

# Digitale Medien



## 9 – DIGITAL AUDIO

# Allgemein

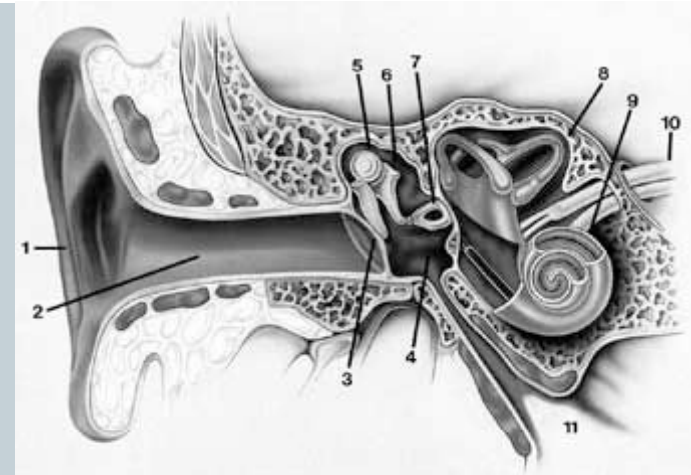
2

- vergangene LVen: Darstellung von Informationen: Bilder
- Heute: Audio
  - Auditive Wahrnehmung
  - Digitalisierung
  - Speicherung von Audiodaten
  - Formate
  - Schnitt

# Das menschliche Ohr

3

- **Außenohr:**
  - Einfangen des Schalls
  - Richtungsbestimmung
- **Mittelohr:**
  - Trommelfell, Hammer, Amboss, Steigbügel: Verstärkung der Kraftwirkung
- **Innenohr:**
  - Schnecke
  - Aufgerollte Röhren (Gänge), gefüllt mit Lymphflüssigkeit



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

# Auditive Wahrnehmung

4

- Hörbereich wird von zwei Parametern bestimmt:
  1. wahrnehmbare Frequenz:  
16Hz bis 20KHz
    - Nicht hörbare Frequenzen <16Hz: Infraschall, >16-20KHz: Ultraschall
    - Höchste Empfindlichkeit: 2...4KHz
  2. Schalldruck:  
Hörschwelle ... Schmerzgrenze



Quelle: <http://de.wikipedia.org/wiki/Ohr>

# Auditive Wahrnehmung - Schalldruck

5

- Umsetzung von Schalldruckschwankungen in Nervenreize
- Dynamik(Hör-)bereich: Schalldruck von  $2 \cdot 10^{-5}$  ... 20Pa
- Grenzen: Hörschwelle und Schmerzgrenze
- logarithmische Verarbeitung durch Gehör
- Schalldruck-Pegelwert logarithmisch, 0dB=Hörschwelle
- Gehörempfindlichkeit hängt von Frequenz ab
  - bei mittleren Frequenzen (1...4KHz) liegt Hörschwelle bei  $2 \cdot 10^{-5}$ Pa=20μPa=0.00002Pa
- Schalldruckpegel: Verhältnis zwischen aktuellem Pegel und Bezugsgröße  $p_0=20\mu\text{Pa}$
- Einheit: Dezibel [dB]
- Berechnung:  $L_p = 20 \log_{10} (p_1/p_0)$ dB

# Auditive Wahrnehmung



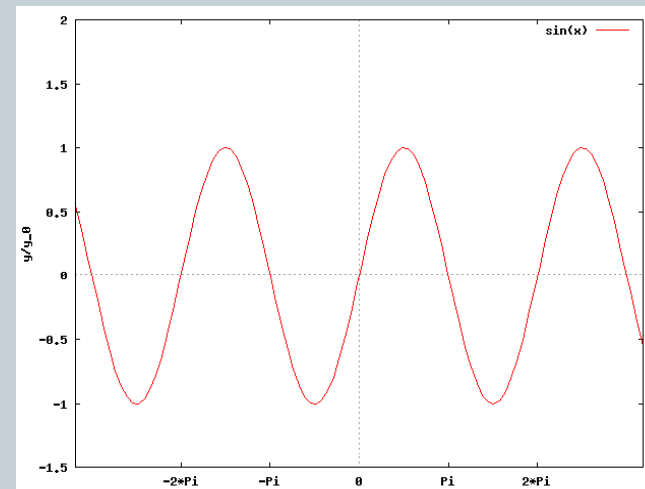
| Situation und Schallquelle                 | Schalldruck $p$                             | Schalldruckpegel $L_p$ |
|--|---|------------------------|
| Hörschwelle bei 2 kHz                      | $2 \cdot 10^{-5}$ Pa                        | 0 dBSPL                |
| Blätterrauschen, ruhiges Atmen             | $6,32 \cdot 10^{-5}$ Pa                     | 10 dBSPL               |
| Sehr ruhiges Zimmer                        | $2 \cdot 10^{-4}$ - $6,32 \cdot 10^{-4}$ Pa | 20 - 30 dBSPL          |
| Normale Unterhaltung, 1 m entfernt         | $2 \cdot 10^{-3}$ - $6,32 \cdot 10^{-3}$ Pa | 40 - 60 dBSPL          |
| Fernseher 1 m entfernt in Zimmerlautstärke | $2 \cdot 10^{-2}$ Pa                        | ca. 60 dBSPL           |
| Pkw, 10 m entfernt                         | $2 \cdot 10^{-2}$ - $2 \cdot 10^{-1}$ Pa    | 60 - 80 dBSPL          |
| Hauptverkehrsstraße, 10 m entfernt         | $2 \cdot 10^{-1}$ - $6,32 \cdot 10^{-1}$ Pa | 80 - 90 dBSPL          |
| Gehörschäden bei langfristiger Einwirkung  | $6,32 \cdot 10^{-1}$ Pa                     | ab 90 dBSPL            |
| Presslufthammer, 1 m entfernt / Diskothek  | 2 Pa  | 100 dBSPL              |
| Düsenflugzeug, 100 m entfernt              | 6,32 - 200 Pa                               | 110 - 140 dBSPL        |
| Gehörschäden bei kurzfristiger Einwirkung  | 20 Pa                                       | ab 120 dBSPL           |
| Schmerzschwelle                            | 63,25 Pa                                    | 130 dBSPL              |

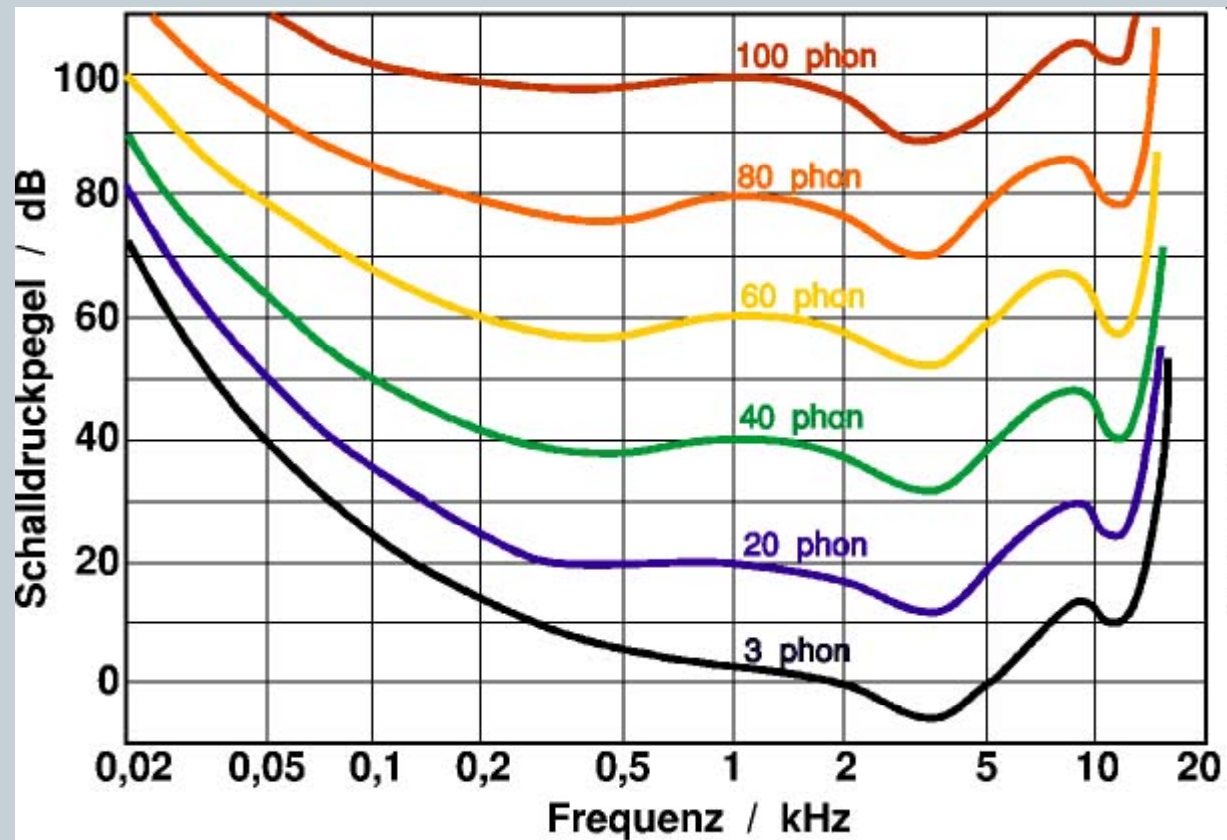
Bezugsschalldruck (Referenz) ist die Hörschwelle  $p_0 = 2 \cdot 10^{-5}$  Pa

# Auditive Wahrnehmung - Lautstärke

7

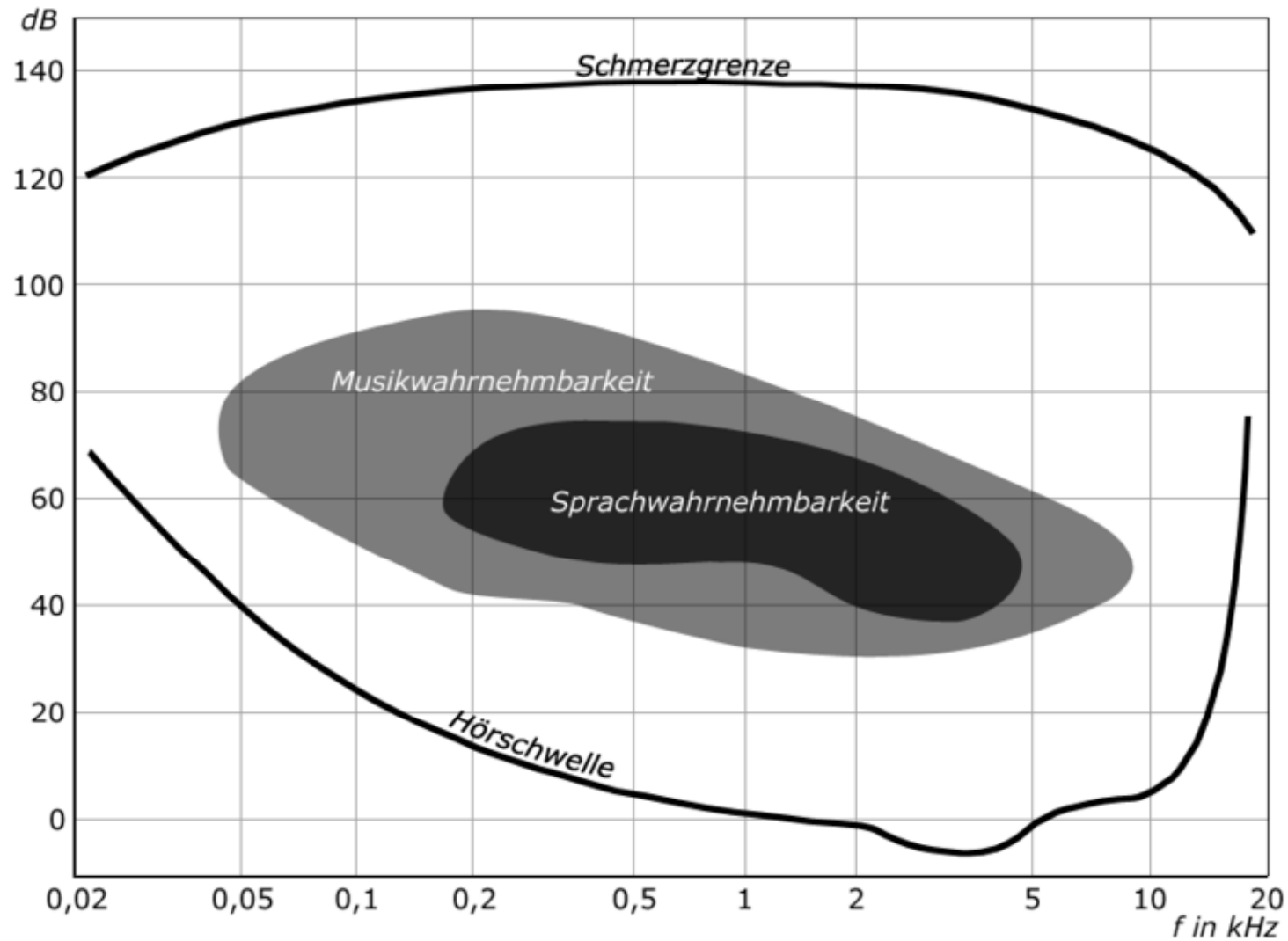
- Einheit: Phon
- Vergleichsmaß: beschreibt, welchen Schalldruckpegel ein Sinuston mit 1kHz haben müsste, um genauso laut wie der betrachtete Schall empfunden zu werden
- Frequenzabhängig







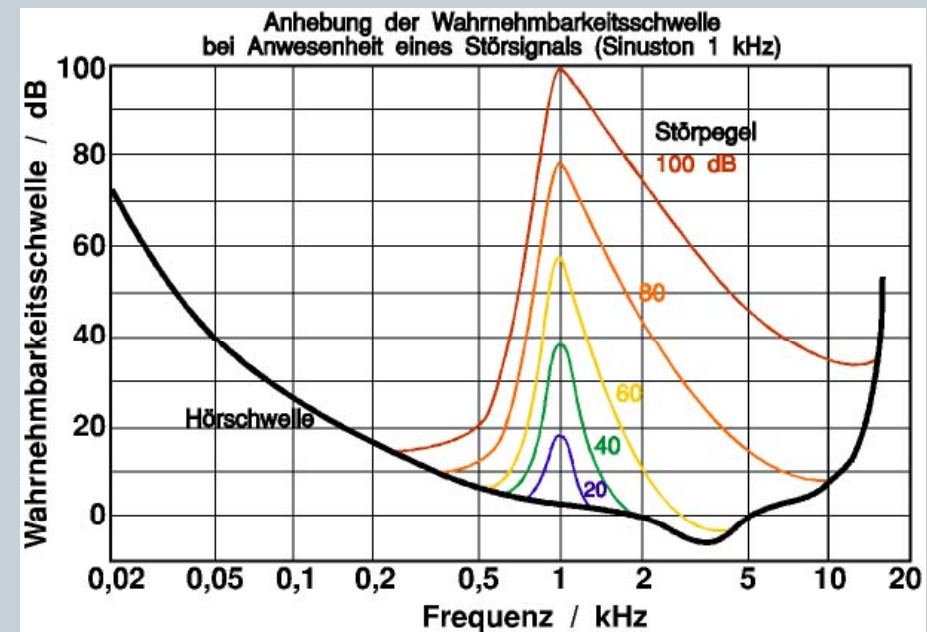
# Hörfläche



# Maskierungseffekt

10

- Verdeckung von Tönen
- Beispiel: laute Bässe bei gleichzeitiger Wiedergabe von mittleren Frequenzen
  - Hörschwelle wird angehoben
  - Mittlere Frequenzen werden nicht oder nur teilweise wahrgenommen



Bildquelle: <http://de.wikipedia.org/wiki/Maskierungseffekt>

# Auditive Wahrnehmung

11

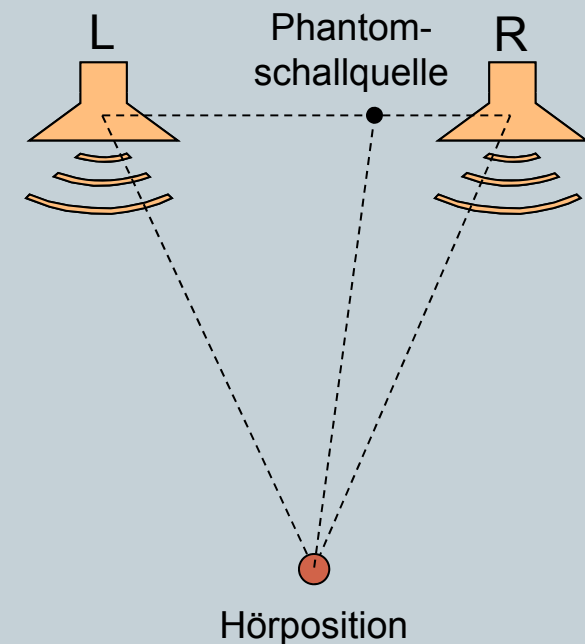
- Richtungswahrnehmung durch Pegel- und Laufzeitdifferenzen zwischen den Ohren

→ Lokalisation: binaurales Hören

- tiefe Frequenzen: Laufzeitdifferenz
  - ✦ keine Richtungsortung möglich
- hohe Frequenzen: Pegelunterschiede
- Gehör lässt sich täuschen und nimmt Phantomschallquellen wahr

→ Ausnutzung bei Stereophonie

- zwei Signalquellen → Kanäle



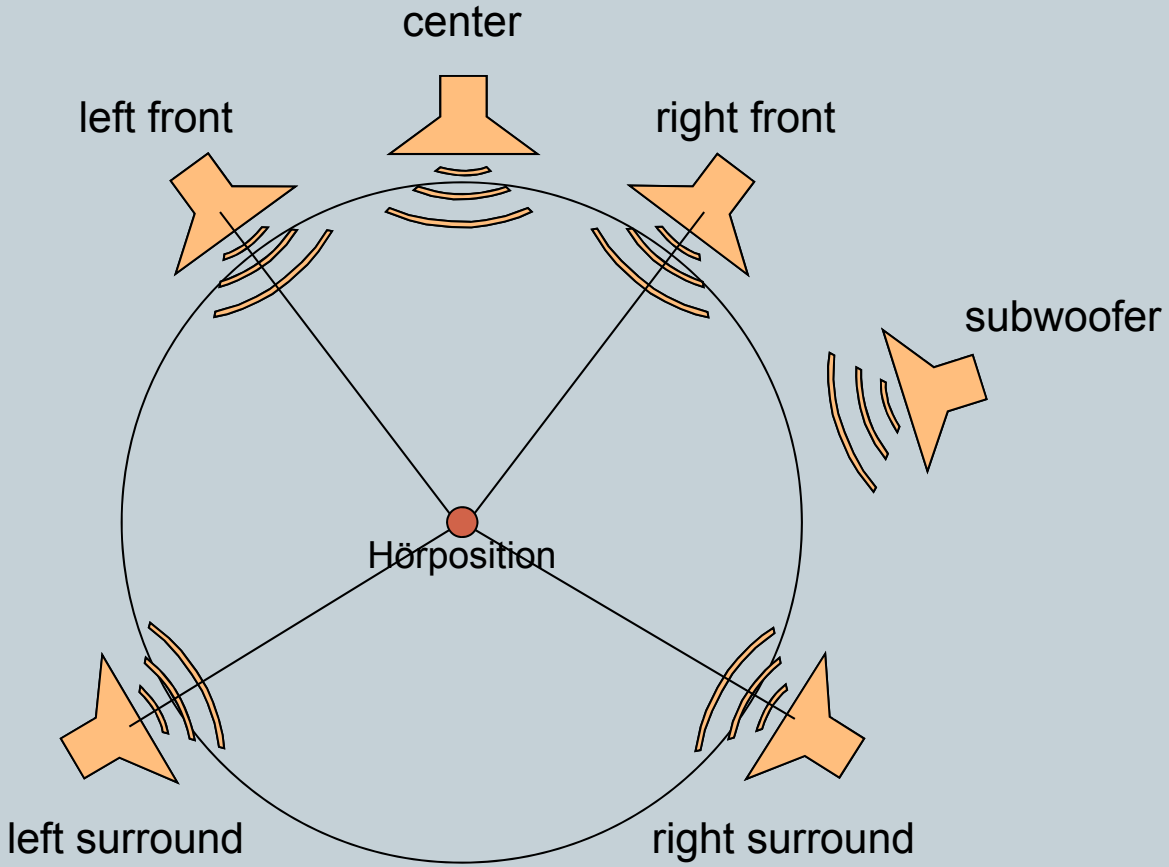
# Kanäle Beispiel

12



- mono: 1 Kanal
- stereo: zwei Kanäle
- 1D Surround: drei Kanäle (rechts, Mitte, links)
- quadrophonie: 4 Kanäle (vorn rechts, vorn links, hinten rechts, hinten links)
- Surround: 5 Kanäle, wie quadrophonie+vorn Mitte
- 5.1 Surround: wie Surround+LFE-Kanal
  - LFE: Low Frequency Effect (Subwoofer)
- 7.1 Surround: Erweiterung von 5.1 um zwei Kanäle hinter dem Hörer

# Surround



# Qualität von Audio

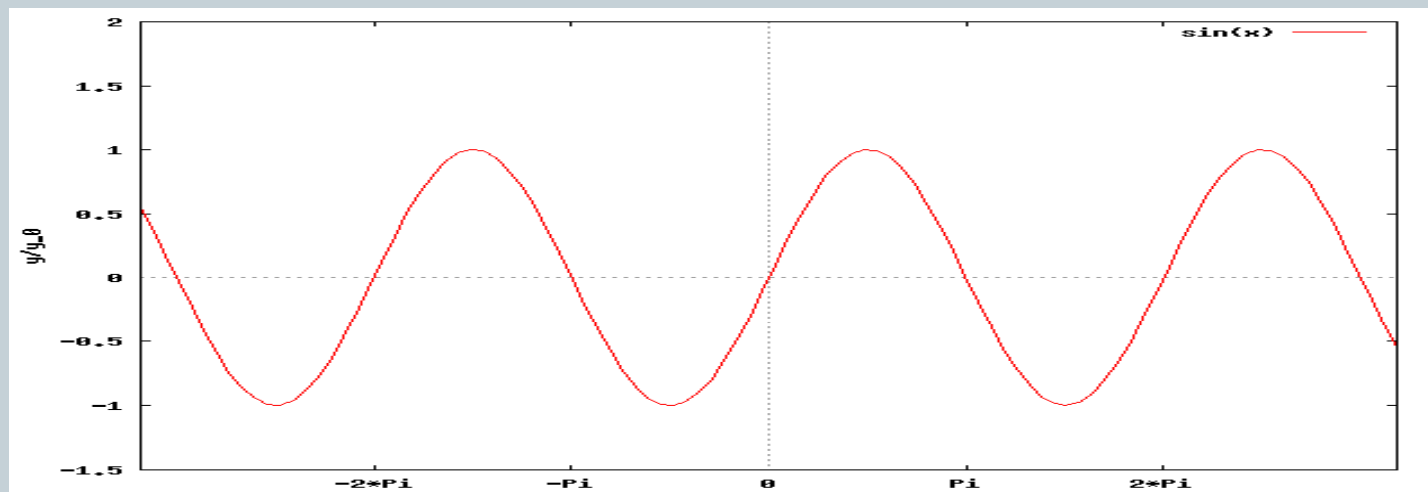
14

- Unterscheidung zwischen Nutz- und Störsignalen
    - Musik und Verkehrslärm
  - Technische Systeme erzeugen Störungen
    - Rauschen
  - wichtig für Qualität ist Signal-Rausch-Abstand (S/N):  
logarithmisch in dB angegeben
  - gut: >70dB
  - Rauschsignale, die mehr als 70dB unter Maximalpegel liegen  
werden nicht mehr wahrgenommen
- Maskierungseffekt, auch Grundidee für MP3

# Signalformen

15

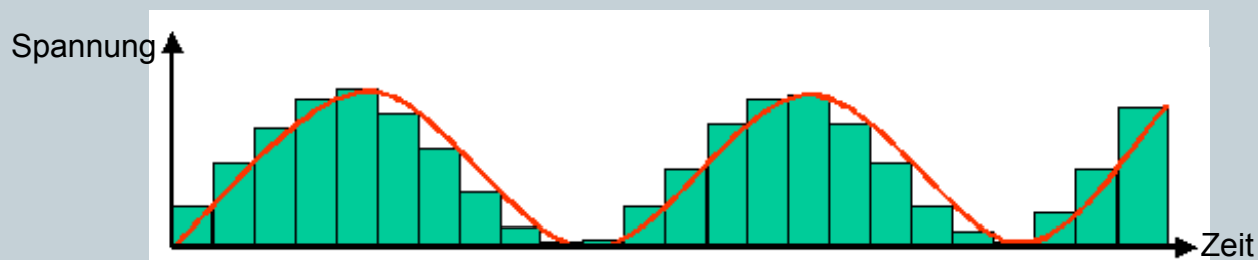
- analoges Signal: wird im Mikrofon erzeugt
- digitale Signale: werden durch Abtastung und Quantisierung der Analogwerte gewonnen
- Analog-Digital Wandlung



# Darstellung: Töne

16

- Töne sind analog – Digitalisierung
- Messung zu bestimmten (diskreten) Zeitpunkten – Sampling
- Qualität ist abhängig von
  - Samplingrate (Messungen pro Sekunde)
    - ✦ sollte doppelt so häufig geschehen, wie höchste Signalfrequenz, entsprechend Abtasttheorem, Nyquist-Shannon
    - ✦  $\approx 20\text{kHz} \rightarrow 44.1$  oder besser  $48\text{kHz}$  Abtastung, pro Kanal
    - ✦ CD-Qualität:  $44100\text{ Hz}$
    - ✦ Abtastwert = Sample
  - Samplingtiefe (Genauigkeit der Messung):
    - ✦ 8 bit: 256 Stufen, 16 bit = 65536 Stufen



Quelle: <http://w5.cs.uni-sb.de/~butz/teaching/hbks-SS00/hbks-ss00-06/sld002.htm>



# Übertragung von Audio: Streaming

17

- **Audio Streaming: netzbasierte Audioübertragung**
  - Vorteil: unbegrenzte Anzahl Radiosender möglich, keine frequenzabhängigkeit
  - Begrenzung: max. Datenübertragungsrate
- **Qualität und Datenrate:**
  - CD:  $44.1\text{KHz} \times 16 \text{ Bit} \times 2 \text{ Kanäle} = 176\text{Kbyte/s} = 1408\text{Kbit/s}$
  - uralter PC:  $22\text{KHz} \times 8 \text{ Bit} \times 1 = 22\text{Kbyte/s}$
  - ISDN:  $8\text{KHz} \times 8 \text{ Bit} \times 1 = 8\text{Kbyte/s}$
  - DSL16000:  $= 2000\text{Kbyte/s} = 16000\text{Kbit/s}$
- **Reduzierung notwendig: Kompression**
  - MPEG3: 1:10 bis 1:20

# Codec

18

- Kunstwort aus **C**oder und **D**ecoder
- Verfahren, oft auch Software zur digitalen (De)Kodierung von Signalen
- Analog-Digital-Analog Wandlung
- Codecs existieren hauptsächlich für:
  - Audio: MP3, Ogg Vorbis, FLAC, Lame, RealAudio
  - Video: TMPGEnc (nur Encoder), Ogg Theodora, DivX
  - Sprache, Telefonie: G.711 (Festnetz), GSM (Mobil)

# WAV

19

- Containerformat zur digitalen Speicherung von Audiodaten
- enthält meist unkomprimierte Daten (PCM)
- auch geeignet für MP3
- Berechnung der Größe einer WAV-Datei:
  - pro Sekunde: Abtastrate · Bytes pro Sample · Anzahl der Kanäle (mono = 1, stereo = 2) Bytes
  - Stereo, Abtastrate 44.1 kHz, 16Bit pro Sample, 5min:
  - $2 \cdot 44100 \cdot 2 \cdot 300 = 52920000$  Byte  $\approx$  50 Mbyte

# Audio CD

20

- Aufzeichnung erfolgt mit
  - Samplingrate: 44.1KHz
  - Samplingtiefe: 16Bit
  - zwei Kanälen
  - $\rightarrow 44100 * 2 \text{Byte} * 2 \text{Kanäle} = 176400 \text{Byte/s}$
- Oft bei älteren CDs anzutreffen: Abkürzungen AAD, ADD, DDD
  - Hinweis ob Aufnahmeschritte (Aufnahme, Mischung, Premaster) analog oder digital erstellt wurden



# Audiokompression

21

- wie bei Bildkompression: verlustfrei, verlustbehaftet
- oft asynchrone Verfahren: Komprimierung aufwendiger als Dekomprimierung
- verlustfrei:
  - kaum benutzt, Datenreduktion von ca. 25-70%
  - Qualität bleibt erhalten
- verlustbehaftet
  - Basis: psychoakustisches Modell
  - leiser Ton nach lautem Ton nicht hörbar, gewisser Unterschied in der Frequenz notwendig
  - Einteilung nach Bitrate, die verwendet werden kann ohne *hörbaren* Qualitätsverlust

# MP3

22

- MPEG-1 Layer 3 Audio Codec (MPEG-1, MPEG-2, Lame)
- ab 1985 bei Fraunhofer in Erlangen entwickelt
- verlustbehaftetes Verfahren
- Ausnutzung des psychoakustischen Modells
- erlaubt Datenraten von 8...320 kBit/s
- Qualität und Datenrate hängen voneinander ab
  - Kodierung mit konstanter Datenrate = schwankender Qualität
  - gleichbleibende Qualität = schwankende Bitrate

# Ogg

23

- freies Containerformat für Audio, Video, Multimedia
- Xiph.org Foundation seit 1993
- wesentliches Ziel: speichern und streamen von Multimedia-Inhalten
- benutzt verschiedene Codecs:
  - Vorbis für Audio
  - Theora für Video
  - Speex für Sprache
  - FLAC für Audio

# Ogg-FLAC

24

- FLAC: Free Lossless Audio Compression
- Erste Veröffentlichung: 20. Juli 2001
- Codec zur verlustfreien Kompression
- Unterschiedliche Kompressionsraten einstellbar: 0 bis 7
  - Je größer die Kompressionseinstellung desto langsamer ist der Vorgang
- Kompression auf 30% bis 75% der Ausgangsgröße
- Freie Software





# Ogg-Vorbis

25

- freier Codec zur verlustbehafteten Audiokompression
  - Entwickelt seit 2000, erste Version: 2002
  - Als freie Alternative zu MP3 entwickelt
- effizienter als MP3 aber weniger verbreitet
  - 25% höhere Kompression bei gleicher Qualität
- unterstützt bis zu 256 Kanäle
- Datenrate: 16-500kBit/s, prinzipiell unbegrenzt
- Hardwareunterstützung noch wenig verbreitet

# Dolby Digital – AC3

26

- Verfahren zur Kodierung von Surround Kanälen hauptsächlich für Kino, DVD gedacht
- AC3 = Audio Code 3
- verlustbehaftete Kompression von Surround 5.1
- Bitraten zwischen 32 und 640KBit/s

# MIDI

27

- Music Instrument Digital Interface
- Protokoll zur Übermittlung, Aufzeichnung und Wiedergabe von Audiodaten
- beinhaltet keine Klänge
- Befehle zur Steuerung digitaler Instrumente oder Soundkarte
  - Note-on, Velocity, Note-off, Instrument
- Vorteil von MIDI:
  - Musikdateien sind sehr klein
  - "Transformationen" sind einfach möglich
- MIDI ist mit digitalem Notenblatt vergleichbar

# Metadaten

28

- für MP3 Dateien entwickelt: ID3-Tag
  - ID3 – Identify an MP3
- erlaubt Einfügen von zusätzlichen Informationen zu einer Audio-Datei
  - Songtitel, Künstler, Album, Erscheinungsjahr, Genre
- abgelegt seit Version 2 am Kopf der Audio-Datei als Header
- Auch für OGG-Dateien benutzbar

# Anwendung

29

- **Internetradio: Audio Streaming:**
  - Unterscheidung zwischen:
    - ✦ Reinen Internetsendern
    - ✦ Internet alternativ zu terrestrischer Verbreitung
  - Verzeichnisse für Internet-Radiostationen
    - ✦ <http://www.surfmusik.de/>
    - ✦ <http://www.shoutcast.com>
- **Privates Audio Streaming:**
  - Verbreitung von Musik über Heimnetz (WLAN)
  - Benötigt Musikserver+entsprechender Streaming-Software:
    - ✦ <http://de.wikipedia.org/wiki/Icecast>
    - ✦ <http://de.wikipedia.org/wiki/SHOUTcast>
- **Podcasts:**

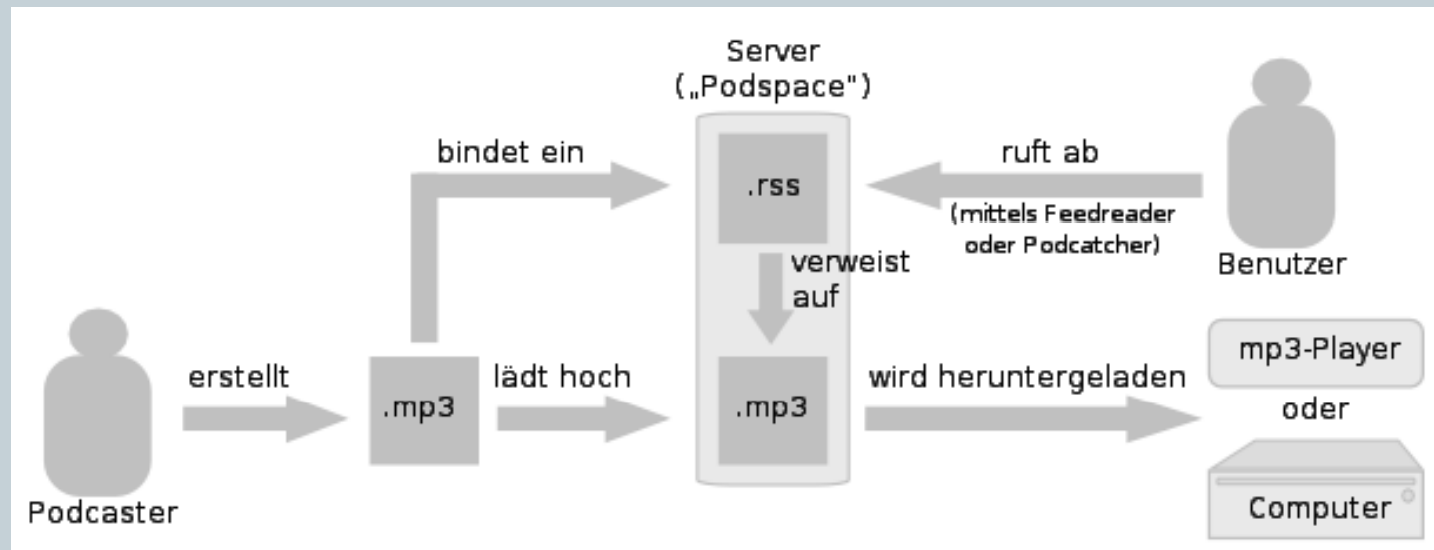
# Podcasts

30

- **Anbieten von Mediendateien im Internet**
  - Kunstwort aus: **IPod** und **Broadcast**
  - Meist für Audio: Video → Videocast
  - Meist kostenlos
  - „Audioblogging“, individualisierter Hörfunk
- **Prinzip:**
  - Anbieter erstellt Mediendatei
  - Stellt diese mittels NewsFeed auf Server im Internet bereit
  - Teilnehmer kann diesen NewsFeed abonnieren
    - ✦ Software überprüft in regelmäßigen Abständen, ob neuer Beitrag vorhanden ist

# Podcasting

31



# Digital Rights Management (DRM)

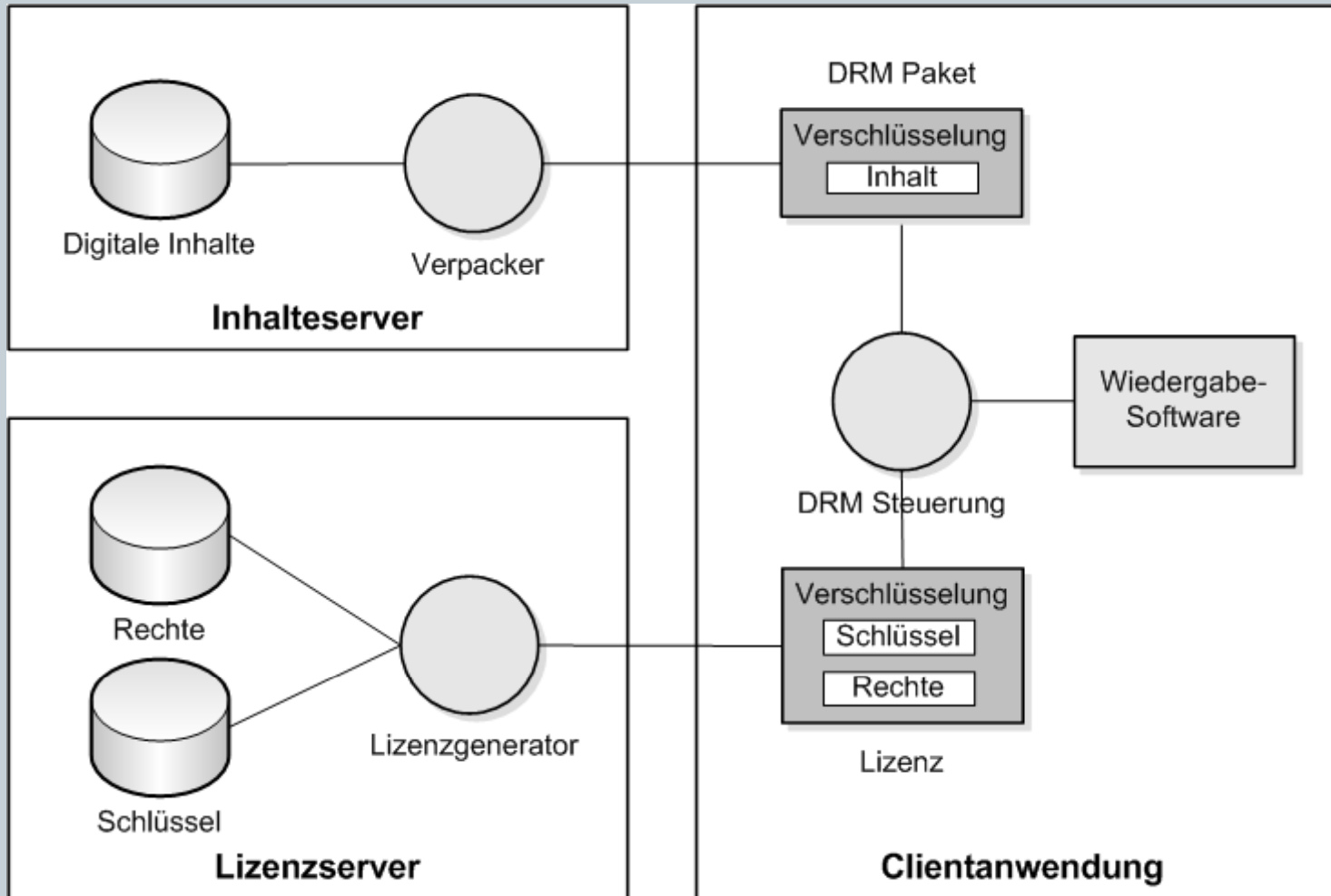
32

- Schutzmechanismen für digitale Informationen
- Ziel: Kontrolle der Nutzung und Lizenzierung geistigen Eigentums
- Beispiele: FairPlay (Apple), OMA DRM (Open Mobile Alliance), Windows Media DRM (Microsoft)
- Realisierung:
  - Information wird verschlüsselt → nicht lesbar für Empfänger
  - Empfänger benötigt Schlüssel zum dechiffrieren der Information
  - Schlüssel wird nach erfolgreicher Authentizitäts- und Integritätsprüfung erteilt



# Digital Rights Management (DRM)

33



# Zusammenfassung

34

- Audio-Signale analog zur rechnergestützten Darstellung → Digitalisierung
  - Abtastung und Quantifizierung
  - typ. Abtastrate: CD-Qualität 44.1 kHz
  - hohe Datenmenge → Kompression
- Verlustfrei: FLAC
- Verlustbehaftet: MP3, OGG, AC3
  - Benutzung des psychoakustischen Modells
  - Original nicht wieder herstellbar!
- Anwendung: Reduzierung der zu übertragenden Datenmenge

# Literatur

wenn nicht auf der entsprechenden Folie angegeben

35

- Kai Bruns, Klaus Meyer-Wegener: "Taschenbuch der Medieninformatik", Fachbuchverlag Leipzig, 2005
- <http://www.fh-wedel.de/~si/seminare/ss02/Ausarbeitung/9.digitalaudio/audio1.htm>
- OGG-Vorbis: <http://www.xiph.org/vorbis>
- MP3: <http://de.wikipedia.org/wiki/Mp3>
- MIDI: <http://de.wikipedia.org/wiki/MIDI>
- ID3-Tag: <http://de.wikipedia.org/wiki/ID3-Tag>
- <http://de.wikipedia.org/wiki/Ac3>