Acoustics II:
sound storage media

Kurt Heutschi
2013-01-18
sound storage media: introduction

- main building blocks of a sound storage device:

```
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>recording</td>
<td>converter</td>
</tr>
<tr>
<td></td>
<td>equalizer</td>
<td>converter</td>
</tr>
<tr>
<td></td>
<td></td>
<td>storage</td>
</tr>
<tr>
<td></td>
<td></td>
<td>medium</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

- sound storage media
- vinyl records
- analog tape recorder
- Dolby noise reduction
- compact disc
- DVD Audio, Super Audio CD
- harddisc recorder
- data compression
- motivation
- lossless compression
- lossy compression
- principal structure
- MPEG-1 coding
- sound examples
- back
Vinyl records
sound storage media

vinyl records

- concept: signal is stored as geometrical form on rotating disc
- basic idea: Phonograph by Edison, 1877 (groove on cylinder with vertical modulation)
- disc: Grammophon by Berliner, 1887 (groove with lateral modulation)
- stereo disc: system 45-45 by Westrex
Vinyl records

- **transducer principle:** electrodynamic
  - **signal voltage** $\sim$ velocity of needle
  - **excursion** $\sim 1/f$
  - **at high frequencies only small excursion amplitudes**
    - **bad signal/noise ratio**
  - **$\rightarrow$ compensation:**
    - **during recording:** amplification of high frequencies
    - **during play-back:** attenuation of high frequencies
    - **$\rightarrow$ RIAA equalization**
Vinyl records: RIAA equalization
Vinyl records: properties

- playing time: LP ca. 20...25 minutes per side
- dynamic range: typ. 50 dB(A) (max. 70 dB(A))
- upper limiting frequency: typ. 12...15 kHz
- lower limiting frequency: typ. 40 Hz
Turntable example: Revox B790

tangential tone arm:
analog tape recorder
analog tape recorder

- concept: signal is stored longitudinally as magnetization on a tape
- first realization (steel wire): Telegraphon by Poulsen, 1898
Analog tape recorder: construction

- **challenges:**
  - constant tape speed
  - geometry of the tape guidance
  - linearization of the magnetization hysteresis
Analog tape recorder: magnetization

- principal relation between magnetic field $H$ and induction $B$:
Analog tape recorder: magnetization

- storage property arises by remanence $B_R$ (remaining induction after discontinuation of $H$):
Analog tape recorder: magnetization

- linearization: operating point $\neq 0$
- → biasing
  - DC-biasing:
    - addition of a DC current
  - AC-biasing:
    - addition of a high-frequency (50... 150 kHz) AC current
Analog tape recorder: biasing
Analog tape recorder: DC-biasing

disadvantages of DC biasing:

- still relative large distortions
- asymmetric characteristic curve $\rightarrow$ odd and even harmonics
- $S/N$ ratio not optimal as only one half of the remanence curve is used
Analog tape recorder: AC-biasing

advantages of AC biasing:

- symmetrical characteristic curve $\rightarrow$ odd harmonics only
- $S/N$ ratio 6 dB higher
Analog tape recorder: frequency response

- longitudinal writing
- wavelength $\lambda$ on tape:

$$\lambda = \frac{v}{f}$$

- $v$: tape speed
- $f$: signal frequency

- upper end of the frequency range:
  - violation of the condition $\lambda \gg$ tape head width
Analog tape recorder: frequency response

- typical tape speeds: 38.1 cm/s, 19.05 cm/s, 9.5 cm/s
- typical tape head widths: 5 µm
- upper limiting frequency: easily > 20 kHz
Analog tape recorder: play-back equalizer

- recording:
  - stored signal on tape $\sim$ audio signal current

- play-back:
  - induced signal voltage $\sim$ change of magnetic flux
  - $\rightarrow \omega$-proportional amplitude response
  - correction *play-back equalizer*
Analog tape recorder: properties

- playing time: order of an hour
- dynamic range: up to 70 dB(A), with Dolby SR > 90 dB(A)
- upper limiting frequency: > 20 kHz
- lower limiting frequency: typ. 30 Hz
Dolby noise reduction
Dolby: introduction

- motivation:
  - insufficient S/N ratio of the analog type recorder
  - → noise becomes audible during low level sections

- strategy:
  - amplification of low level sections before recording
  - attenuation of low level sections during play-back
  - compander-expander-system
Dolby: principle

- $G_1, G_2$: signal steered amplifiers
- for $G_1 = G_2 \rightarrow$ audio OUT = audio IN
Dolby: variants

- differences of various systems:
  - signal steered amplifiers $G$
  - number of frequency bands processed individually:
    - consumer systems usually use single-band (e.g. Dolby B)
    - professional systems use multi-band (e.g. Dolby A, SR)
Dolby: type B

- signal path through variable filter:
  - amplification of low level mid- and high frequency components
Dolby: type B

characteristics of variable filter:

difficulty: time constant for adjustment of amplification
Dolby: improvement of dynamics

- Dolby B: approx. 10 dB
- Dolby SR: approx. 25 dB
compact disc
compact disc

- concept: signal stored digitally as pattern of depressions in a plastic disc
- sampling frequency: 44.1 kHz
- resolution: 16 Bit
Compact Disc: construction

- medium: plastic disc of 1.2 mm thickness and 12 cm diameter
- composition of several layers (top down):
  - protection layer with imprint
  - metal layer of 50...100 nm thickness
  - transparent layer

![Diagram of Compact Disc layers](image)
Compact Disc: construction

- information coding:
  - grooves of different lengths arranged on a spiral starting at the center of the disc:
    - width: 0.5 \( \mu \)m
    - length: 1 or 3 \( \mu \)m
    - depth: 125 nm
  - scanning electron microscope picture:
Compact Disc: reading process

- disc rotates at 500 r.p.m. (center) and slows down to 200 r.p.m (circumference)
- grooves sensed optically from the bottom
- groove depth $\lambda/4 \rightarrow$ destructive interference with reflection at top of surface
- reading head tracks the grooves spiral
- constant data stream: 4.3218 Mbit/s
Compact Disc: data coding: error correction

- typical error rate: $10^{-6}$
- error correction by introduction of redundancy:
  - CIRC (cross-interleave Reed-Solomon code)
  - very efficient (3/4 information, 1/4 redundancy)
  - up to 1 mm gaps can be reconstructed
Compact Disc: data coding: modulation

- requirements:
  - no too high-frequency components (interference between symbols)
  - not too low frequency components (inaccurate clock reconstruction)
- solution: EFM (Eight to Fourteen Modulation):
  - 1 Byte data coded as 14 Bit symbol
  - transition between symbols by 3 additional Bits
Compact Disc: properties

- Playing time: 74 minutes
- Dynamic range: > 90 dB(A)
- Upper limiting frequency: 20 kHz
- Lower limiting frequency: 20 Hz
DVD Audio, Super Audio CD
DVD Audio, Super Audio CD

- storage capacity: 4.7 GByte (7 × CD)
- usable for:
  - increased amplitude resolution → quantization with 24 Bit
  - increased sampling frequency → 192 kHz
  - increased number of channels → surround sound, 6 channels
DVD Audio

- single-, double-layer (4.7, 8.5 Gbyte)
- no data reduction or lossless coding (Meridian Lossless Packing (25...50% reduction))
- perceptual coding possible (Dolby Digital, MPEG...)
- quantization: 16, 20, 24 Bit
- sampling frequency: 44.1...192 kHz
- channels: 2...6
DVD Audio

- playing time single-layer:
  - Stereo, 24 Bit, 192 kHz → 64 minutes
  - 5 channel Surround, 20 Bit, 96 kHz → 61 minutes
  - Stereo, 16 Bit, 44.1 kHz → 8 hours
  - 5 channel Surround Dolby Digital → 35 hours

- DVD-Audio will disappear, no now releases
Super Audio CD

- development Sony/Philips
- capacity 4.7 Gbyte (identical to DVD)
- up to 6 channels, up to 100 kHz bandwidth, up to 120 dB dynamic range
- hybrid version with additional layer for reproduction in ordinary CD player
harddisc recorder
harddisc recorder

- concept: signal stored digitally on harddisc or memory card
- Laptop based solution with audio-interface
- stand-alone solutions
- sampling frequencies: up to 192 kHz
- resolution: up to 24 Bit
Harddisc Recorder: example stand-alone device
data compression
Data compression: motivation
Data compression: motivation

- digital audio material is quite demanding regarding memory space
- compression (data reduction) is of interest in the context of
  - storage
  - transmission
- types of compression:
  - lossless
    - avoids redundancy
    - perfect signal reconstruction (sample by sample) is possible
  - lossy
    - elimination of inaudible components
    - only reasonably well sounding reconstruction is possible
data compression: lossless compression
lossless compression

- elimination of redundant information
  - redundancy: samples so far give some information about future samples
    - example for high redundancy: sinusoidal signal
    - example for zero redundancy: white noise
  - typical audio signals: relatively little redundancy → potential in order of about 50%
Lossless compression

formats:

**Apple Lossless**  ALAC: proprietary format by Apple for the compression of WAV or AIFF Dateien.

**Free Lossless Audio**  FLAC: freely available format of Xiph.Org

**Meridian Lossless Packing**  MLP: proprietary format by Meridian Audio, also known as Dolby Lossless.

data compression:
lossy compression
lossy compression

- requirement: conservation of listening impression $\rightarrow$ significantly higher compression rates possible (in order of 1:10)
- reasons why this can work:
  - ear only evaluates certain signal attributes (e.g. insensitive to phase)
  - temporal and frequency masking $\rightarrow$ certain parts of the signal can be omitted
  - necessary amplitude resolution depends on actual signal value (linear quantization is overkill)
- $\rightarrow$ Perceptual Coding
Lossy compression

formats:

Adaptive Transform Acoustic Coding  ATRAC: proprietary format by Sony, MiniDisc

MPEG-1 Layer 2,3  MP2, MP3: developed by MPEG (Moving Picture Experts Group) MP3: internet applications, portable music players

Advanced Audio Coding  AAC: developed by MPEG as successor of MP3

Windows Media Audio  WMA: proprietary format by Microsoft
principal structure of perceptual coders
principal structure of perceptual Coders

- analysis filter bank → filtered signals $x_1$ to $x_N$
- psychoacoustic model (by FFT) → necessary quantization and coding of $x_1$ to $x_N$
- data organization in frames → data stream $x_c$
Principal structure of Perceptual Coders

- data stream $x_c$ demultiplexing and decoding $\rightarrow$ filtered time signals $x'_1$ to $x'_N$
- synthesis filter bank $\rightarrow$ audio signal $y[n]$
MPEG-1 coding
MPEG-1: overview

- audio coding format developed by the Moving Pictures Expert Group
- variants of MPEG-1:
  - Layer 1
  - Layer 2
  - Layer 3 → mp3
- specified decoders
- variable encoders (respecting the decoder specifications)
MPEG-1 Coding

- analysis filter bank: signal separation into 32 bands of constant absolute bandwidth
  - layer 3: further separation into 6/18 subbands
- amplitude quantization in each frequency band depending on a psychoacoustic model:
  - make usage of masking effect:
    - within a band
    - neighbor bands
  - dynamic bit allocation
MPEG-1 Coding

quantization noise for different resolutions in one frequency band:

- noise due to 16 Bit quantization
- shifted hearing threshold
- noise due to 8 Bit quantization
MPEG-1 Coding

- further compression possibilities:
  - multichannel signals:
    - make usage of the similarity between channels
    - → coding of differences only
  - typical Stereo bit rates for high quality:
    - 128 kBit/s . . . 160 kBit/s
    - → compression: 1:12 . . . 1:10
sound examples
sound examples

examples of 1:10 (192 kBit/s) mp3 coded sounds:

| female voice | A | B |
| guitar       | A | B |
| yello        | A | B |
| chris jones  | A | B |
| drums        | A | B |
| harpsichord  | A | B |
## sound examples

Examples of 1:10 (192 kBit/s) mp3 coded sounds:

<table>
<thead>
<tr>
<th></th>
<th>A: mp3</th>
<th>B: original</th>
</tr>
</thead>
<tbody>
<tr>
<td>female voice</td>
<td>mp3</td>
<td>original</td>
</tr>
<tr>
<td>guitar</td>
<td>original</td>
<td>mp3</td>
</tr>
<tr>
<td>yello</td>
<td>original</td>
<td>mp3</td>
</tr>
<tr>
<td>chris jones</td>
<td>mp3</td>
<td>original</td>
</tr>
<tr>
<td>drums</td>
<td>original</td>
<td>mp3</td>
</tr>
<tr>
<td>harpsichord</td>
<td>original</td>
<td>mp3</td>
</tr>
<tr>
<td>drums</td>
<td>original</td>
<td>mp3</td>
</tr>
<tr>
<td>harpsichord</td>
<td>original</td>
<td>mp3</td>
</tr>
</tbody>
</table>

back
sound examples

examples of 1:12 (128 kBit/s) mp3 coded sounds:

| female voice | A | B |
| guitar      | A | B |
| yello       | A | B |
| chris jones | A | B |
| drums       | A | B |
| harpsichord | A | B |
### sound examples

Examples of 1:12 (128 kBit/s) mp3 coded sounds:

<table>
<thead>
<tr>
<th></th>
<th>A: mp3</th>
<th>A: original</th>
<th>B: mp3</th>
<th>B: original</th>
</tr>
</thead>
<tbody>
<tr>
<td>female voice</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>guitar</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>yello</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>chris jones</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>drums</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>harpsichord</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
sound examples

examples of 1:24 (64 kBit/s) mp3 coded sounds:

<table>
<thead>
<tr>
<th>female voice</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>guitar</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>yello</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>chris jones</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>drums</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>harpsichord</td>
<td>A</td>
<td>B</td>
</tr>
</tbody>
</table>
## sound examples

Examples of 1:24 (64 kBit/s) mp3 coded sounds:

<table>
<thead>
<tr>
<th></th>
<th>A: mp3</th>
<th>B: original</th>
</tr>
</thead>
<tbody>
<tr>
<td>female voice</td>
<td>A: mp3</td>
<td>B: original</td>
</tr>
<tr>
<td>guitar</td>
<td>A: mp3</td>
<td>B: mp3</td>
</tr>
<tr>
<td>yello</td>
<td>A: original</td>
<td>B: mp3</td>
</tr>
<tr>
<td>chris jones</td>
<td>A: original</td>
<td>B: original</td>
</tr>
<tr>
<td>drums</td>
<td>A: original</td>
<td>B: mp3</td>
</tr>
<tr>
<td>harpsichord</td>
<td>A: mp3</td>
<td>B: original</td>
</tr>
</tbody>
</table>
eth-acoustics-2