Compression: Audio Compression (MPEG and others)

As with video a number of compression techniques have been applied to audio.

Simple Audio Compression Methods

RECAP (Already Studied)

Traditional lossless compression methods (Huffman, LZW, etc.) usually don't work well on audio compression

- For the same reason as in image and video compression: Too much change variation in data over a short time

Some Simple But Limited Practical Methods

- Silence Compression - detect the ”silence”, similar to run-length encoding (seen examples before)
- Differential Pulse Code Modulation (DPCM)

  Relies on the fact that difference in amplitude in successive samples is small then we can used reduced bits to store the difference (seen examples before)

Simple But Limited Practical Methods (Cont.)

- Adaptive Differential Pulse Code Modulation (ADPCM) e.g., in CCITT G.721 – 16 or 32 Kbits/sec.

  (a) Encodes the difference between two consecutive signals but a refinement on DPCM,

  (b) Adapts at quantisation so fewer bits are used when the value is smaller.

  – It is necessary to predict where the waveform is heading –> **difficult**

  – Apple had a proprietary scheme called ACE/MACE. Lossy scheme that tries to predict where wave will go in next sample. About 2:1 compression.

Simple But Limited Practical Methods (Cont.)

- Adaptive Predictive Coding (APC) typically used on Speech.
  - Input signal is divided into fixed segments (**windows**)
  - For each segment, some sample **characteristics** are computed, **e.g. pitch, period, loudness**.
  - These characteristics are used to predict the signal
  - Computerised talking (Speech Synthesisers use such methods) but low bandwidth:

Acceptable quality at 8 kbits/sec

Simple But Limited Practical Methods (Cont.)

- Linear Predictive Coding (LPC) fits signal to speech model and then transmits parameters of model as in APC.

  Speech Model:

  - Speech Model:

    Pitch, period, loudness, vocal tract parameters (voiced and unvoiced sounds).

  - Synthesised speech
  - More prediction coefficients than APC – lower sampling rate
  - Still sounds like a computer talking,
  - Bandwidth as low as 2.4 kbits/sec.

Simple But Limited Practical Methods (Cont.)

- Code Excited Linear Predictor (CELP) does LPC, but also transmits error term.
  - Based on more sophisticated model of vocal tract than LPC
  - Better perceived speech quality
  - Audio conferencing quality at 4.8 kbits/sec.

Psychoacoustics or Perceptual Coding

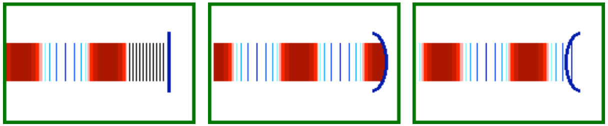
Basic Idea: Exploit areas where the human ear is **less sensitive** to sound to achieve compression

E.g. MPEG audio, Dolby AC.

How do we hear sound?

Sound Revisited

- Sound is produced by a vibrating source.
- The vibrations disturb air molecules
- Produce variations in air pressure: lower than average pressure, **rarefactions**, and higher than average, **compressions**. **This produces sound waves.**
- When a sound wave impinges on a surface (**e.g.** eardrum or microphone) it causes the surface to vibrate in sympathy:



- In this way **acoustic energy** is transferred from a source to a receptor.

Human Hearing

- Upon receiving the the waveform the eardrum vibrates in sympathy
- Through a variety of mechanisms the acoustic energy is transferred to nerve impulses that the brain interprets as sound.

The ear can be regarded as being made up of 3 parts:

- The outer ear,
- The middle ear,
- The inner ear.

Human Ear

We consider:

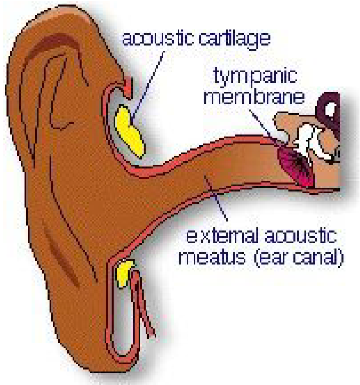
- The function of the main parts of the ear
- How the transmission of sound is processed.

⇒ FLASH EAR DEMO (Lecture ONLY)

Click Here to Run Flash Ear Demo over the Web

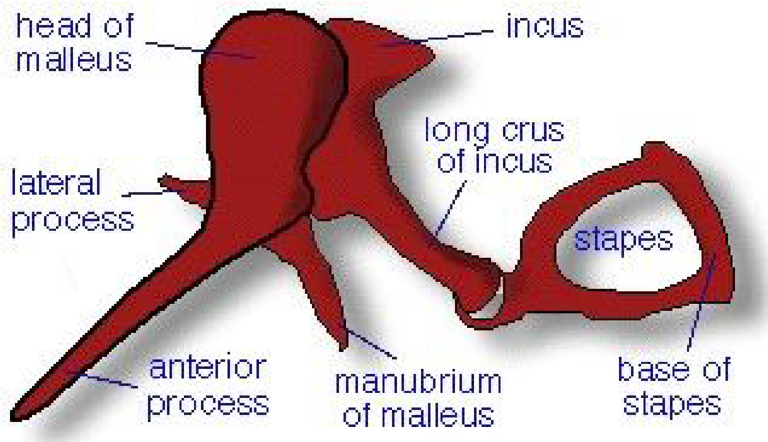
(Shockwave Required)

The Outer Ear



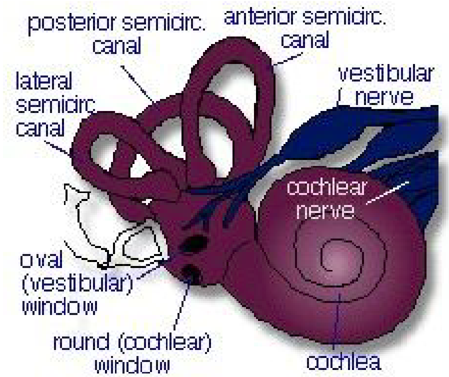
- **Ear Canal**: Focuses the incoming audio.
- **Eardrum (Tympanic Membrane)**:
  - Interface between the external and middle ear.
  - Sound is converted into mechanical vibrations via the middle ear.
  - Sympathetic vibrations on the membrane of the eardrum.

The Middle Ear



- 3 small bones, the **ossicles**:
  Malleus, **Incus**, and **Stapes**.
- Form a system of levers which are linked together and driven by the eardrum
- Bones amplify the force of sound vibrations.

The Inner Ear



The Cochlea:

- Transforms mechanical ossicle forces into hydraulic pressure,
- The cochlea is filled with fluid.
- Hydraulic pressure imparts movement to the cochlear duct and to the organ of Corti.
- Cochlea which is no bigger than the tip of a little finger!!

Semicircular canals

- Body's balance mechanism
- Thought that it plays no part in hearing.

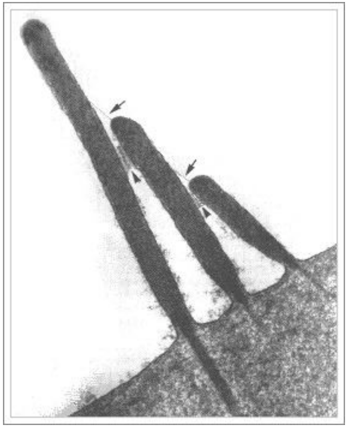
How the Cochlea Works

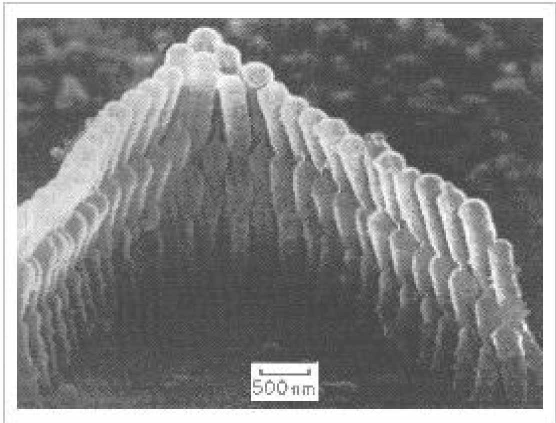
- Pressure waves in the cochlea exert energy along a route that begins at the oval window and ends abruptly at the membrane-covered round window
- Pressure applied to the oval window is transmitted to all parts of the cochlea.

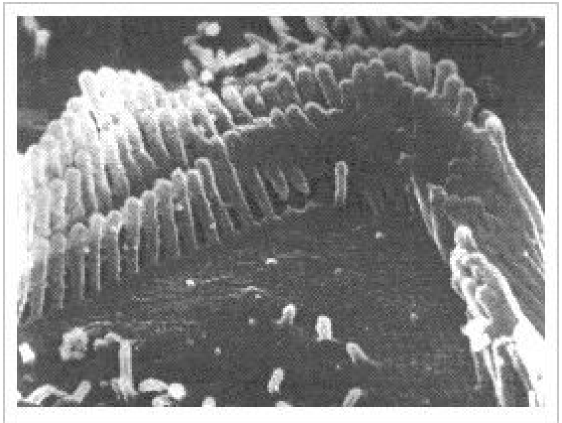
Stereocilia

- Inner surface of the cochlea (**the basilar membrane**) is lined with over 20,000 hair-like nerve cells — **stereocilia**,
- One of the most critical aspects of hearing.

Stereocilia Microscope Images







Hearing different frequencies

- Basilar membrane is tight at one end, looser at the other
- High tones create their greatest crests where the membrane is tight,
- Low tones where the wall is slack.
- Causes resonant frequencies much like what happens in a tight string.
- Stereocilia differ in length by minuscule amounts
- they also have different degrees of resiliency to the fluid which passes over them.

Finally to nerve signals

- Compressional wave moves middle ear through to the cochlea
- Stereocilia will be set in motion.
- Each stereocilia sensitive to a particular frequency.
- Stereocilia cell will resonate with a larger amplitude of vibration.
- Increased vibrational amplitude induces the cell to release an electrical impulse which passes along the auditory nerve towards the brain.

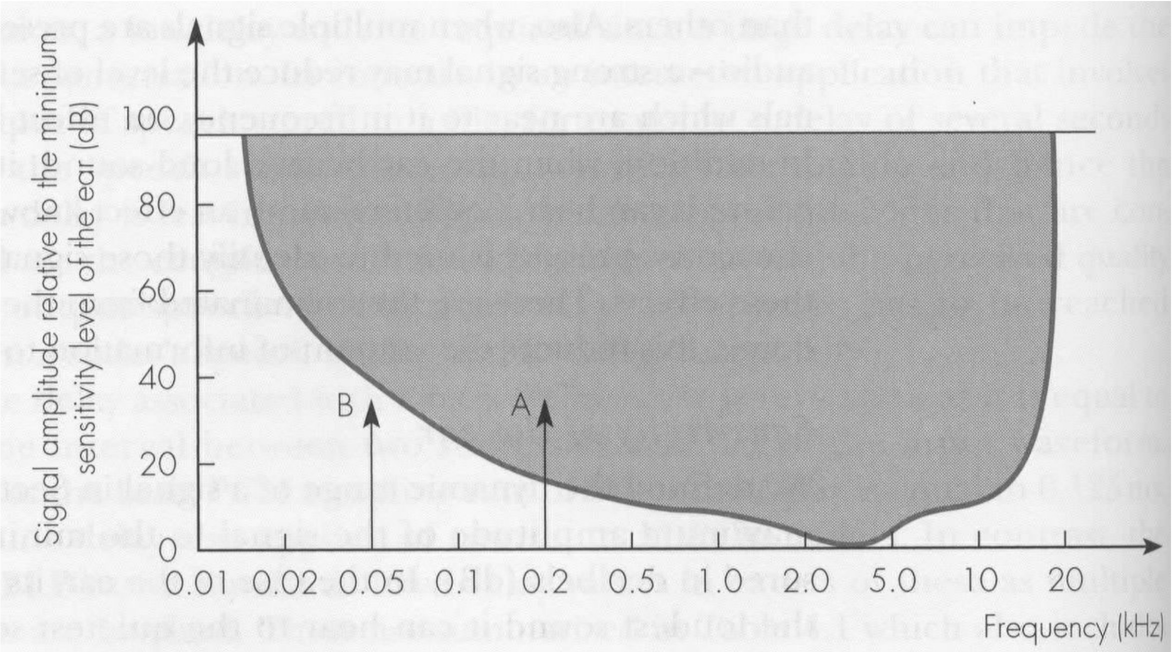
In a process which is not clearly understood, the brain is capable of interpreting the qualities of the sound upon reception of these electric nerve impulses.

Sensitivity of the Ear

- Range is about 20 Hz to 20 kHz, most sensitive at 2 to 4 KHz.
- Dynamic range (quietest to loudest) is about 96 dB
- Approximate threshold of pain: 130 dB
- Hearing damage: $>$ 90 dB (prolonged exposure)
- Normal conversation: 60-70 dB
- Typical classroom background noise: 20-30 dB
- Normal voice range is about 500 Hz to 2 kHz
  - Low frequencies are vowels and bass
  - High frequencies are consonants

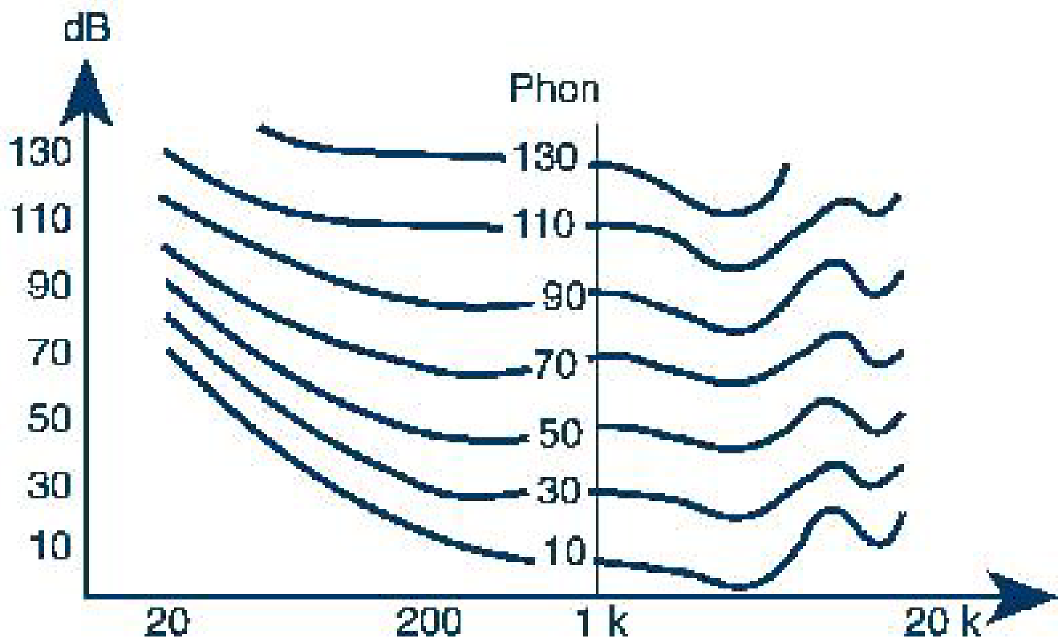
Question: How sensitive is human hearing?

The sensitivity of the human ear with respect to frequency is given by the following graph.

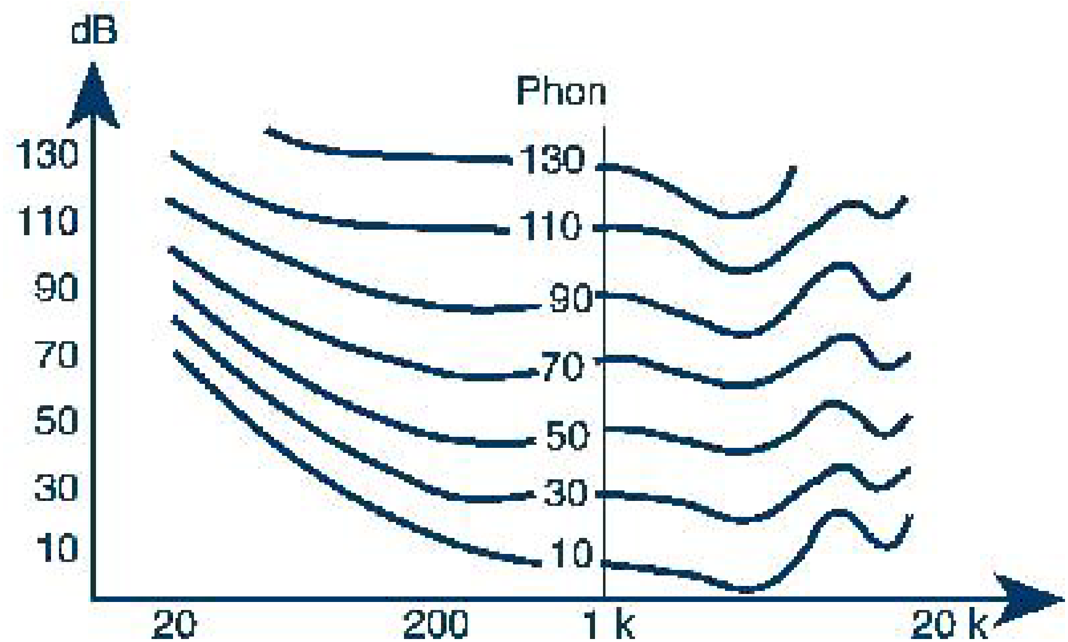


Frequency dependence is also level dependent!

- Ear response is even more complicated.
- Complex phenomenon to explain.
- Illustration : Loudness Curves or Fletcher-Munson Curves:



Loudness Curves or Fletcher-Munson Curves

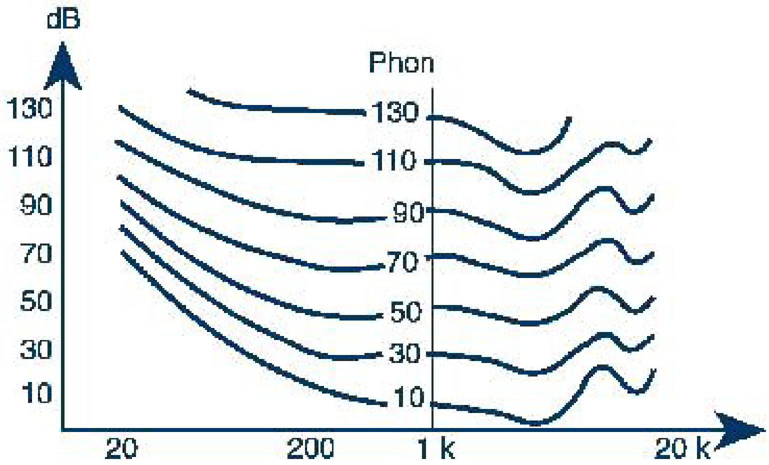


- Connecting pure tone stimuli producing the same perceived loudness ('Phons', in dB).

Perceptual Audio Demos [4]

[4] Also demos relevant for all this section here

What do the curves mean?



- Curves indicate perceived loudness is a function of both the frequency and the level (sinusoidal sound signal)
- Equal loudness curves. Each contour:
  - Equal loudness
  - Express how much a sound level must be changed as the frequency varies, **to maintain a certain perceived loudness**

Physiological Implications

Why are the curves accentuated where they are?

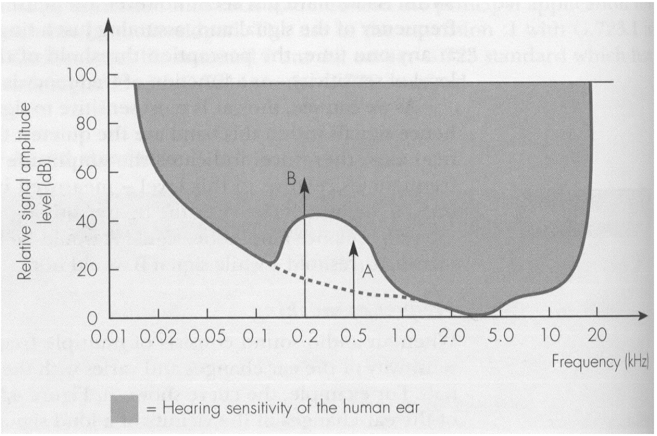
- Accentuates of frequency range to coincide with speech.
- Sounds like **p** and **t** have very important parts of their spectral energy within the accentuated range
- Makes them more easy to discriminate between.

The ability to hear sounds of the **accentuated** range (around a few kHz) is thus vital for speech communication.

Traits of Human Hearing

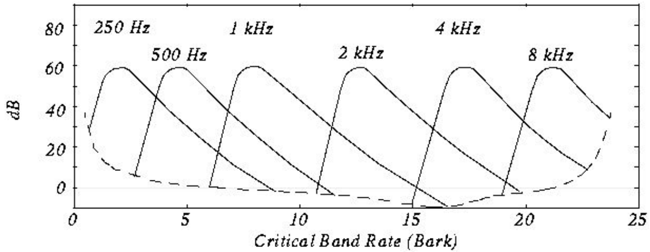
Frequency Masking

- **Multiple frequency audio changes the sensitivity with the relative amplitude of the signals.**
- If the frequencies are close and the amplitude of one is less than the other close frequency then the second frequency may not be heard.



Critical Bands

- Range of closeness for frequency masking depends on the frequencies and relative amplitudes.
- Each **band** where frequencies are masked is called **the Critical Band**
- Critical bandwidth for average human hearing varies with frequency:
  - Constant 100 Hz for frequencies less than 500 Hz
  - Increases (approximately) linearly by 100 Hz for each additional 500 Hz.
- Width of critical band is called a **bark**.



Critical Bands (cont.)

First 12 of 25 critical bands:

Band #	Lower Bound (Hz)	Center (Hz)	Upper Bound (Hz)	Bandwidth (Hz)
1	-	50	100	-
2	100	150	200	100
3	200	250	300	100
4	300	350	400	100
5	400	450	510	110
6	510	570	630	120
7	630	700	770	140
8	770	840	920	150
9	920	1000	1080	160
10	1080	1170	1270	190
11	1270	1370	1480	210
12	1480	1600	1720	240

What is the cause of Frequency Masking?

- **The stereocilia** are excited by air pressure variations, transmitted via outer and middle ear.
- **Different stereocilia** respond to **different ranges** of frequencies — the **critical bands**

Frequency Masking occurs because after excitation by one frequency further excitation by a less strong similar frequency of the same group of cells is not possible.

Click here to hear example of Frequency Masking.

See/Hear also: Click here (in the Masking section).

Temporal masking

After the ear hears a loud sound: **It takes a further short while before it can hear a quieter sound.**

Why is this so?

- **Stereocilia** vibrate with corresponding force of input sound stimuli.
- If the stimuli is strong then stereocilia will be in a high state of excitation and **get fatigued**.
- **Hearing Damage**: After extended listening to loud music or headphones this sometimes manifests itself with ringing in the ears and even temporary deafness.
- **Hearing Damage**: Prolonged exposure to noise permanently damages the **Stereocilia**.

Temporal masking Explained

Temporal Masking occurs because the hairs take time to settle after excitation to respond again.

- **Any loud tone will cause the hearing receptors in the inner ear to become saturated and require time to recover.**

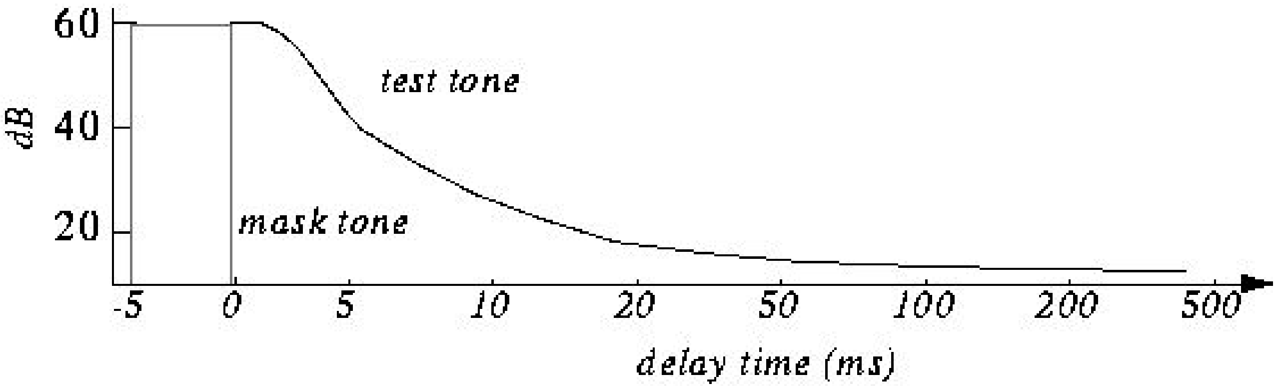
Example of Temporal Masking

- Play 1 kHz *masking tone* at 60 dB, plus a *test tone* at 1.1 kHz at 40 dB. Test tone can't be heard (it's masked).

  Stop masking tone, then stop test tone after a short delay.

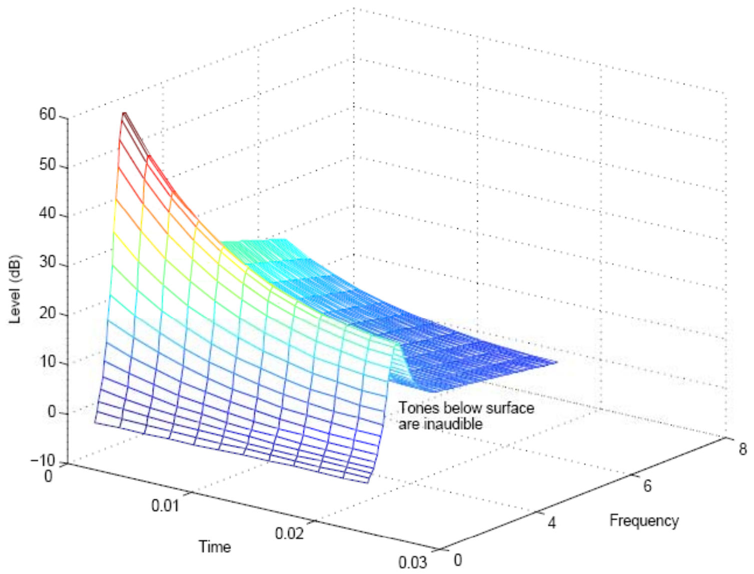
  Adjust delay time to the shortest time that test tone can be heard (e.g., 5 ms).

  Repeat with different level of the test tone and plot:



Example of Temporal Masking (Cont.)

- Try other frequencies for test tone (masking tone duration constant). Total effect of masking:



Compression Idea: How to Exploit?

- If we have a loud tone at, say at 1 kHz, then nearby quieter tones are masked.
- Best compared on **critical band scale** – range of masking is about 1 critical band
- Two factors for masking – **frequency masking** and **temporal masking**
- Question: How to use this for compression?

  Two examples:

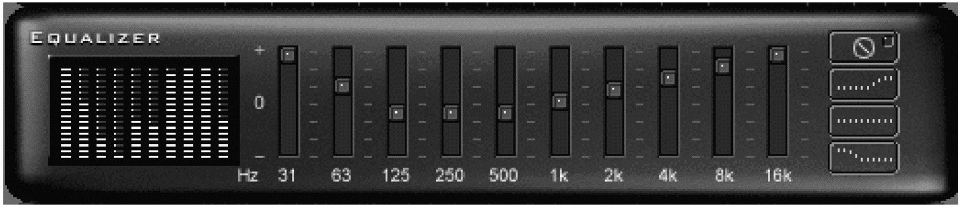
  – **MPEG Audio**

  – **Dolby**

How to compute?

We have met basic tools:

- Bank Filtering with **IIR/FIR** Filters
- **Fourier** and **Discrete Cosine Transforms**
  - Work in **frequency space**
- (Critical) **Band Pass Filtering** — Visualise a graphic equaliser



MPEG Audio Compression

- Exploits the psychoacoustic models above.
- **Frequency masking** is always utilised
- More complex forms of MPEG also employ **temporal masking**.

Basic Frequency Filtering Bandpass

MPEG audio compression basically works by:

- Dividing the audio signal up into a set of frequency subbands
- Use Filter Banks to achieve this.
- Subbands approximate **critical bands**.
- Each band quantised according to the **audibility of quantisation noise**.

Quantisation is the key to MPEG audio compression and is the reason why it is lossy.

How good is MPEG compression?

Although (data) lossy
MPEG claims to be **perceptually lossless**:

- Human tests (part of standard development), Expert listeners.
- 6-1 compression ratio, stereo 16 bit samples at 48 Khz compressed to 256 kbits/sec
- **Difficult**, real world examples used.
- **Under Optimal listening conditions** no statistically distinguishable difference between original and MPEG.

Basic MPEG: MPEG audio coders

- Set of standards for the use of video **with** sound.
- Compression methods or **coders** associated with audio compression are called **MPEG audio coders**.
- MPEG allows for a variety of different coders to employed.
- **Difference** in level of sophistication in applying perceptual compression.
- Different **layers** for levels of sophistication.

An Advantage of MPEG approach

Complex psychoacoustic modelling only in **coding phase**

- Desirable for real time (Hardware or software) decompression
- Essential for broadcast purposes.
- Decompression is independent of the psychoacoustic models used
- Different models can be used
- If there is enough bandwidth no models at all.

Basic MPEG: MPEG Standards

Evolving standards for MPEG audio compression:

- MPEG-1 is by the most prevalent
- So called `mp3` files we get off Internet are members of MPEG-1 family.
- Standards now extends to MPEG-4 (structured audio) — **Earlier Lecture**.

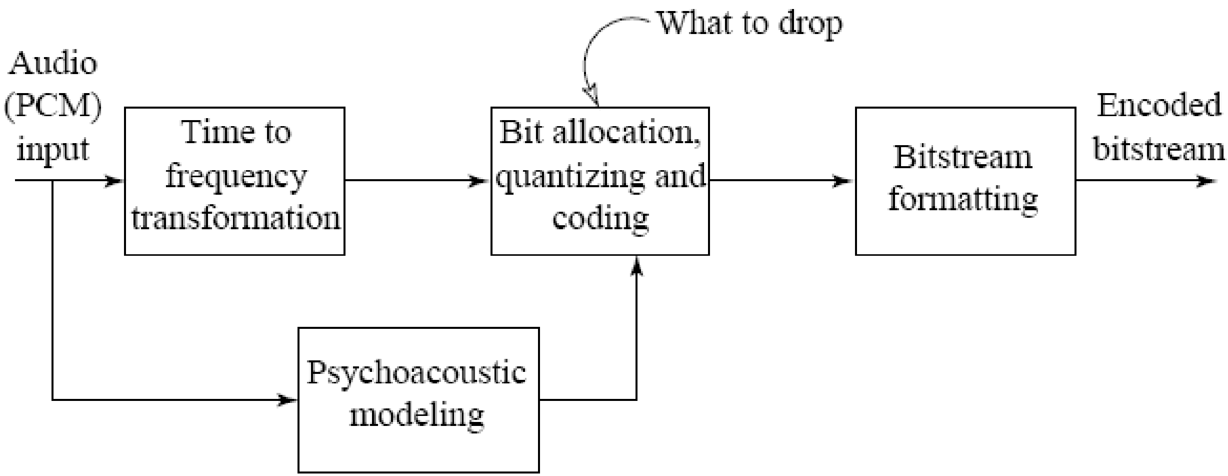
For now we concentrate on MPEG-1

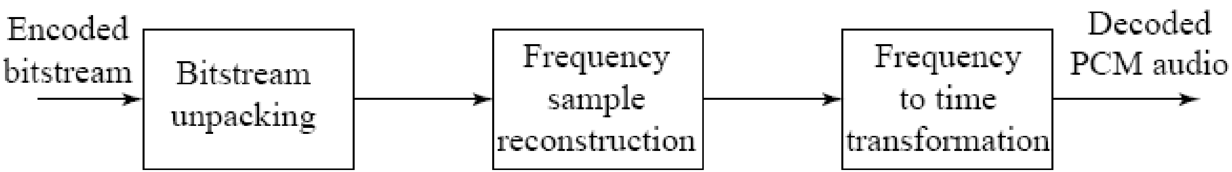
Basic MPEG: MPEG Facts

- MPEG-1: 1.5 Mbits/sec for audio and video
  About 1.2 Mbits/sec for video, 0.3 Mbits/sec for audio
  (Uncompressed CD audio is
  44,100 samples/sec * 16 bits/sample * 2 channels $>$ 1.4 Mbits/sec)
- Compression factor ranging from 2.7 to 24.
- MPEG audio supports sampling frequencies of 32, 44.1 and 48 KHz.
- Supports one or two audio channels in one of the four modes:
  1. Monophonic – single audio channel
  2. Dual-monophonic – two independent channels
     (functionally identical to stereo)
  3. Stereo – for stereo channels that share bits, but not using joint-stereo coding
  4. Joint-stereo – takes advantage of the correlations between stereo channels

Basic MPEG-1 Encoding/Decoding algorithm

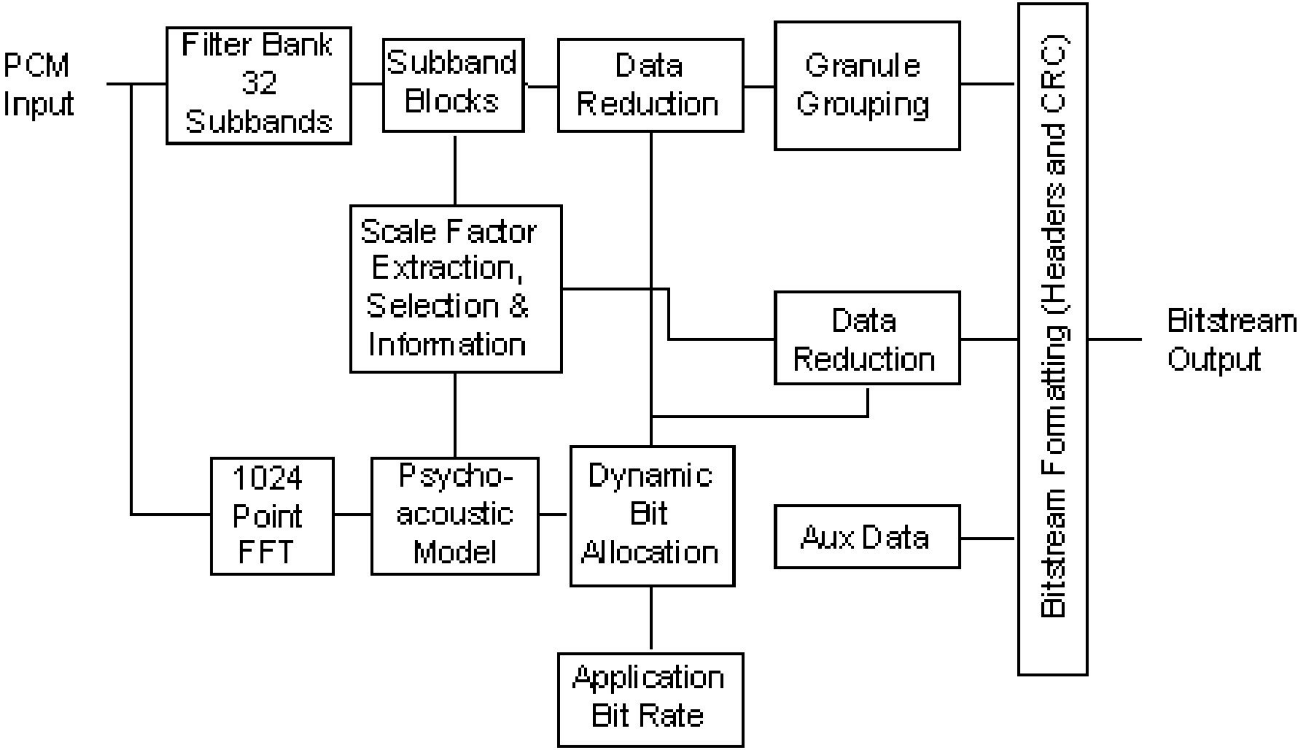
Basic MPEG-1 encoding/decoding maybe summarised as:





Basic MPEG-1 Compression algorithm (1)

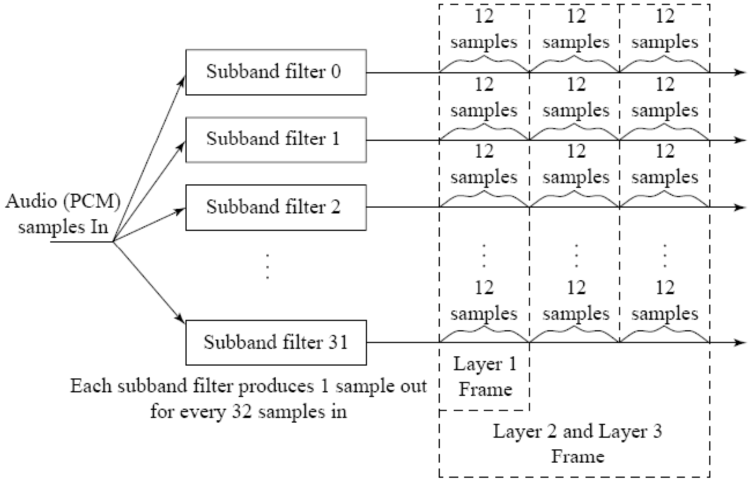
Basic encoding algorithm in more detail summarised below:



Basic MPEG-1 Compression algorithm (2)

The main stages of the algorithm are:

- The audio signal is first samples and quantised use PCM
  - Application dependent: Sample rate and number of bits
- The PCM samples are then divided up into a number of **frequency subband** and compute **subband scaling factors**:



Basic MPEG-1 Compression algorithm (3)

Analysis filters

- Also called **critical-band filters**
- Break signal up into equal width subbands
- Use Use Filter Banks (modified with discrete cosine transform (DCT) Level 3)
- Filters divide audio signal into frequency subbands that approximate the 32 critical bands
- Each band is known as a *sub-band sample*.
- **Example**: 16 kHz signal frequency, Sampling rate 32 kbits/sec gives each subband a bandwidth of 500 Hz.
- Time duration of each sampled segment of input signal is time to accumulate 12 successive sets of 32 PCM (subband) samples, **i.e.** 32*12 = 384 samples.

Basic MPEG-1 Compression algorithm (4)

Analysis filters (cont)

- In addition to filtering the input, analysis banks determine
  - Maximum amplitude of 12 subband samples in each subband.
  - Each known as the **scaling factor** of the subband.
  - Passed to *psychoacoustic model* and **quantiser blocks**

Basic MPEG-1 Compression algorithm (5)

Psychoacoustic modeller:

- Frequency Masking and may employ temporal masking.
- Performed concurrently with filtering and analysis operations.
- Uses Fourier Transform (FFT) to perform analysis.
- Determine amount of masking for each band caused by nearby bands.
- Input: set hearing thresholds and subband masking properties (model dependent) and scaling factors (above).

Basic MPEG-1 Compression algorithm (6)

Psychoacoustic modeller (cont):

- Output: a set of **signal-to-mask** ratios:
  - Indicate those frequencies components whose amplitude is below the audio threshold.
  - If the power in a band is below the masking threshold, don't encode it.
  - Otherwise, determine number of bits (from scaling factors) needed to represent the coefficient such that noise introduced by quantisation is below the masking effect (Recall that 1 bit of quantisation introduces about 6 dB of noise).

Basic MPEG-1 Compression algorithm (7)

Example of Quantisation:

- Assume that after analysis, the first levels of 16 of the 32 bands are:

Band	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Level (db)	0	8	12	10	6	2	10	60	35	20	15	2	3	5	3	1

- If the level of the 8th band is 60 dB,

  then assume (according to model adopted) it gives a masking of 12 dB in the 7th band, 15 dB in the 9th.

  Level in 7th band is 10 dB ($<$ 12 dB), so ignore it.

  Level in 9th band is 35 dB ($>$ 15 dB), so send it.

  –$>$ Can encode with up to 2 bits (= 12 dB) of quantisation error.

- **More on Bit Allocation soon.**

MPEG-1 Output bitstream

The basic output stream for a basic MPEG encoder is as follows:

Header	SBS Forrmat	12x32 subband	Ancillary data

- **Header**: contains information such as the sample frequency and quantisation,.
- **Subband sample (SBS) format**: Quantised scaling factors and 12 frequency components in each subband.
  - Peak amplitude level in each subband quantised using 6 bits (64 levels)
  - 12 frequency values quantised to 4 bits
- **Ancillary data**: Optional. Used, for example, to carry additional coded samples associated with special broadcast format (**e.g surround sound**)

Decoding the bitstream

- Dequantise the subband samples after demultiplexing the coded bitstream into subbands.
- **Synthesis bank** decodes the dequantised subband samples to produce PCM stream.
  - This essentially involves applying the inverse fourier transform (**IFFT**) on each substream and multiplexing the channels to give the PCM bit stream.

MPEG Layers

MPEG defines 3 levels of processing layers for audio:

- Level 1 is the basic mode,
- Levels 2 and 3 more advance (use temporal masking).
- Level 3 is the most common form for audio files on the Web
  - Our beloved MP3 files that record companies claim are bankrupting their industry.
  - Strictly speaking these files should be called **MPEG-1 level 3** files.

Each level:

- Increasing levels of sophistication
- Greater compression ratios.
- Greater computation expense

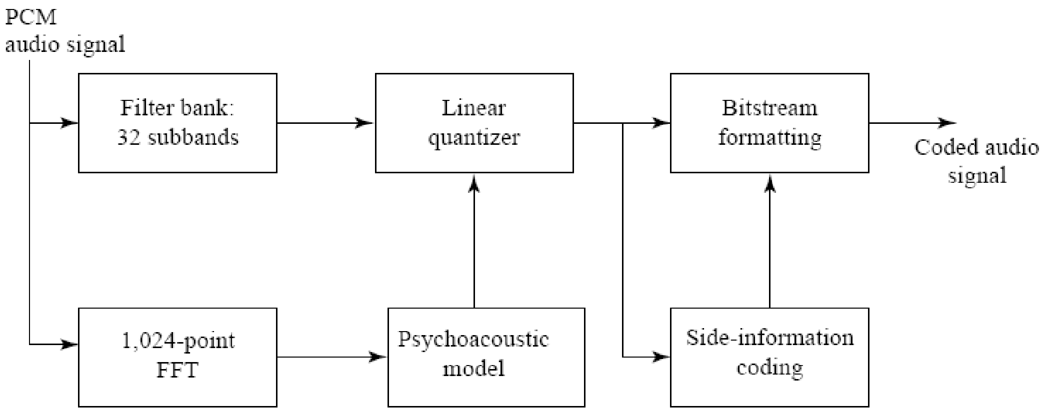
Level 1

- Best suited for bit rate bigger than 128 kbits/sec per channel.
- Example: Phillips Digital Compact Cassette uses Layer 1 192 kbits/sec compression
- Divides data into frames,
  - Each of them contains 384 samples,
  - 12 samples from each of the 32 filtered subbands as shown above.
- Psychoacoustic model only uses frequency masking.
- Optional Cyclic Redundancy Code (CRC) error checking.

Level 1 (and Level 2) Audio Layers

Note: Mask Calculations done in parallel with subband filtering

- Accurate frequency decomposition via Fourier Transform.



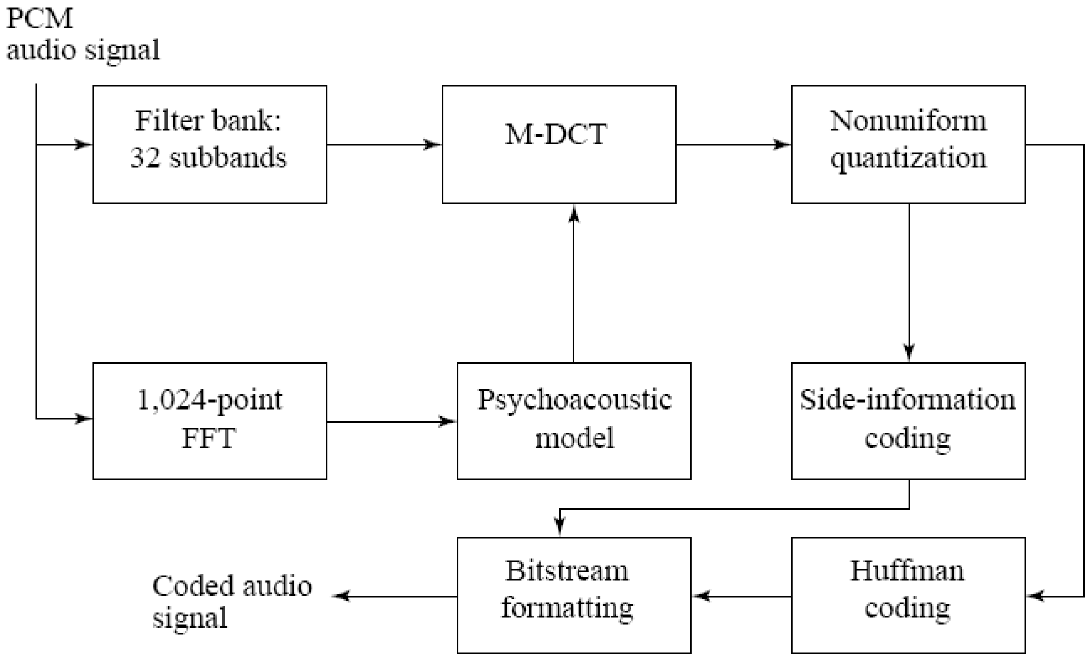
Layer 2

- Targeted at bit rates of around 128 kbits/sec per channel.
- Examples: Coding of Digital Audio Broadcasting (DAB) on CD-ROM, CD-I and Video CD.
- Enhancement of level 1.
- Codes audio data in larger groups:
  - Use three frames in filter:
    before, current, next, a total of 1152 samples.
  - This models a little bit of the temporal masking.
- Imposes some restrictions on bit allocation in middle and high subbands.
- More compact coding of scale factors and quantised samples.
- Better audio quality due to saving bits here so more bits can be used in quantised subband values

Layer 3

- Targeted at bit rates of 64 kbits/sec per channel.
- Example: audio transmission of ISDN or suitable bandwidth network.
- Much more complex approach.
- Psychoacoustic model includes temporal masking effects,
- Takes into account stereo redundancy.
- Better critical band filter is used (non-equal frequencies)
- Uses a modified DCT (MDCT) for lossless subband transformation.
- Two different block lengths: 18 (long) or 6 (short)
- 50% overlap between successive transform windows gives window sizes of 36 or 12 — **accounts for temporal masking**
- Greater frequency resolution accounts for poorer time resolution
- Uses Huffman coding on quantised samples for better compression.

Level 3 Audio Layers



Bit Allocation

- Process determines the number of code bits for each subband
- Based on information from the psychoacoustic model.

Bit Allocation For Layer 1 and 2

- Compute the mask-to-noise ratio (MNR) for all subbands:

  $$MNR_{dB} = SNR_{dB} - SMR_{dB}$$

  where

  MNR_{dB} is the mask-to-noise ratio,
  SNR_{dB} is the signal-to-noise ratio (SNR), and
  SMR_{dB} is the signal-to-mask ratio from the psychoacoustic model.

- Standard MPEG lookup tables estimate SNR for given quantiser levels.

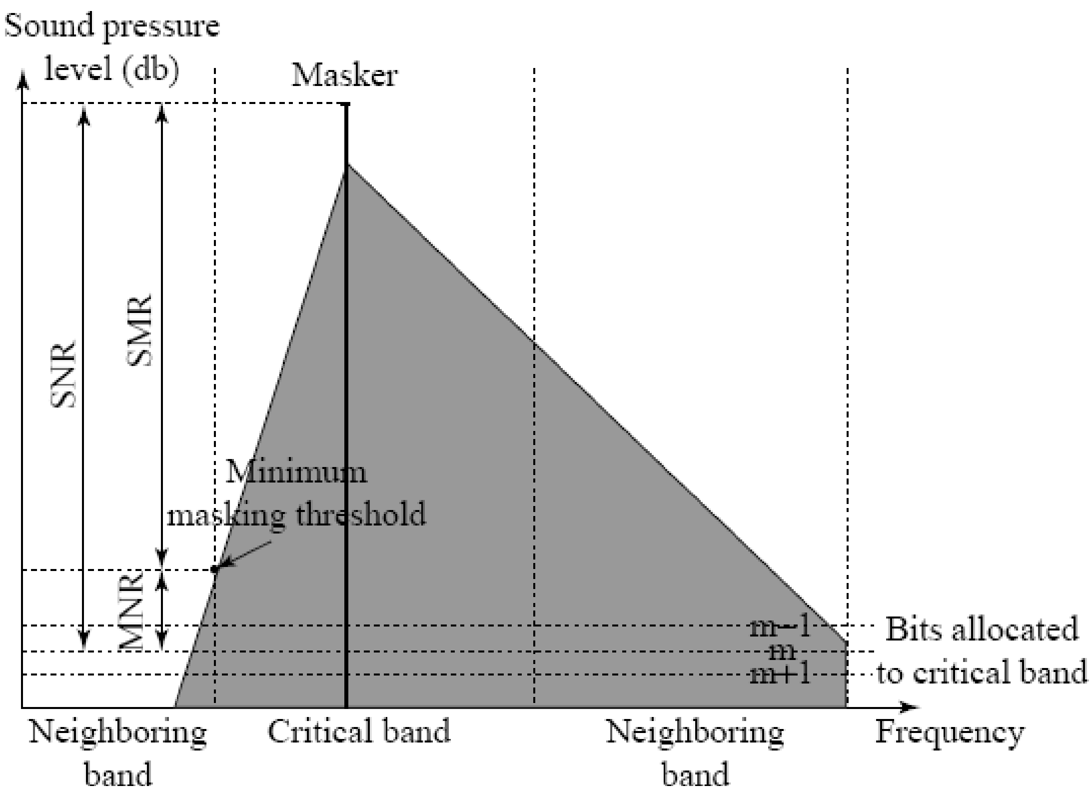
- Designers are free to try other methods SNR estimation.

Bit Allocation For Layer 1 and 2 (cont.)

Once MNR computed for all the subbands:

- Search for the subband with the lowest MNR
- Increment code bits to that subband.
- When a subband gets allocated more code bits, the bit allocation unit:
  - Looks up the new estimate for SNR
  - Recomputes that subband's MNR.
- The process repeats until no more code bits can be allocated.

Bit Allocation For Layer 1 and 2 (cont.)



Bit Allocation For Layer 3

- Uses **noise allocation**, which employs Huffman coding.
- Iteratively varies the quantisers in an orderly way
  - Quantises the spectral values,
  - Counts the number of Huffman code bits required to code the audio data
  - Calculates the resulting noise in Huffman coding.

    If there exist scale factor bands with more than the allowed distortion:

    - Encoder amplifies values in bands
    - **To effectively decreases** the quantiser step size for those bands.

Bit Allocation For Layer 3 (Cont.)

After this the process repeats. The process stops if any of these three conditions is true:

- None of the scale factor bands have more than the allowed distortion.
- The next iteration would cause the amplification for any of the bands to exceed the maximum allowed value.
- The next iteration would require all the scale factor bands to be amplified.

Real-time encoders include a time-limit exit condition for this process.

Stereo Redundancy Coding

Exploit redundancy in two couple stereo channels?

- Another perceptual property of the human auditory system
- Simply stated at low frequencies, the human auditory system can't detect where the sound is coming from.
  - So save bits and encode it mono.
- **Used in MPEG-1 Layer 3.**

Two types of stereo redundancy coding:

- Intensity stereo coding — **all layers**
- Middle/Side (MS) stereo coding — **Layer 3 only** stereo coding.

Intensity stereo coding

Encoding:

- Code some upper-frequency subband outputs:
  - A single summed signal instead of sending independent left and right channels codes
  - Codes for each of the 32 subband outputs.

Decoding:

- Reconstruct left and right channels
  - Based only on a single summed signal
  - Independent left and right channel scale factors.

With intensity stereo coding,

- The spectral shape of the left and right channels is the same within each intensity-coded subband
- **But** the magnitude is different.

Middle/Side (MS) stereo coding

- Encodes the left and right channel signals in certain frequency ranges:
  - **Middle** — sum of left and right channels
  - **Side** — difference of left and right channels.
- Encoder uses specially tuned threshold values to compress the side channel signal further.

MATLAB MPEG Audio Coding Code

MPEGAudio (DIRECTORY)
MPEGAudio.zip (All Files Zipped)

Dolby Audio Compression

Application areas:

- FM radio Satellite transmission and broadcast TV audio (DOLBY AC-1)
- Common compression format in PC sound cards (DOLBY AC-2)
- High Definition TV standard **advanced television** (ATV) (DOLBY AC-3). *MPEG a competitor in this area.*

Differences with MPEG

- MPEG perceptual coders control quantisation accuracy of each subband by computing bit numbers for each sample.
- MPEG needs to store each quantise value with each sample.
- MPEG Decoder uses this information to dequantise: **forward adaptive bit allocation**
- **Advantage of MPEG?**: no need for psychoacoustic modelling in the decoder due to store of every quantise value.
- **DOLBY**: Use **fixed bit rate allocation** for each subband.
  - No need to send with each frame — as in **MPEG**.
  - **DOLBY** encoders and decoder need this information.

Fixed Bit Rate Allocation

- Bit allocations are determined by known sensitivity characteristics of the ear.

Different Dolby standards

DOLBY AC-1 :

Low complexity psychoacoustic model

- 40 subbands at sampling rate of 32 kbits/sec or
- (Proportionally more) Subbands at 44.1 or 48 kbits/sec
- Typical compressed bit rate of 512 kbits per second for stereo.
- Example: FM radio Satellite transmission and broadcast TV audio

DOLBY AC-2 :

Variation to allow subband bit allocations to vary

- **NOW** Decoder needs copy of psychoacoustic model.
- Minimised encoder bit stream overheads at expense of transmitting encoded frequency coefficients of sampled waveform segment — known as the **encoded spectral envelope**.
- Mode of operation known as
  backward adaptive bit allocation mode
- HIgh (hi-fi) quality audio at 256 kbits/sec.
- Not suited for broadcast applications:
  - encoder cannot change model without changing (remote/distributed) decoders
- Example: Common compression format in PC sound cards.

DOLBY AC-3 :

Development of AC-2 to overcome broadcast challenge

- Use **hybrid backward/forward adaptive bit allocation mode**
- Any model modification information is encoded in a frame.
- Sample rates of 32, 44.1, 48 kbits/sec supported depending on bandwidth of source signal.
- Each encoded block contains 512 subband samples, with 50% (256) overlap between successive samples.
- For a 32 kbits/sec sample rate each block of samples is of 8 ms duration, the duration of each encoder is 16 ms.
- Audio bandwidth (at 32 kbits/sec) is 15 KHz so each subband has 62.5 Hz bandwidth.
- Typical stereo bit rate is 192 kbits/sec.
- Example: High Definition TV standard **advanced television** (ATV). MPEG competitor in this area.

Further Reading

- A tutorial on MPEG audio compression
- AC-3: flexible perceptual coding for audio trans. & storage