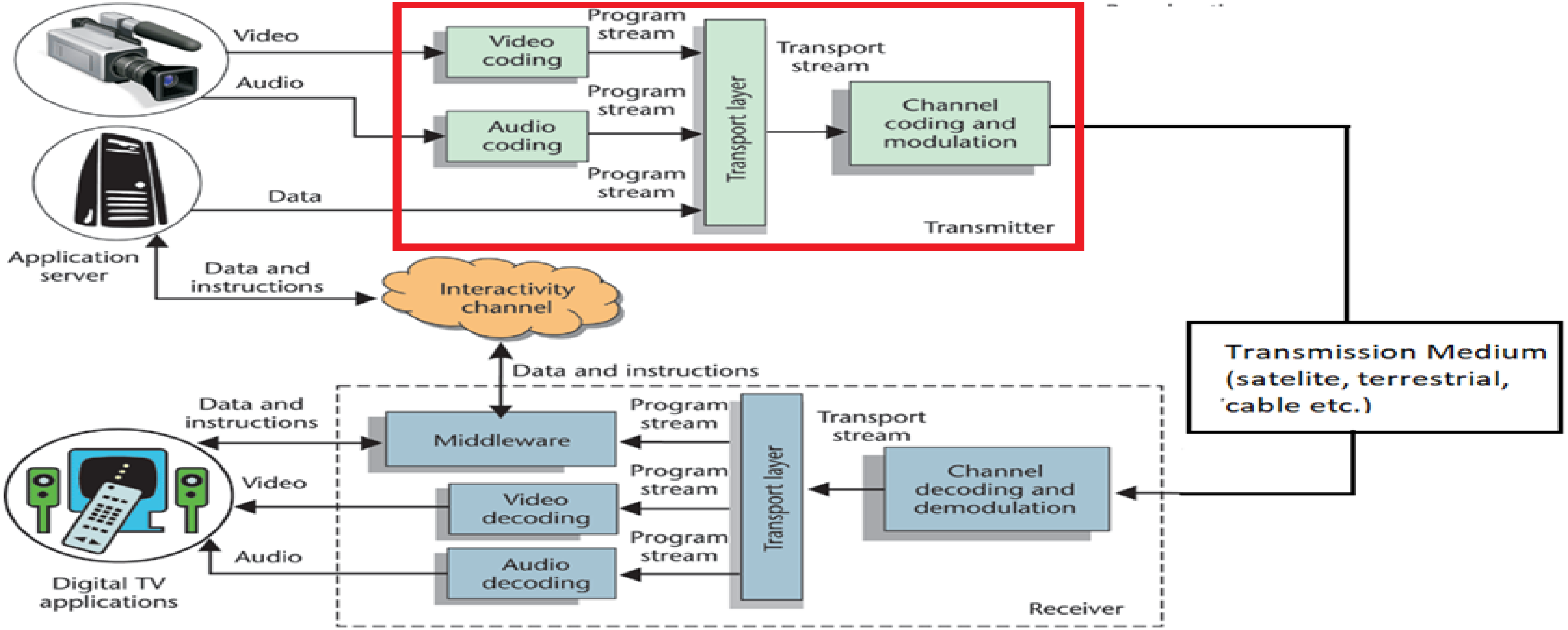
Lecture 7

Outline

- Last 6 lectures were focused mostly on standard related issues.
- From this one, we will be more focused on underlying technologies
- However, we will not go to the details of the topics that are covered by Telecoms related modules
- Roughly we will cover
    - Multiplexing
    - Transmission
    - DB Modulation
    - Psychoacoustics and Audio Coding
    - MPEG Layer I, II, III


Video
Audio
Data
Application server
Video coding
Audio coding
Program stream
Program stream
Program stream
Transport layer
Transport stream
Channel coding and modulation
Transmitter
Data and instructions
Interactivity channel
Data and instructions
Transmission Medium (satelite, terrestrial, cable etc.)
Data and instructions
Middleware
Program stream
Program stream
Program stream
Transport layer
Transport stream
Channel decoding and demodulation
Video
Video decoding
Audio
Audio decoding
Digital TV applications
Receiver

Data Streams characteristics

- Audio/visual content needs compression
  - To enable feasible transmission of high-resolution audio/video data rates
- Source multiplex and transport
  - To transport several programmes ‘simultaneously’
- Channel coding (error protection)
  - To avoid corruption of data by noise during transmission
- Modulation
  - To use available bandwidth, beyond baseband

<u>Inverse operations (in reverse order) at receiver</u>: demodulation, channel decoding, de-multiplexing, audio/visual decompression.

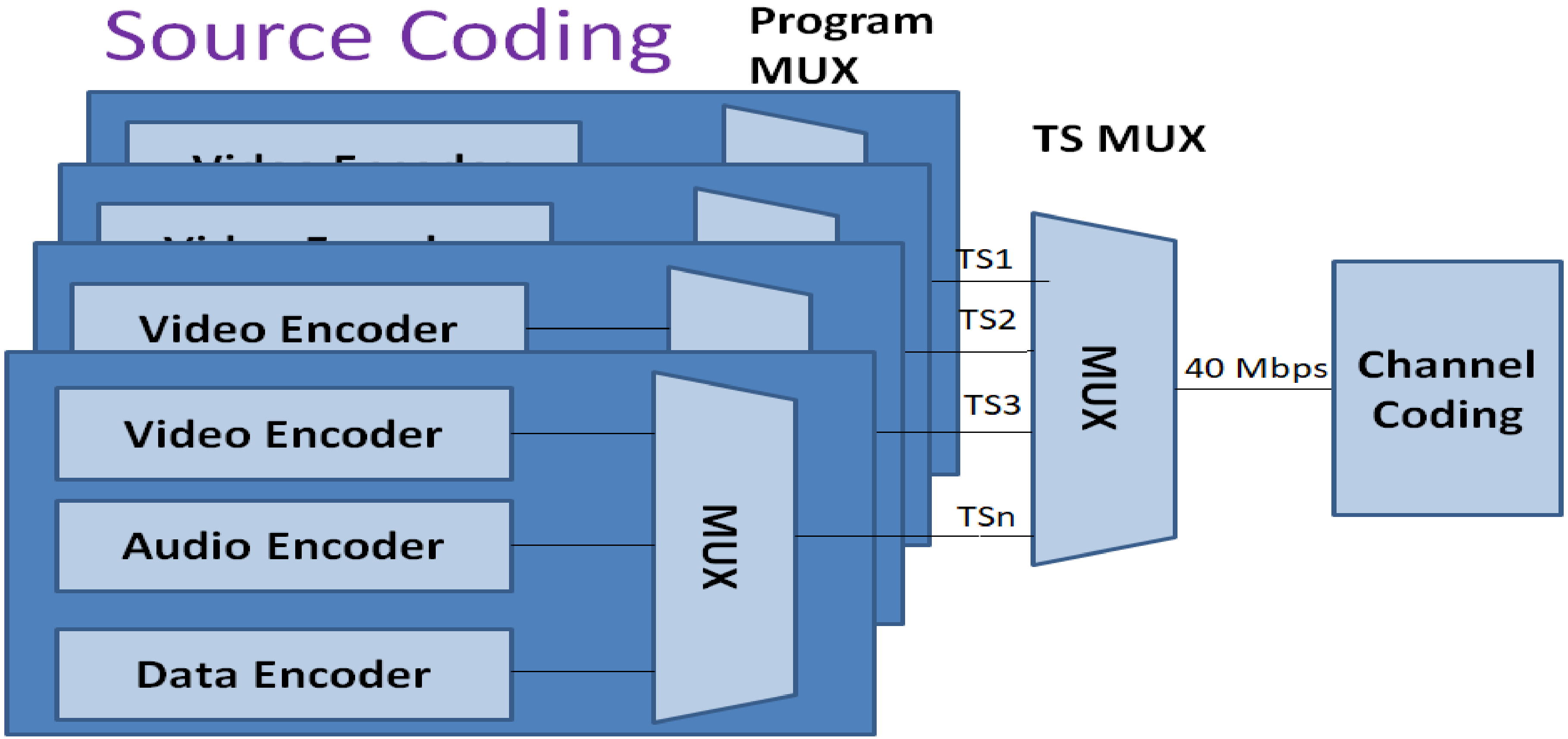
Source Multiplexing and Transport Stream

- Audio/visual encoders produce ‘packets’ of compressed data – **Packetised Elementary Stream (PES)**
  - Pure or (typically) mixed video/audio/data
  - 1 PES ~ 1 broadcast program
  - SDTV: 270 Mbps uncompressed; MPEG-2 compresses to 2…7 Mbps
  - Stereo audio: 1.5 Mbps uncompressed; MPEG-2 compresses to 192 kbps
- Information is added to each packet that allows for identification of packet (time stamp, program ID, etc.)
- PES packets have variable length (≤ 64 kB), depending on instantaneous content

Source Multiplexing and Transport Stream

- PES packets are subdivided into fixed-length “transport stream packet” units of **188 B** each
- This PES is then multiplexed with other audio/visual/data PESs to form the **Transport Stream (TS)**
    - TS contains up to 20 independent DTV channels
    - Data rate of TS is $\leq$ 40 Mbps
- Depending on available bandwidth, several TSs are then further multiplexed together to create a ‘bouquet’ or ‘ensemble’ of services.

Source Multiplexing and Transport Stream



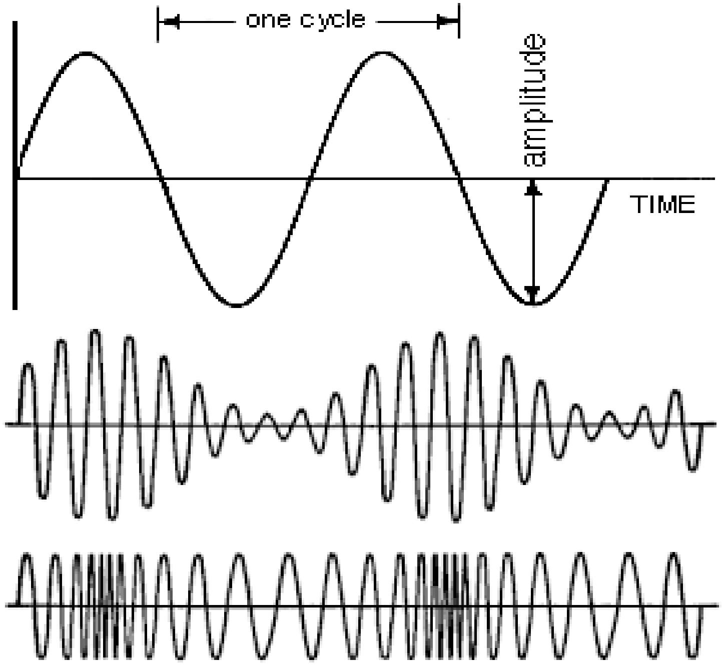
Transmitting Digital TV

- The transmission medium influences the way we process the signal prior to transmission
    - Satellite
        - Long distance
        - Low power
        - Higher noise
    - Cable
        - Physical Connection
        - Low noise
- Basic technology is the same:
    - Sinusoid at a particular (or number of) frequency
    - Phase, amplitude and frequency is altered to encode information

Modulation

A brief review of modulation techniques

Modulation

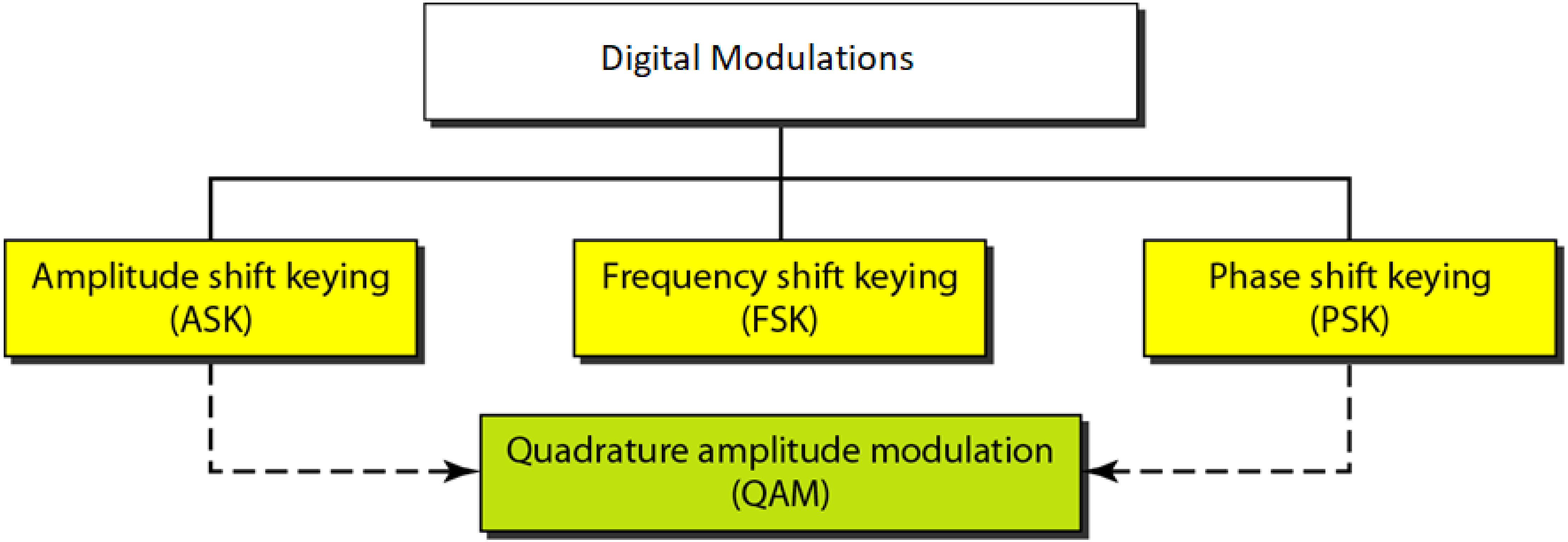


  - no changing parameter $\Rightarrow$ no information
- Modulation (signal):
  - changes amplitude, phase or frequency: allows for transmitting *information*
  - increases bandwidth
- Modulation has software and hardware aspects:
  - (SW, HW) Signal processing: easier at low frequencies
  - (SW) Signal multiplexing: combines different signals to single baseband waveform (FDM, TDM)
  - (SW) Frequency up-shift: larger bandwidths available
  - (HW) Antenna directivity: larger at higher frequencies, more compact HF antennas (~ quarter wavelength)
  - (HW) Physical channel characteristics and transfer (water absorption bands)

Digital: Pulse Modulation

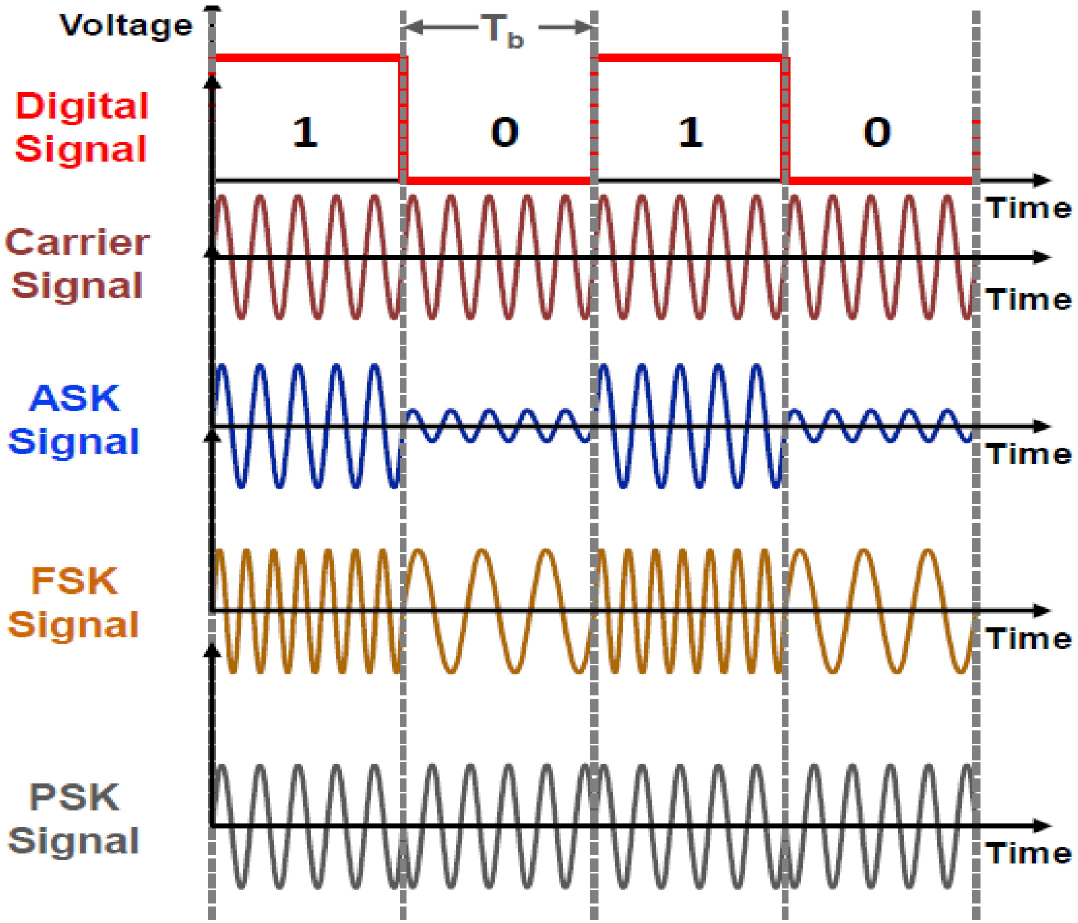
- Unmodulated carrier is a succession of pulse waveforms (digital carrier)
- Information is conveyed by modulation of some waveform parameter:
    - Amplitude (PAM), duration/position (PWM/PPM), time of occurrence (PSK), frequency (FSK), shape, …
- *Digital (binary) modulated signals*: on/off is governed by amplitude of baseband signal.

Digital Modulations


Digital Modulations
Amplitude shift keying (ASK)
Frequency shift keying (FSK)
Phase shift keying (PSK)
Quadrature amplitude modulation (QAM)

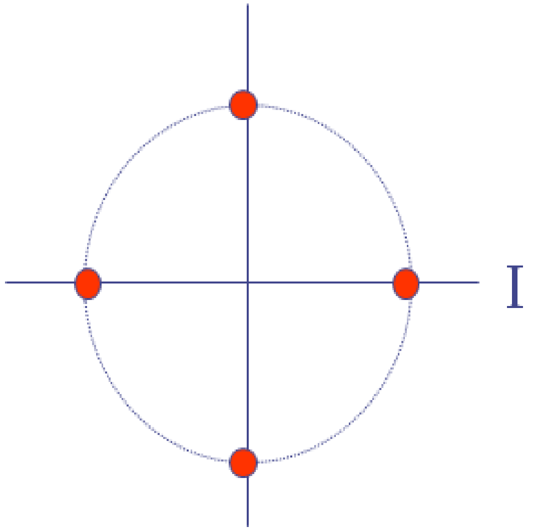
Digital Modulations

1. **Baseband digital message signal**: $m(t)$
2. **Analog sinusoidal carrier signal**:
   A. Carrier signal: $A_c \cos(2\pi f_c t + \phi_c)$
3. **ASK: Amplitude Shift Keying**.
   A. Message signal changes the carrier's **amplitude** : $A_i(t)$.
4. **FSK: Frequency Shift Keying**.
   A. Message signal changes the carrier's **frequency** : $f_i(t)$.
5. **PSK: Phase Shift Keying**.
   A. Message signal changes the carrier's **phase** : $\phi_i(t)$.

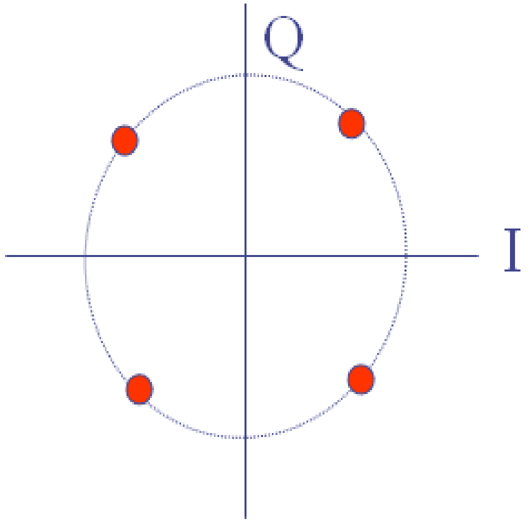


QPSK

- Quadrature Phase Shift Keying (QPSK) can be interpreted as two independent BPSK systems:
  - one on the I-channel and one on Q-channel
  - Thus the same performance but twice the bandwidth (spectrum) efficiency.
- The bit error probability of QPSK is identical to BPSK, but twice as much data can be sent in the same bandwidth.

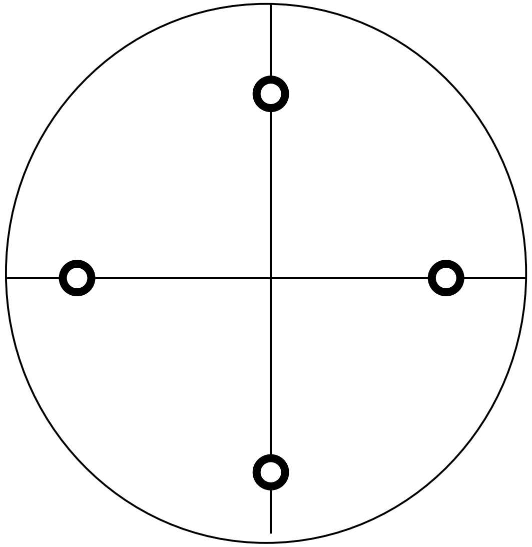


Carrier phases
$\{0, \pi/2, \pi, 3\pi/2\}$

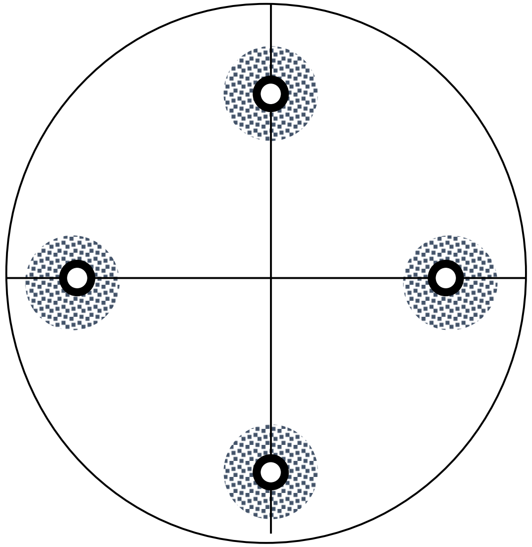


Carrier phases
$\{\pi/4, 3\pi/4, 5\pi/4, 7\pi/4\}$

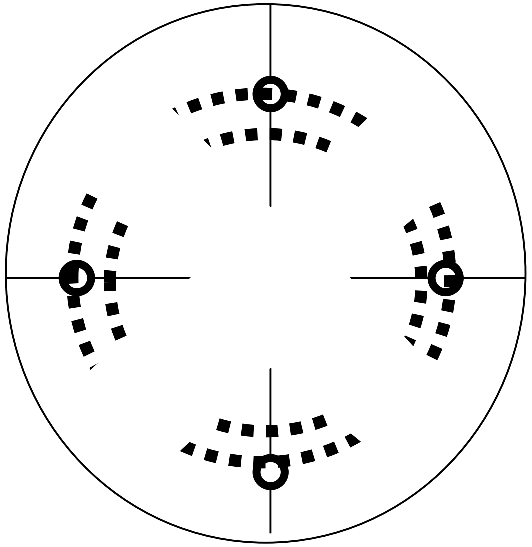
Distortions



Perfect channel



White noise

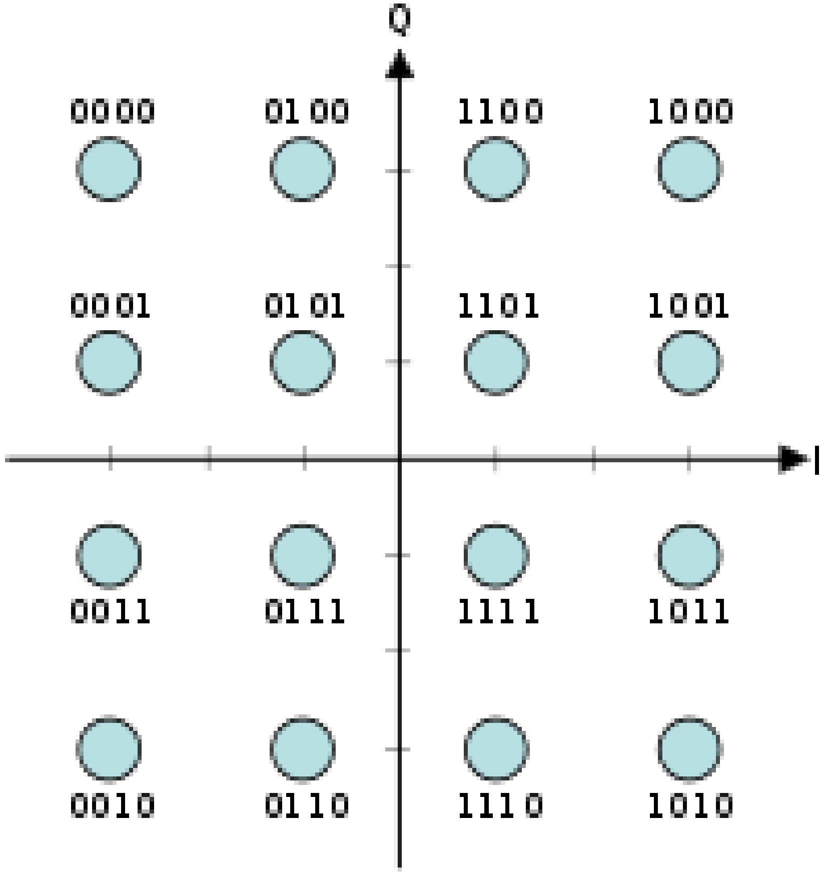


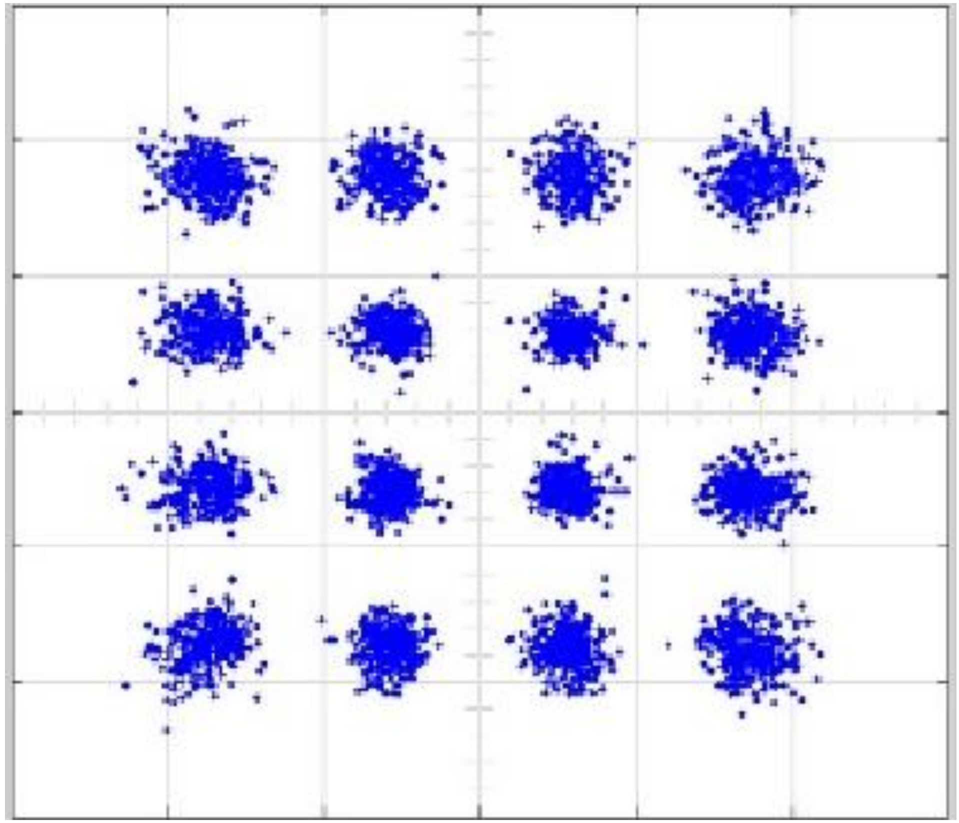
Phase jitter

M-QAM

- It’s a Hybrid modulation
- As we allow the amplitude to also vary with the phase, a new modulation scheme called quadrature amplitude modulation (QAM) is obtained.
- The constellation diagram of 16-ary QAM consists of a square lattice of signal points.
- In M-ary QAM energy per symbol and also distance between possible symbol states is not a constant.
- Efficiency:
  - Power efficiency of QAM is superior to M-ary PSK.
  - Bandwidth efficiency of QAM is identical to M-ary PSK.

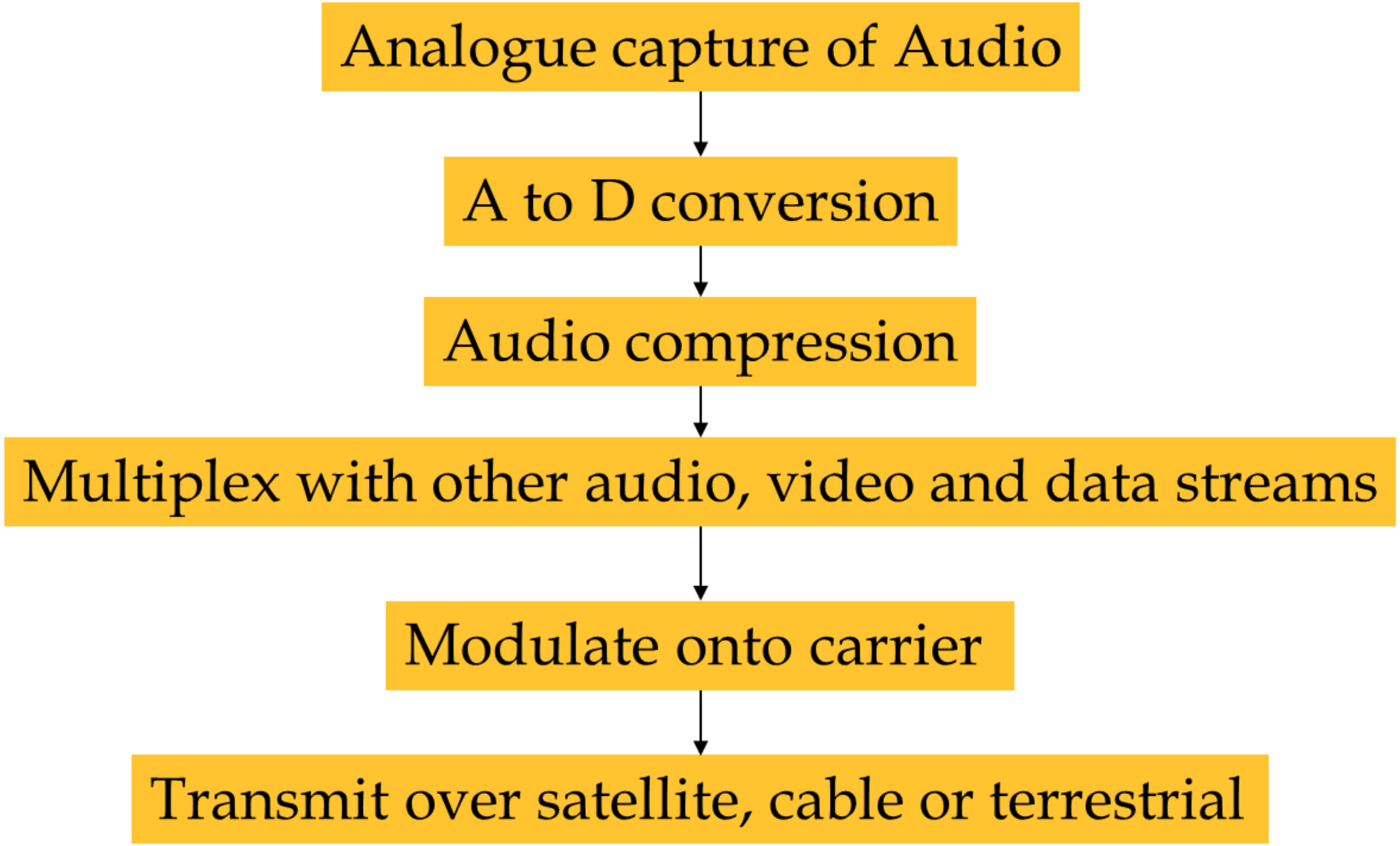
16-QAM


Q
0000
0100
1100
1000
0001
0101
1101
1001
I
0011
0111
1111
1011
0010
0110
1110
1010



Psychoacoustics and Audio Coding

Overview: System at Audio Broadcaster



Analogue capture of Audio

A to D conversion

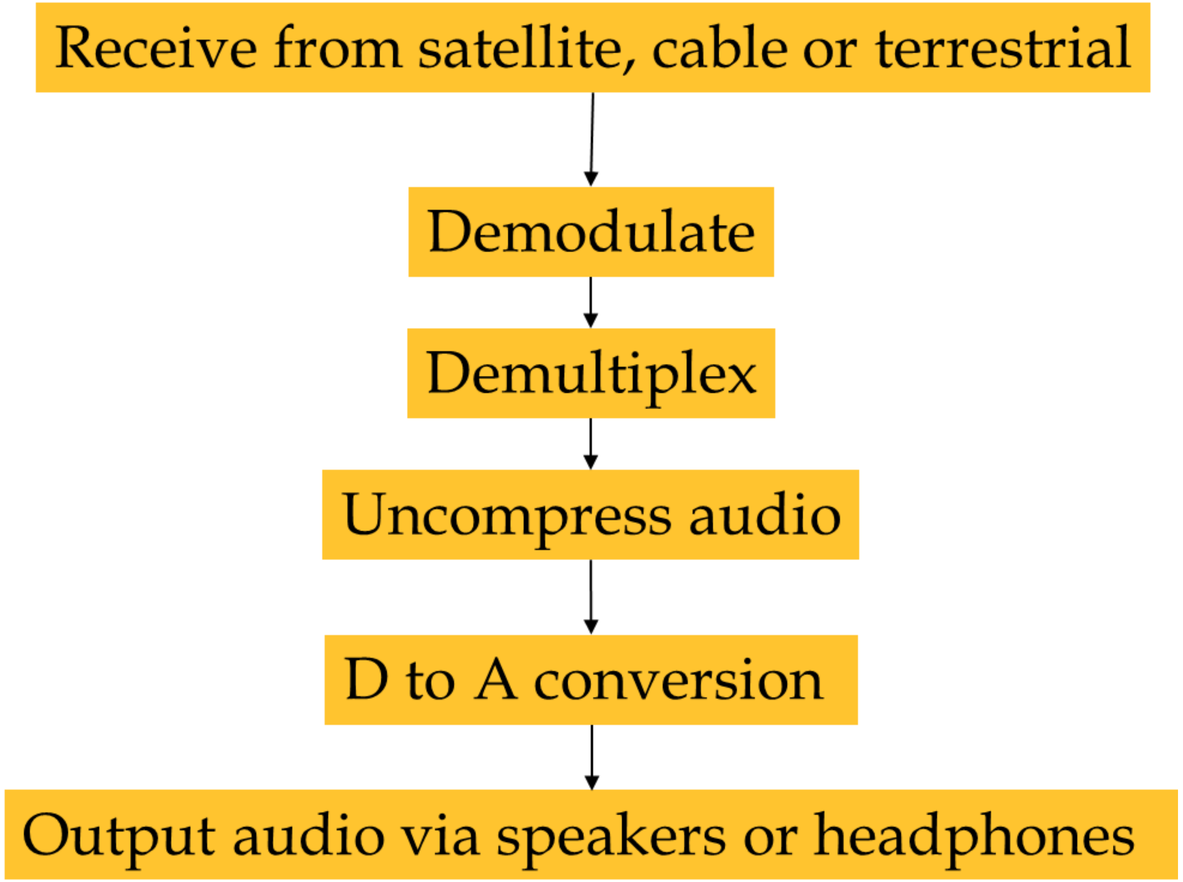
Audio compression

Multiplex with other audio, video and data streams

Modulate onto carrier

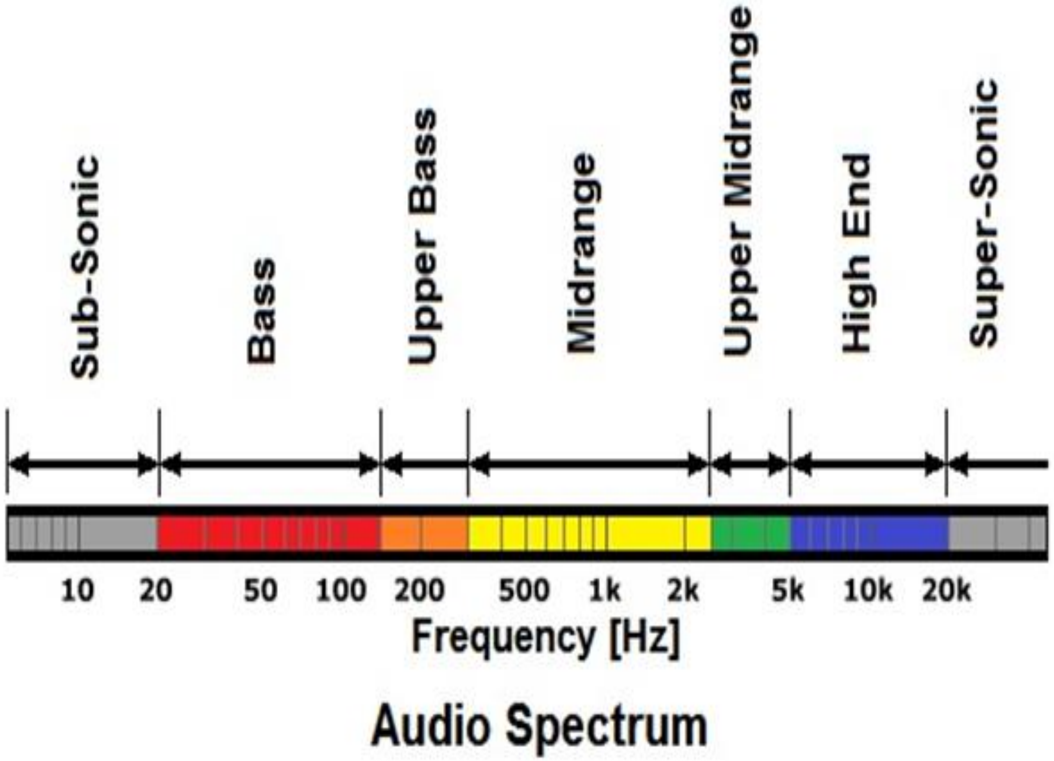
Transmit over satellite, cable or terrestrial

Overview: System at Audio Broadcaster


Receive from satellite, cable or terrestrial
Demodulate
Demultiplex
Uncompress audio
D to A conversion
Output audio via speakers or headphones

Audio Signals

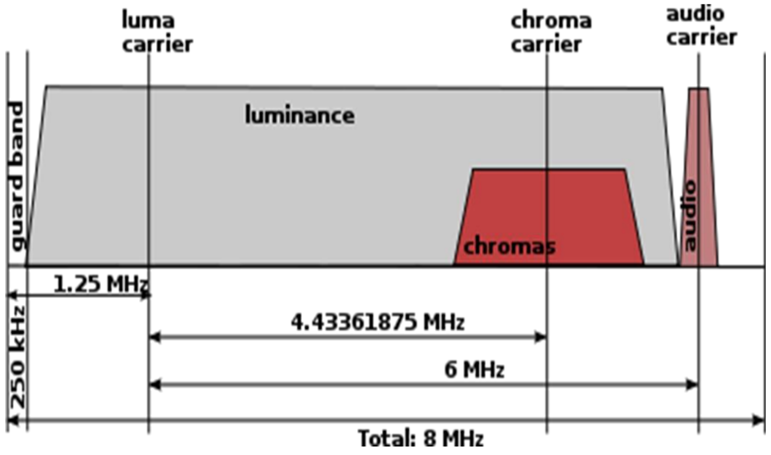
- What is an audio signal? Anything we can hear, including audible noise
- Compression: to discard unnecessary information
  - Redundancy (overhead)
  - Irrelevance (imperceptible)
- Start by excluding acoustic information outside the range of human hearing
  - audible range: 20 – 20,000 Hz
  - in practice: 50 – 15,000 Hz



Audio Signals in TV

- The diagram shows vestigial sideband modulation
- A **vestigial sideband** (in radio communication) is a sideband that has been only partly cut off or suppressed. It may also be used in digital transmission, such as the ATSC standardized 8-VSB.
- In vestigial sideband, the full upper sideband of bandwidth W2 = 4 MHz is transmitted, but only W1 = 1.25 MHz of the lower sideband is transmitted, along with a carrier.

- Analogue TV: separated audio carrier at 6 MHz
  - requires separation from video band ⇒ wasted bandwidth (gap)
  - different modulations: e.g. PAL (AM), SECAM (FM)



Audio Signals: Analogue vs. Digital

- Signals are digitized by sampling and quantization
  - Uncoded digital audio: 48 kHz sampling, 16-bit quantization, 2-channel stereo = 2 × 768 kbps ≈ 1.5 Mbps
  - Source coding: compress this to 100…400 kbps
- Decoding back to PCM signal, then D/A conversion
- For lossless digitization:
  - minimum sampling rate is twice the maximum frequency (Nyquist Theorem)
- Fourier Theorem: "*Any **periodic** waveform can be decomposed into a **series** of harmonic (sine and cosine) functions with associated real amplitudes and phases (or complex amplitude). It* has a **discrete** spectrum.
- For **aperiodic** continuous waveforms: representation is through a Fourier transformation (**integral**). It has a **continuous** spectrum

Audio Coding Objectives

- Aims:
  - (i) making the perceived audio *indistinguishable* from the original signal
  - (ii) using as *few bits* as possible for representation
- Trade-off: quality vs. compression
  - Compact disk standard does not deliberately discard any audio information, but uses a high resolution & data rate
- Standards such as MP3 and AAC discard bits that represent sounds which most people do not hear
  - Imperceptive = irrelevant

Audio Coding Principle

- Compression – reduction of bits
  - Redundancy reduction (lossless)
    - Identify patterns in the bit stream, so that the number of bits used to transfer the information can be reduced
      - E.g., long continuous harmonic ~ amplitude + duration
    - Removes repetition (contains no new information)
    - Huffman coding: most frequent signal ~ shortest code
  - Irrelevance reduction (lossy)

Audio Compression Principles

- Redundancy
  - Measure of predictability of signal
  - Removal of redundancy still allows for perfect reconstruction (lossless)
  - Based on knowledge of statistical properties of signal
    - Well characterized for speech (strong correlation)
    - Poorly characterized for music/generic audio

    ⇒Redundancy has no strong potential for audio compression

    ⇒Better to focus on irrelevance compression in audio, but this is more challenging than speech compression

Audio Compression Principles

- Irrelevance
  - ‘Art’ of identifying and removing signals that are not perceptible by audio recipient (= human hearing)
  - Represents audio ‘overhead’ to human listener
  - Relating to amplitude, time, frequency
  - Lossy (irreversible) process: loss of information

Psychoacoustics: Perceptual Dimensions of Audio

- Pitch
  - higher frequencies are perceived as higher pitch
  - humans hear sounds in range 20 Hz – 20 kHz (decreasing with age)
- Loudness
  - higher amplitude of sound wave results in louder sounds
  - <u>perceived</u> intensity
  - measured in decibels (dB or dBA)
    - x bel = $\log_{10}(x)$

    $\Rightarrow$ x decibel (power) = $10 \log_{10}(x)$; x decibel (pressure) = $20 \log_{10}(x)$;
    - Sound pressure level:

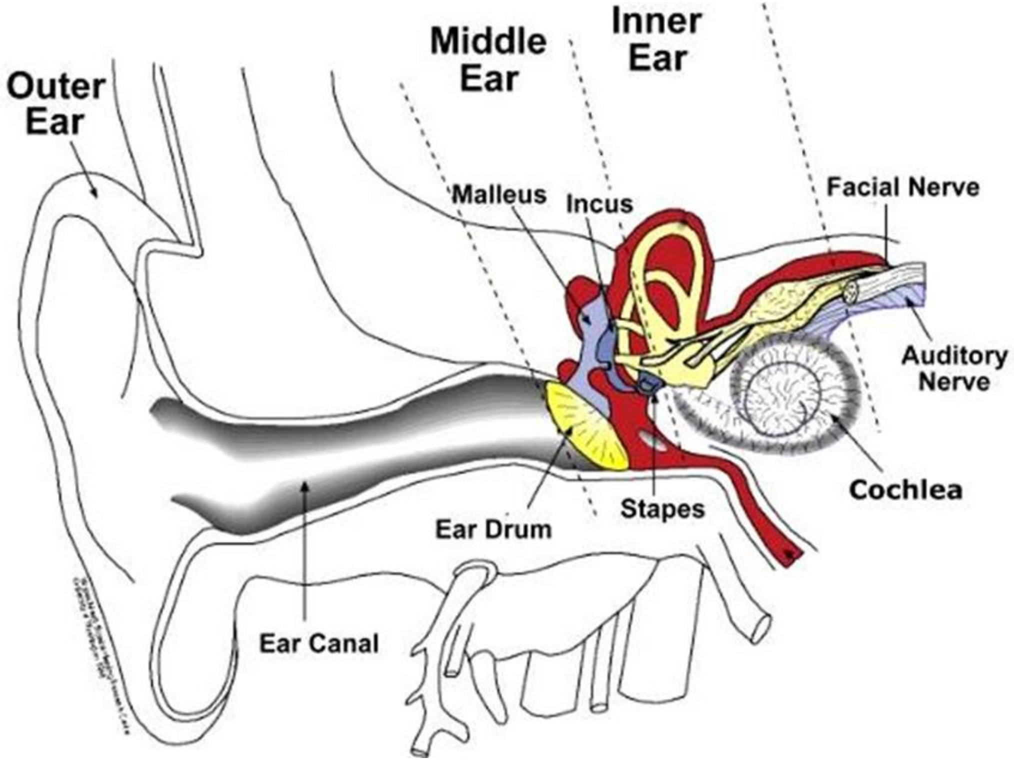
      $L\,(\text{dBA}) = 20 \log_{10}(p/p_0)$, where $p_0 = 2\times10^{-5}\,\text{N/m}^2 = 20\ \mu\text{Pa}$
    - 0 dB = hearing threshold (~ $p=p_0$ at 2 kHz), 130 dB = pain level

Psychoacoustics: Perceptual Dimensions of Audio

- Basis of most audio and video compression: exploitation of human perception characteristics
  - In audio: called "psychoacoustics"
  - Idea: sound that cannot be heard (by majority of healthy population) need not be encoded and can be removed
- Observations (from extensive experimental study):
  - Hearing threshold (= minimum pressure of audible sound signal) varies with frequency
  - Sound signal can be masked by louder sounds
    - Masking effect also varies with duration of masking signal

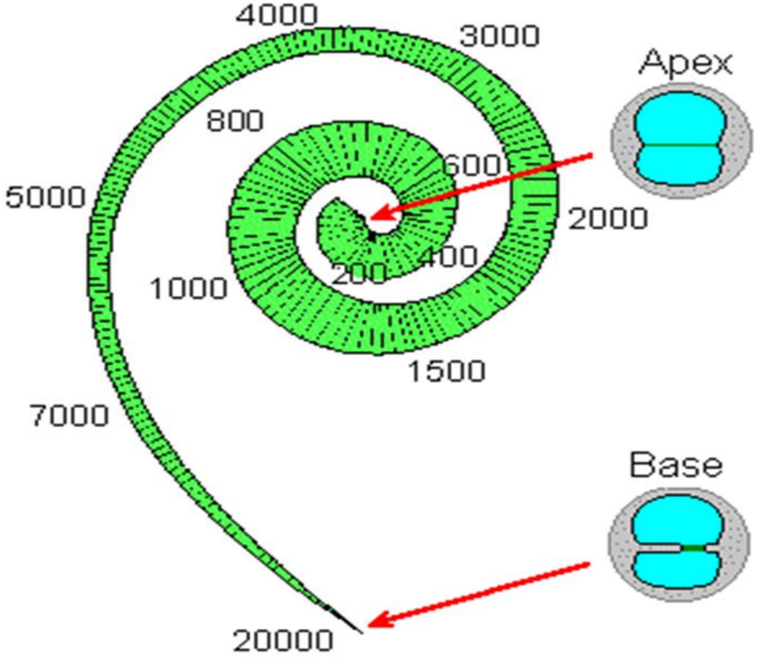
Ear Physiology

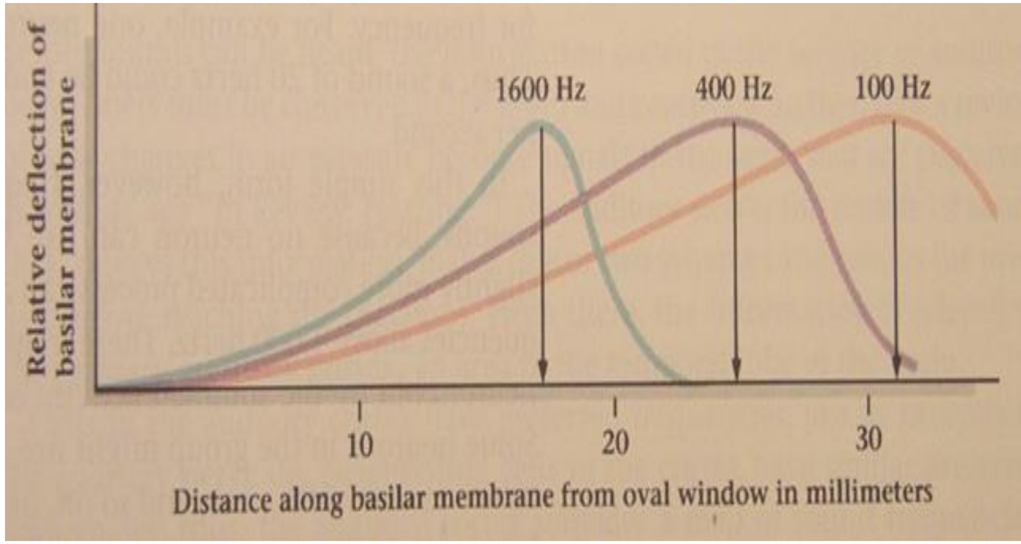
- Human ear: 3 main parts: (1) outer ear (‘channel’), (2) middle ear (‘mechanical transformer’), (3) inner ear (‘electromechanical transducer’).
- The outer ear directs speech pressure variations toward the eardrum where the middle ear transforms these variations into mechanical motion.
- The inner ear converts these vibrations into electrical firings in the auditory neurons, which lead to the brain.
- Label Eustachian tube (red pipe between staples label and cochlea label).



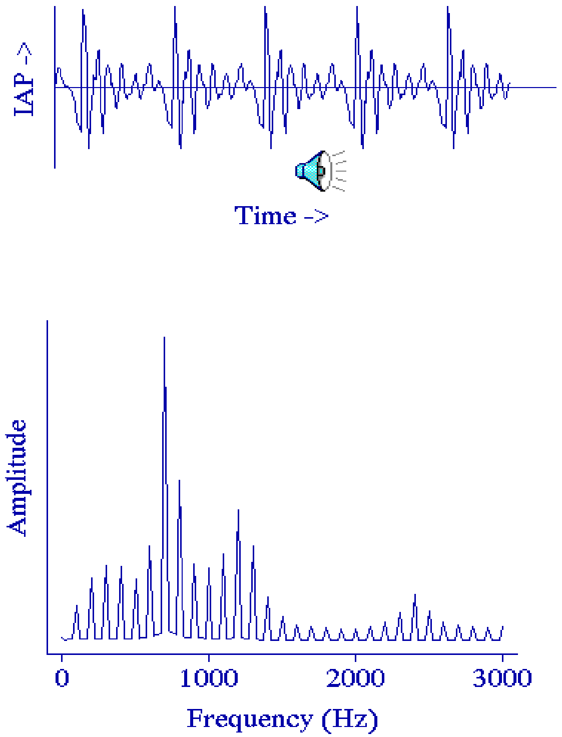
Basilar Membrane

- The basilar membrane **varies gradually in tautness and shape** along its length and so it has a **varying frequency response**.
- BM is **stiff and thin at the basal end** (base – connects to oval window) but **compliant and massive at the apex** (deep inside the snail).
- Each location has a **characteristic frequency** at which it vibrates maximally.
- For a specific location the response curve is that of a **bandpass filter** with almost **constant Q** (ratio of centre frequency to bandwidth). Can think of BM as a **filter bank**.
- Because of constant Q, **frequency resolution along the Basilar membrane is best at low frequencies**. (Δf = f / Q)
- Note that the distance of the point of maximal vibration on the BM from the apex is roughly proportional to the logarithm of the frequency.

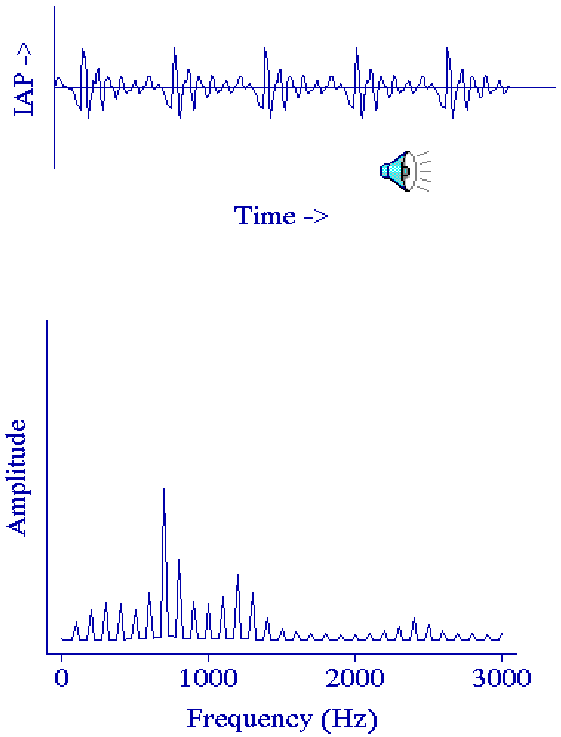




Loudness and Intensity



Higher intensity, higher loudness



Lower intensity, lower loudness

Nonlinear: doubling the intensity does not double the loudness

To double loudness, intensity must be increased by a factor ***10***, i.e., 10 dB [10 $\log_{10}$ (10) = 10 dB].

Example: two signals differing by 10 dB:

Loudness perception

- Loudness is strongly dependent on frequency
- - For constant intensity, ***mid-frequency*** signal (in the range 1000-6000 Hz) is perceived as ***louder*** than lower- or higher-frequency signals

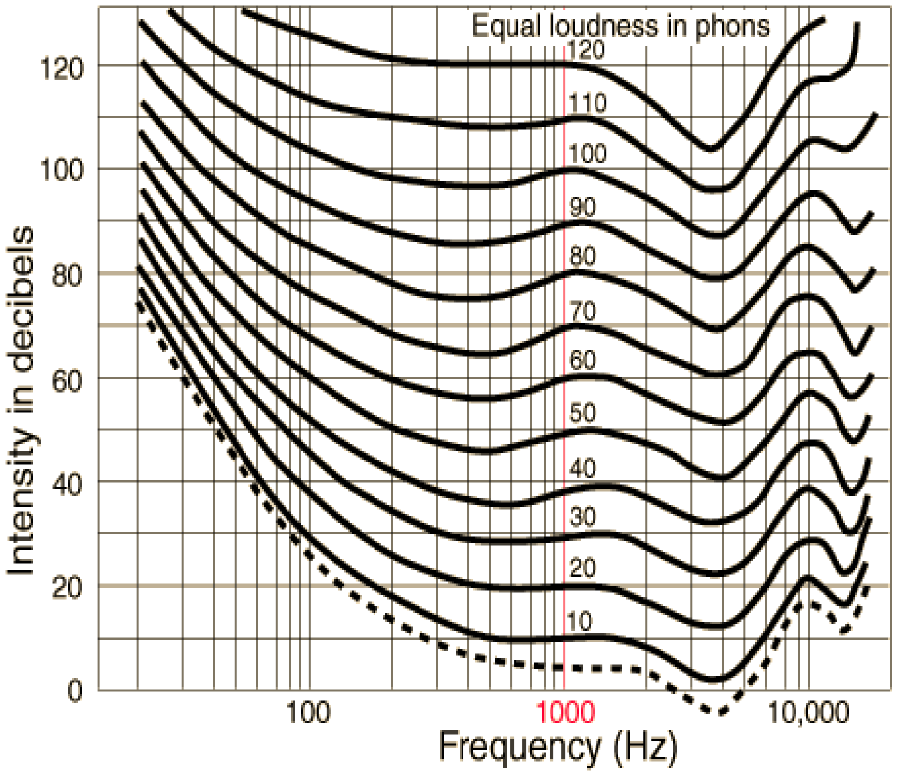
    - **125 Hz, 3000 Hz, 8000 Hz**

- 3000 Hz signal appears louder than 125 Hz or 8000 Hz signals, even though intensities are equal
- Greatest sensitivity to loudness is at mid-frequencies

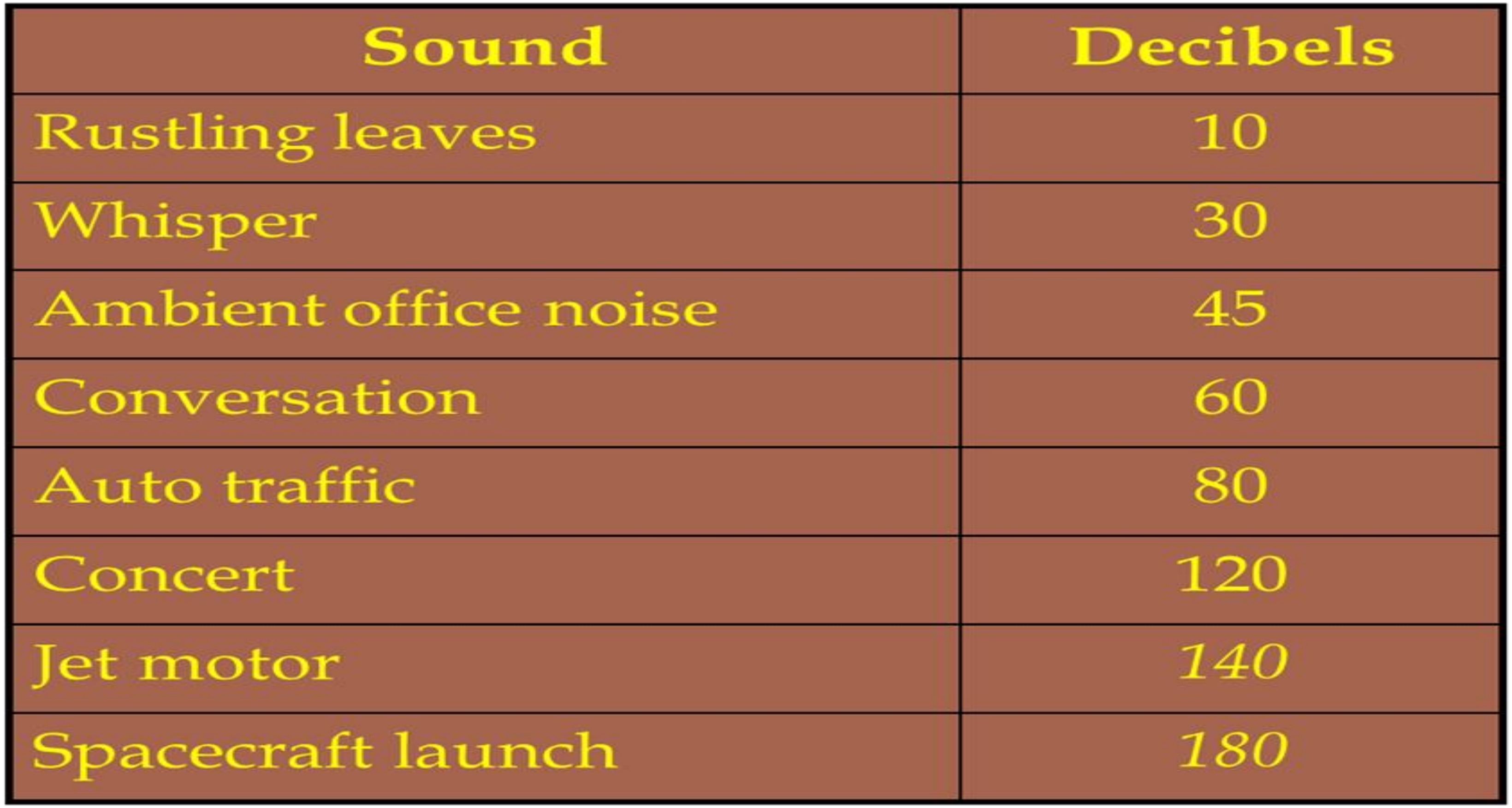
    - basilar membrane reacts more to intermediate frequencies than other frequencies

Equiloudness

- Saying that two sounds have **equal intensity** is **not** the same thing as saying that they have **equal loudness**. Due to the frequency variation in auditory thresholds, the **perceptual loudness of a sound is frequency dependent** and is specified via its **relative intensity above the threshold**.
- **Equal-loudness curves** show the **variation between intensity and equal loudness** for the human ear.
- A sound's loudness is often defined in terms of **how intense a reference 1 kHz** must be in order to be heard as an equally loud sound.
- This measures loudness in **phons** eg. if a given sound is as loud as a 60 dB at 1kHz then it has a loudness of 60 phons. At **1 kHz**, the **phon scale equals the dB scale**.
- The curves show that a sound at eg. 20 Hz tone (taking the bottom solid curve) needs to have an intensity of 75 dB to sound equally as loud as a 1 kHz tone with an intensity of 10 dB or 10 phons.
- **Threshold of hearing (dashed curve)**
- Shows the sound intensity required of a sound to be **heard**. There is a marked **discrimination against low frequencies**.
- The **maximum sensitivity** at about 3.5 kHz to 4 kHz is related to the **resonance of the auditory canal** (length is approximately 2.7 cm long, giving a first resonance near 3 kHz)



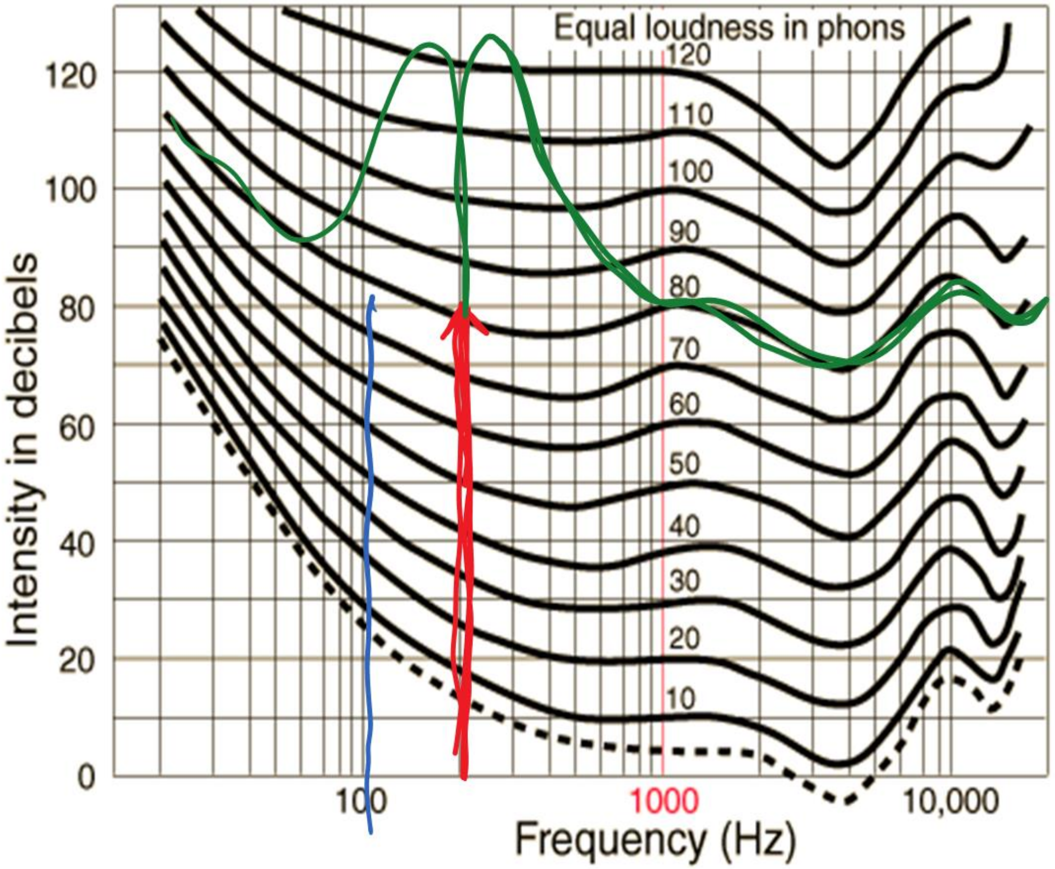
Some Everyday Sounds



Sound	Decibels
Rustling leaves	10
Whisper	30
Ambient office noise	45
Conversation	60
Auto traffic	80
Concert	120
Jet motor	*140*
Spacecraft launch	*180*

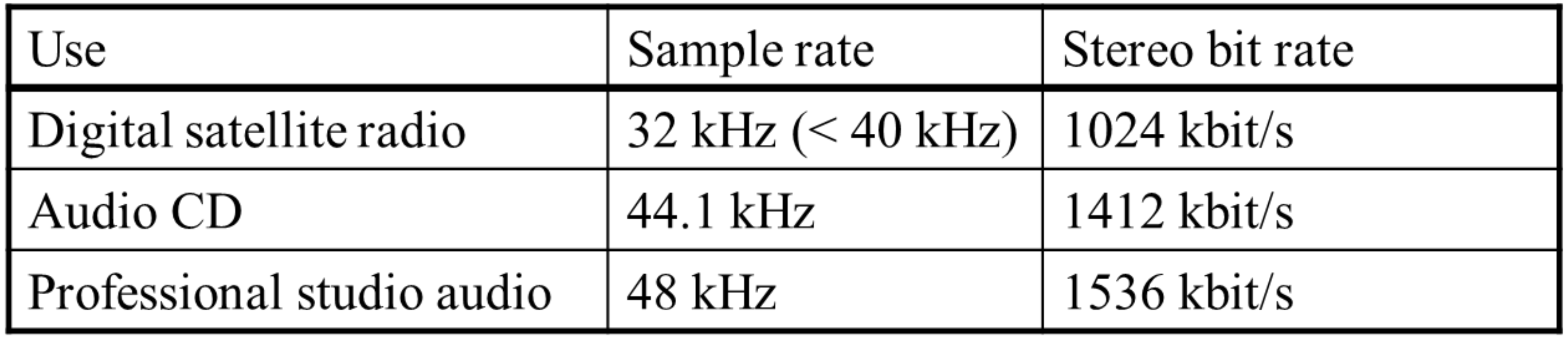
Masking and Equiloudness

- The presence of a sound at one frequency will alter our perception of the loudness of another sound.
- Presence of 200 Hz tone warps the frequency characteristic of a 100 Hz tone:
- 100 Hz tone then sounds significantly quieter
  - e.g. 85 dB at 100 Hz needs to increase to 105 dB to sound equally loud again
- However, if both tones are playing simultaneously, the 100 Hz tone will sound significantly quieter than 80 phons.



Audio Coding: Bit rate requirement

- Human audio perception requires minimum sampling frequency of 40 kHz per channel (why?) (studio quality: 48 kHz - why?)
- 16-bit resolution for amplitude of audio in 2-channel stereo broadcast thus requires $2 \times 16 \times 40 \times 10^3 = 1.28$ Mbit/s.
    - Compare with non-broadcast audio: MP3 files: 60 Mbit/s after compression
- In practice: extra bits needed for synchronization and error correction.
- Oversampling (1-bit $\Sigma\Delta$M in ADC): smears quantization noise power over frequency; exchange performance for speed



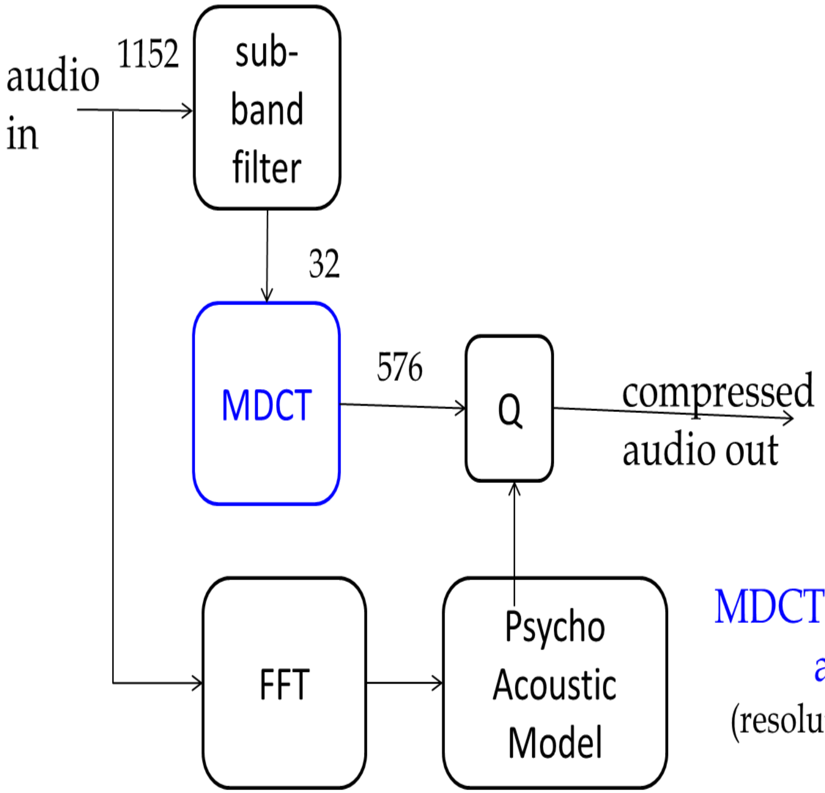
Use	Sample rate	Stereo bit rate
Digital satellite radio	32 kHz (< 40 kHz)	1024 kbit/s
Audio CD	44.1 kHz	1412 kbit/s
Professional studio audio	48 kHz	1536 kbit/s

MPEG Layers: I & II

- They are almost same
- Several audio source coding standards were developed within the principal standardization body, the *Motion Picture Experts Group (MPEG)*
    - MPEG-1: for audio/video on physical media only (CD)
    - MPEG-2 introduced different possible audio coders known as layer I, layer II, layer III and AAC
        - AAC (Advanced Audio Coding) standardized in 1997 as part MPEG-2; used e.g. in iPods
- Other audio coding can be used within digital broadcasting
    - e.g. Dolby AC-3 and Digital Theatre System (DTS) surround sound (= more than 2 channels)

MPEG Layer III

- Resolution of MDCT performed on each subband is frequency dependent, so each subband will give a different number of spectral power values.
- High frequencies – low resolution of hearing – few spectral values
- Low frequencies – high resolution of hearing – many spectral values



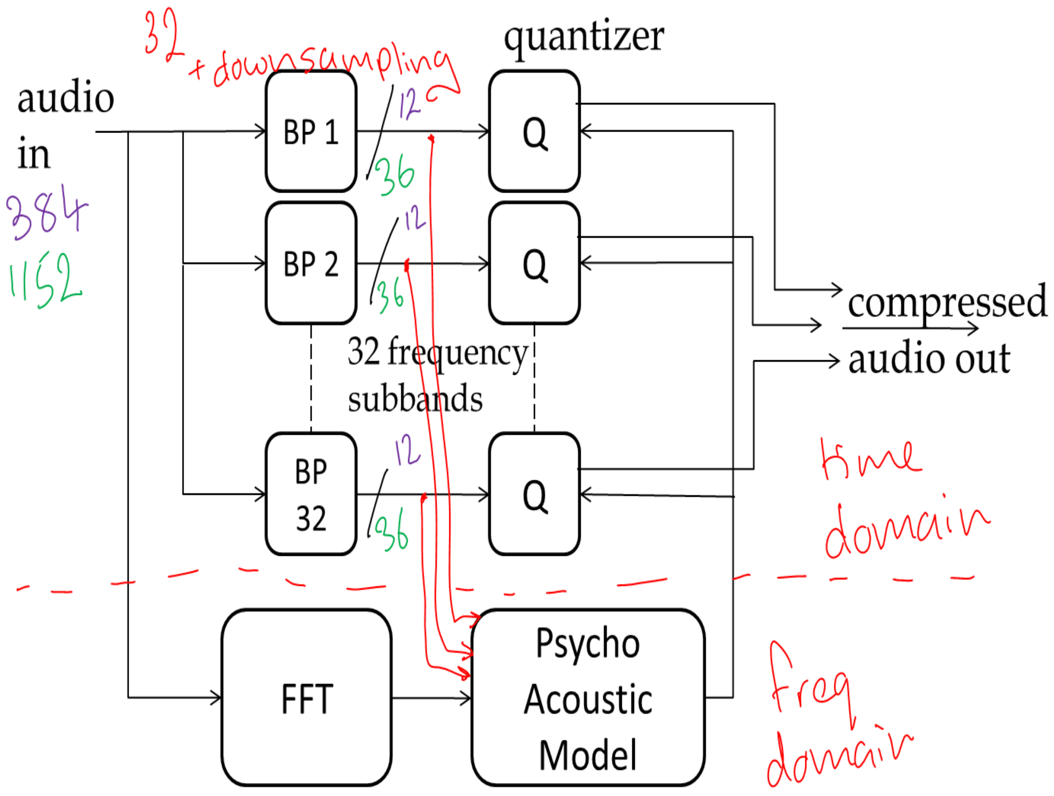
MDCT = frequency-adaptive DCT
(resolution depends on subband)

MPEG Layer III

- Divides 1152 time samples into 576 sub-bands by a DCT (Discrete Cosine Transform) technique
  - For transient signal, only 192 samples instead of 1152
- Defined for decoding MP3, but encoding quality can vary
  - one that works well for high bit rates may not be good for low bit rates
- Quantization is nonlinear operation adapted to basilar membrane characteristics of human hearing (logarithmic)
- Huffman encoding is used to reduce number of bits in encoded signal (redundancy reduction)

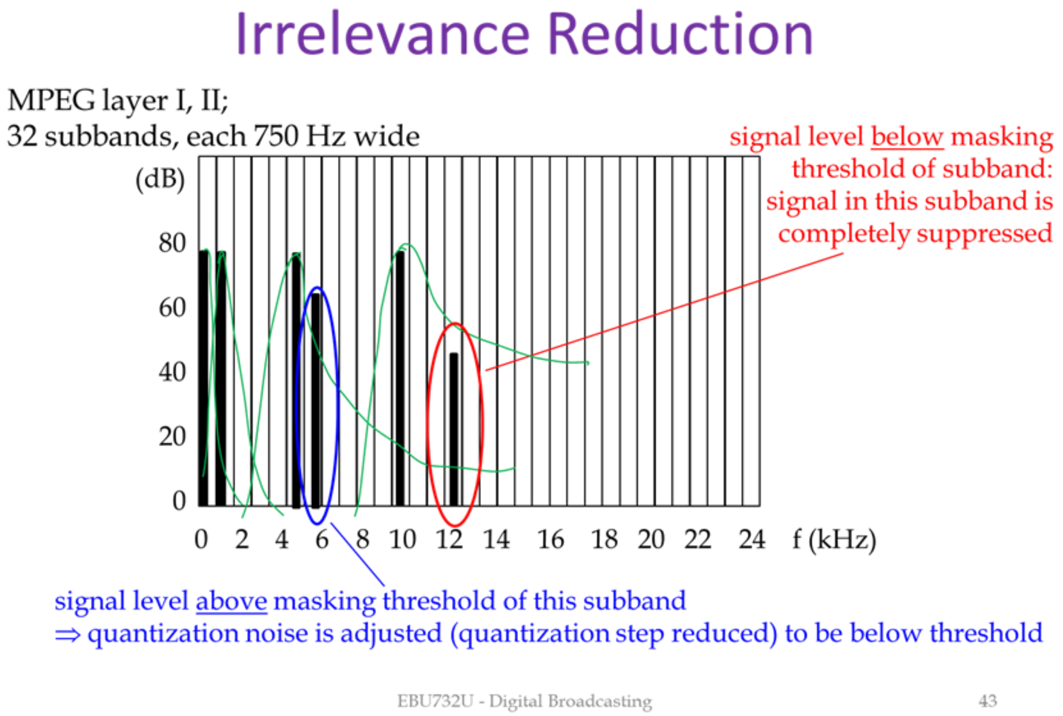
MPEG Layer I and II Subband Coding

- Before the coder can remove or compress different audio frequencies (in accordance with the sensitivity of the human auditory system), the audio is segmented into different frequency bands.
- The audio signal is passed through a filter bank of 32 filters that split the 24 kHz audio signal into frequency subbands, each of 750 Hz (overlapping).
- The signal is also sent in parallel through an FFT filter, which transforms the signal to the frequency domain.
- The psychoacoustic model uses this frequency representation to determine what elements of the signal are irrelevant.
- Each subband has a separate quantizer controlled by the psychoacoustic model.
- The quantizer either completely suppresses the subband or reduces the number of quantization steps according to the instructions given by the psychoacoustic model.
- Quantization must be finest at low frequencies where the ear is most sensitive, and can be coarser at high frequencies.
- In the case of layer II coding, FFT is carried out every 24 ms on 1024 samples.
- The signal is treated as short-time stationary within 24 ms.



Irrelevance Reduction

- Two examples are shown.
- In one subband there is a signal at approx. 5 kHz whose level is **above** the masking threshold.
- In this subband, only the number of quantization steps can be reduced.
- In another subband, a signal is at 10 kHz with a level **below** the masking threshold.
- This subband is hence fully masked by signals of neighbouring subbands and can therefore be suppressed completely.



Irrelevance Reduction

MPEG layer I, II;
32 subbands, each 750 Hz wide

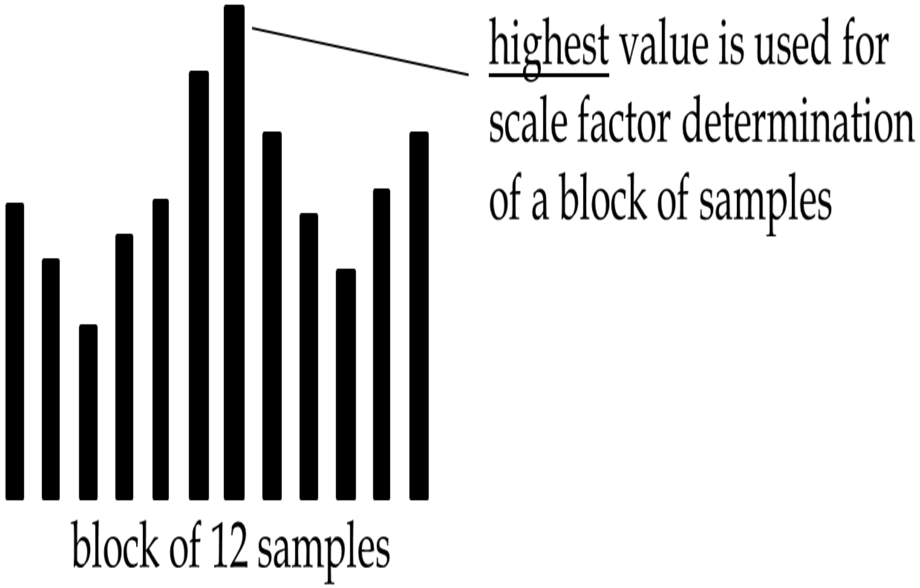
signal level below masking threshold of subband: signal in this subband is completely suppressed

signal level above masking threshold of this subband
⇒ quantization noise is adjusted (quantization step reduced) to be below threshold

EBU732U - Digital Broadcasting 43

Scale Factor

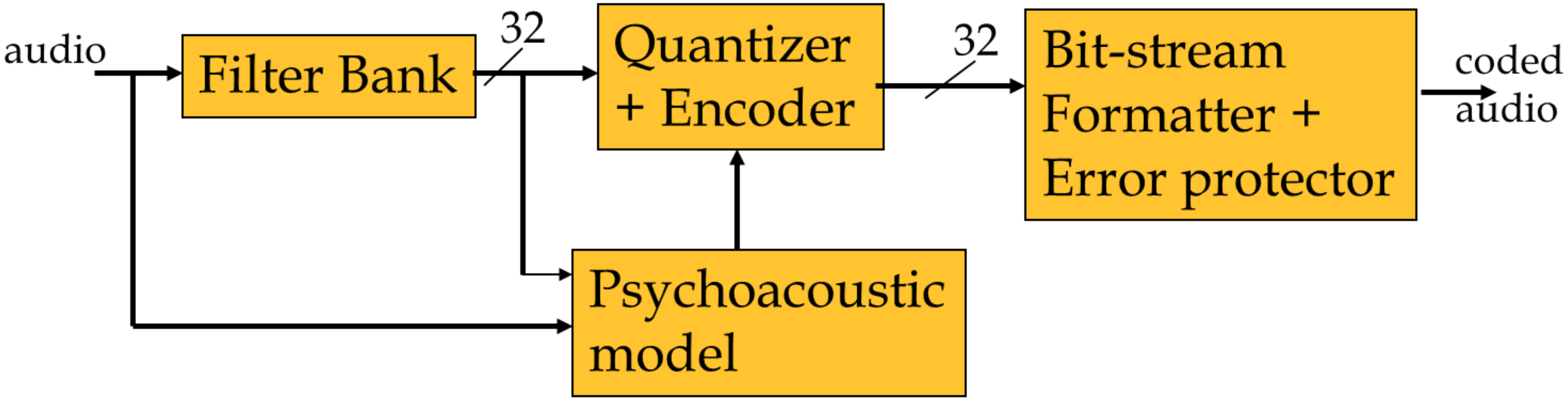
- Maximum amplitude of every *n* samples (in time) defines scale factor, sent to psychoacoustic model
    - MPEG layer I: *n* = 12 samples for each subband
    - MPEG layer II: *n* = 36 samples for each subband
- Model then uses this scale factor to calculate quantization level for each sample in subband
- 12 sample sub-frames (blocks).
- Look at each 12 samples to form a scale factor to determine the magnitude of our signal.
- This is a multiplication factor that allows us to normalise the signal.
- Don’t have to code the actual magnitude, but the normalised magnitude (smaller range) so we only send the difference from the scale factor.

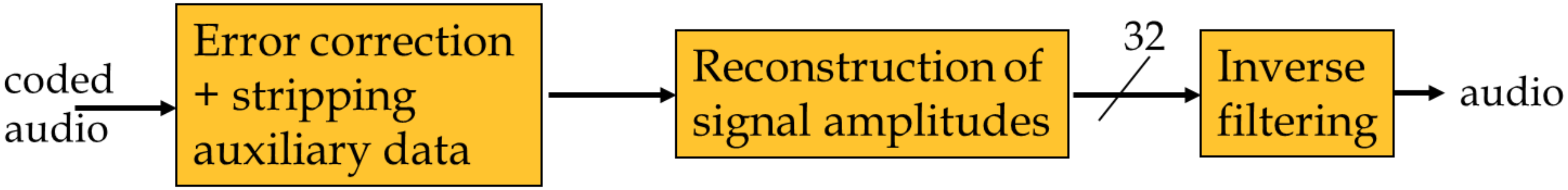

highest value is used for scale factor determination of a block of samples
block of 12 samples

MPEG-2 Layer II

- This is minor extension of MPEG-1 layer II: MP2 or MUSICAM used for much of digital audio broadcasting . 36 samples combined into 1 block
- Each subband is *critically subsampled* by factor 32. hence at 48 kHz sampling frequency, audio block size = 24 ms
- 48 kHz sampling means a maximum frequency of 24 kHz.
- Implication: single scaling factor may not be adequate, because major temporal signal changes (e.g., drumbeats) would not be reproduced correctly when masking lasts only for max. 20 ms
- The number of scale factors can vary between one to three per block, depending on the nature of audio number of scale factors has to be included in frame as extra information

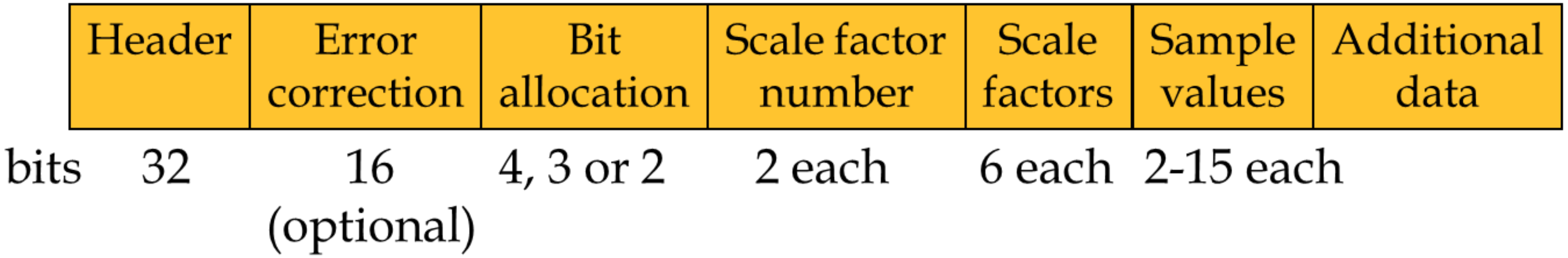
Encoder and Decoder: MPEG-2 Layer II Block Scheme




coded audio
Error correction + stripping auxiliary data
Reconstruction of signal amplitudes
32
Inverse filtering
audio

MPEG Layer II Frame Structure

- Data redundancy
- Bit allocation is variable and data dependent. Could fix the number of bits used, but this wastes bits (redundancy)
- Each frame is 1152 audio samples long
- Each scale factor is quantised using 6 bits.
- Sample values – amplitude of each sample in each subband. Depending on the quantisation level used for a subband, each sample in the subband can be represented by 2, 3, 4 … 15 bits. The more bits, the finer the quantisation step.



	Header	Error correction	Bit allocation	Scale factor number	Scale factors	Sample values	Additional data
bits	32	16 (optional)	4, 3 or 2	2 each	6 each	2-15 each	

MPEG Layer III

- *Hybrid coding*: uses a filterbank for subband filtering and then uses either the Discrete Cosine Transform (DCT) or modified DCT (MDCT) on each subband for finer frequency resolution
- 1152 time samples (as with Layer II) $\rightarrow$ up to 576 power levels of spectral (harmonic) components
- Simultaneously, a high-resolution FFT is performed to inform the psychoacoustic model (as with Layers I and II)
- Psychoacoustic model controls quantization and suppression of spectral power levels (masking)
- For good quality MP3 encoder: most music encoded at 192 kbit/s
    - Output audio then indistinguishable from uncompressed input audio

Time and Frequency Resolution

- Time resolution: determines how closely on/off the beat, like metronome (NB: each frequency band individually filtered => may result in instruments or overtones sounding out of synchronism)
- Frequency resolution: for ‘long’ tone, how ‘crisp’ the sound is, i.e., how much fine detail of spectrum is captured and reproduced
- One weakness of FT is its fixed resolution
- Width of the window relates to representation of signal — good frequency resolution (frequency components close together can be separated) vs. good time resolution (the time at which frequencies change can be accurately identified)
  - Wider window: better frequency resolution, poorer time resolution

How to Increase Frequency Resolution

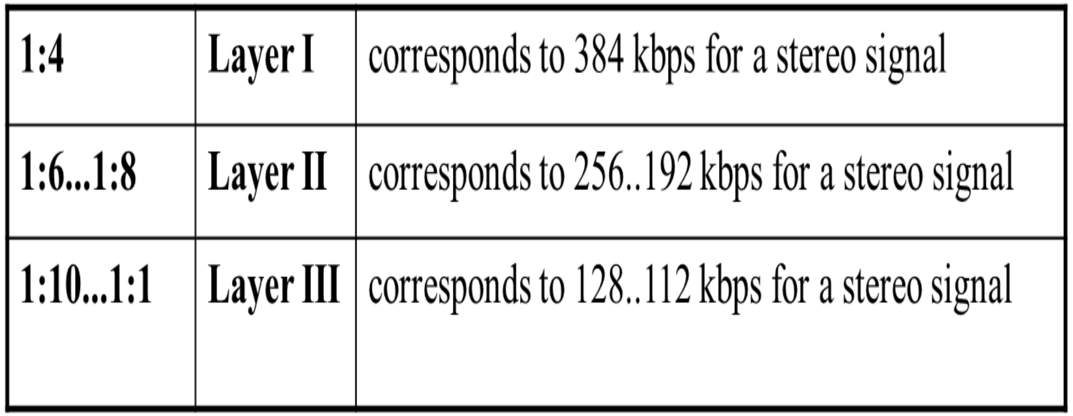
- To increase the frequency resolution of the window, this δf needs to be reduced (by definition)
  - decreasing f_s (and keeping N constant): causes time window to enlarge, because fewer samples per unit time
  - increasing N: also increases window size
- Thus, increasing frequency resolution always causes larger window and, hence, reduced time resolution (and vice versa)
  - reason why 192 samples for transient instead of 1152

MP3 Limitations

- Bitrate limited to max. 320 kbit/s
- Time resolution can be too low for rapidly rising transients
  - e.g. smearing of percussive sounds
- Frequency resolution is limited by small window size, which decreases the coding efficiency
  - may lead to masked frequencies becoming inseparable from wanted frequencies
- No scale factor band above 15.5/15.8 kHz
- Overall delay in encoder/decoder is undefined
  - implies lack of official provision for gapless playback
  - some encoders attach metadata allowing some players to produce gapless playback.

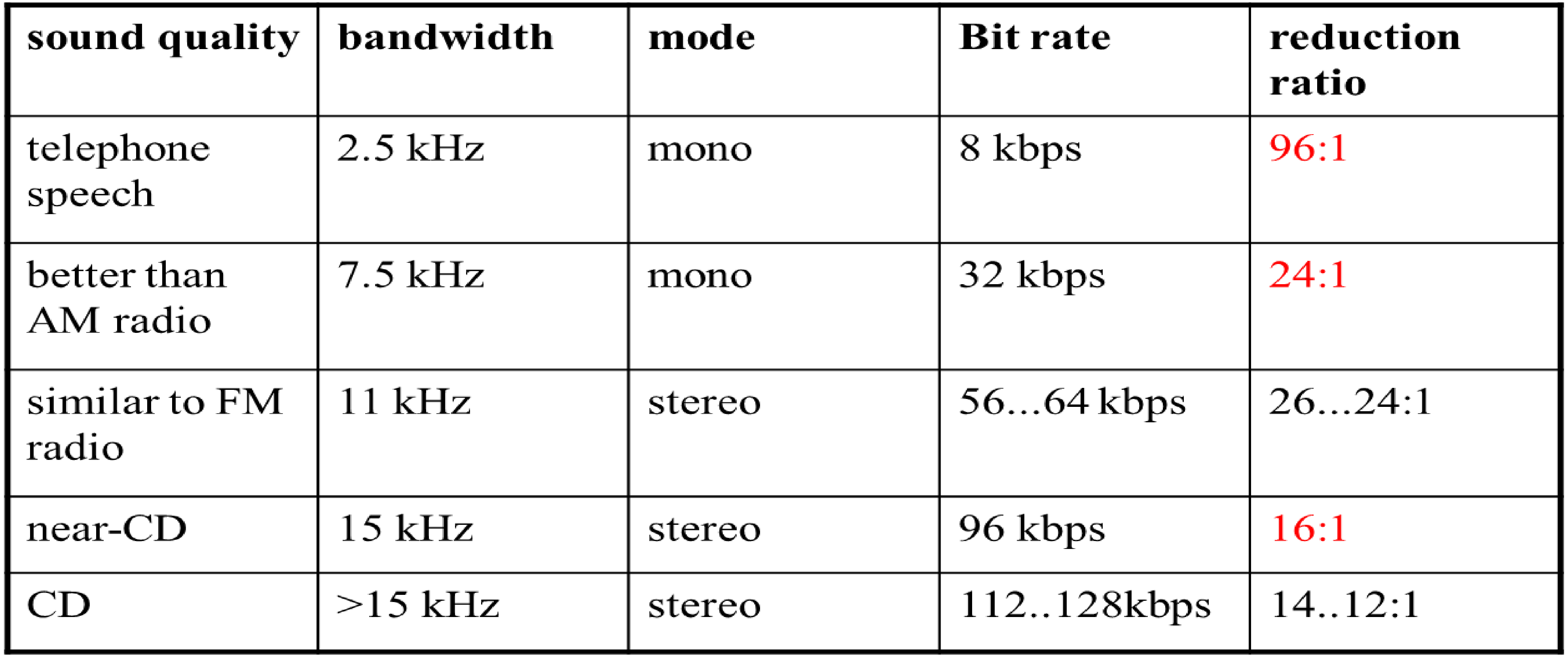
Comparing Compression Efficiency

- Current digital video broadcasting (DVB) implementations: typically with layer II and 128 kbit/s
- DAB audio broadcasting: also uses layer II, but needs 192 kbit/s for CD-like quality (classical music)
- Most broadcasters use 128 kbit/s



1:4	**Layer I**	corresponds to 384 kbps for a stereo signal
1:6...1:8	**Layer II**	corresponds to 256..192 kbps for a stereo signal
1:10...1:1	**Layer III**	corresponds to 128..112 kbps for a stereo signal

Compression Efficiency Layer III



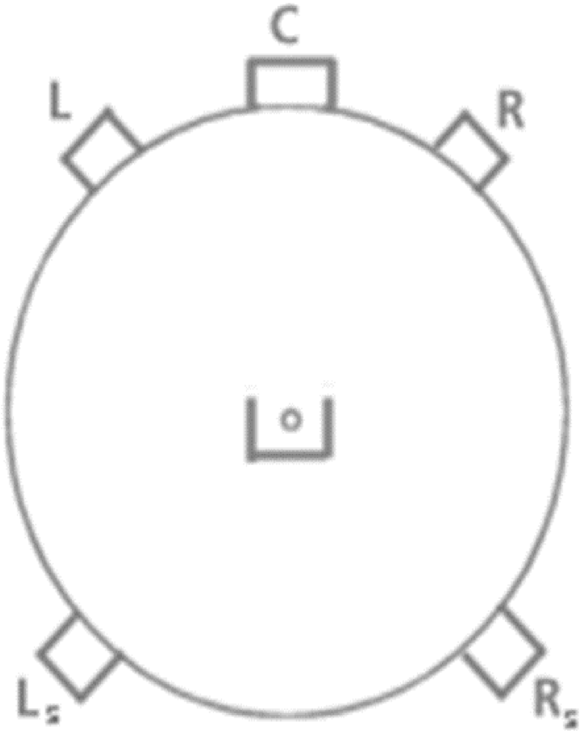
sound quality	bandwidth	mode	Bit rate	reduction ratio
telephone speech	2.5 kHz	mono	8 kbps	96:1
better than AM radio	7.5 kHz	mono	32 kbps	24:1
similar to FM radio	11 kHz	stereo	56...64 kbps	26...24:1
near-CD	15 kHz	stereo	96 kbps	16:1
CD	>15 kHz	stereo	112..128kbps	14..12:1

Advanced Audio Coding (AAC)

- Developed jointly within MPEG by several companies, including Sony and Dolby Labs
- Part of MPEG-2 and MPEG-4 part III
- Codec for iTunes
- Sampling frequencies: between 8 kHz and 96 kHz; number of channels: between 1 and 48
- Unlike MP3 hybrid filter bank, AAC uses MDCT with increased window lengths (2048 points)
- AAC is much more effective than MP3 or MP2 for encoding complex pulses and square waves

5.1 Surround-Sound Coding

- Part of MPEG-2
- "5": left, centre, right, left-surround and right-surround channels;
- ".1": subwoofer for f < 20 Hz
- For compatibility with stereo receivers (= 2 channels), Left and Right are computed as:
- Left = L + 0.71 C + 0.71 L_s
- Right = R + 0.71 C + 0.71 R_s
- Because of correlation between channels: fewer bits are needed for the same channels



Audio Signal: Summary

- Audio signal is anything we (humans) can hear; not all sounds that are present are sound waves
- Audio coding exploits the imperfect human auditory system to remove irrelevant information
- Most effective compression algorithms (to date) use both time and frequency domain approaches to determine what is relevant and what is irrelevant
- Performance of the different layers of MPEG audio coding varies greatly