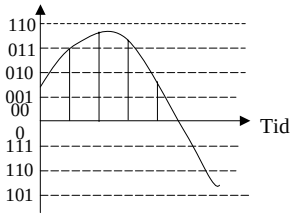
UNIVERSITETET
I OSLO

MPEG audio
Psykoakustikk som grunnlag for lydkompresjon

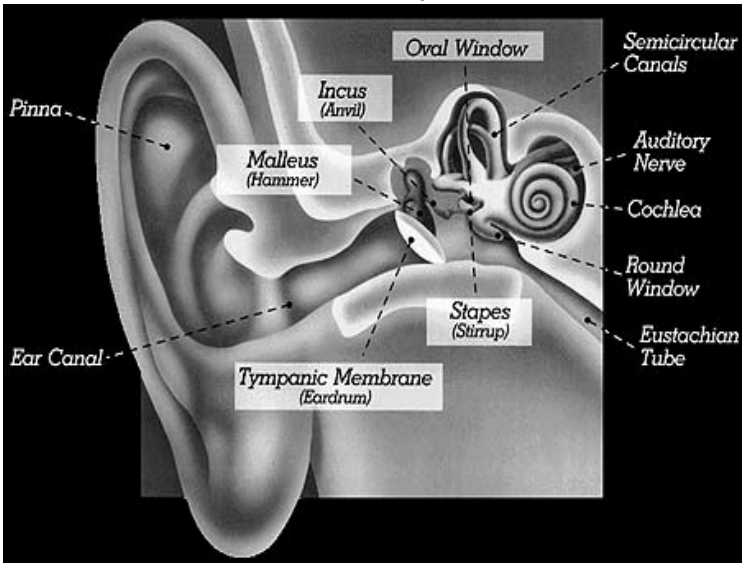
Torbjørn Ekman og Nils Christophersen

- Hvordan oppfatter vi lyd
- Digital representasjon av lyd
- Lydkompresjon
  - Rett frem og dumt
  - Lurt med psykoakustikk



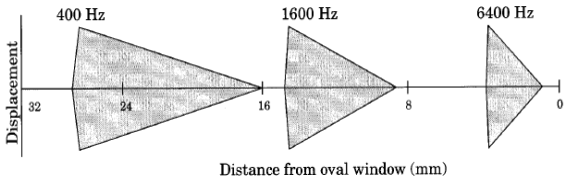
UNIVERSITETET
I OSLO

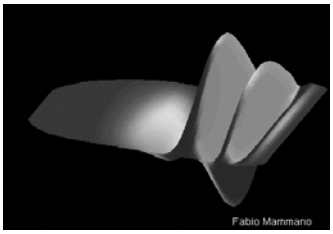
Det indre øret



UNIVERSITETET
I OSLO

Ørets båndpassfilter – ”frekvens til sted”

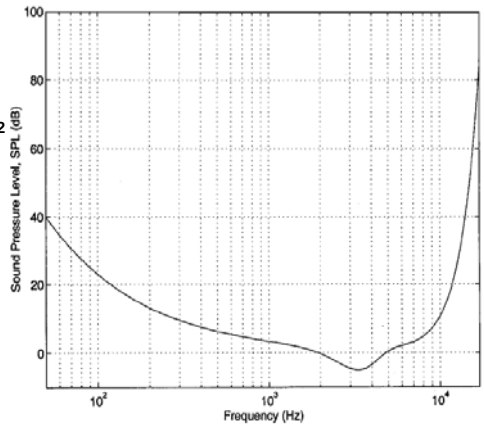




UNIVERSITETET
I OSLO

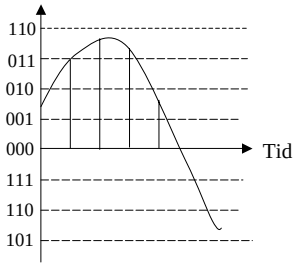
Terskel for hørbar lyd

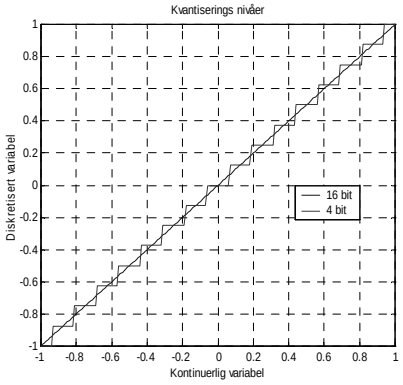
Referansen
20 µPa = 2 10^{-5} N/m^2



UNIVERSITETET
I OSLO

Digital representasjon av lyd





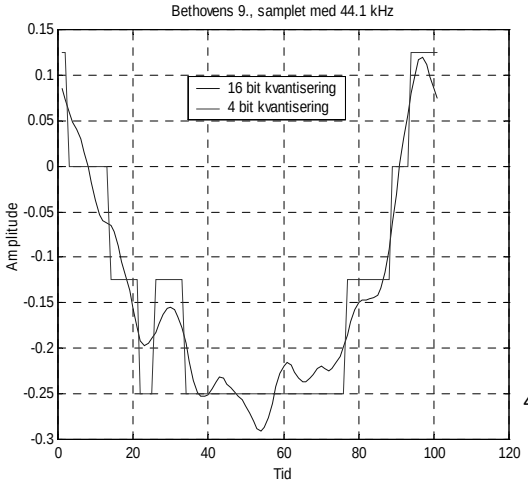
Med 16 bit får vi

$2*2*2*\ldots2 = 2^{16} =$ 65 536 nivåer

Dette betyr 16*44100 = 705 600 bits/sek

UNIVERSITETET
I OSLO

Beethovens 9. symfoni

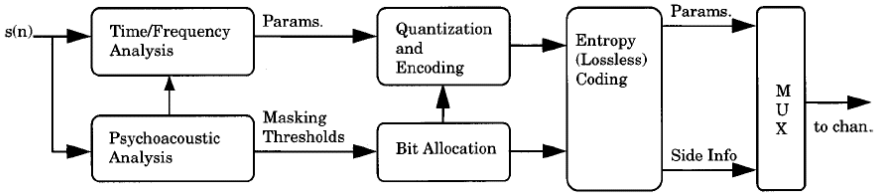


16 bit kvantisering

4 bit kvantisering (dum og dårlig)

UNIVERSITETET
I OSLO

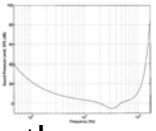
What is this Psychoacoustics that is used in the Encoder ?



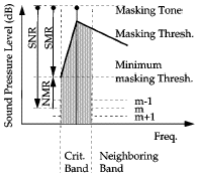
UNIVERSITETET
I OSLO

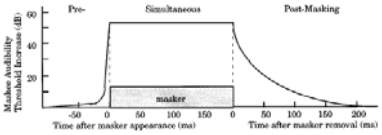
Masking

We do not hear all sounds.



- Absolute threshold of hearing.
- 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)

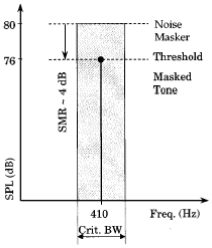




UNIVERSITETET I OSLO

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 



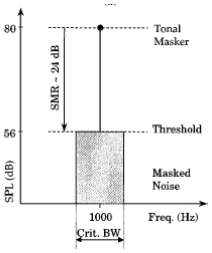
You can not hear a sinusoid that lies in the same critical band as a filtered noise if the soundpreasure level is below a certain threshold.

This effect also stretches out beyond the critical band.

UNIVERSITETET I OSLO

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 



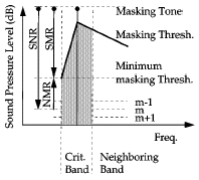
You can not hear a filtered noise that lies in the same critical band as a sinusoid if the soundpreasure level is below a certain threshold.

This effect also stretches out beyond the critical band.

UNIVERSITETET I OSLO

Exploit Masking

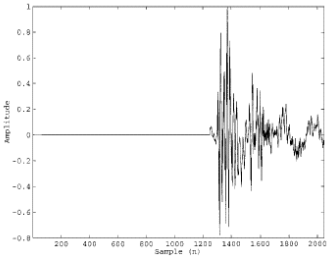
- If a sound is masked we can't hear it.

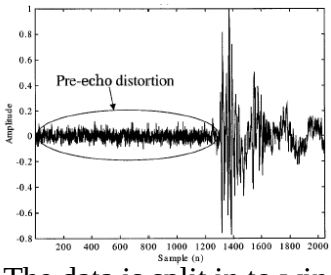


- Make a frequency analyze of the signal and find the masking threshold.
- Put the quantization noise under the masking threshold and we don't hear the quantization.

UNIVERSITETET I OSLO

Pre echo distortion





- The original sound of a castanet.
- The abruptness in time domain result in that all frequencies are involved.

- The data is split in to windows of finite length.
- The quantization noise is spread over a whole window.
- This makes the castanets sound less distinct.
- Audible effects can be avoided with shorter windows, exploiting premasking.

Scale factors and Quantization

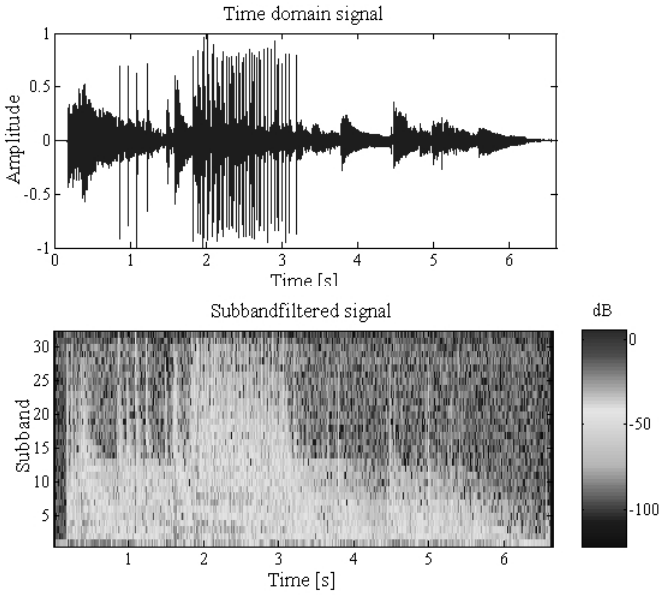
- When the dynamics change over time, only a small subset of the quantization steps are used in regions with low magnitudes.
- Use scale factors instead:
  - Take a window of data.
  - Find the max magnitude in this window.
    - Use the next larger scale factor from a table.
  - Normalize with the scale factor.
  - Quantize.
    - Now the whole dynamic range of the quantizer is used.
  - Send scale factor and quantized samples.

Bit Allocation and Masking

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

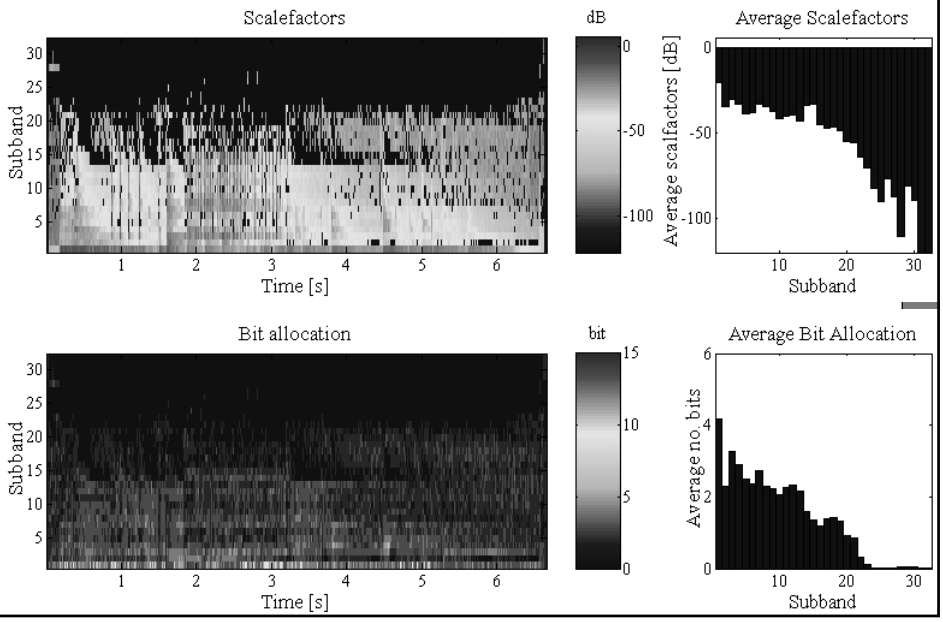
UNIVEI
I OSLO

Castanets and Guitar



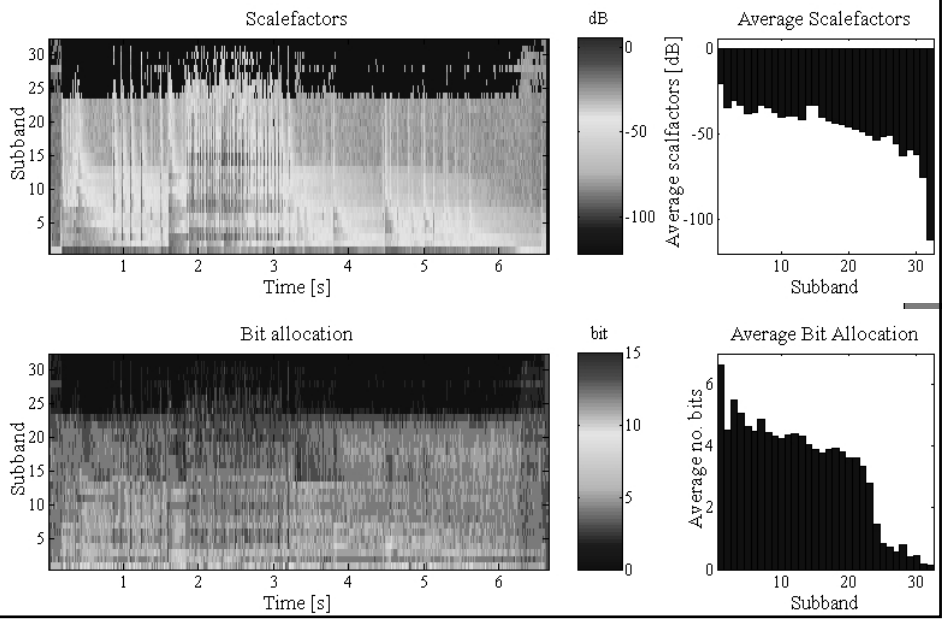
UNIVERSITETET
I OSLO

Bit allocation with 2 bits per sample



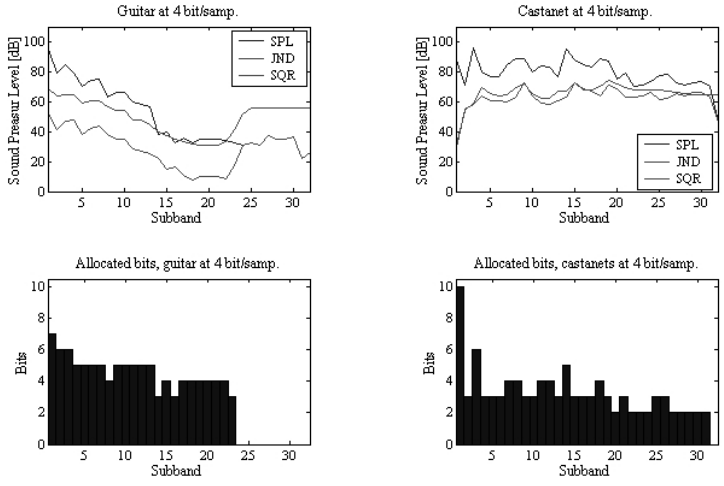
UNIVERSITETET I OSLO

Bit allocation with 4 bits per sample



UNIVERSITETET I OSLO

Signal to Quantization Noise Ratio and the Just Noticeable Distortion



Frame at t=0.6 s

Frame at t=1.1 s

UNIVERSITETET I OSLO

Examples on compression

Compression	2	4	8
MP1		4 bit	2 bit
MP1 error (SQR)		22 dB	11 dB
Direct Quantization	8 bit	4bit	2 bit
Direct Quantization Error (SQR)	31 dB	7.8 dB	1.1 dB
Downsampling to 22 kHz bandwidth and quantization	16 bit	8 bit	4 bit