MPEG-1 lag 1, 2 og lag 3

Sverre Holm

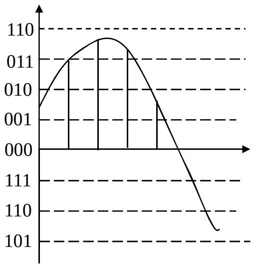
Basert på presentasjon laget av
Torbjörn Ekman, 2005 (nå på NTNU)

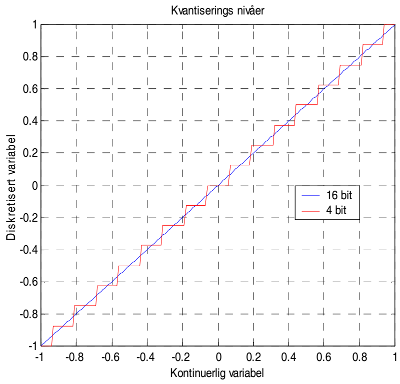
MPEG audiokoding

- Motivasjon for de fleste kapitlene i Ambardar, Digital signal processing: A Modern Introduction, Thomson, 2007.

UNIVERSITETET
I OSLO

Digital representation of Sounds
Pulse Coded Modulation (PCM)





16 bit kvantisering gir

$2*2*2*\ldots 2 = 2^{16} =$ 65 536 nivåer

Ved 44100 samples per sek, blir bitraten:

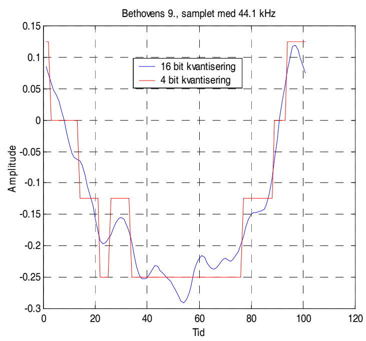
16*44100 = 705 600 bits/s = halv CD-rate

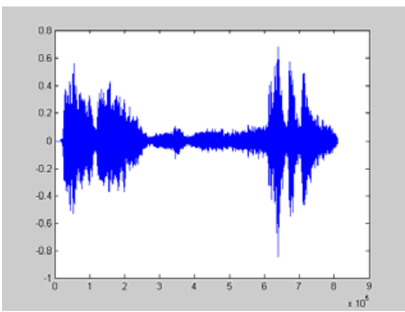
UNIVERSITETET
I OSLO

Beethovens 5. symfoni

16 bit kvantisering
2^{16}= 65536 nivåer

4 bits kvantisering
2^{4}=16 nivåer





UNIVERSITETET
I OSLO

Hvorfor høres det så ille ut?

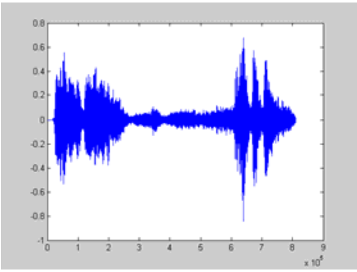
- Problem: Bare noen få kvantiseringsnivåer => for stor avrundingsfeil. Ofte at signalet settes til null da nivået var lavere enn laveste kvantiseringsnivå
- Kvantisering og sampling:
  Kap 7: Digital behandling av analoge signaler

- Mulig løsning: Skaler blokker av data slik at maximumsverdien alltid utnytter hele dynamikkområdet til kvantisereren
- Kostnad: Må sende over skalafaktorer
  - NICAM, Near Instantaneous Companded Audio Multiplex: format for digital lyd over analog TV.
  - Blokklengde 32 samples, 3 bit pr blokk sideinfo. Stereo kodes med 10 av 14 bit ved samplingsrate 32 kHz => 728 kbit/s.
  - Variant av adaptiv differensiell puls kode modulasjon

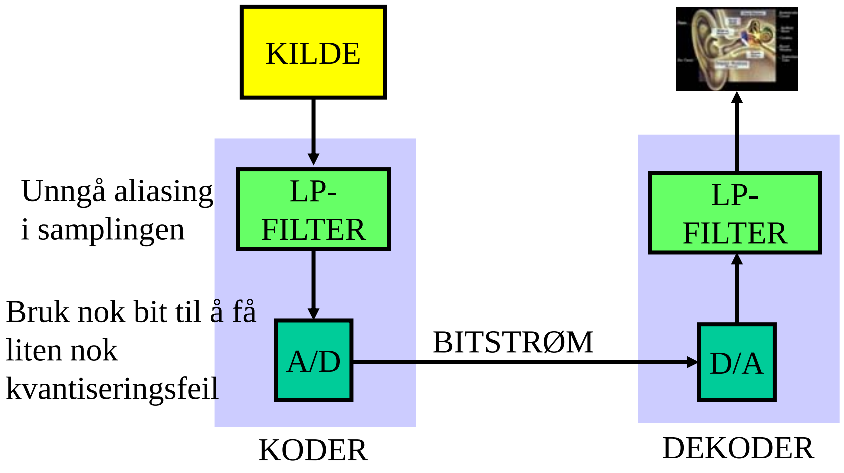
UNIVERSITETET
I OSLO

Stasjonæritet - tidsinvarians

- = Egenskaper varierer ikke med tiden, **kap 3: Tids-domene analyse**
- Forutsettes i det meste av analyser
- Tale er korttids stasjonær, dvs bare over ca 20 ms,
  - Endres 1/20e-3=50 ganger pr sekund



Direktesampling (PCM)

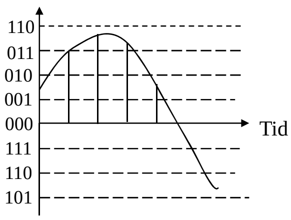


Bitrater

- CD: 44.1 * 2 * 16 = 1.411 Mbit/s
  – 4 bit: 25% => 350 kbit/s låter forferdelig
- MP3, AAC etc: 128 kbit/s ~ CD/12
- Hva er det lure trikset?

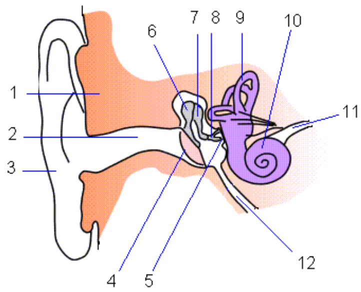
MPEG-1 Audio
Psychoacoustics in sound compression

- How do we hear?
- Digital representation of sounds
- Sound compression
- Psychoacoustics
  - Masking
  - Adaptive quantization
  - Bit allocation
- Filterbanks



Øret

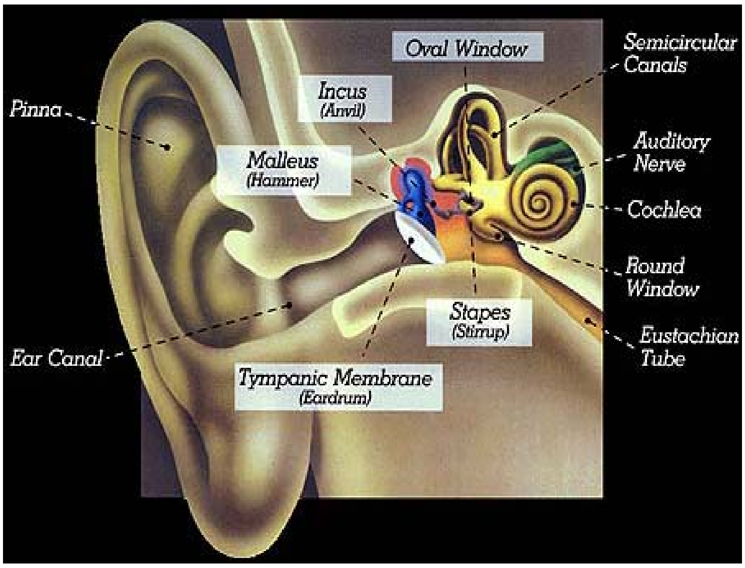
1. Tinning
2. Øregang
3. Ytre øre
4. Trommehinne
5. Ovale vindu
6. Hammeren
7. Ambolt
8. Stigbøylen
9. Bueganger
10. Sneglehuset
11. Hørselnerve
12. Øretrompeten



Wikimedia Commons

UNIVERSITETET
I OSLO

The Ear

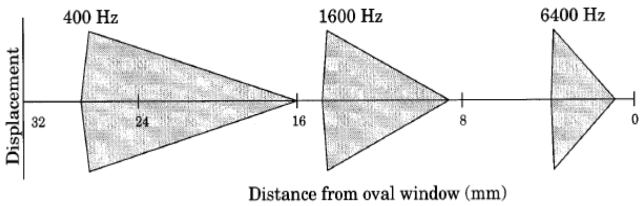


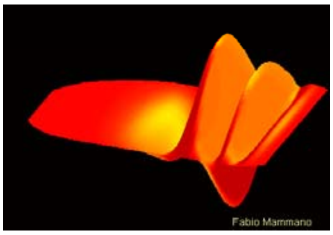
UNIVERSITETET
I OSLO

The frequency filters of the ear: Mapping frequency to a location

Unwound

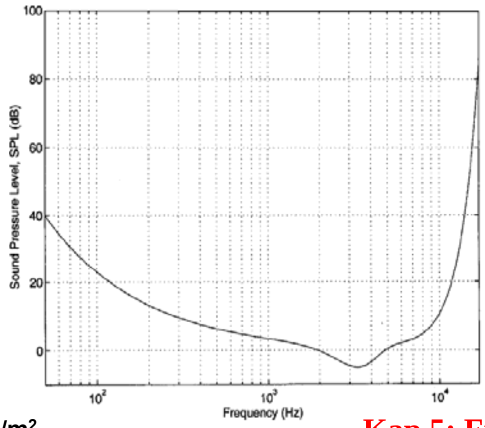
cochlea





Kap 5: Frekvensanalyse

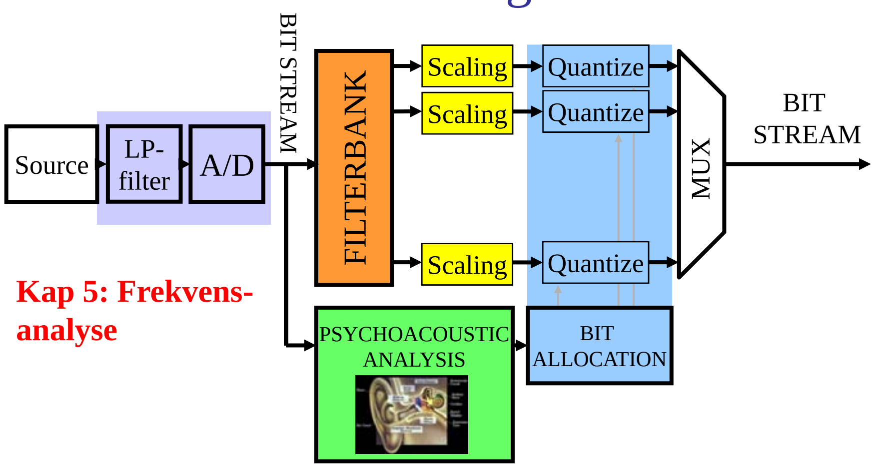
Threshold for audible sounds



Reference 0 dB:
20 μPa = $2{\cdot}10^{-5}$ N/m^2

Kap 5: Frekvensanalyse

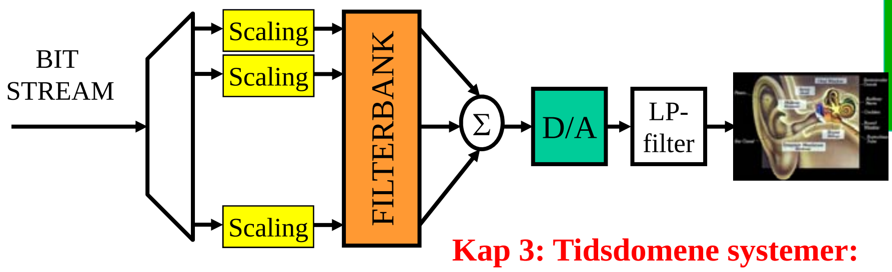
Filterbank Approach Encoding



Kap 5: Frekvens-analyse

UNIVERSITETET
I OSLO

Decoding is much simpler

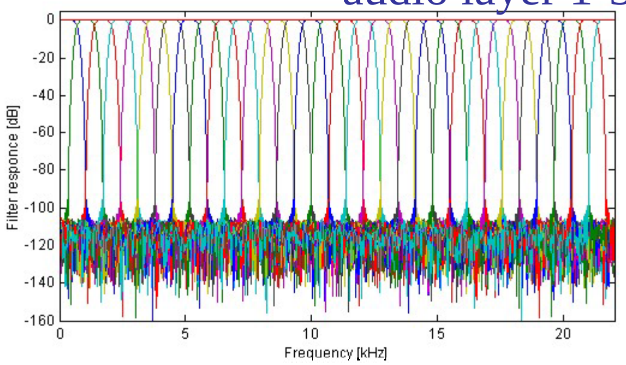


Kap 3: Tidsdomene systemer: linearitet

Kap 3: Inverse systemer

UNIVERSITETET
I OSLO

Filterbanks in MPEG-1 audio layer 1-3



Kap 5: Frekvens-analyse av systemer

Kap 6: Digitale filtre

Kap 10: FIR Filterdesign

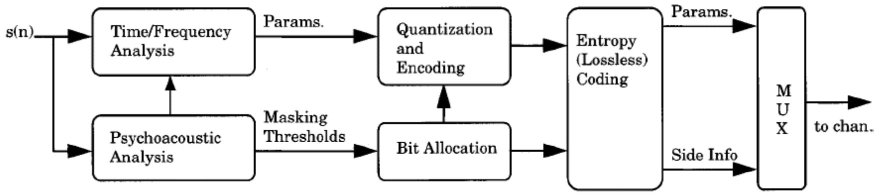
- Polyphase filterbank
- 32 subbands, e.g. bw 44100/2/32 = 689 Hz
- 512 tap FIR-filters
- 80 adds and mults per output

- Equal width
- Not perfect reconstruction
- Frequency overlap

Kap 4: z-transform

UNIVERSITETET
I OSLO

What is this Psychoacoustics that is used in the Encoder ?



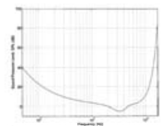
Kap 8: Diskret Fourier Transform; Estimering av effektspektrum

UNIVERSITETET
I OSLO

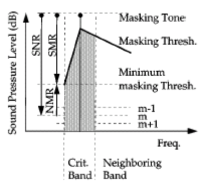
Masking

We do not hear all sounds.

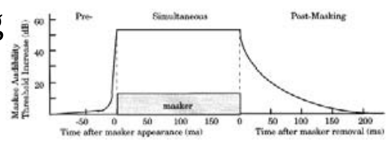
1. Absolute threshold of hearing: →



2. Masking: One sound is inaudible in the presence of another sound.
   1. Simultaneous masking
      - Noise Masking Tone
      - Tone Masking Noise
      - Noise Masking Noise

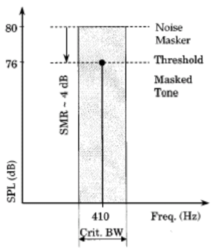


   2. Nonsimultaneous masking
      - Pre masking (2 ms)
      - Post masking (100 ms)



Noise Masking Tone

Filtered Noise Center 410 Hz Width 111 Hz	Tone 1, 820 Hz 5 dB below noise	Tone 2, 410 Hz 5 dB below noise	Noise + Tone 1	Noise + Tone 2
			Not masked	Masked

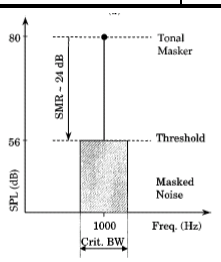


You can not hear a sinusoid that lies in the same critical band as a filtered noise if the sound pressure level is below a certain threshold.

This effect also stretches out beyond the critical band.

Tone Masking Noise

Filtered Noise Center 1 kHz Width 162 Hz 15 dB below	Tone 1, 2 kHz	Tone 2, 1 kHz	Noise + Tone 1	Noise + Tone 2
			Not masked	Masked



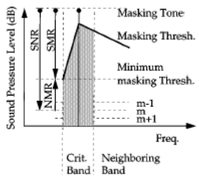
You can not hear a filtered noise that lies in the same critical band as a sinusoid if the sound pressure level is below a certain threshold.

This effect also stretches out beyond the critical band.

UNIVERSITETET
I OSLO

Exploit Masking

- If a sound is masked we can’t hear it.



- Make a frequency analysis of the signal and find the masking threshold.
- Put the quantization noise under the masking threshold and we won’t hear the quantization.

Kap 8: DFT, Fast Fourier transform, Estimering av effektspektrum

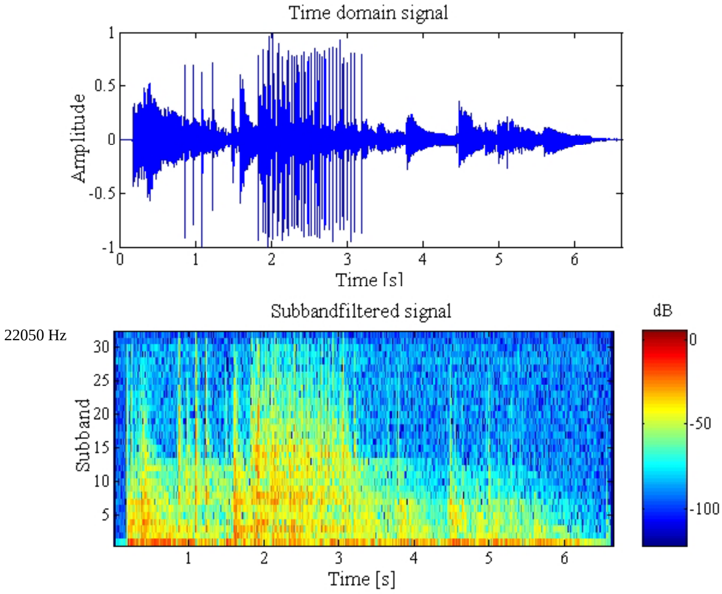
UNIVERSITETET
I OSLO

Bit Allocation and Masking

- The masking threshold in each subband gives the Just Noticeable Distortion (JND) limit for that band.
- Bits are assigned to subbands so that the quantization noise falls below or as little over the JND as possible.
- Then the Signal to Quantization noise Ratio (SQR) falls below JND

UNIVERSITETET
I OSLO

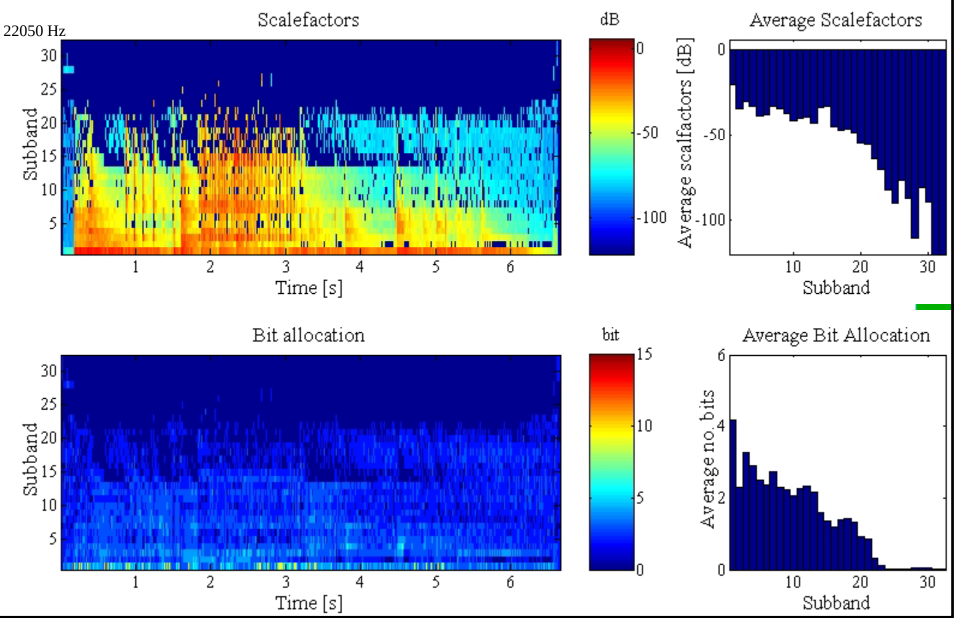
Castanets and Guitar



INSTITUTT FOR INFORMATIKK

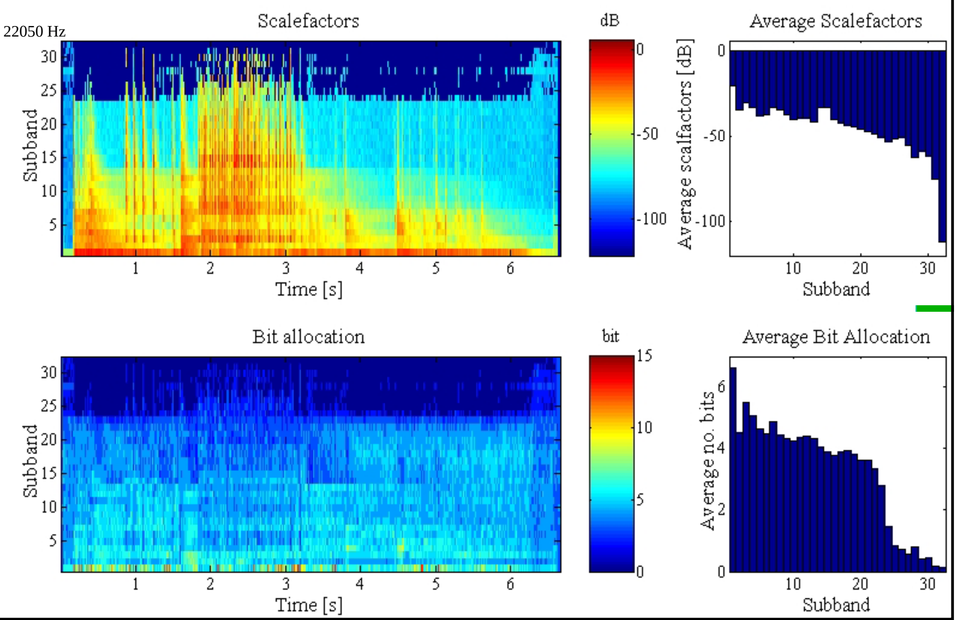
UNIVERSITETET
I OSLO

Bit allocation with 2 bits per sample



UNIVERSITETET I OSLO

Bit allocation with 4 bits per sample



UNIVERSITETET I OSLO

Kapitler i læreboka – MP3-koding

1	Overview	• Praktiske eksempler som mp1/mp3
2	Discrete Signals	• Musikk og talesignal
3	Time-Domain Analysis	• Linearitet: kan dele i frekvensbånd, behandle for seg og så sette sammen igjen • Tidsinvarians gjelder bare over kort tid for musikk og tale. • Differanseligninger: FIR filtre
4	z-Transform Analysis	• Analyse av filter i filterbanken: nøkkel til å få til filterdesign
5	Frequency Domain Analysis	• Frekvensdomene er sentralt i modell av hørsel • Frekvensselektive filtre: båndpassfiltre • Inverse systemer: kan dele i bånd i koder og addere sammen igjen i dekoder
6	Filter Concepts	• Filterstrukturer, hvordan implementere filterbank i koder og dekoder
7	Digital Processing of Analog Signals	• A/D-analyse: kvantiseringsstøy ved direkte sampling • Mutirate system: Hvert delfilter kan nedsamples pga bare 1/32 av total båndbredde => trenger bare 1/32 samplerate per filter
8	The Discrete Fourier Transform and Its Applications.	• Frekvensanalyse av signaler • FFT brukes i estimering av spektrum i koder. • Må estimere korttidsspektrum for å gjøre adaptiv bittildeling
9	Design of IIR Filters.	
10	Design of FIR Filters.	• Hvordan finne koeffisienter til bp-filtrene i filterbanken?
11	MATLAB Examples	
A	Useful Concepts from Analog Theory	