Physics of Music PHY103

Loudness

Physical units

For a sound wave

- The amplitude of the pressure variation
- The amplitude of velocity variations
- The power in the wave --- that means the energy carried in the wave

Decibel scale

- A sound that is ten times more powerful is 10 decibels higher.
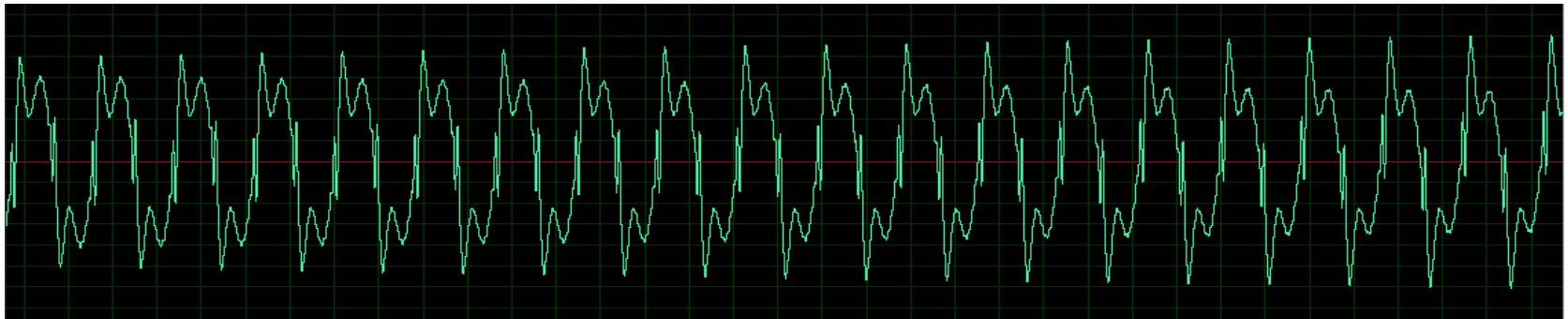
- What do we mean by ten times more powerful?

Ten times the power.

- Decibel is abbreviated dB

Amplitude in pressure or velocity would be related to the square root of the power.
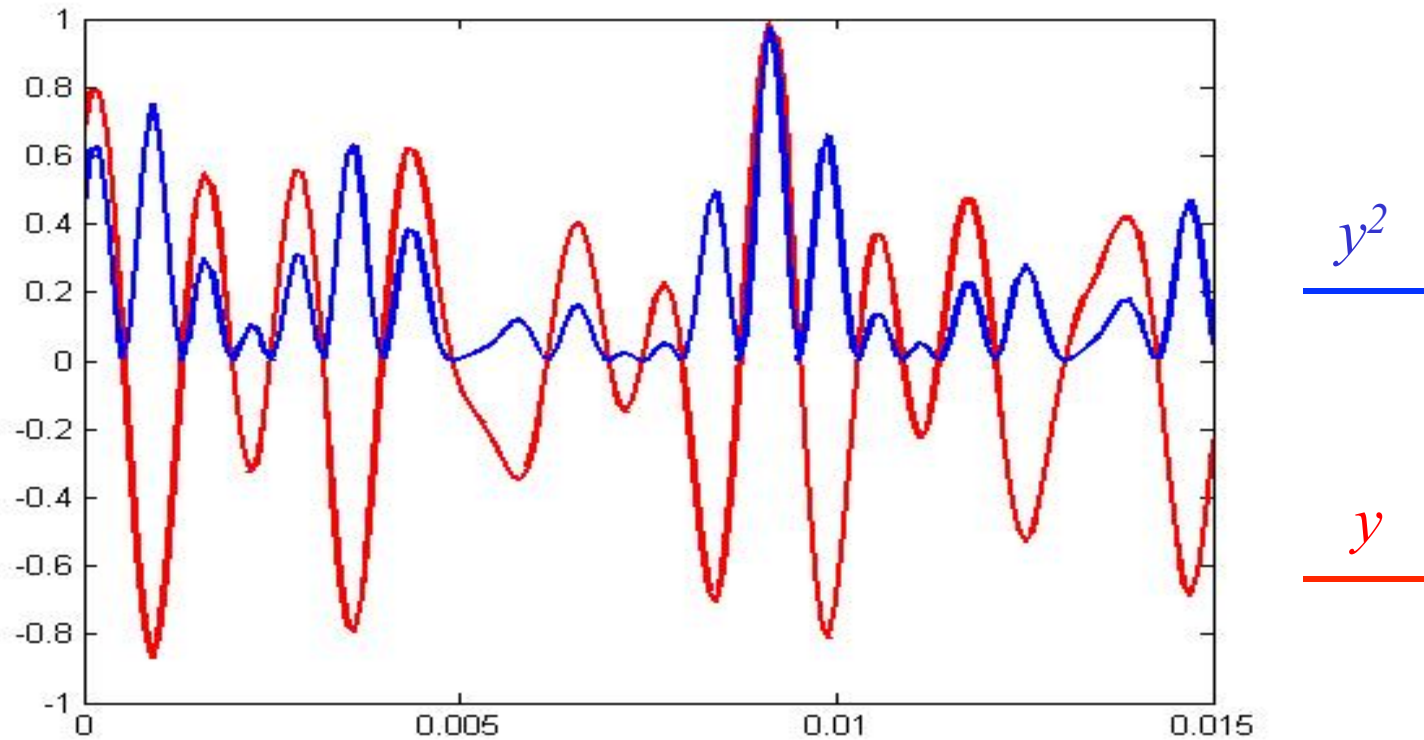
dB, amplitude and power

Energy of a harmonic oscillator depends on the square of the amplitude

Average value of pressure or air velocity is zero for a sound wave

The mean is not a good estimate for power or loudness

Square of sound level and power

The square of the air pressure or air velocity is never zero. Its average gives you the power in the signal
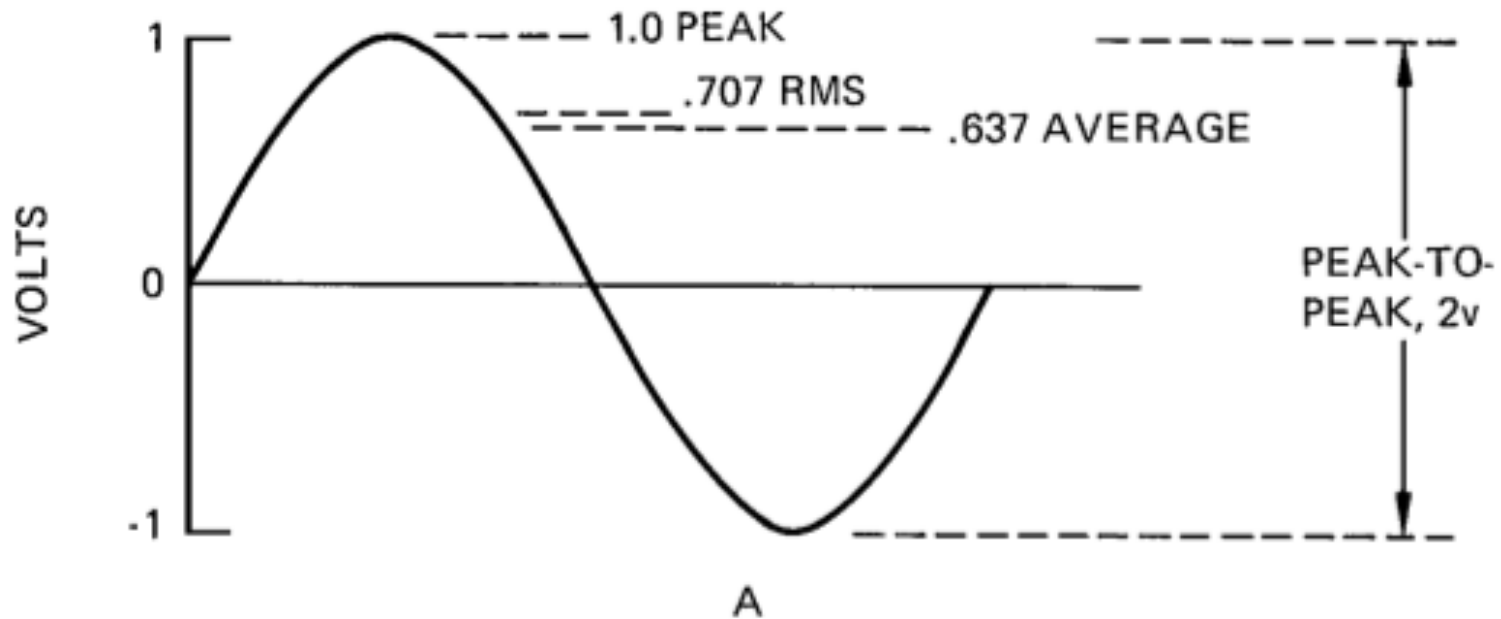
Peak and RMS
(Root Mean Square)

Figure from JBL Sound System Design Reference Manual

dB scale is logarithmic

SIL = Sound intensity level in dB

P = sound power

SIL = 10 $\log_{10}$ (P) + constant

Logarithmic scale like astronomical magnitudes or the earthquake Richter scale

A sound that is 10^7 times more powerful than the threshold of hearing has what db level?
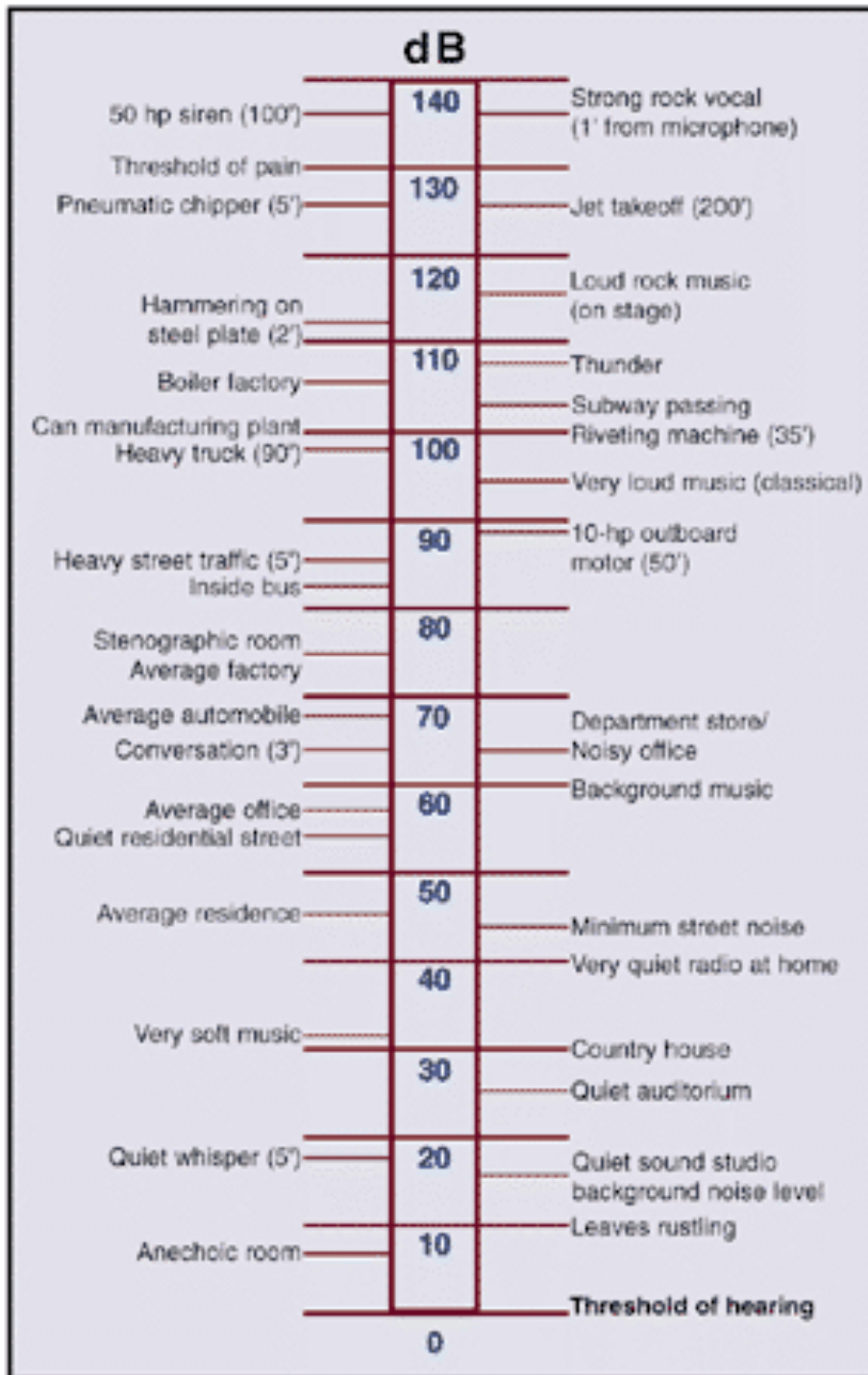
dB is a relative scale

Answer: $10 \log_{10}(10^7) = 70\text{dB}$

Since the dB scale is defined with respect to the threshold of hearing, it is easier to use it as a relative scale, rather than an absolute scale.

You can describe one sound as x dB above another sound and you will know how much more powerful it is.

$P_1/P_2 = 10^{x/10}$

dB	
50 hp siren (100')	140 — Strong rock vocal (1' from microphone)
Threshold of pain	130
Pneumatic chipper (5')	— Jet takeoff (200')
	120 — Loud rock music (on stage)
Hammering on steel plate (2')	
	110 — Thunder
Boiler factory	— Subway passing
Can manufacturing plant	— Riveting machine (35')
Heavy truck (90')	100
	— Very loud music (classical)
	90 — 10-hp outboard motor (50')
Heavy street traffic (5')	
Inside bus	
Stenographic room	80
Average factory	
Average automobile	70 — Department store/
Conversation (3')	— Noisy office
	— Background music
Average office	60
Quiet residential street	
Average residence	50
	— Minimum street noise
	— Very quiet radio at home
	40
Very soft music	
	30 — Country house
	— Quiet auditorium
Quiet whisper (5')	20 — Quiet sound studio background noise level
	— Leaves rustling
Anechoic room	10
	Threshold of hearing
	0

Over what range of sound power are our ears sensitive to?
In other words: how much more powerful are deafening sounds than barely audible sounds?

Combining sound levels

- Supposing you have two speakers and each produces a noise of level 70dB at your location.

- What would the combined sound level be?

Combining sound levels

1) The two sounds have the same frequency and are in phase:

Amplitudes add, the total amplitude is twice that of the original signal.

The power is then 4 times original signal.

The change in dB level would be $10\log_{10}(4)=6.0$

70dB + 6dB = 76dB for total signal

Combining Sound Levels

2) The two signals don't have the same frequency

They go in and out of phase. In this case the power's add instead of the amplitudes.

We expect an increase in dB level of $10\log_{10}(2)=3.0$

70dB + 3dB=73dB for total signal

Combining signals

1) Summation of two signals with the same frequency and amplitude

The signal $\qquad A\cos(2\pi f t)$

   -its square $\qquad A^2 \cos^2(2\pi f t)$

   -the time average $\qquad A^2/2$
   of the square

   -the power $\qquad A^2/2$

Now the sum of two identical signals $\qquad 2A\cos(2\pi f t)$

   -The square $\qquad 4A^2 \cos^2(2\pi f t)$

   -the power is four times as big $\qquad 2A^2$

Combining Signals

2) Two signals with different frequencies

$A \cos(2\pi f t)$ with power A²/2

$A \cos(2\pi g t)$ with power A²/2

Add the two and take the square

$$A^2 \left[\cos^2(2\pi f t) + \cos^2(2\pi g t) + 2\cos(2\pi f t)\cos(2\pi g t) \right]$$

| averages to 1/2 | averages to 1/2 | averages to 0 |

Power is then A^2 which is twice the power of one signal

Phase and is it important?

- For low frequencies the wavelength is large
- Consider a120Hz signal
- The wavelength is $\lambda = v / f$

$$\lambda = v / f = \frac{334\text{m/s}}{120\text{Hz}} = 2.78\text{m}$$

- Constructive and destructive interference can be noticeable when frequencies are low and so wavelengths are large
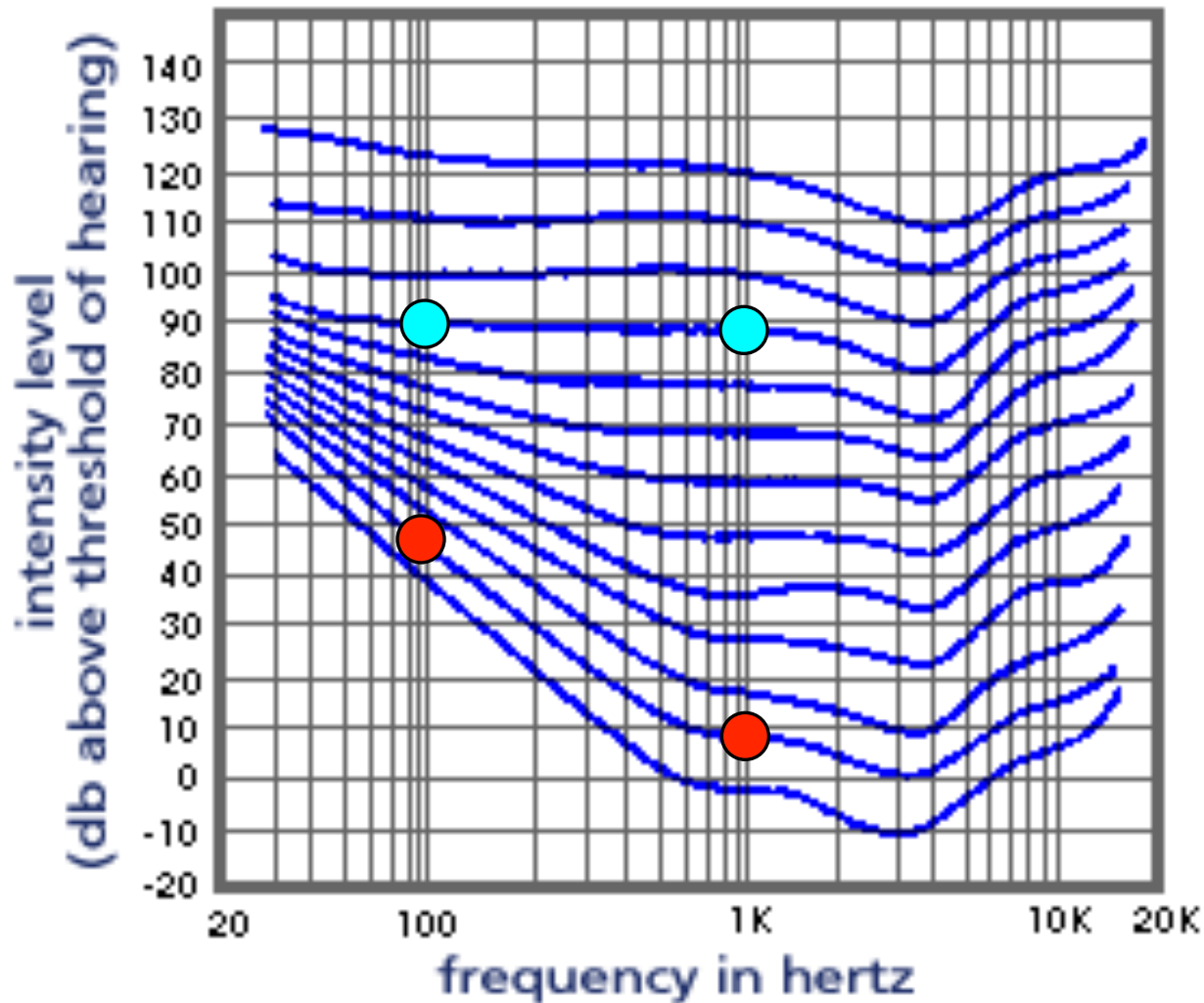
equal loudness curves
(based on Fletcher and Munson)
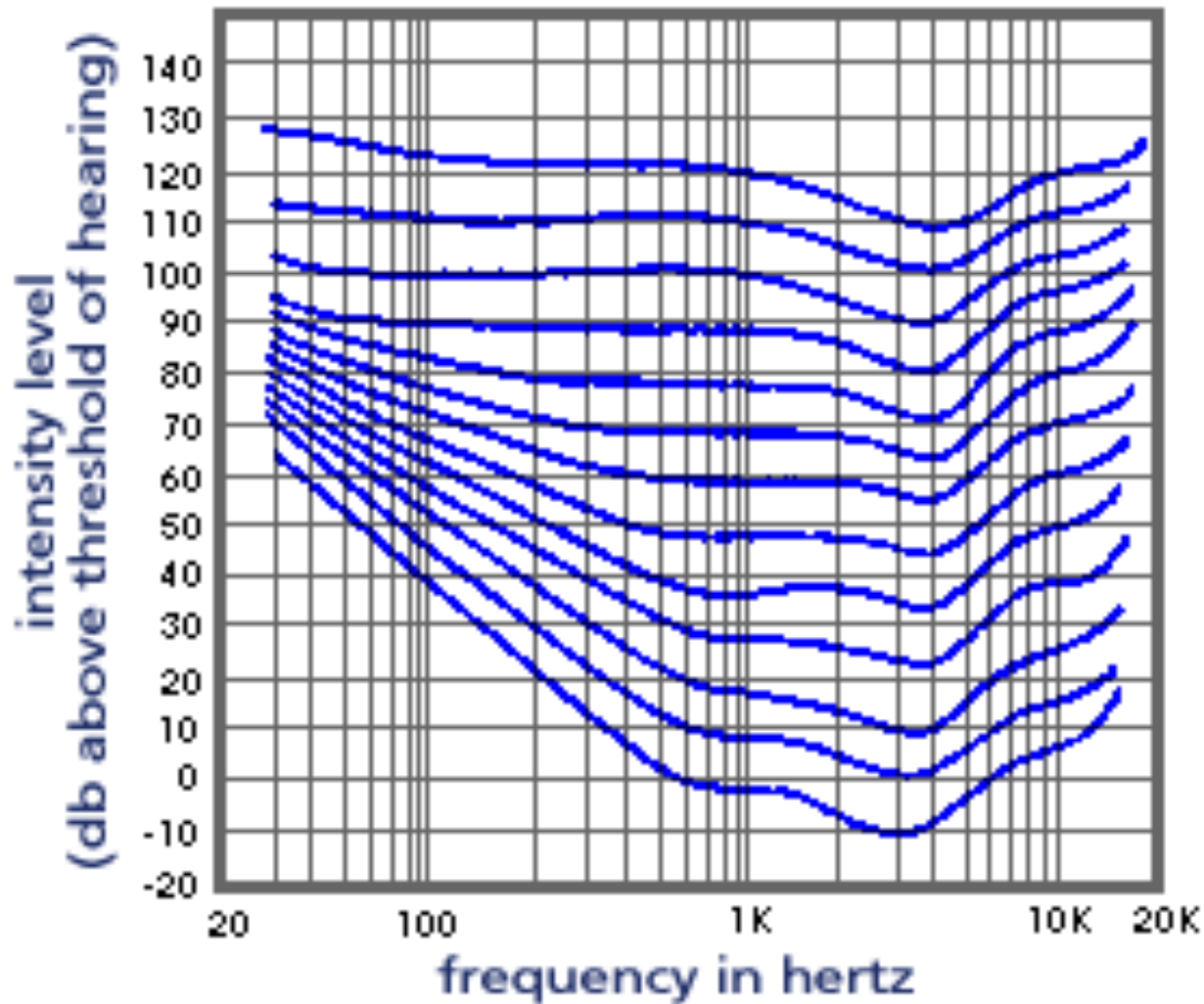
Perceived loudness depends on pitch

equal loudness curves
(based on Fletcher and Munson)
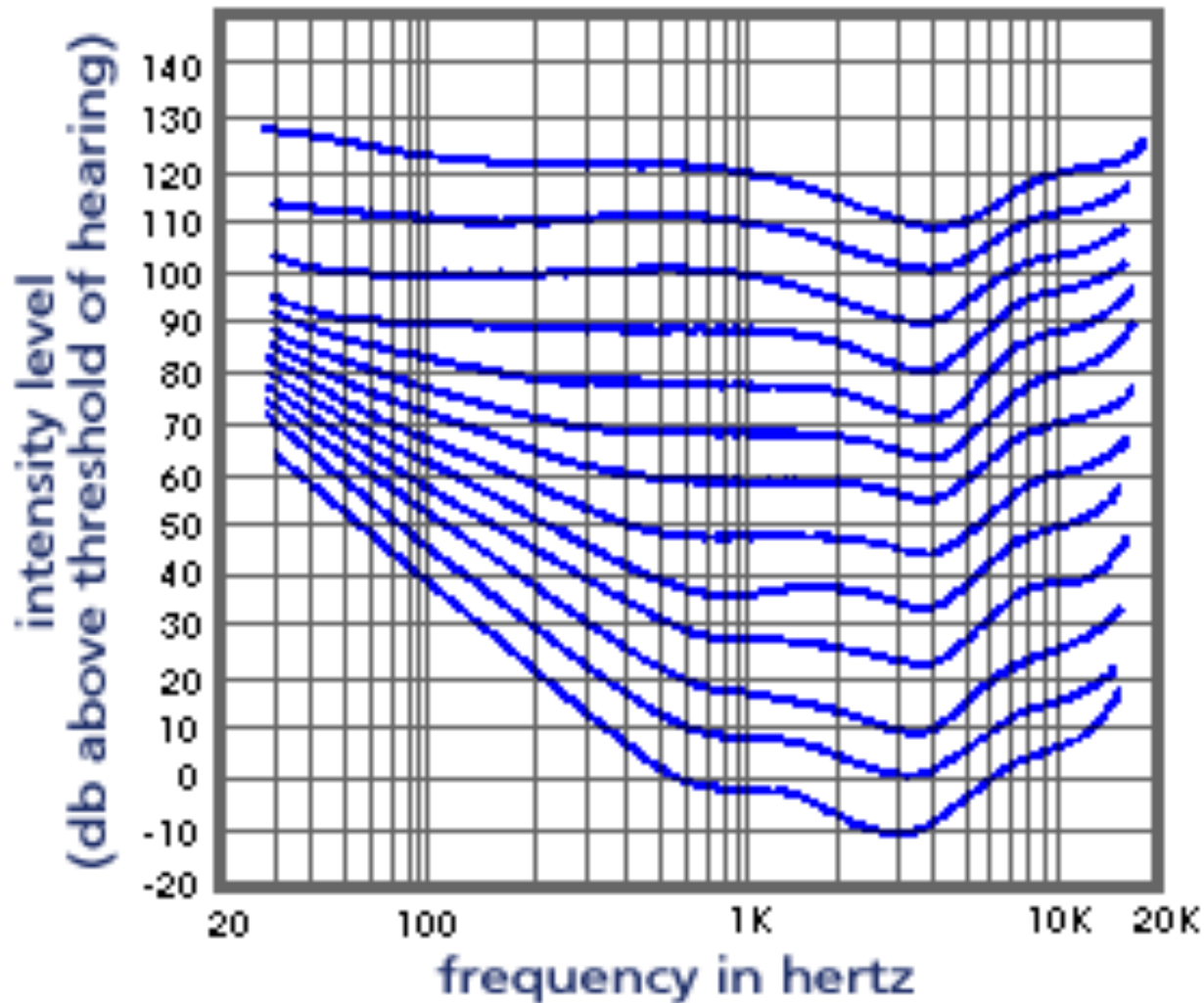
Points along one curve are all the same perceived loudness

equal loudness curves
(based on Fletcher and Munson)

Loud music need less bass.

The mix you need depends on the db level.

If you change the volume you need to change the mix.

Fletcher-Munson diagram
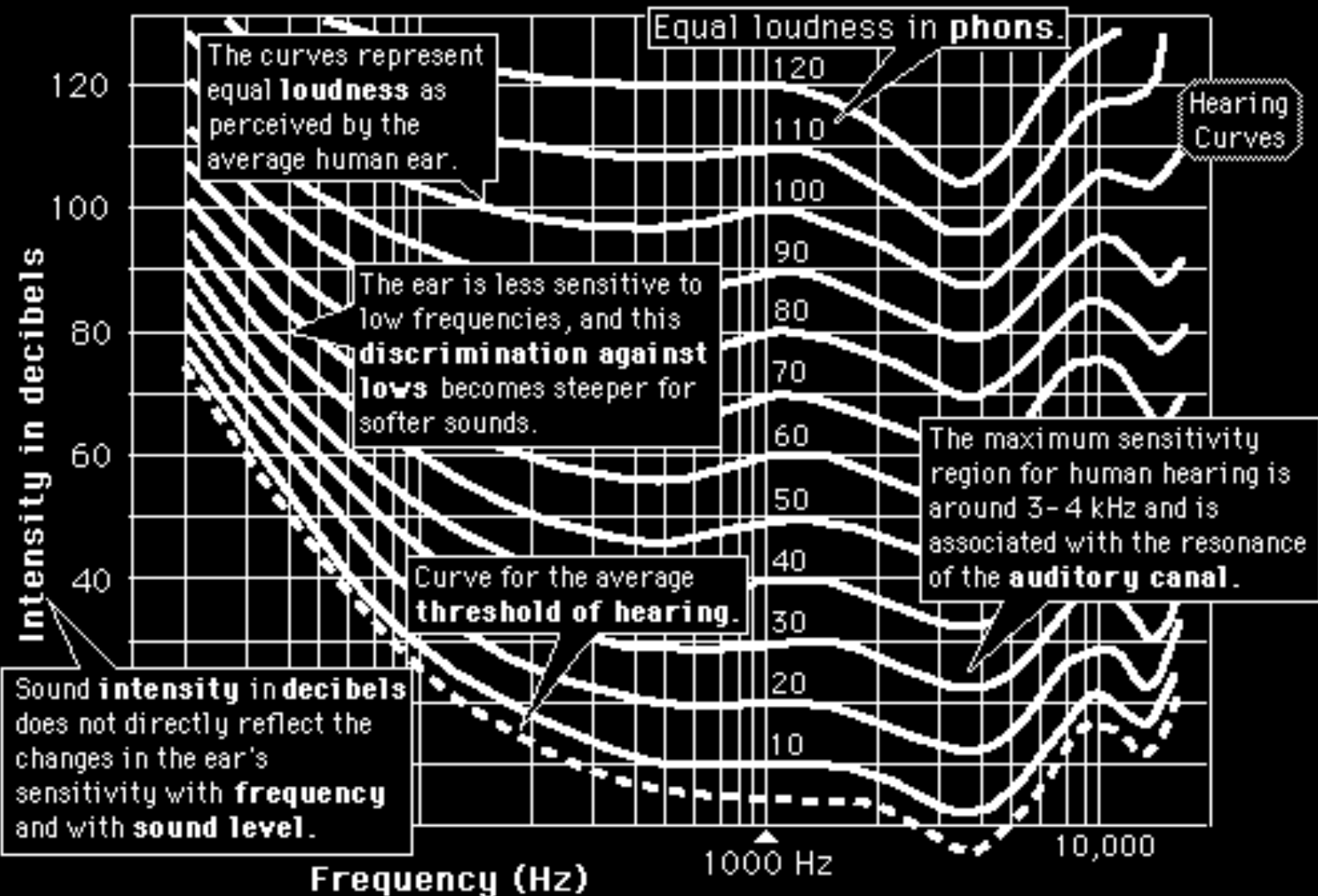
Phons

- The 10 phon curve is that which passes through 10db at 1000Hz.

- The 30 phon curve is that which passes through 30db at 1000Hz.

- Phons are the perceived loudness level equivalent to that at 1000Hz

How is the Fletcher-Munson diagram measured?

- It is a "perceived loudness" diagram?
- So how would you carry out an experiment to measure this?

Done with sine waves.

Psychoacoustic experiments:

1) Matching the volume of one sound to that of another with a fixed frequency and volume (reference sound)

2) Rating sounds on a numerical scale (magnitude estimation

Models of Loudness

- Loudness is somehow related to total neural activity ---
- Loudness is related to a summation of neural activity across different frequency channels
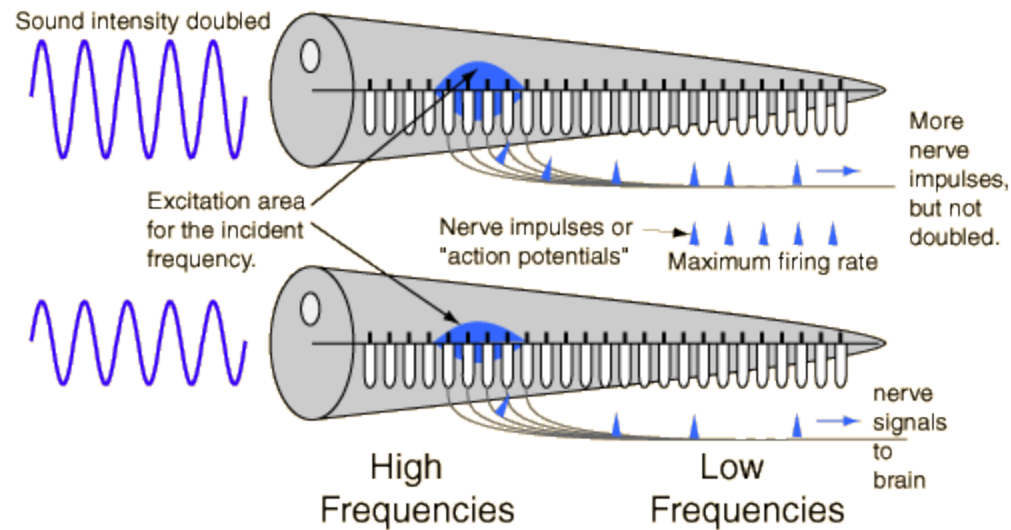
Perceived loudness scale-Sones

- A sound that we say is twice as loud is not twice as powerful.
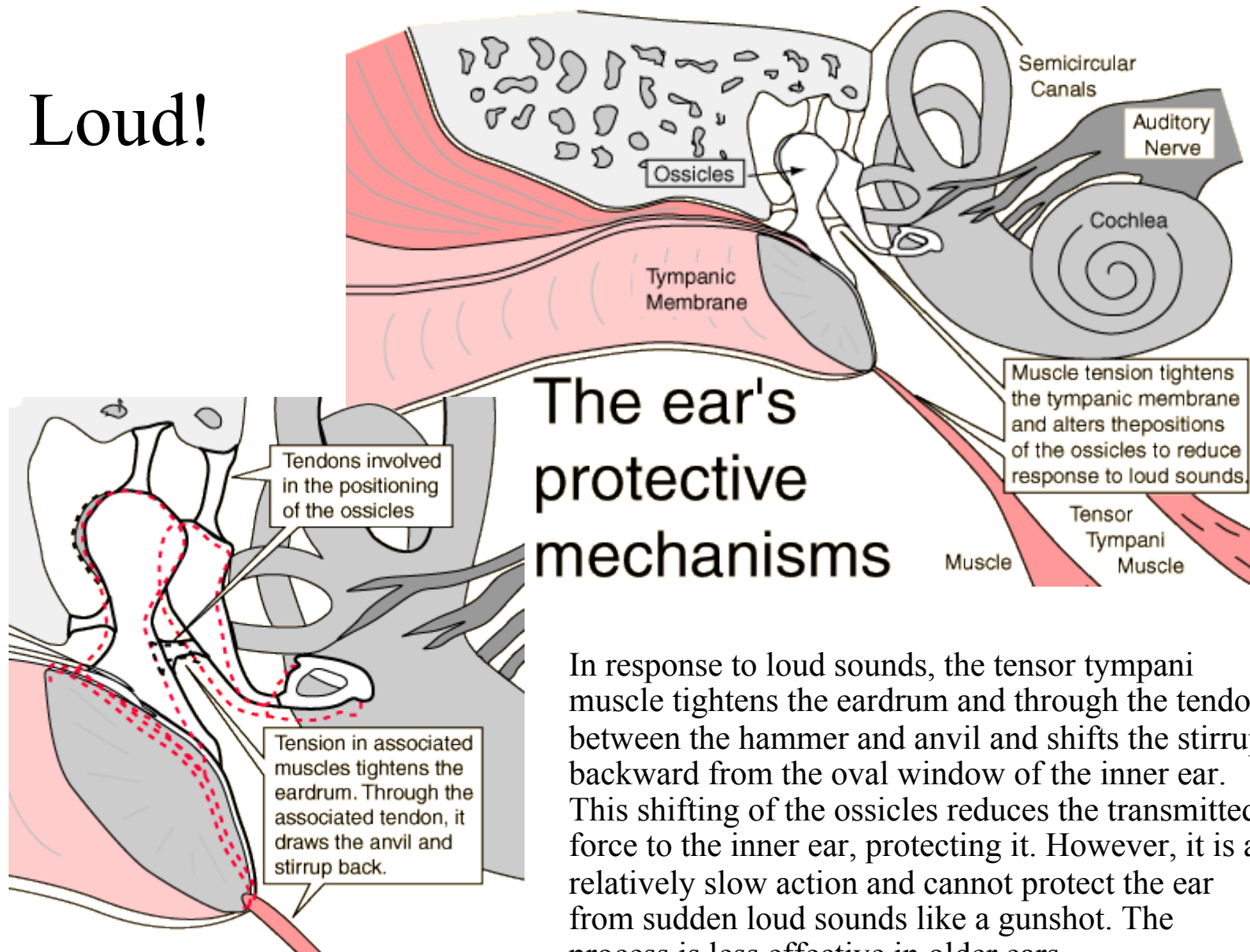- Rule of thumb: twice as loud (perceived) is equivalent to 10 times in power or 10 dB
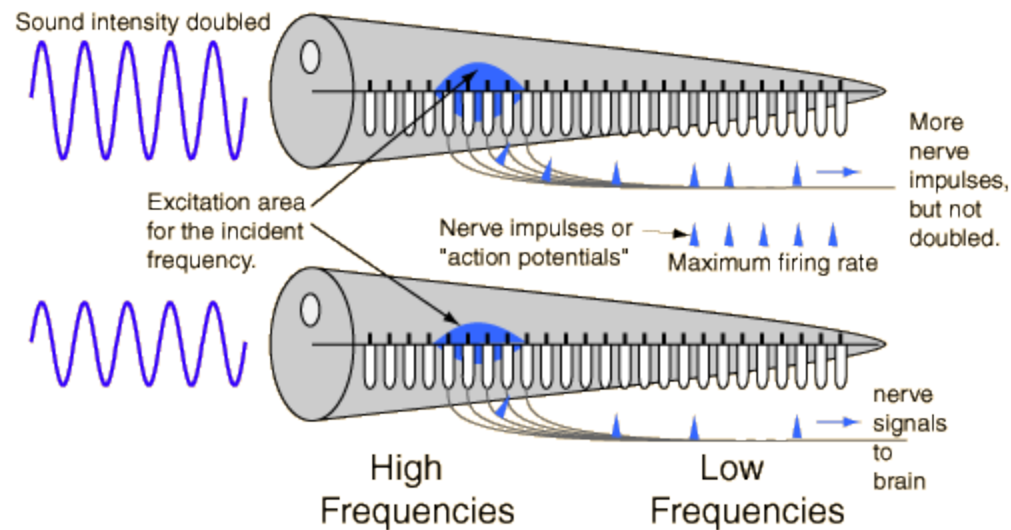
Violin sections in orchestra

- How many violins are needed to make a sound that is "twice as loud" $\sqrt{10}$ as a single violin?

  – 10 in power or 10dB

  – If all playing in phase $\sqrt{10} \sim 3$
     or about 3 violins.

  – But that is not likely so may need about 10 violins

  – Many violins are needed to balance sound and a few don't make much of a difference!

Dynamic range

- Threshold of pain is about 140 dB above the threshold of hearing.

- A difference of 10^{14} in power.

- How is this large range achieved?

Loud!

Semicircular Canals

Auditory Nerve

Ossicles

Cochlea

Tympanic Membrane

Muscle tension tightens the tympanic membrane and alters thepositions of the ossicles to reduce response to loud sounds.

Tensor Tympani Muscle

Muscle

The ear's protective mechanisms

Tendons involved in the positioning of the ossicles

Tension in associated muscles tightens the eardrum. Through the associated tendon, it draws the anvil and stirrup back.

In response to loud sounds, the tensor tympani muscle tightens the eardrum and through the tendon between the hammer and anvil and shifts the stirrup backward from the oval window of the inner ear. This shifting of the ossicles reduces the transmitted force to the inner ear, protecting it. However, it is a relatively slow action and cannot protect the ear from sudden loud sounds like a gunshot. The process is less effective in older ears.
http://hyperphysics.phy-astr.gsu.edu

The outer ear (pinna) collects more sound energy than the ear canal would receive without it and thus contributes some area amplification.
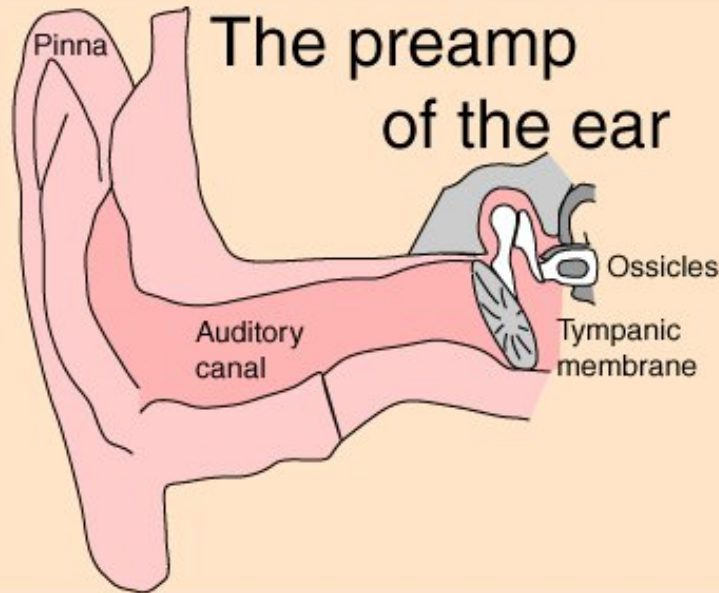
The numbers here are just representative ... not precise data.

The preamp of the ear

Pinna

Auditory canal

Ossicles

Tympanic membrane

Closed tube resonance of the auditory canal enhances 2000-5000 Hz	Tympanic membrane (eardrum) has some 15x area of oval window contributing an area amplification.	Ossicles (hammer, anvil and stirrup) contribute a lever-type amplification when listening to soft sounds.
Outer ear 2x	Tympanic membrane 15x	Ossicles 3x

The outer and middle ears contribute something like a factor of 100 or about 20 decibels of amplification under optimum conditions.

Outer and middle ear contribute to the sensitivity and dynamic range of the ear.

The Ossicles

The three tiniest bones in the body form the coupling between the vibration of the eardrum and the forces exerted on the oval window of the inner ear.

With a long enough lever, you can lift a big rock with a small applied force on the other end of the lever. The amplification of force can be changed by shifting the pivot point.
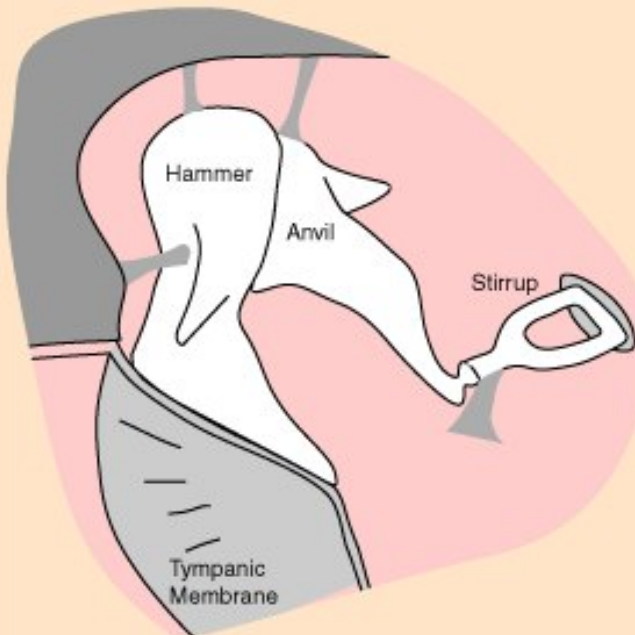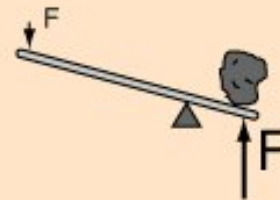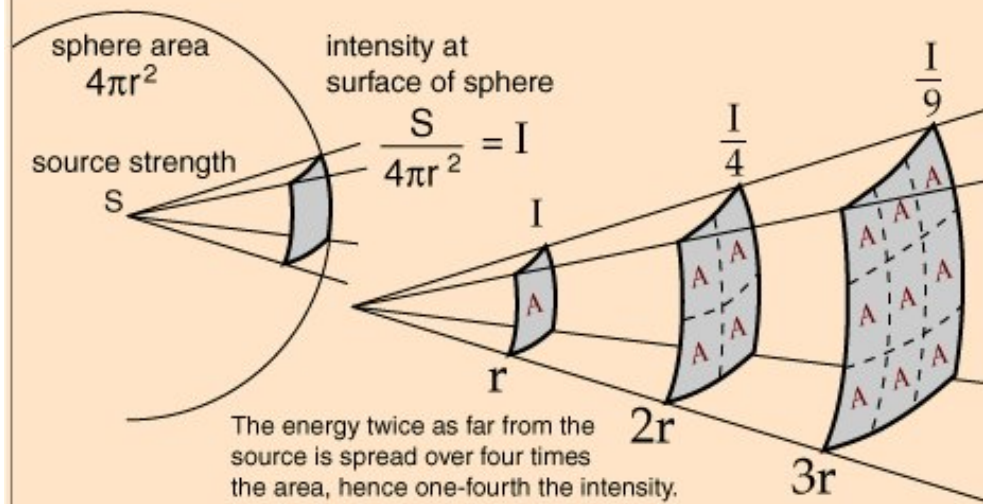
The ossicles can be thought of as a compound lever which achieves a multiplication of force. This lever action is thought to achieve an amplification by a factor of about three under optimum conditions, but can be adjusted by muscle action to actually attenuate the sound signal for protection against loud sounds.

Hammer

Anvil

Stirrup

Tympanic Membrane

A small motion on the eardrum causes a larger motion on the oval window connecting to the inner ear.

Inverse Square Law, General

Any point source which spreads its influence equally in all directions without a limit to its range will obey the inverse square law. This comes from strictly geometrical considerations. The intensity of the influence at any given radius r is the source strength divided by the area of the sphere. Being strictly geometric in its origin, the inverse square law applies to diverse phenomena. Point sources of gravitational force, electric field, light, sound or radiation obey the inverse square law. It is a subject of continuing debate with a source such as a skunk on top of a flag pole; will it's smell drop off according to the inverse square law?

sphere area $4\pi r^2$

intensity at surface of sphere

$$\frac{S}{4\pi r^2} = I$$

source strength S

I

$\frac{I}{4}$

$\frac{I}{9}$

The energy twice as far from the source is spread over four times the area, hence one-fourth the intensity.

r

2r

3r

Sound in space

Power goes as distance^{-2}

How would you expect the dB level to depend on distance?

Inverse square law in dB

$P \propto d^{-2}$ P = power

d = distance

$SIL = 10\log_{10}(P) + \text{constant}$

$SIL = 10\log_{10}(d^{-2}) + \text{constant}$

$= -20\log_{10}(d) + \text{constant}$

A factor of 10 in distances leads

to a change of what in dB?
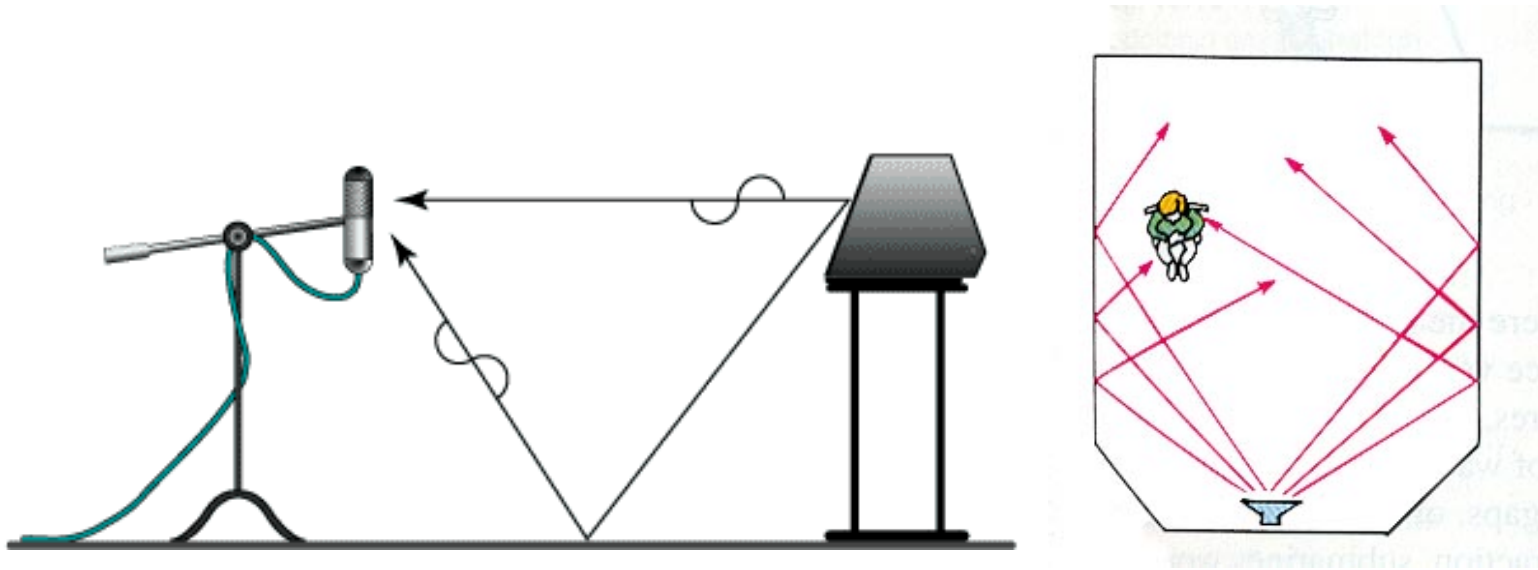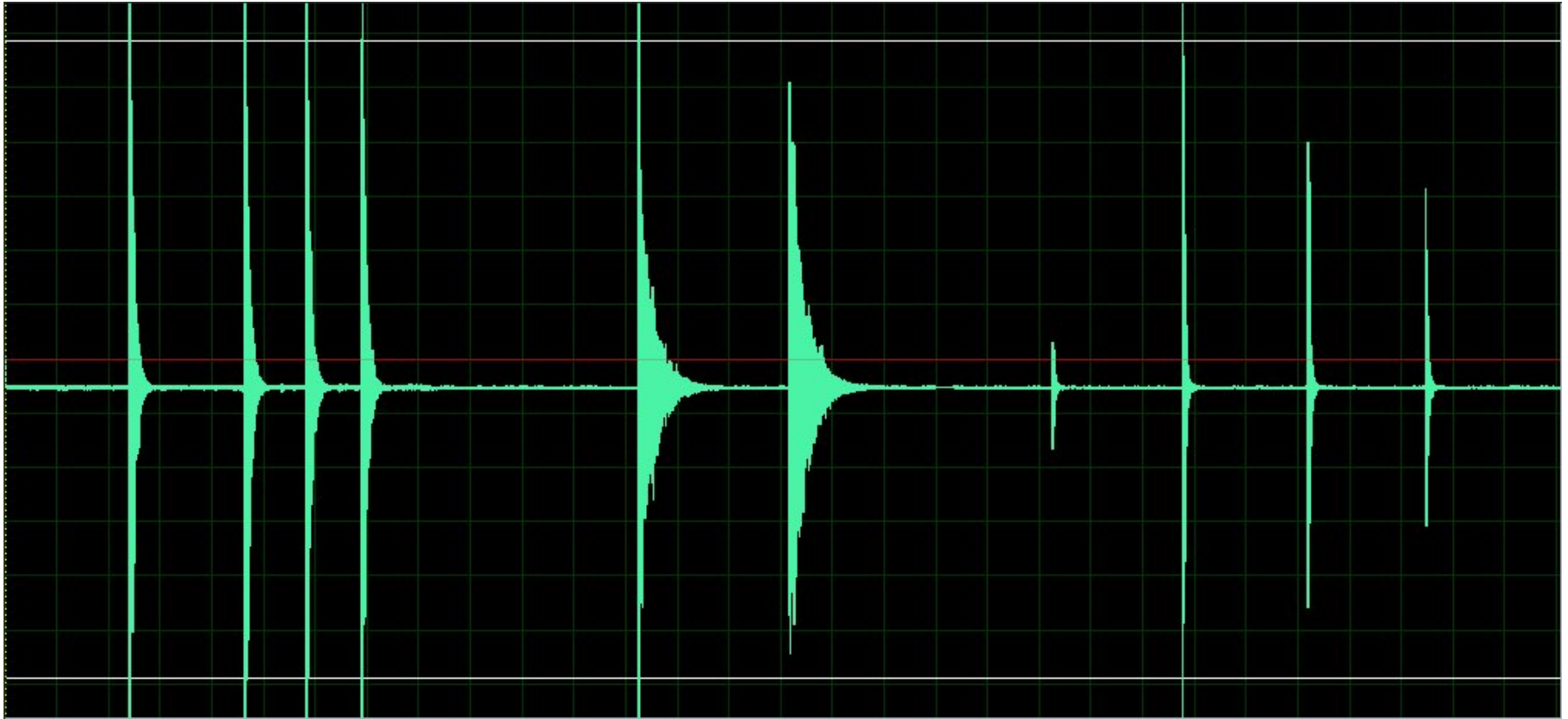
Is the inverse square law relevant for room acoustics?

Figure from JBL Sound System Design Reference Manual

What is relevant to room acoustics?

- Locations and angles of reflections
- Timing of reflections
- Quality of reflections:
  - as a function of frequency
  - what amount absorbed
- Number of reflections
- Modes in the room that are amplified

🔊 3 rooms

time →

Timing of echos

- Speed of sound is 340m/s.
- 1/340 ~ 0.003 ~ 3ms
- 3 milliseconds per meter.
- Echo across a room of 10m is 30ms.
- Decay rates related to the travel time across a room.

Effect of Echoes

- ASAdemo 35 Speech in 3 rooms and played backwards so echoes are more clearly heard

- Echo suppression

How do we describe a decay?

- A timescale.
- How about a halftime? Like a half life?
  After $t_{1/2}$ the sound is ½ the power
- We could use a function to describe the power
  We can write this $P(t) = P_0 2^{-t/t\_1/2}$
- Or we could use
  $P(t) = P_0 \exp(-t/t_{decay})$ exponential drop
- Or we could use something like this
- $P(t) = P_0 10^{-t/td}$.
- Actually if you adjust td or t_{decay} these three expressions can be the same function.
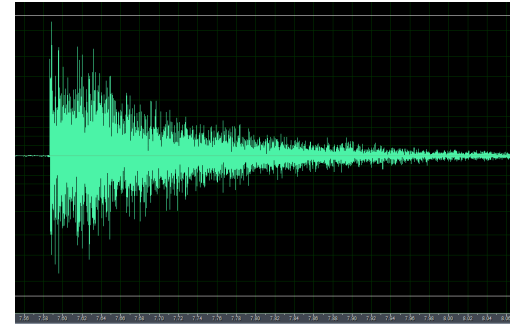
An exponential decay

becomes a line
on a log plot→

drops to half

drops to half again

drops to half again
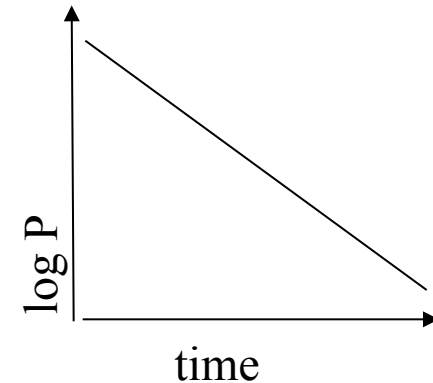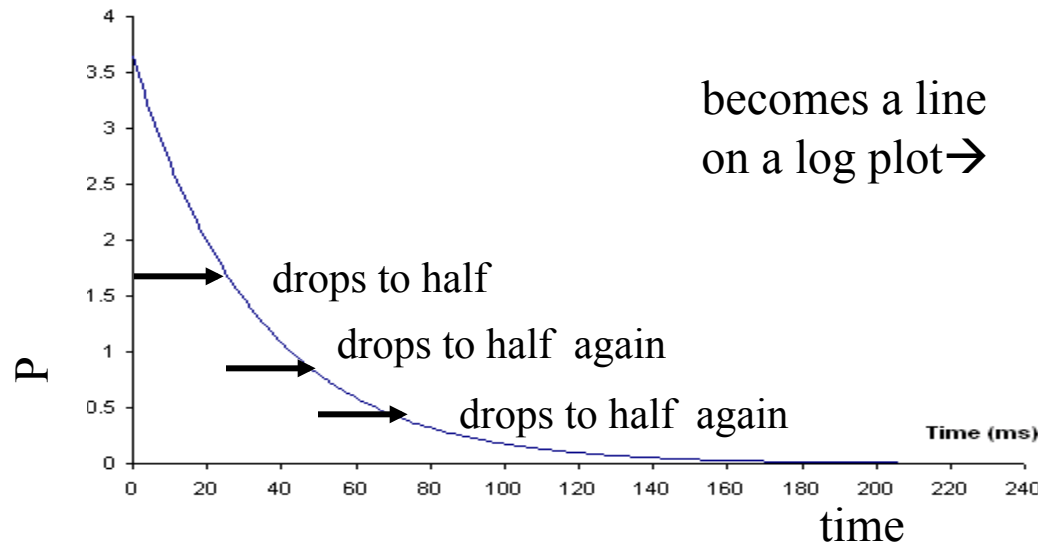
- **The height decreases very quickly at first, and then decreases more slowly**.
- A curve with this exact shape is called an **exponential curve**. So, since the current is decreasing (or decaying) along a curve of this shape, we call this **exponential decay**.

Measuring a half life

- Suppose we measure power in dB. How many dB change corresponds to twice the power?

- $10\log_{10}2 \sim 3.0\text{dB}$

- Measure the time it takes to drop in power by 3dB and that corresponds to the $t_{1/2}$.

Room acoustics

- It is now recognized that the most important property of a room is its reverberation time.
- This is the timescale setting the decay time, or the half time of acoustic power in the room
- Surfaces absorb sound so the echoes get weaker and weaker.
- The reverberation time depends on the size of the room and the way the surfaces absorb sound.
- Larger rooms have longer reverberation times.
- More absorptive rooms have shorter reverberation times.

Reverberation time

- The reverberation time, RT_{60}, is the time to drop 60 dB below the original level of the sound.

- The reverberation time can be measured using a sharp loud impulsive sound such as a gunshot, balloon popping or a clap.

- Why use 60dB to measure the reverberation time? the loudest crescendo for most orchestral music is about 100 dB and a typical room background level for a good music-making area is about 40 dB.

- 60dB corresponds to a change in power of a million!

- It is in practice hard to measure sound volume over this range. However you can estimate the timescale to drop by 20 dB and multiply by ......?

What is a good reverberation time for a room?

- If you are using the room for lectures (speech) then a long reverberation time makes it difficult for the audience to understand words.

- However long reverberation times are desirable for example in churches for organ music.

- Rooms that are good for both speech and music typically have reverberation times between 1.5 and 2 seconds.

- If the direct sound is week compared to the echo, then speech sounds garbled.

Estimating the reverberation time

Sabine's formula

$$RT_{60} = 0.16 \text{ s m}^{-1} \frac{V}{S_e}$$ V/S_e in m, RT_{60} in seconds

$$RT_{60} = 0.049 \text{ s ft}^{-1} \frac{V}{S_e}$$ V/S_e in feet, RT_{60} in seconds

V is the volume of the room and S_e the effective area

$$S_e = a_1 S_1 + a_2 S_2 + a_3 S_3 + \ldots$$

S_i surface area for a type of surface

a_i absorption coefficient for this surface

People or seats also can be given effective areas.

Sabine's formula

$$RT_{60} = 0.16 \text{ s m}^{-1} \frac{V}{S_e}$$

- A decay timescale proportional to a Volume/Area is a length
- Time between reflections depends on length
- Decay time depends on length
- Bigger rooms have longer decay times

Sabine's formula

$$RT_{60} = 0.16 \text{ s m}^{-1} \frac{V}{S_e}$$

$$S_e = a_1 S_1 + a_2 S_2 + a_3 S_3 + \ldots$$

S_i surface area for a type of surface

a_i absorption coefficient for this surface

- If each more energy is absorbed each reflection then the decay timescale is shorter
- Higher absorption on walls means shorter decay timescale

Standing waves in a Room

- x,y,z wave-numbers
- loud in corners
- problem at low frequencies/large wavelengths
- corner foams to damp low frequencies
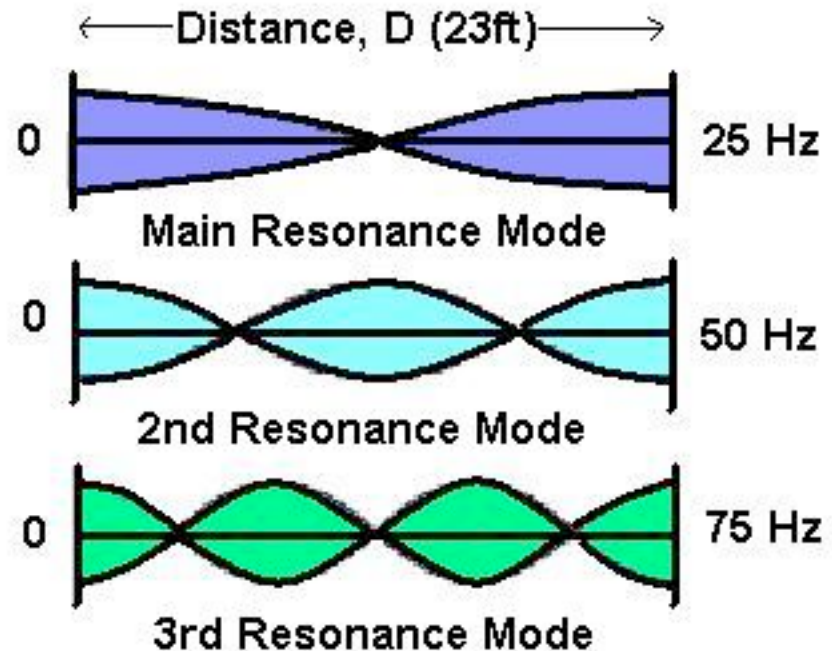
Just Noticeable Differences

- **JND in Sound Intensity**

- A useful general reference is that the just noticeable difference in sound intensity for the human ear is about 1 decibel.

- JND = 1 decibel

- In fact, the use of the factor of 10 in the definition of the decibel is to create a unit which is about the least detectable change in sound intensity.

JND as a function of loudness
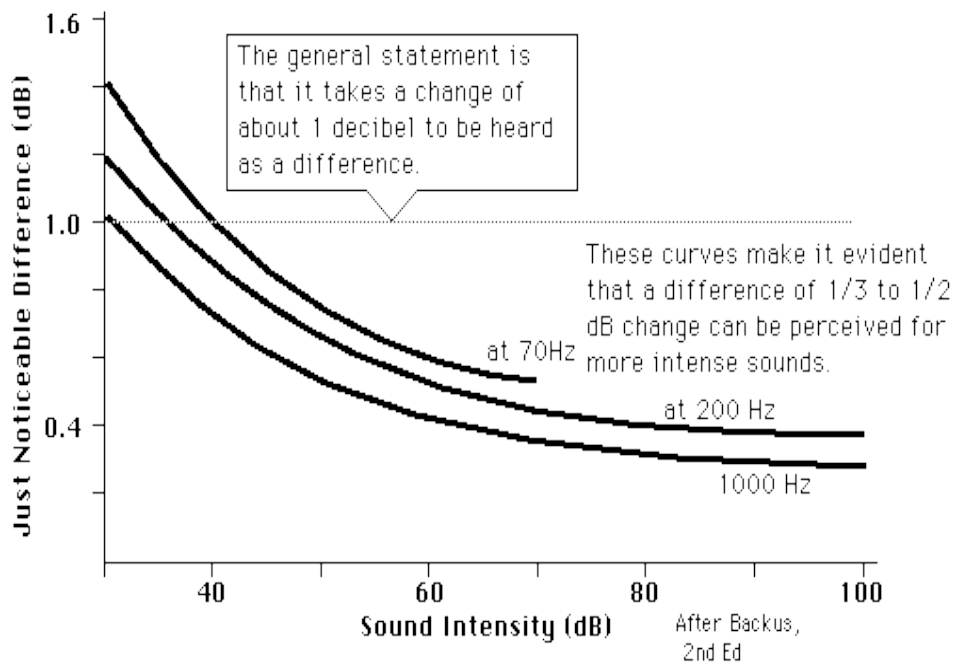
There are some variations. The JND is about 1 dB for soft sounds around 30-40 dB at low and midrange frequencies. It may drop to 1/3 to 1/2 a decibel for loud sounds.

Acoustic Gain and Feedback

Gain: increase in dB at the position of the listener. What is the maximum gain without feedback?
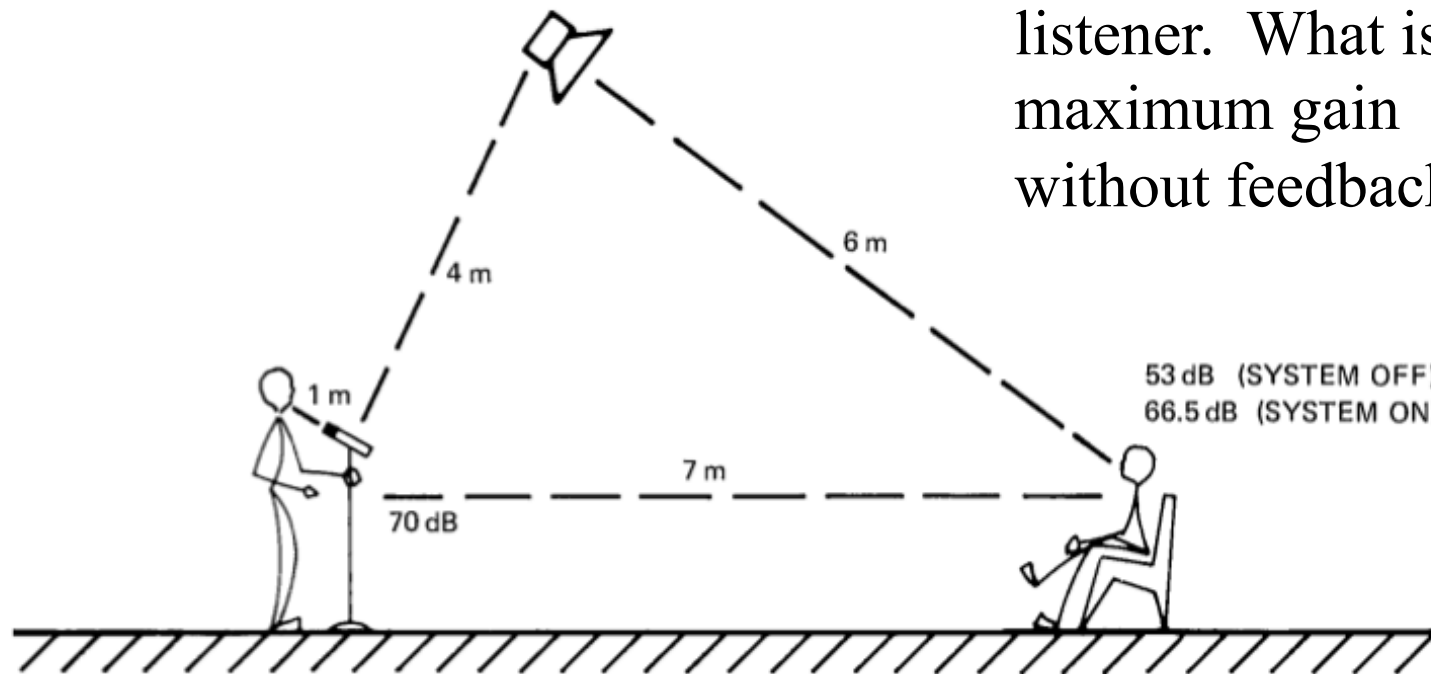
6 m

4 m

1 m

53 dB (SYSTEM OFF)
66.5 dB (SYSTEM ON)

7 m

70 dB

Figure 4-3. System gain calculations, loudspeaker and microphone both omnidirectional

Figure from JBL Sound System Design Reference Manual
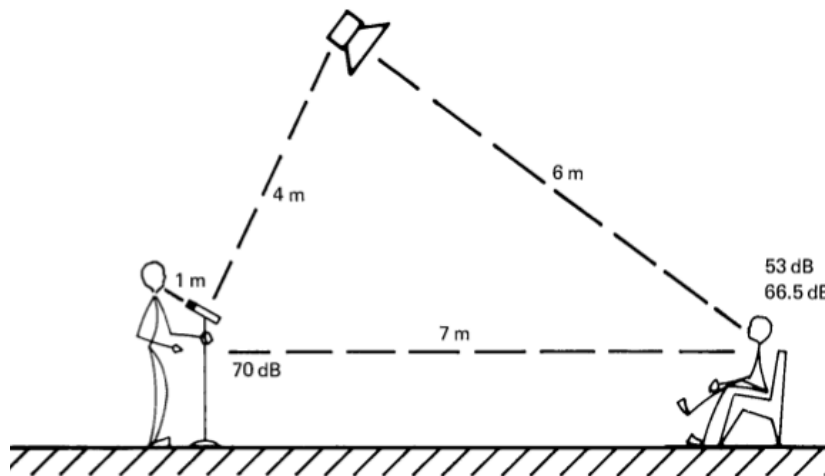
Acoustic Gain and Feedback

Can turn up the speaker until the signal at the microphone from the speaker is as loud as the voice
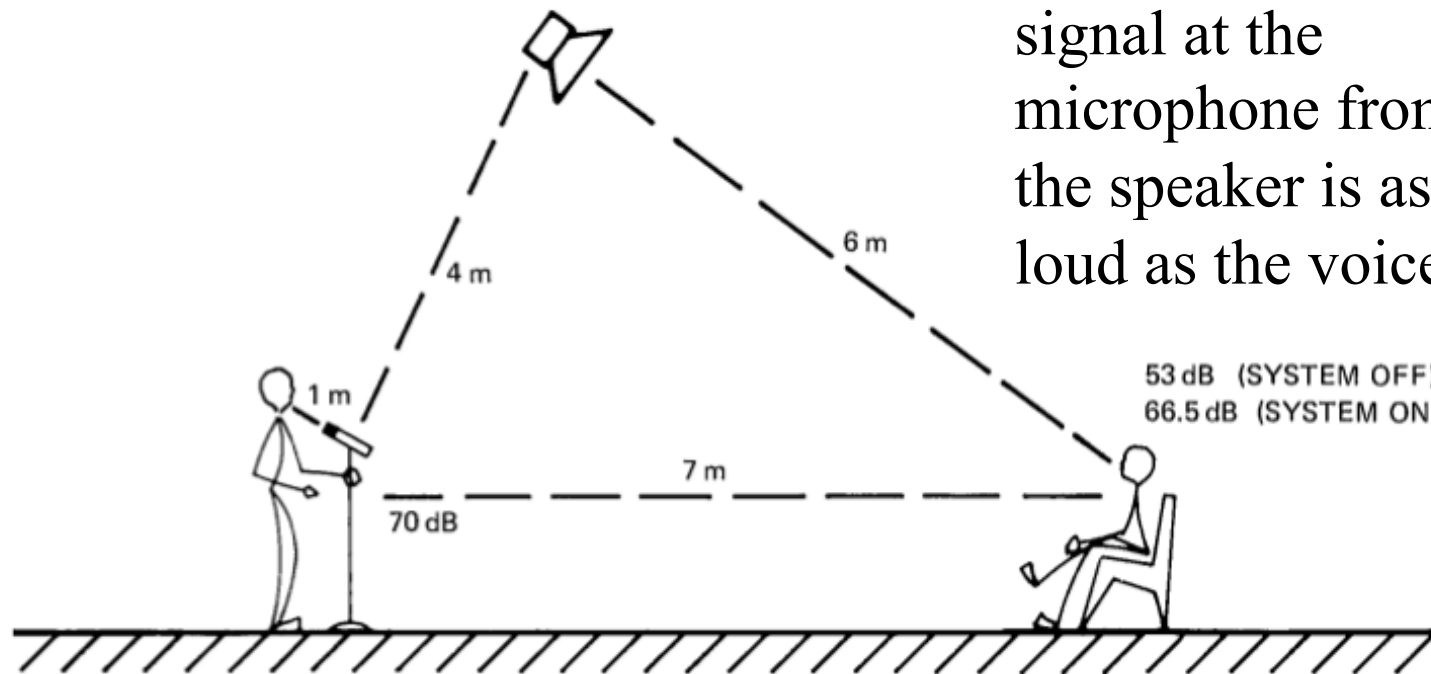
6 m

4 m

1 m

53 dB (SYSTEM OFF)
66.5 dB (SYSTEM ON)

7 m

70 dB

Figure 4-3. System gain calculations, loudspeaker and microphone both omnidirectional

Figure from JBL Sound System Design Reference Manual

Acoustic Gain and Feedback

Figure 4-3. System gain calculations, loudspeaker and microphone both omnidirectional

Figure from JBL Sound System Design Reference Manual

Difference in dB for direct voice between speaker and mic locations
$20 \log_{10} 4 =$ 12.0 dB
Here 4/1 is ratio of distance between speaker and mic and mic and person.
You can amplify the voice by up to x4 the amplitude (with no feedback safety margin).
Listener has
$20 \log_{10} 7 =$ 16.9 dB for direct voice
$20 \log_{10} 6 =$ 15.6 dB for speaker
Volume 12.0-15.6=-3.6 speaker
Volume -16.9 dB direct
Difference is 16.9-3.6=**13.3** dB.
Maximum gain is **13.3** dB.

Acoustic Gain and Feedback

Max gain $20 \log_{10} \left[\dfrac{D_{ml} D_{ms}}{D_{sl} D_{vm}} \right]$

D_{sl}

6 m

D_{ms} 4 m

D_{vm}

1 m

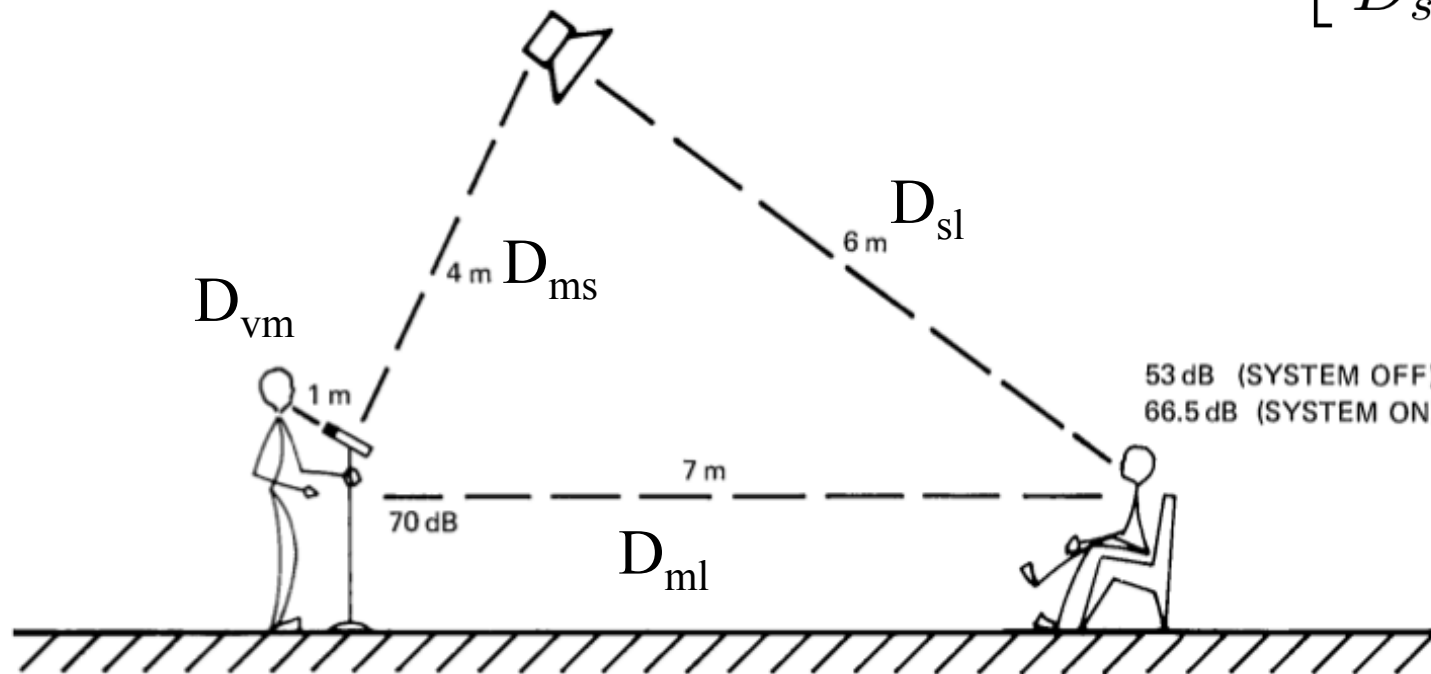53 dB (SYSTEM OFF)
66.5 dB (SYSTEM ON)

7 m

70 dB

D_{ml}

Figure 4-3. System gain calculations, loudspeaker and microphone both omnidirectional

Critical Band

Two sounds of equal loudness (alone) but close together in pitch sounds only slightly louder than one of them alone.

They are in the same critical band competing for the same nerve endings on the basilar membrane of the inner ear.

According to the place theory of pitch perception, sounds of a given frequency will excite the nerve cells of the organ of Corti only at a specific place. The available receptors show saturation effects which lead to the general rule of thumb for loudness by limiting the increase in neural response.

Outside the critical band

- If the two sounds are widely separated in pitch, the perceived loudness of the combined tones will be considerably greater because they do not overlap on the basilar membrane and compete for the same hair cells.

Critical Band continued

Critical Band Measurement

For low frequencies the critical band is about 90 Hz wide. For higher frequencies, it is between a whole tone and 1/3 octave wide.
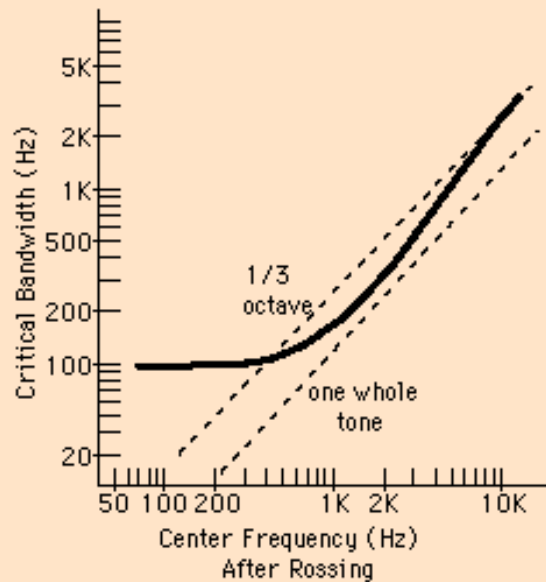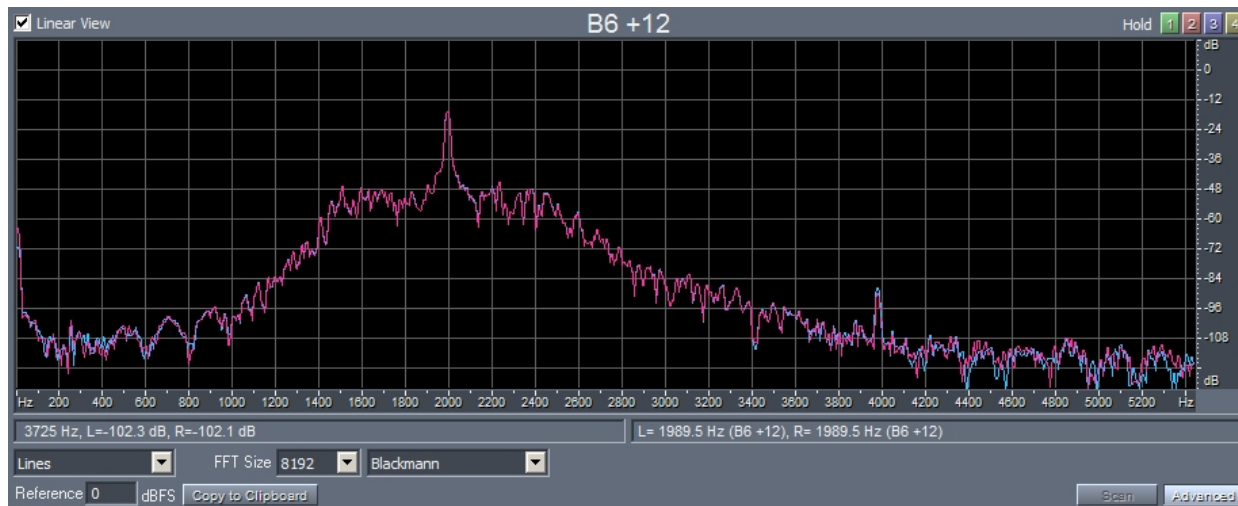
Center Freq (Hz)	Critical bandwidth (Hz)
100	90
200	90
500	110
1000	150
2000	280
5000	700
10000	1200

Rossing 2nd Ed p74

Critical Bands by Masking

ASAdemo2   Tone in presence of broad band noise

- The critical band width at 2000Hz is about 280Hz so you can hear more steps when the noise bandwidth is reduced below this width.

Critical Bands by Loudness comparison

- A noise band of 1000Hz center frequency.
- The total power is kept constant but the width of the band increased.
- When the band is wider than the critical band the noise sounds louder.
- ASA demo 3

Frequency Selectivity and Masking

critical band
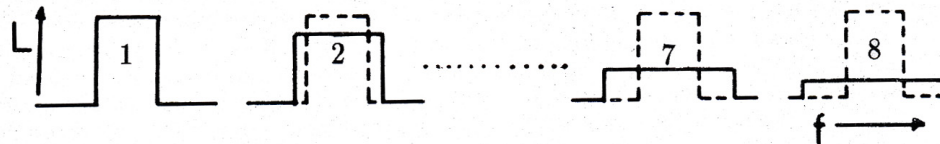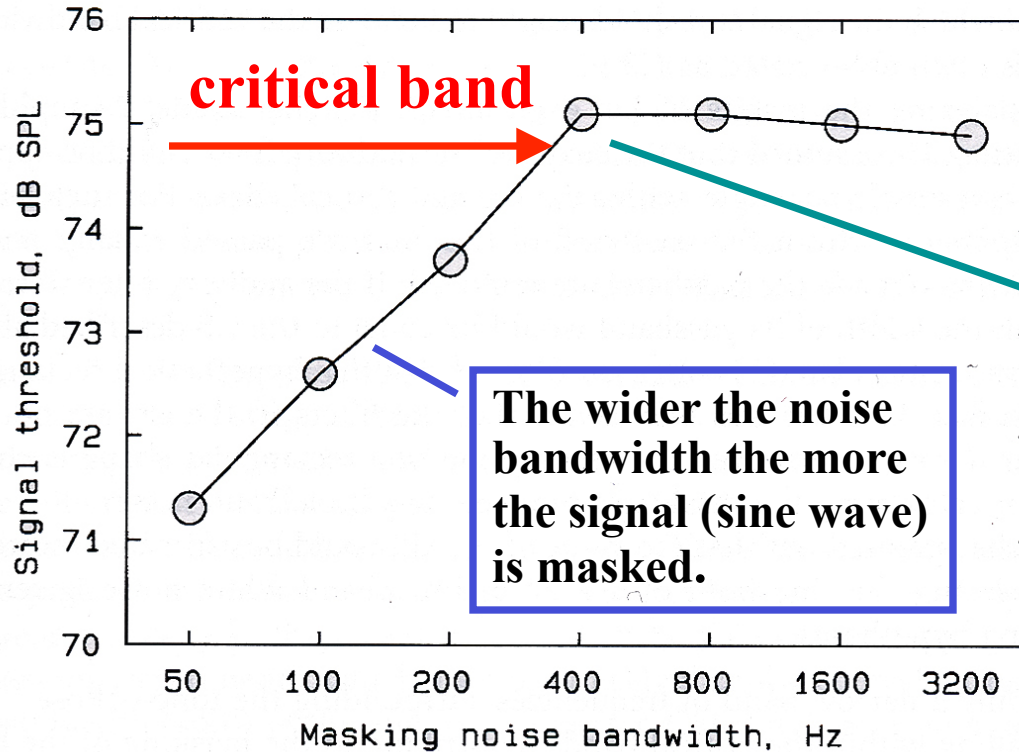
67 **A sine (signal) in the presence of noise that has a band width (in frequency) centered around the signal.**

The wider the noise bandwidth the more the signal (sine wave) is masked.

Past a particular frequency width the masking doesn't increase.

FIGURE 3.1 The threshold of a 2000-Hz sinusoidal signal plotted as a function of the bandwidth of a noise masker centered at 2000 Hz. Notice that the threshold of the signal at first increases with increasing masker bandwidth and then remains constant. From Schooneveldt and Moore (1989), by permission of the authors.

critical band

Masking

If a dominant tone is present then noise can be added at frequencies next to it and this noise will not be heard. Less precision is required to store nearby frequencies.

Energy

Dominant Tone

13dB

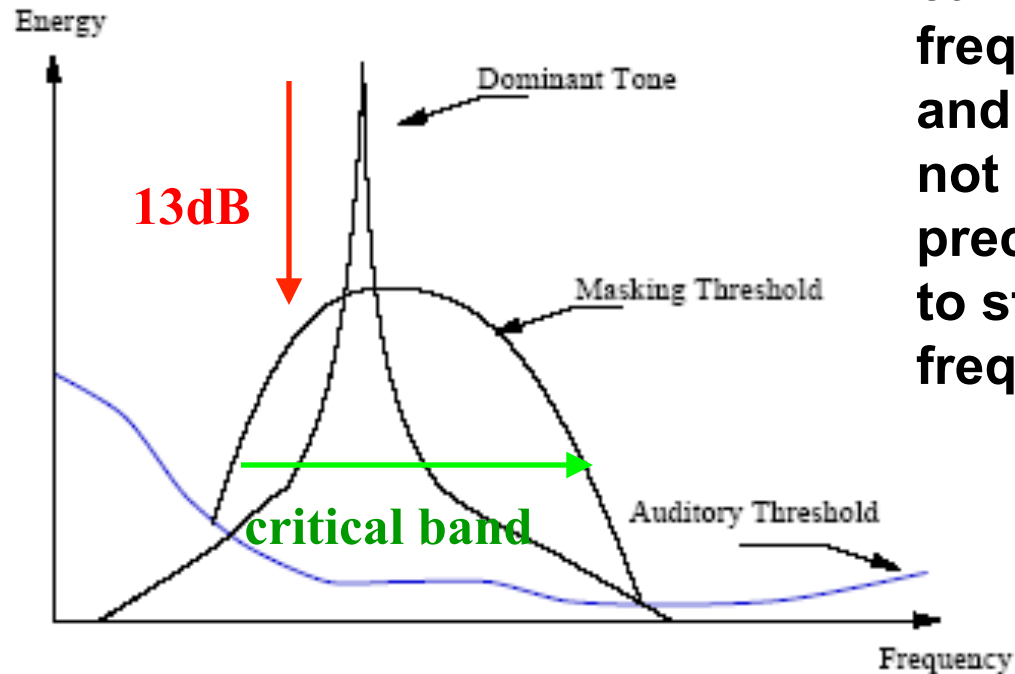Masking Threshold

critical band

Auditory Threshold

Frequency

Illustration of a dominant tone masking out the signal energies in the vicinity with an energy below the masking threshold. The Masking threshold curve shown is normalized to an arbitrary hearing threshold.

13 dB miracle

- If the signal is 13 dB louder than the noise then the noise can't be heard (within a band).

- Each sub-band is quantized differently depending upon the masking threshold estimated in that band

Critical band width as a function of frequency
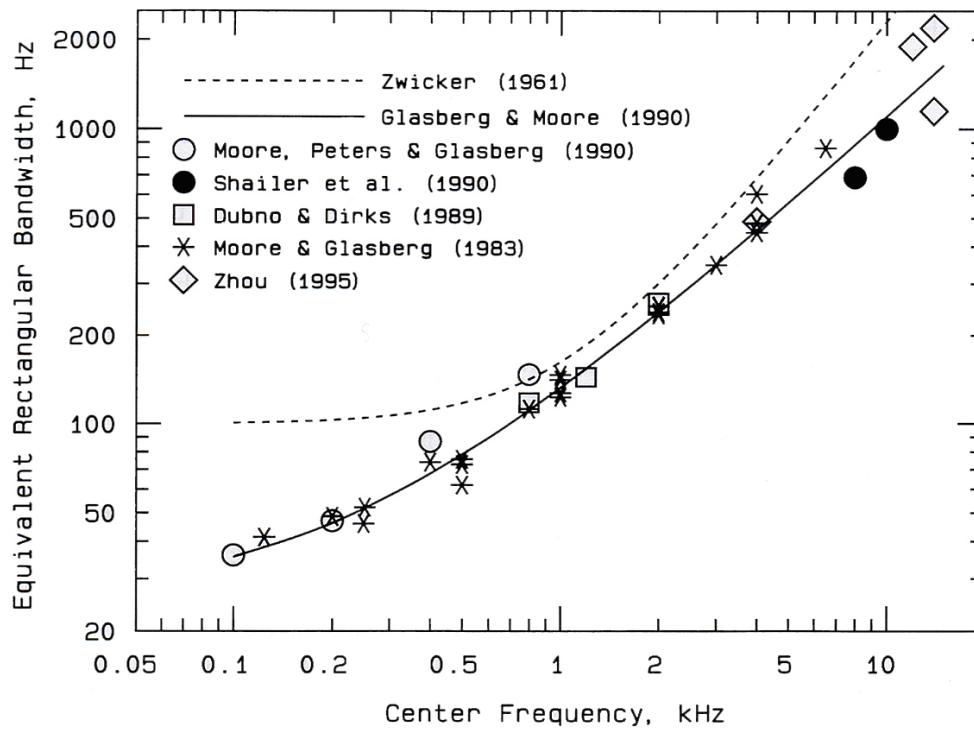
Size of critical band is typically one tenth of the frequency

FIGURE 3.5   The dashed curve shows the "old" value of the critical bandwidth as a function of frequency (Zwicker, 1961). The solid curve shows the value of ERB_N of the auditory filter as a function of frequency. The solid curve was obtained by combining the results of several experiments using Patterson's notched-noise method of estimating the auditory filter shape.

Critical band concept

- Only a narrow band of frequencies surrounding the tone – those within the critical band contribute to masking of the tone

- When the noise just masks the tone, the power of the tone divided by the power of the noise inside the band is a constant.

The nature of the auditory filter

- The auditory filter is not necessarily square – actually it is more like a triangle shape
- Critical band width is sometimes referred to as ERB (equivalent rectangular bandwidth)
- Shape difficult to measure in psychoacoustic experiments because of side band listening affects -some innovative experiments (notched filtered noise + signal) designed to measure the actual shape of the filter).

Physiological reasons for the masking

- Basal membrane? The critical bandwidths at different frequencies correspond to fixed distances along the basal membrane.

- However the masking could be a result of feedback in the neuron firing instead. Negative reinforcement or suppression of signals. Or swamping of signals.

Temporal integration

- ASA demo8
- Bursts are perceived to be **louder** if they are **longer** –up to a particular threshold time
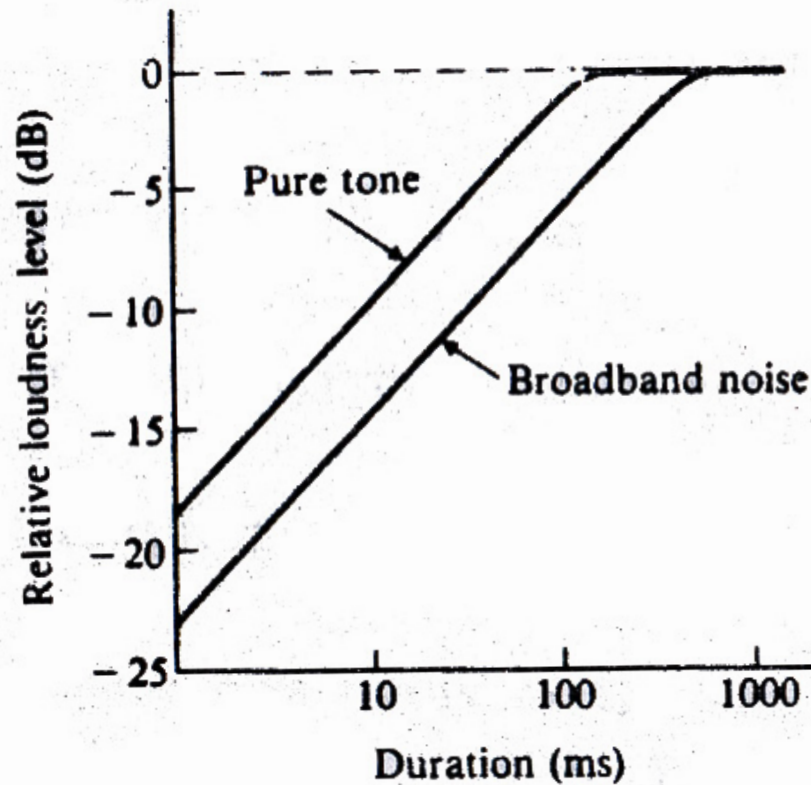
Temporal effects - non-simultaneous masking

- The peak ratio of the masker is important -- that means its variations in volume as a function of time compared to its rms value. Short loud peaks don't necessarily contribute to the masking as much as a continuous noise.

- Both forward and backward masking - masking can occur if a loud masker is played just after the signal!

- Masking decays to 0 after 100-200ms

Physiological explanations for temporal masking

- Basal membrane is ringing preventing detection in that region for a particular time
- Neurons take a while to recover - neural fatigue

Outside the critical band

- If the two sounds are widely separated in pitch, the perceived loudness of the combined tones will be considerably greater because they do not overlap on the basilar membrane and compete for the same hair cells.
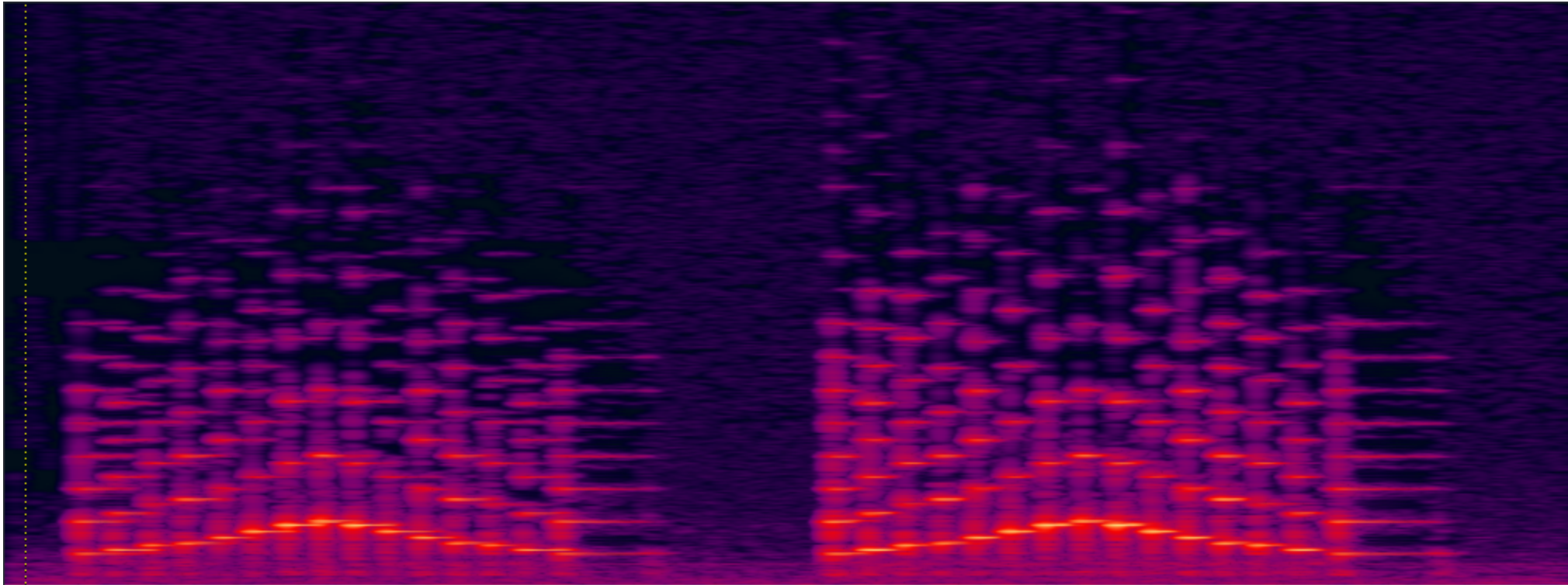
Pushing MP3 to its limits
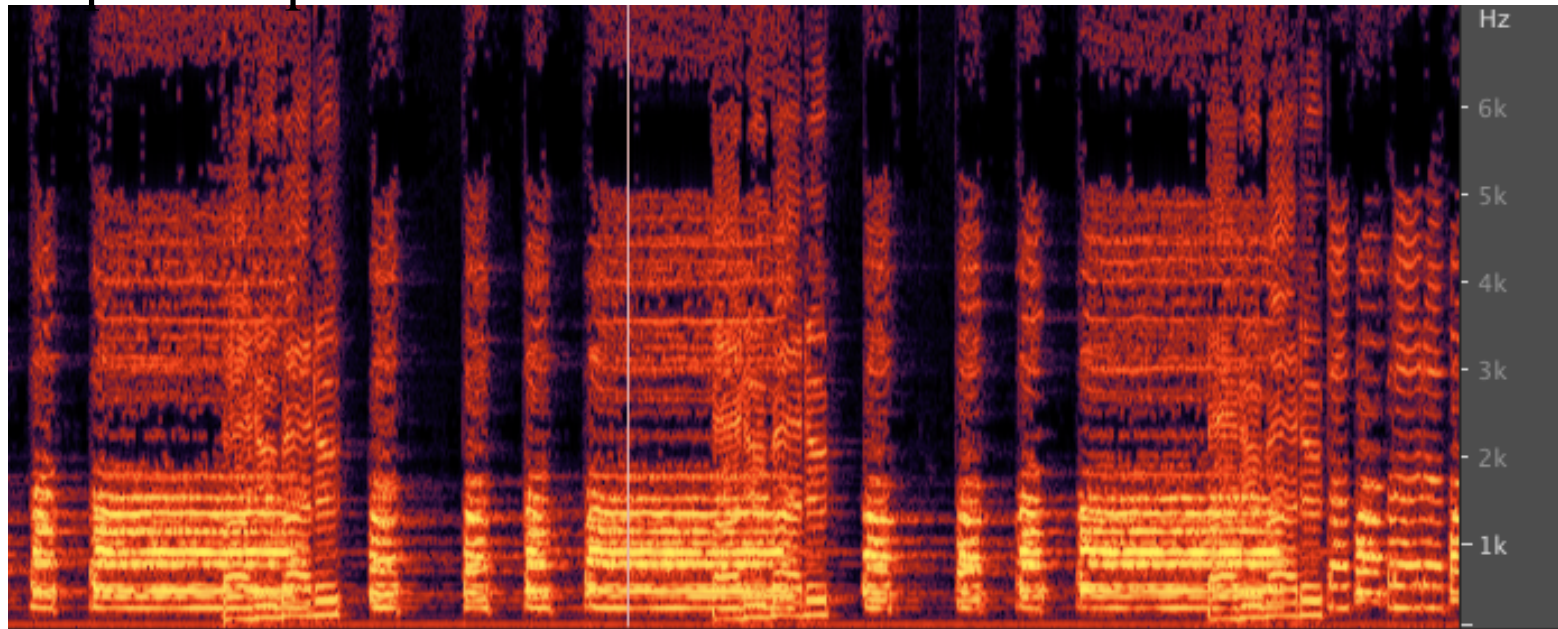
🔊 -uncompressed
🔊 -over compressed mp3

- Above compressing to 60kbps
- Using home.c4.scale.AIFF show mp3 **options** DEMO
  with Audition to experiment

mp3 compressions   .wav 11.9MBytes

- 192 Kbps (44100Hz)   414KBytes
- 56 Kbps(44100Hz)   246KBytes
- 32 Kbps(22050Hz)   143KBytes
- 8Kbps(8000Hz)   45KBytes
  bps=bits per second

Recommended Reading

- Moore Chap 4 on the Perception of Loudness

- Berg and Stork Chap 6 pages 144-156 on the dB scale and on the Human Auditory System or Hall chapters 5+6

- On Room acoustics: Berg and Stork Chap 8 or Hall Chap 15.