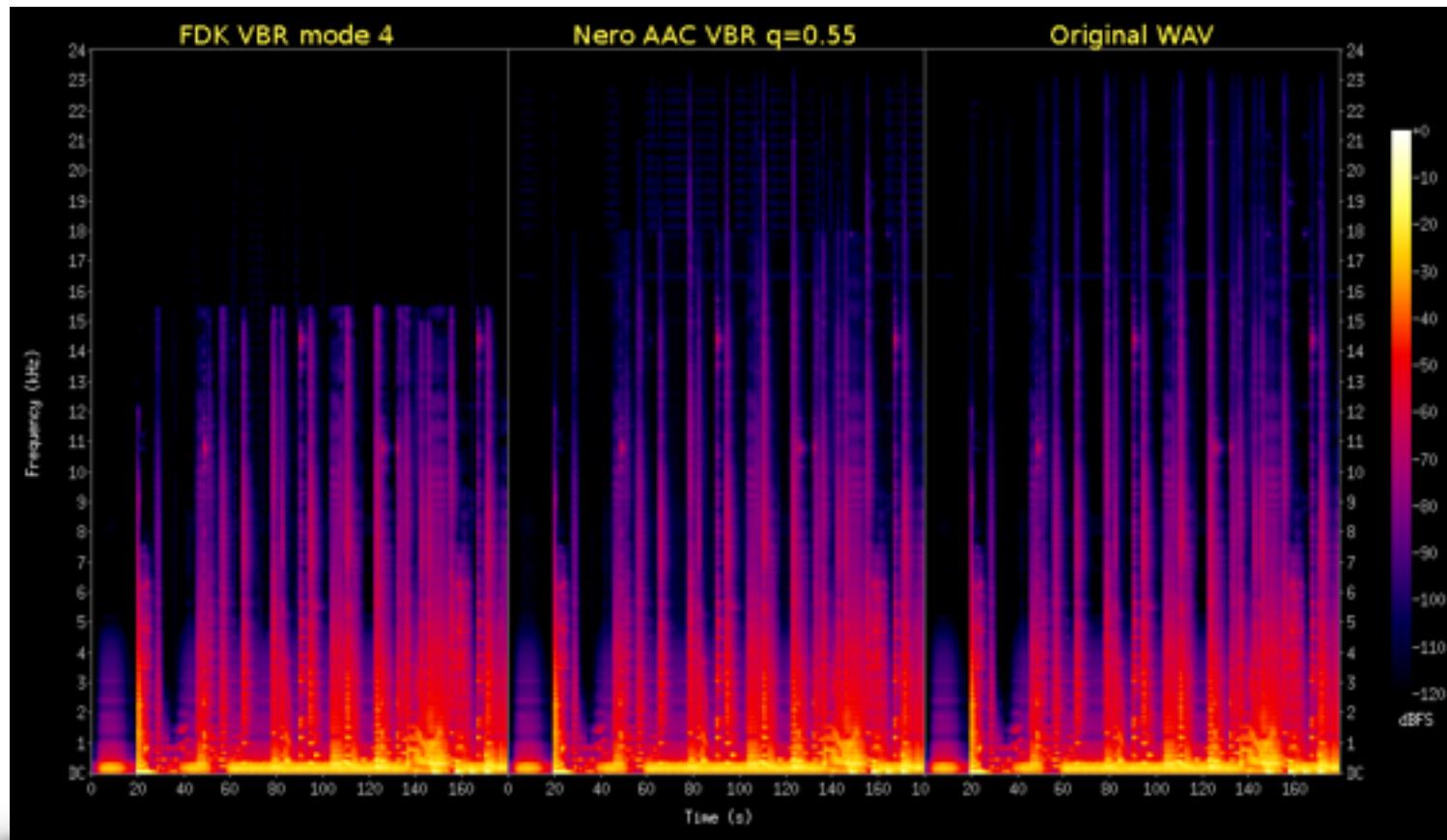
## Generelle Idee

- Frequenzanteile oberhalb Grenzfrequenz (z.B. 8 kHz) werden nicht direkt kodiert, sondern aus kodierten tieferen Signalbereichen durch ähnlich klingende Signale rekonstruiert



# Spektralbandreplikation

## Beispiele



# Advanced Audio Coding

Bsp: AAC-Profile (HE) (5.1 MB)



# Advanced Audio Coding

Bsp: AAC-Profile (HE) (995 KB)



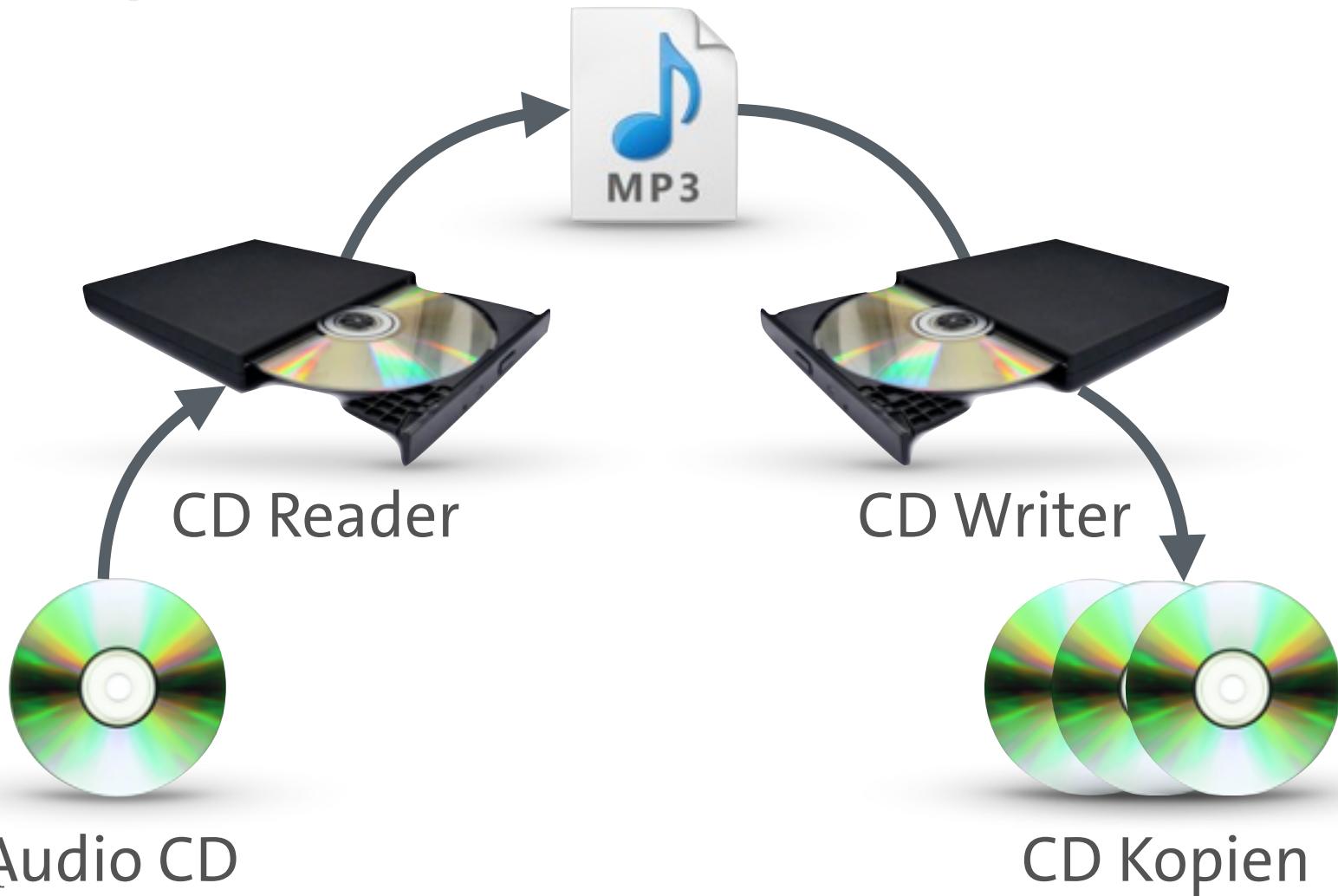
# Ogg/Vorbis

- Ogg/Vorbis (OGG)
- verlustbehaftetes komprimiertes Audiodatenformat entwickelt von Xiph.Org Foundation
- (patent-)freies Audioformat (im Gegensatz zu MP3 oder AAC)



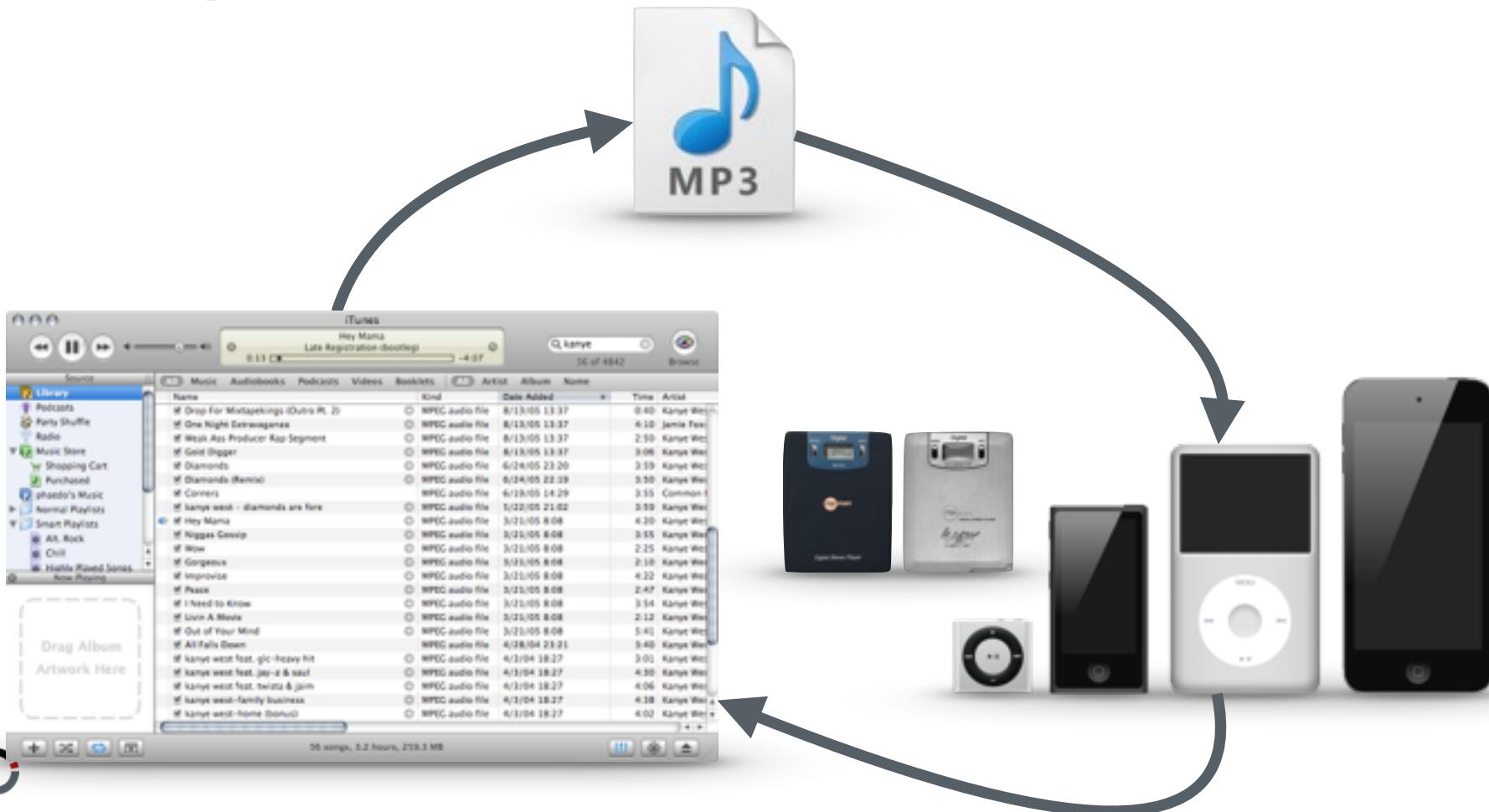
# MP3

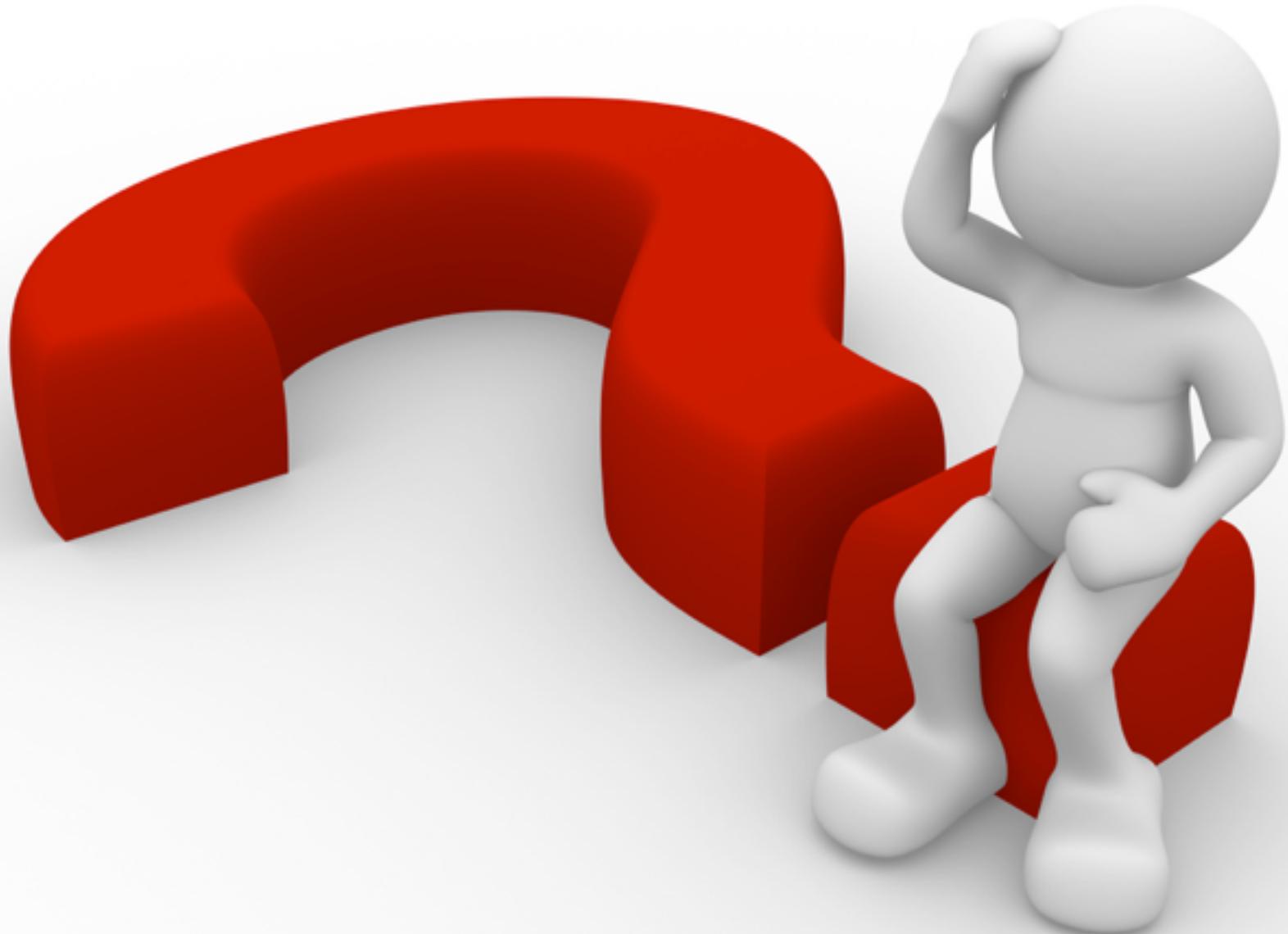
## Implikationen



# MP3

## Implikationen







# Interaktive Medien

## Kapitel Audio

### Audioverarbeitung

# Verarbeitung

- Nach Digitalisierung und Kompression beginnt eigentliche Arbeit mit Audiodateien
- Prinzipielle Vorgehensweisen der Audioverarbeitung
  - Audioverarbeitung ohne Kontext
  - Audioverarbeitung mit Berücksichtigung benachbarter Signalwerte

# Verarbeitung

- **zeit-invariante Audioverarbeitung** ändert nicht Länge des Signals
- Beispiele:
  - Veränderung der Amplitude
  - Veränderung der Hüllkurve
  - Filter
  - ...

# Verarbeitung

- **nicht-zeitinvariante Audioverarbeitung**
  - ändert Länge des Signals
- Beispiele:
  - Veränderung der Frequenz und/oder Phase
  - Echo und Hall
  - Schneiden von Audiomaterial
  - ...

# A/D-Wandlung

- Digitalisierung erfordert, dass analoger Signalwert quantisiert wird

Signalwert  
z.B. 120 dB



8-Bit-Quantisierung  
z.B. 255

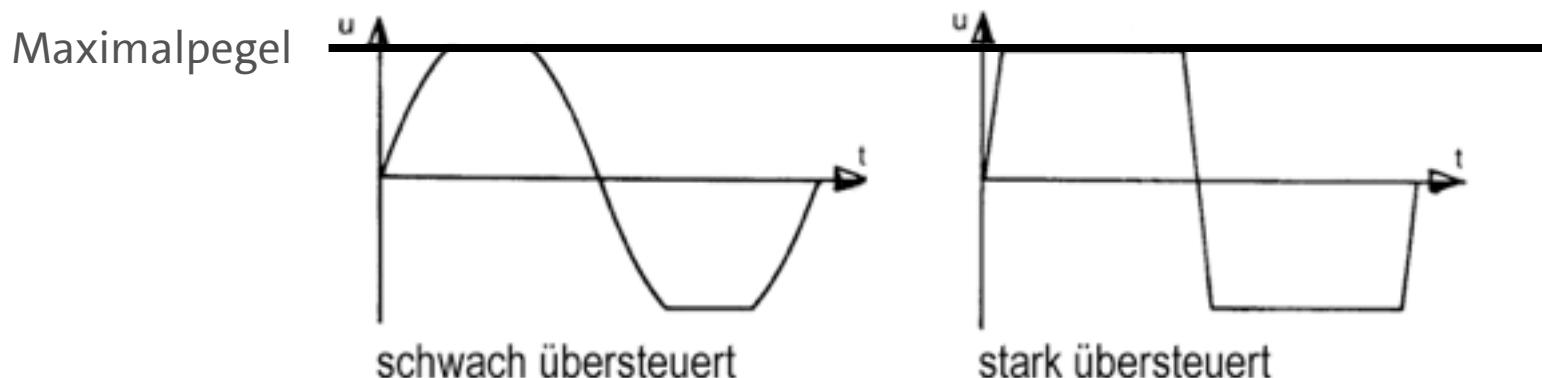


?

- Frage: Wie sieht Zuweisung aus?

# Pegel

- **Maximalpegel** ist höchster darstellbarer Signalwert (=größte darstellbare Zahl im Quantisierungsbereich)
- **Arbeitspegel** wird i.d.R. deutlich unter Maximalpegel (als 0dB) festgelegt

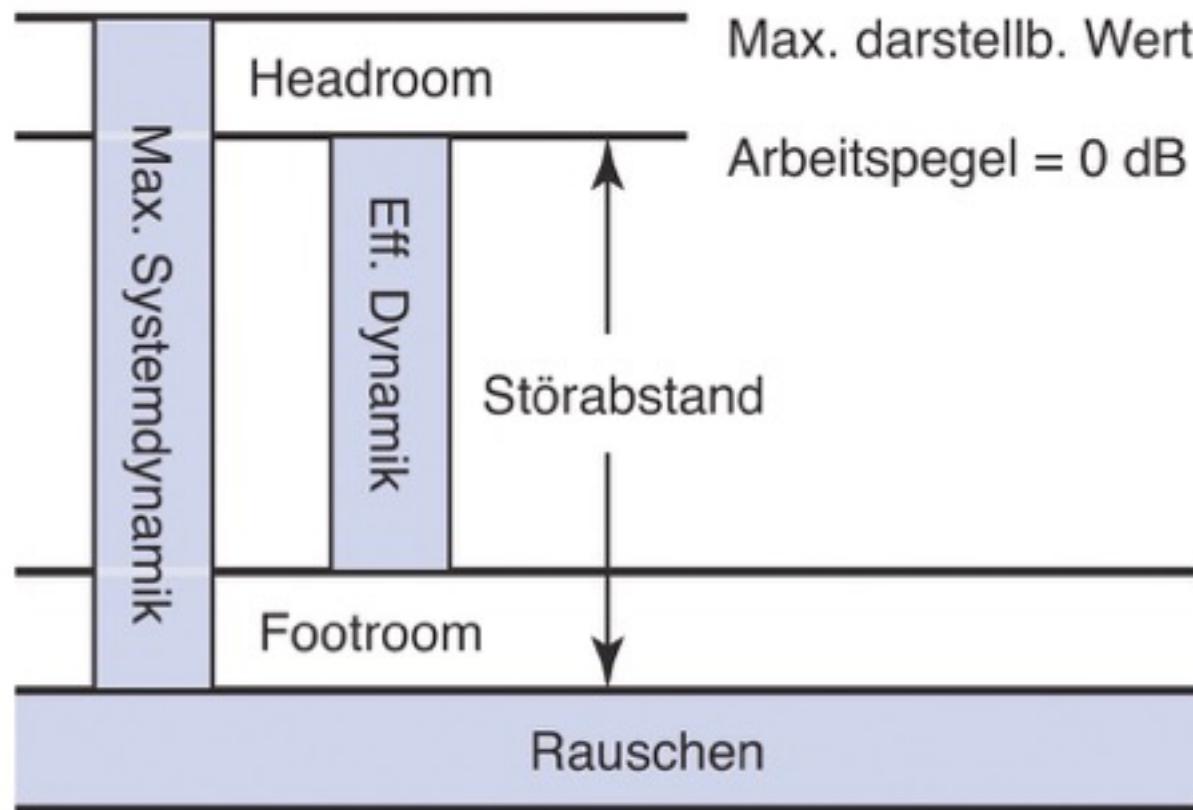


# Pegel

## Begriffe

- **Headroom** bezeichnet Abstand zwischen Arbeitspegel und Maximalpegel
- **Footroom** bezeichnet Abstand zwischen leisesten auftretendem Signal und Rauschen des Systems
- **Störabstand** oder **Signal-Rausch-Abstand** bezeichnet Abstand zwischen Arbeitspegel und Rauschen

# Pegel



# Pegelanpassung

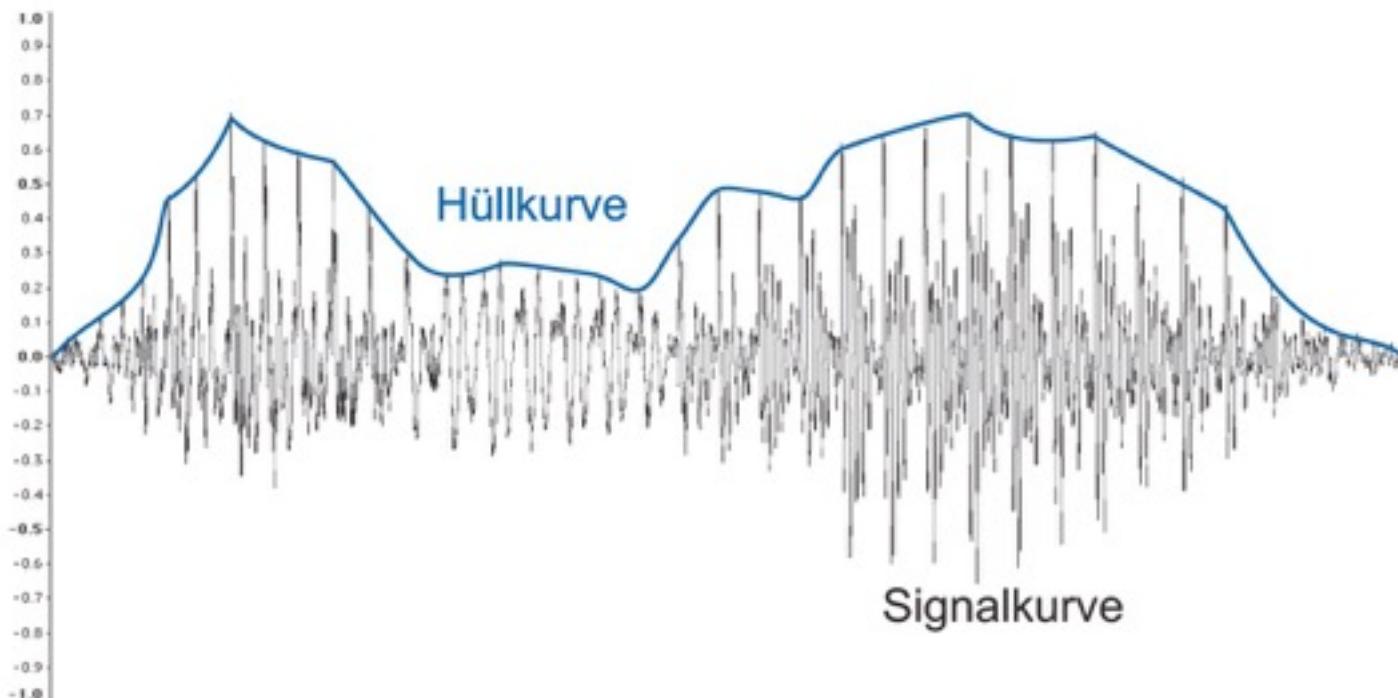
- Pegelanpassung bezeichnet Veränderung aller Signalwerte um konstanten Faktor
  - Übersteuerung
    - ▶ höchsten Signalwerte liegen außerhalb des darstellbaren Bereichs
  - Untersteuerung
    - ▶ nutzt nicht gesamten Bereich der Quantisierung aus

# Normalisierung

- Um Über- bzw. Untersteuerung entgegen zu wirken, wird Pegel eines Audiosignals vor Verarbeitung auf **0dB (Arbeitspegel)** angehoben bzw. gesenkt
- Kompensation eines eventuell vorhandenen **DC-Offset**, so dass nach Kompensation Signal symmetrisch um Null-Linie verteilt ist

# Hüllkurve

- **Hüllkurve** verbindet Spitzen eines (periodischen) Signals



# Audioverarbeitung

## Hüllkurve

- **Hüllkurve** wird oft verwendet, um zeitlich beschränkte Pegelveränderungen zu beschreiben
  - **Fading**
    - ▶ Multiplikation mit (logarithmisch, linear, oder exponentiell) Faktor zwischen 0 und 1 zum Ein- und Ausblenden
    - ▶ Hüllkurve wird vertikal auseinander- bzw. zusammengezogen

# Dynamik

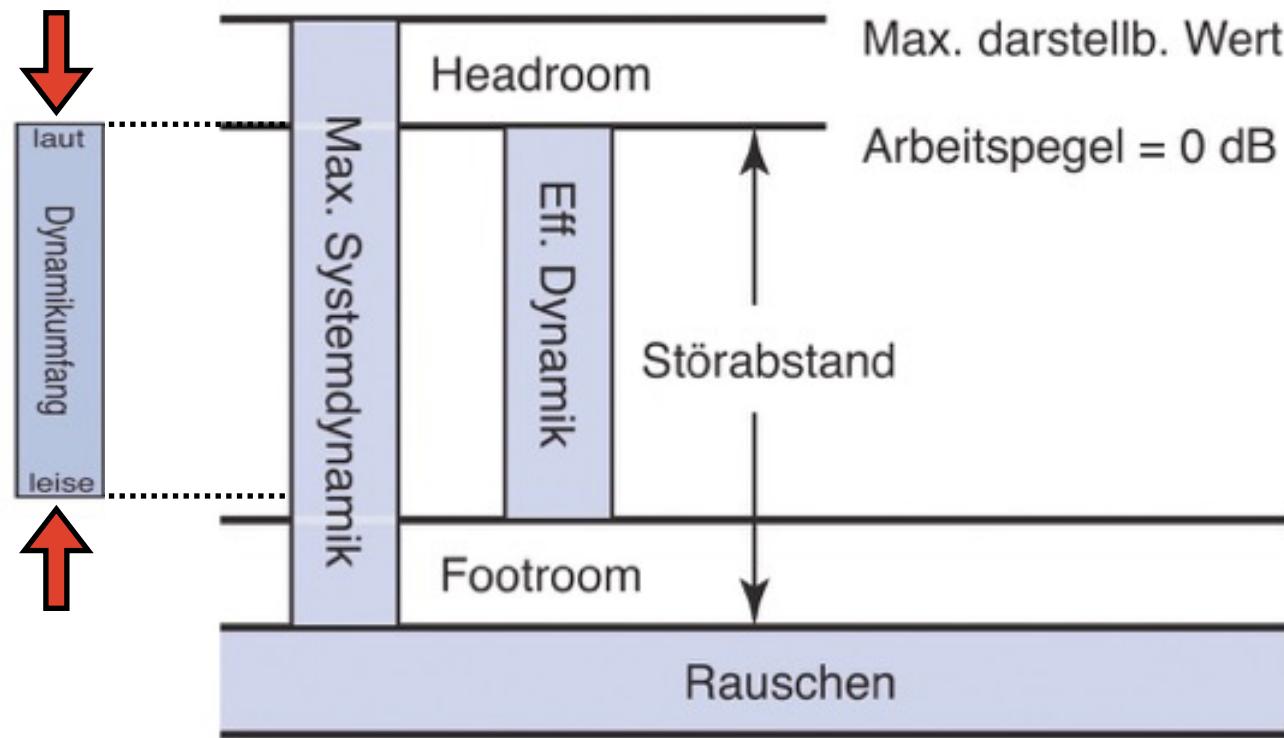
- **Dynamik** einer Musikpassage ist Verhältnis vom größten zum kleinsten Amplitudenwert
  - **Mikrodynamik** (Sekundenbruchteile)
    - ▶ Charakter eines Instruments
  - **Makrodynamik** (Sekunden und Minuten)
    - ▶ Lautstärkenaufbau gesamter Musikstücke (Pianissimo bis Fortissimo)

# Dynamikkompression

- bei **Dynamikkompression** werden leise Stellen im Pegel angehoben und laute Stellen herabgesenkt
- wird Signal danach **normalisiert** sind laute Stellen gleich laut, aber leise Stellen lauter

# Dynamikkompression

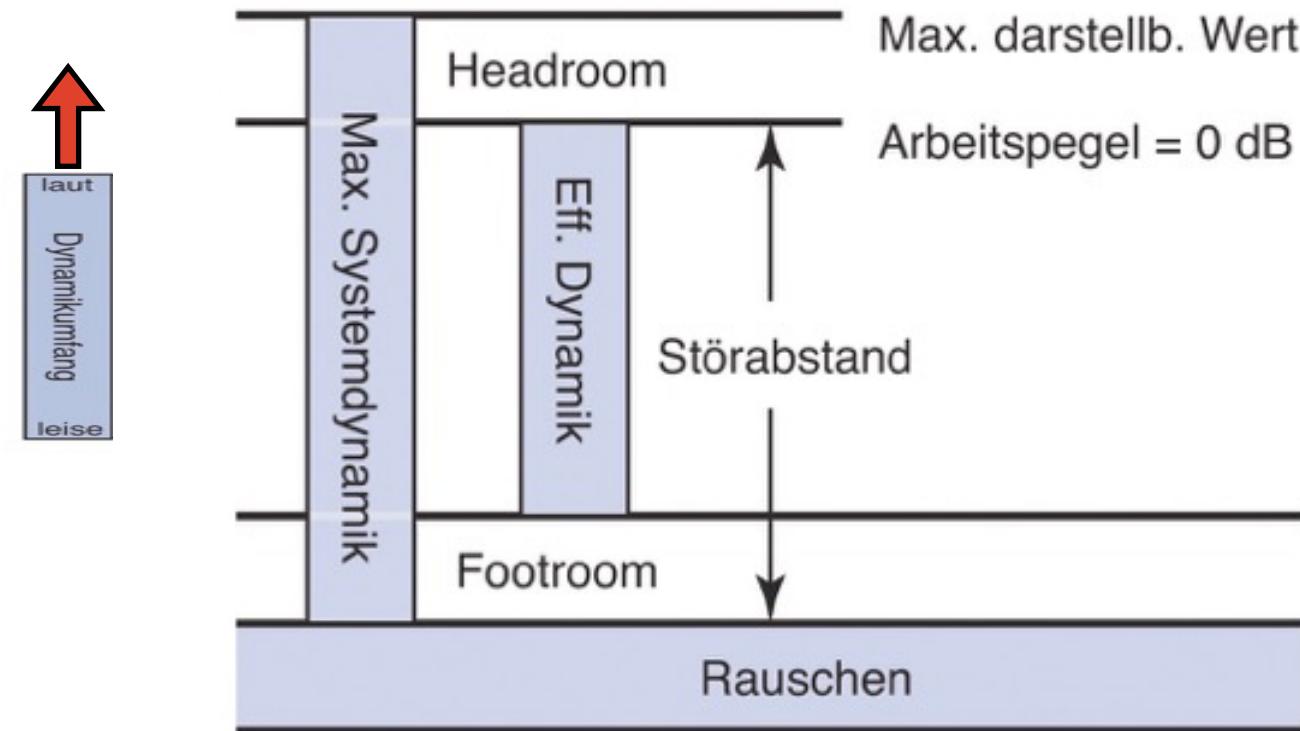
## Beispiel



1. Schritt: Kompression

# Dynamikkompression

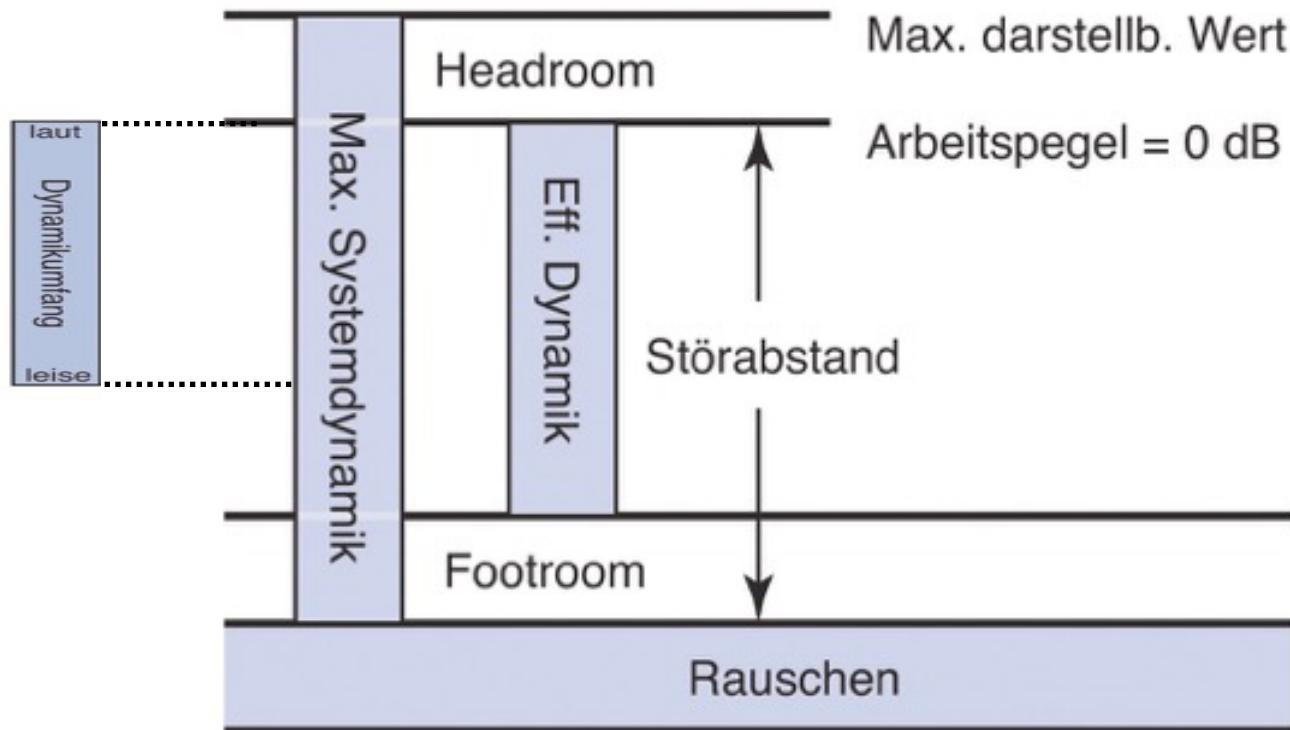
## Beispiel



2. Schritt: Normalisierung

# Dynamikkompression

## Beispiel



# Diskussion



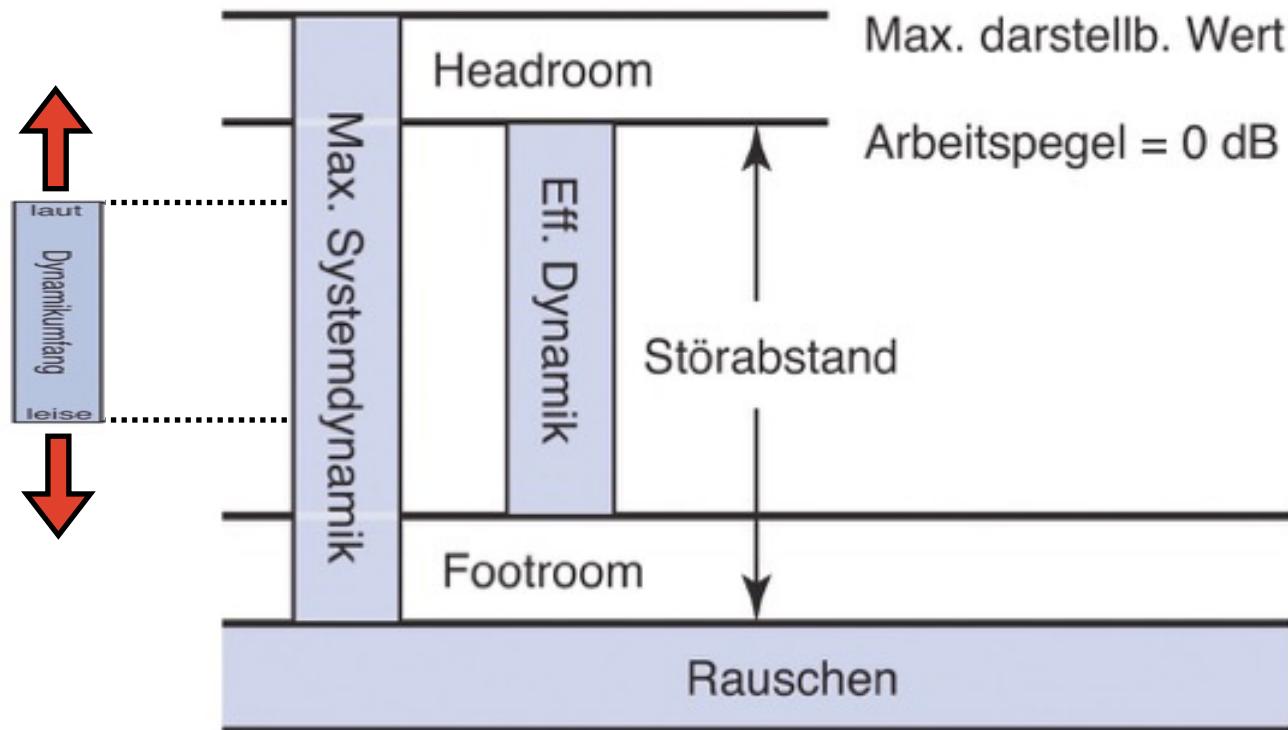
Wann ist eine Dynamikkompression sinnvoll?

# Dynamikexpansion

- bei **Dynamikexpansion** werden leise Stellen im Pegel herabgesenkt und laute Stellen angehoben
  - laute Stellen werden lauter
  - leise Steller werden leiser

# Dynamikexpansion

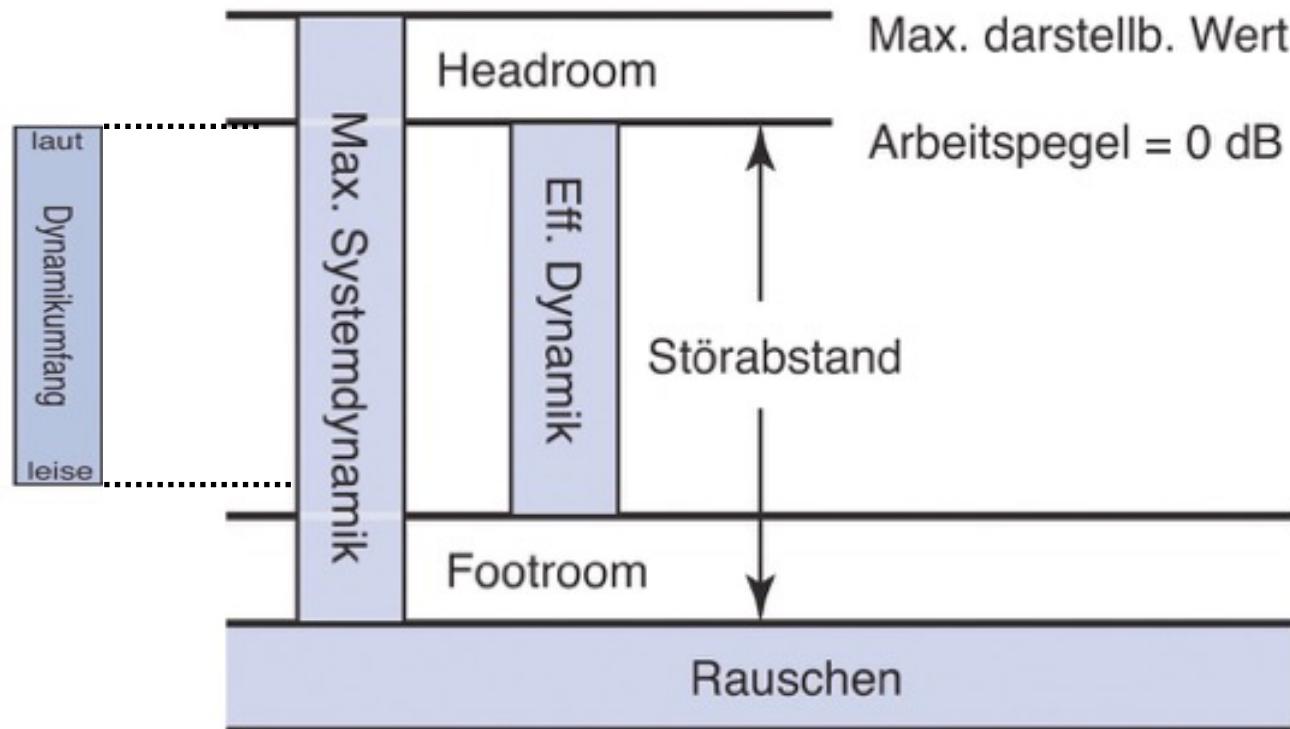
## Beispiel



1. Schritt: Expansion

# Dynamikexpansion

## Beispiel



# Diskussion



Wann ist eine Dynamikexpansion sinnvoll?

# Filter

- **Filter** berechnen Signalwerte auf Basis aktueller sowie zeitlich benachbarter Signalwerte
  - erlauben i.d.R. keine zeit-invarianten Bearbeitungen
- digitalen Audiotechnik:
  - **Digital Signal Processors (DSP)**

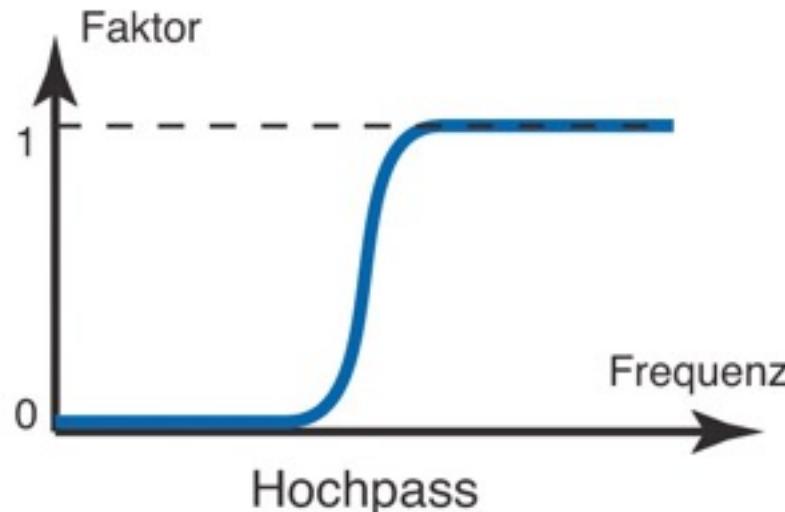
# Frequenz-Filter

- **Frequenz-Filter** heben bestimmte Frequenzanteile an oder schwächen sie ab
- Filter werden durch **Filterkurve** beschrieben
- Steilheit der Filterkurve wird **Güte des Filters** genannt

# Frequenz-Filter

## Hochpass-Filter

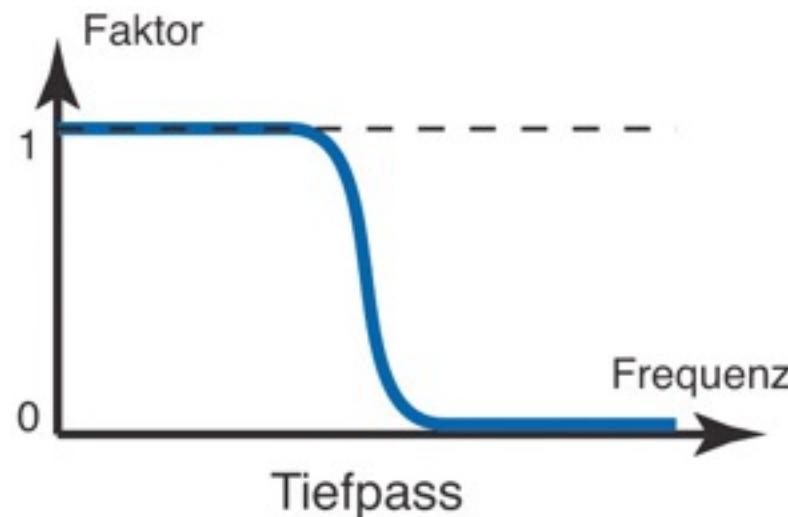
- **Hochpass-Filter** lässt hohe Frequenzen durch, d.h. passieren
- Frequenzen unterhalb bestimmten Grenzfrequenz werden abgeschwächt



# Frequenz-Filter

## Tiefpass-Filter

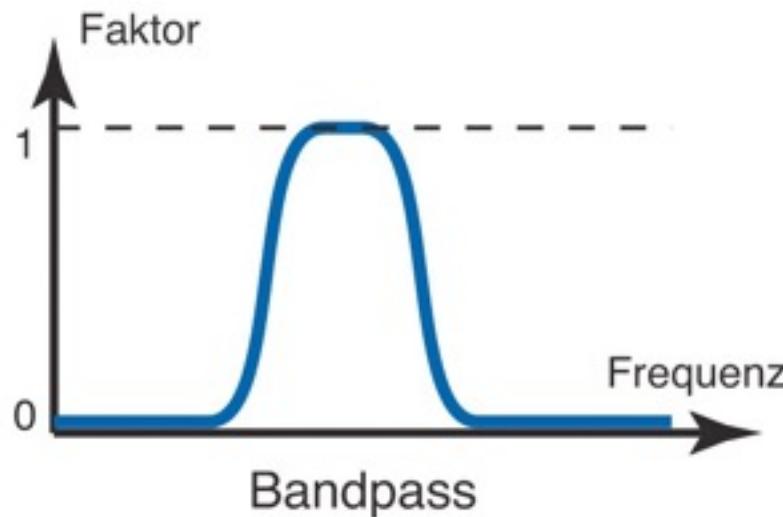
- **Tiefpass-Filter** lässt tiefen Frequenzen durch, d.h. passieren
- Frequenzen oberhalb bestimmter Grenzfrequenz werden abgeschwächt



# Frequenz-Filter

## Bandpass-Filter

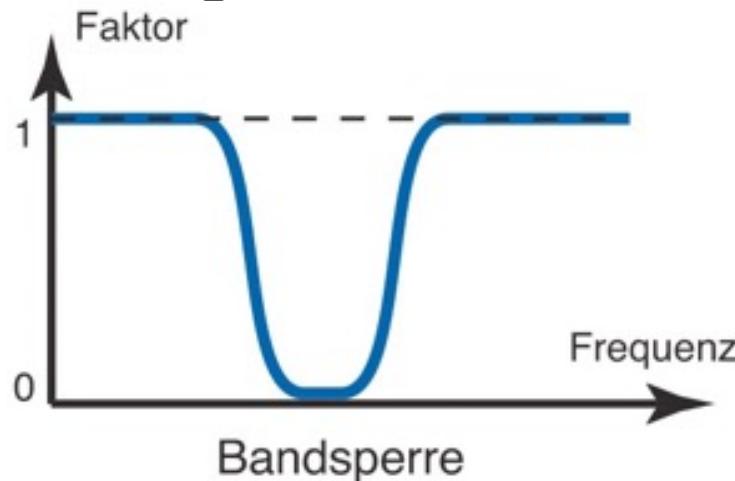
- **Bandpass-Filter** wendet Hoch- und Tiefpass-Filter nacheinander an
- Frequenzen zwischen Grenzfrequenzen werden durchgelassen, andere abgeschwächt



# Frequenz-Filter

## Bandsperre-Filter

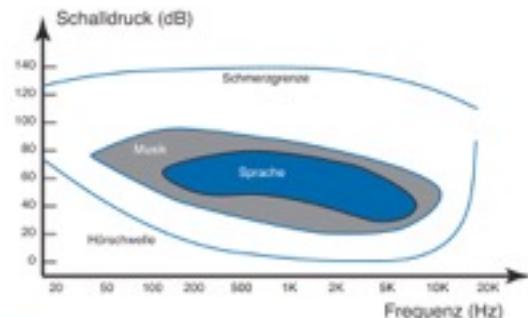
- **Bandsperre-Filter** wendet Hoch- und Tiefpass-Filter parallel an und addiert Signale
- Frequenzen zwischen beiden Grenzfrequenzen werden abgeschwächt, andere durchgelassen



# Fokus Mediendesign

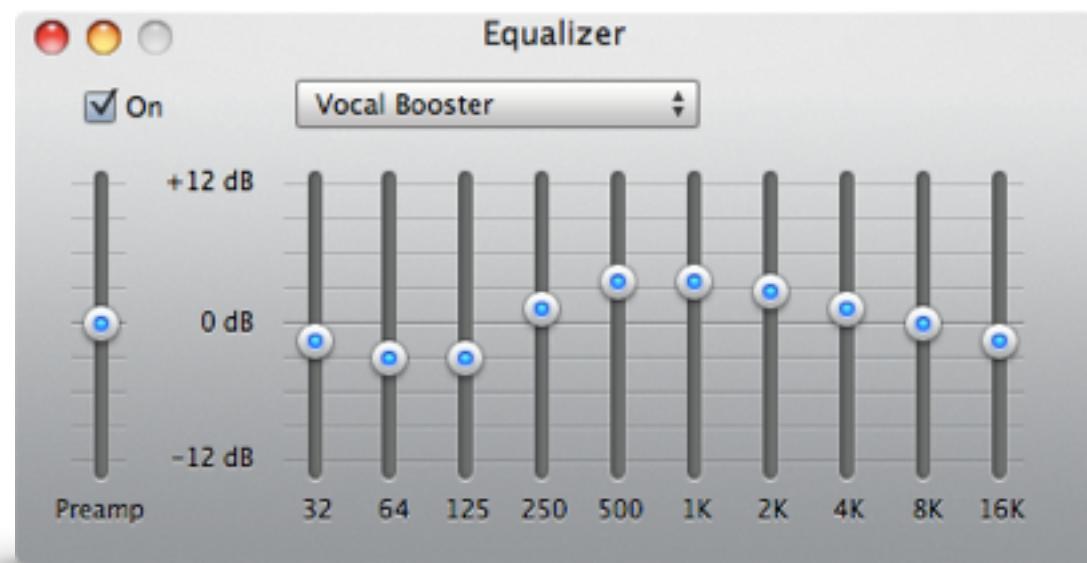
## Beispiel: Loudness-Funktion

- bei leisen Signalen hören wir vor allem mittlere Frequenzen
- Loudness-Funktion erhöht entsprechend tiefen und hohen Frequenzen



# Equalizer

- Kombination mehrerer Frequenzfilter
- grafischer Equalizer, z.B. iTunes Equalizer mit 10 Bändern



# Frequenzmodifikation

- **Frequenzmodifikation** ändert zeitlichen Verlauf
  - **Resampling:** Abspielen mit anderer Samplingrate führt zur Änderung der wahrgenommenen Tonhöhe und Abspiellänge
  - **Time Stretching:** Veränderung der Abspiellänge ohne Veränderung der Tonhöhe

# Echo und Hall

- **Echo** spielt Originalsignal zeitverzögert und abgeschwächt ab
  - Beispiele:
    - ▶ kurze Verzögerung klingt wie Keller
    - ▶ lange Verzögerung klingt wie Gebirge
- **Hall** spielt Originalsignal mehrfach mit verschiedenen Verzögerungen ab

# Restauration

- Restauration findet bspw. Anwendung bei Aufbereitung (alter) analoger Tonaufnahmen
- Unterscheidung der Störungen vom Signal
- Typische Störungen sind
  - Rauschfehler (Noise, Hiss)
  - Clickfehler/Signallücken (Clicks)
  - Knistern (Crackles)

# Restauration

- beim **Denoising** wird Fingerprint (typisches Spektrum) des Rauschens bestimmt und diese Frequenz herausgefiltert
- beim **Deklicking** werden Signallücken durch Interpolation (oder Daten aus zweitem Stereokanal) ersetzt
- beim **Decracking** handelt es sich um ein wiederholt ausgeführtes Deklicking an entsprechenden Stellen

# Audioschnitt

- Terminologie Cut-and-Paste-Operationen
- Übergänge sollten natürlich klingen,  
daher in Pausen oder im Nulldurchgang  
Schneiden

# Fokus Mediendesign

## Beispiel

*„Zu Risiken und Nebenwirkungen lesen Sie die Packungsbeilage und fragen Sie Ihren Arzt oder Apotheker.“*

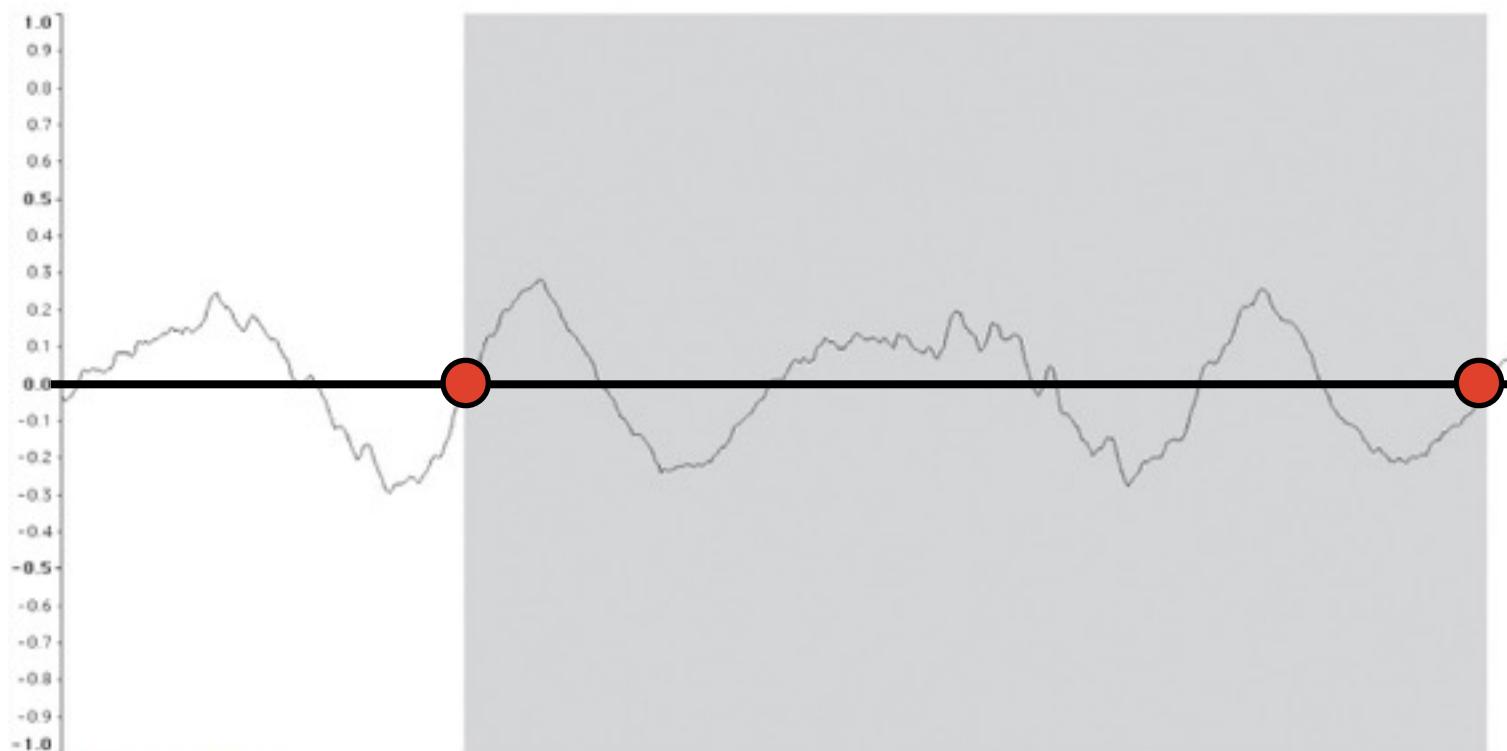
- Kostenersparnis durch Verkürzung
- bei gleicher Tonhöhe Geschwindigkeit soweit erhöht, dass Text noch verständlich bleibt

# Schnitte

- **harter Schnitt:** Aneinanderhängen zweier unterschiedlicher Signale
- **Überblenden:** gleichzeitiges Ausblenden des ersten bzw. Einblenden des zweiten Signals
  - **Kreuzblende:** langsamer Übergang
  - **Sturzblende:** schneller Übergang

# Audioschnitt

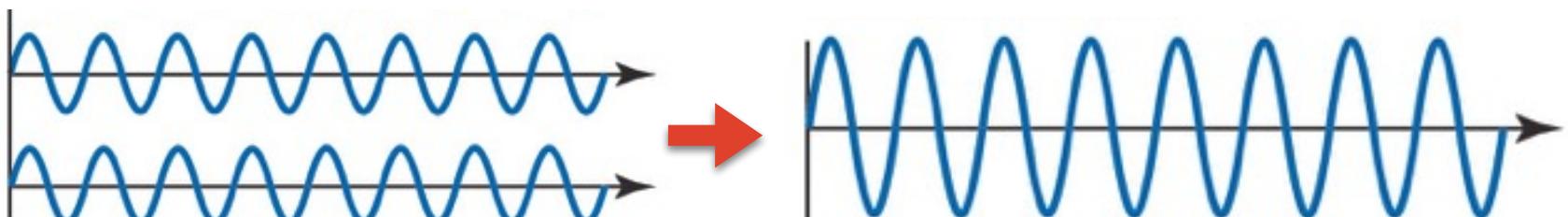
## Beispiel: Nulldurchgänge



# Interferenzen

## Konstruktive Interferenz

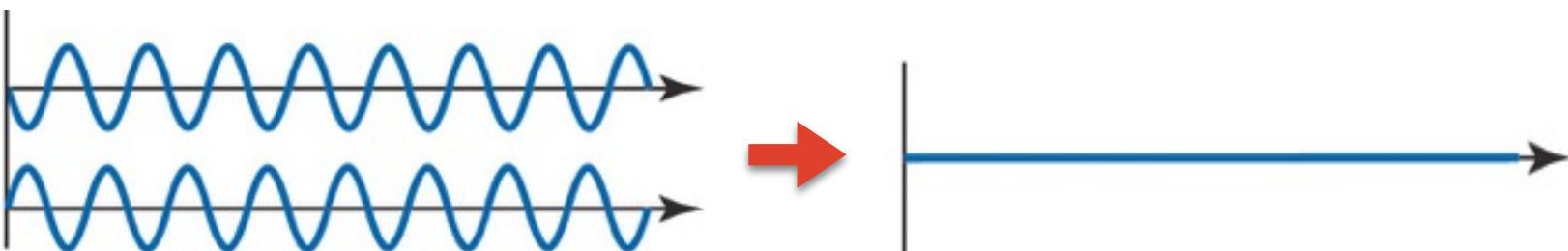
- werden zwei Signale gemischt, addieren sich Signalwerte
- wird Signal mit sich selbst gemischt, verstärkt sich Amplitude
- Effekt heißt **konstruktive Interferenz**



# Interferenzen

## Destruktive Interferenz

- werden zwei inverse Signale (phasenverschoben um 180 Grad) gemischt, löschen sich Signalwerte gegenseitig aus
- Effekt heißt **destruktive Interferenz**



# Fokus Mediendesign

## Geräuschmindernder Kopfhörer

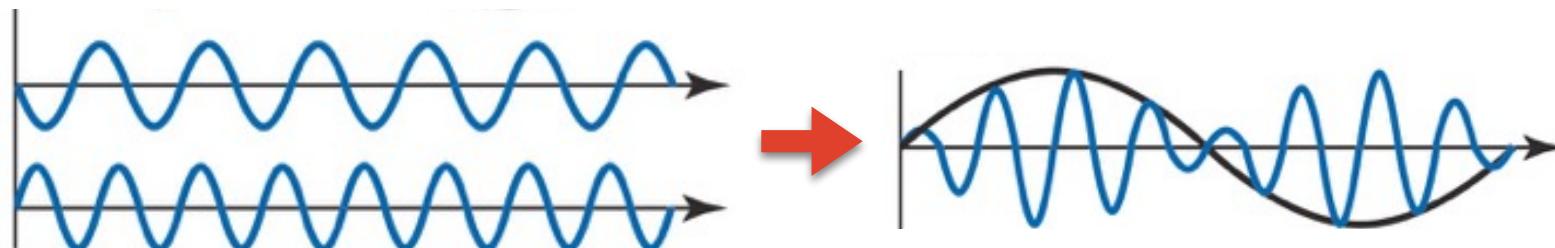
- Kopfhörer mit **aktiver Geräuschunterdrückung**
  - Mikrofon misst Umgebungsgeräusche
  - Verwendung destruktiver Interferenz, um Umgebungsgeräusche zu reduzieren



# Interferenzen

## Schwebung

- werden zwei periodische Signale gemischt, deren Frequenzen nahe beieinander liegen, kommt es zur **Schwebung**, d.h. konstruktive und destruktive Interferenz wechseln sich ab



# Phasenverschiebung

- bei **Phasenverschiebung** (engl. **Phase Shift**) wird Kopie des Signals mit Original vermischt, um verschiedene Audio-Effekte hervorzurufen
  - **Chorus**
    - ▶ Signale voller klingen lassen durch Abspielen von Signalkopien mit zufälligen Verzögerungen (10-25ms)

# Phasenverschiebung

- **Flanging**
  - ▶ ähnlich wie Chorus, aber nur eine phasen-verschobene (verzögerte) Signalkopie
- **Phasing**
  - ▶ Kombination aus Originalsignal und variierender Phasenverschiebung einzelner Signale Frequenzen

# Phasenverschiebung

## Beispiel: Drum - Original



# Phasenverschiebung

## Beispiel: Drum - Flanging



# Phasenverschiebung

## Beispiel: Drum - Phasing



# Audioverarbeitung

## Beispiel: Audacity

