PHYS 1406: Physics of Sound & Music (Additional Lecture Notes)

J.D. Romano

Spring 2020

Abstract

Disclaimer: These notes are meant to supplement the textbook "Physics of Sound & Music" by Prof. Borst. Please send corrections, comments, criticisms, suggestions to: `joseph.d.romano@ttu.edu`.

Contents

Part I
Introduction

1 Preliminaries

1.1 Questions we'd like to answer

- What distinguishes music from noise?
- Why does a clarinet sound different from a flute?
- Why do 10 violins sound only twice as loud as a single violin?
- Why do you sound better when you sing in the shower?
- Why do some notes harmonize while others clash?
- What's equal temperament and why do we use it?
- Others??

1.2 Topics we will cover

- Preliminaries: basic math, physics, and music terminology
- Physics of oscillations and waves
- Production of sound (instruments, voice)
- Perception of sound (hearing, loudness, pitch)
- Room acoustics; reproduction and synthesis of sound
- Musical scales and tuning systems

1.3 What is sound? What is music?

- Sound consists of pressure waves in air or in some other medium.
- Sounds are characterized by their loudness, duration, and pitch (if they have one).
- Sounds can be measured and represented in different ways:

  (i) A *sound-level meter* measures the loudness or intensity of a sound.

  (ii) A *microphone attached to an oscilloscope* measures the pressure wave as it changes in time (the waveform or shape of the wave).

  (iii) A *spectrum analyzer* measures the different frequency components of the sound (high pitch, low pitch, or some combination).

  (iv) A *spectrogram* shows how the frequency components of a sound change over time.

- <u>Demonstration</u>: Illustrate these different representations of sound using the Faber Electroacoustics Toolbox.
- <u>Demonstration</u>: Illustrate the difference between a musical sound and other sounds (e.g., noise) by looking at the various representations of somebody speaking, singing, whistling; play various instruments; have students applaud; crumple a piece of paper, etc.
- A musical note typically has a characteristic pitch (fundamental frequency). Hence, its associated pressure wave repeats over time. Noise, on the other hand, doesn't have a characteristic pitch.
- Not all sounds are within our range of hearing (the pitch might be too low or too high, or the sound might not be loud enough to hear).

- Demonstration: Produce pure tones having different frequencies. Determine the range of frequencies that the class can hear. (The nominal range of human hearing is between 20 Hz and 20,000 Hz.)
- Ultrasound: sound waves whose frequencies lie *above* the upper end of the audible range for humans.

  Example: Dog whistle, echo-localization by bats, non-ionizing medical imaging, sonograms, $\cdots$
- Infrasound: sound waves whose frequencies lie *below* the lower end of the audible range for humans.

  Example: Seismic waves, earthquakes, $\cdots$

1.4 Basic math review

- Basic operations: Addition, subtraction, multiplication, division
- Entering numbers on a calculator: To evaluate $\frac{1}{2\pi}$, enter $1/(2\cdot\pi)$ and not $1/2\cdot\pi$. The correct answer is 0.1592, not 1.5708.
- Dividing fractions: Multiply by reciprocal—e.g.,

$$\frac{2}{3/2} = 2\cdot\frac{2}{3} = \frac{4}{3} \tag{1.1}$$

- Powers (exponential notation):

$$2^4 = 2\cdot 2\cdot 2\cdot 2 = 16 \tag{1.2}$$

$$10^3 = 10\cdot 10\cdot 10 = 1000 \tag{1.3}$$

$$10^{-2} = \frac{1}{10^2} = .01 \tag{1.4}$$

- Prefixes:

nano	micro	milli	centi	kilo	mega	giga	tera
10^{-9}	10^{-6}	10^{-3}	10^{-2}	10^{3}	10^{6}	10^{9}	10^{12}

  For example, nanometer, centimeter, kilometer, millisecond, microsecond, etc.

- Comparing two numbers: First need same units, then subtract, take ratio, or percent difference.
- Example: I'm 5.5 ft tall while Prof. Borst is 72 in tall.

  Convert units:

$$5.5\text{ ft}\cdot\frac{12\text{ in}}{1\text{ ft}} = 66\text{ in} \tag{1.5}$$

  Subtract:

$$72\text{ in} - 66\text{ in} = 6\text{ in} \tag{1.6}$$

  Ratio:

$$\frac{72\text{ in}}{66\text{ in}} = 1.09 \tag{1.7}$$

  Percent difference:

$$\frac{72\text{ in} - 66\text{ in}}{66\text{in}}\cdot 100 = 9\% \tag{1.8}$$

  Thus, one can say that Prof. Borst is 6 inches taller than I am, 1.09 times taller than I am, or 9% taller than I am.

- In music, when comparing two numbers (e.g., frequencies or loudness levels), using ratios is more convenient.

  Example: We know that A_2 has a frequency of 110 Hz (or 110 cycles/sec). A_3 has a frequency of 220 Hz, which is twice that of A_2, or an octave higher. A_4 has a frequency of 440 Hz, which is four times that of A_2, or two octaves higher.

- Converting units: multiply by the appropriate conversion factor written as a fraction so that the unwanted units cancel out.

- Exercise: How many centimeters are in a foot? (Use the conversion factor 1 in = 2.54 cm.)

  Answer:

$$1 \text{ ft} = 12 \text{ in} \cdot \frac{2.54 \text{ cm}}{1 \text{ in}} = 30.48 \text{ cm} \tag{1.9}$$

- Exercise: Sound travels at $v_s = 346$ m/s in dry air at a temperature of 25 °C. Calculate the speed of sound in feet per second and miles per second. (Use the conversion factors 1 m = 3.28 ft and 1 mile = 5280 ft.)

  Answer:

$$v_s = 346 \,\frac{\text{m}}{\text{s}} \cdot \frac{3.28 \text{ ft}}{1 \text{ m}} = 1135 \,\frac{\text{ft}}{\text{s}} \approx 1000 \,\frac{\text{ft}}{\text{s}} \tag{1.10}$$

$$v_s = 1135 \,\frac{\text{ft}}{\text{s}} \cdot \frac{1 \text{ mile}}{5280 \text{ ft}} = 0.21 \,\frac{\text{mi}}{\text{s}} \approx \frac{1}{5} \,\frac{\text{mile}}{\text{s}} \tag{1.11}$$

- This last approximation allows us to estimate the distance to a thunderstorm if we note the time between seeing a lightning strike and hearing the thunder. (Light travels so fast (3×10^8 m/s) that we see the lightning strike almost instantaneously.) For example, if we hear thunder 5 seconds after we see a lightning strike, the storm is approximately 1 mile away.

1.5 Logarithms

- Since ratios of frequencies play such an important role in music, the mathematical concept of a logarithm is an indispensible tool to have.

- Logarithms:

$$y = \log x \quad \Leftrightarrow \quad x = 10^y \tag{1.12}$$

  The above definition is for a base-10 logarithm, $\log x = \log_{10} x$.

- Since

$$10 = 10^1\,, \quad 100 = 10^2\,, \quad 1,000,000 = 10^6\,, \quad 1 = 10^0\,, \quad 0.001 = 10^{-3} \tag{1.13}$$

  it follows that

$$\log 10 = 1\,, \quad \log 100 = 2\,, \quad \log 1,000,000 = 6\,, \quad \log 1 = 0\,, \quad \log 0.001 = -3 \tag{1.14}$$

  Thus,

$$\log 10^n = n \tag{1.15}$$

  which means that each multiplicative factor of 10 increase in x corresponds to an additive increase of $y = \log x$ by 1. (See Figure 1.)

- Key property of logarithms:

$$\log(ab) = \log a + \log b \tag{1.16}$$

- Examples of logarithmic frequency scales: Piano keyboard, musical staff, basilar membrane of the human ear (which we shall describe later in the semester).

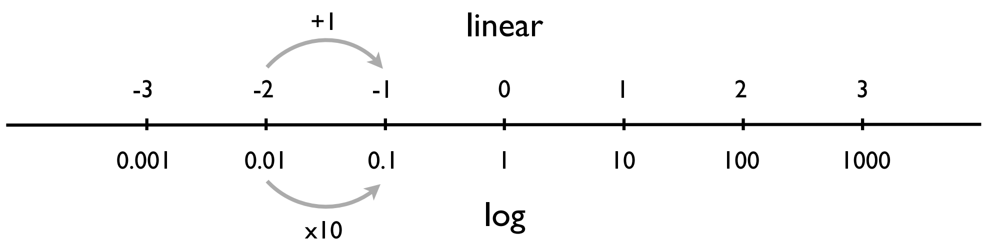


Figure 1: Difference between linear and logarithmic scales. The tick marks on a linear scale are separated by a constant additive term (here +1); the tick marks on a logarithmic scale are separated by a constant multiplicative factor (here $\times 10$).

- Equal divisions on the piano keyboard correspond to musical notes whose frequencies are related by the same constant multiplicative factor—e.g., neighboring keys are 1 semitone apart, corresponding to a frequency ratio of $2^{1/12} = 1.05946$. Thirteen keys on a piano keyboard are an octave apart, corresponding to a frequency ratio of 2.

- Useful logarithms to remember:

$$\log 2 \approx 0.3\,, \quad \log 3 \approx 0.5\,, \quad \log 4 \approx 0.6\,, \quad \log 5 \approx 0.7\,, \quad \log 10 = 1 \tag{1.17}$$

  Note that $\log 4 = \log(2 \cdot 2) = 2 \log 2$ and $\log 5 = \log(10/2) = \log 10 - \log 2$.

1.6 Music terminology

- Pitch: the fundamental frequency (i.e., the number of repetitions per second) of a musical note—e.g., concert A_4 has a fundamental frequency of 440 Hz.

- Timbre: the distinctive quality or *color* of a musical sound, which is due to the presence of overtones or harmonics of the fundamental frequency. (It's what distinguishes A_4 played on a piano and A_4 played on a guitar.)

- Interval: the separation or jump between two musical notes, usually expressed as a ratio of the fundamental frequencies of the two notes.

- Octave: a frequency interval corresponding to a factor of 2 difference in frequency.

- Chromatic scale: a division of the octave into 12 half-steps (or semitones). Figure 64 shows the corresponding notes in a chromatic scale on a piano keyboard:

$$\mathrm{C} - \mathrm{C}^\sharp - \mathrm{D} - \mathrm{E}^\flat - \mathrm{E} - \mathrm{F} - \mathrm{F}^\sharp - \mathrm{G} - \mathrm{A}^\flat - \mathrm{A} - \mathrm{B}^\flat - \mathrm{B} - \mathrm{C}'$$

- Equal temperament: a tuning system for which all semitones in the octave have the same frequency ratio

$$2^{1/12} = 1.05946 \tag{1.18}$$

  In equal-temperament, the sharps and flats are equal to one another—e.g., $\mathrm{C}^\sharp$ and $\mathrm{D}^\flat$ are tuned to the same frequency. These are called *enharmonic* notes.

- Diatonic scale: divides the octave into 7 intervals consisting of both tones and semitones. The order of tones and semitones defines the *major* and *minor* interval orders.

- The diatonic major interval order is T-T-S-T-T-T-S (2-2-1-2-2-2-1). This is the standard

$$\text{do} - \text{re} - \text{mi} - \text{fa} - \text{sol} - \text{la} - \text{ti} - \text{do}$$

  interval order.

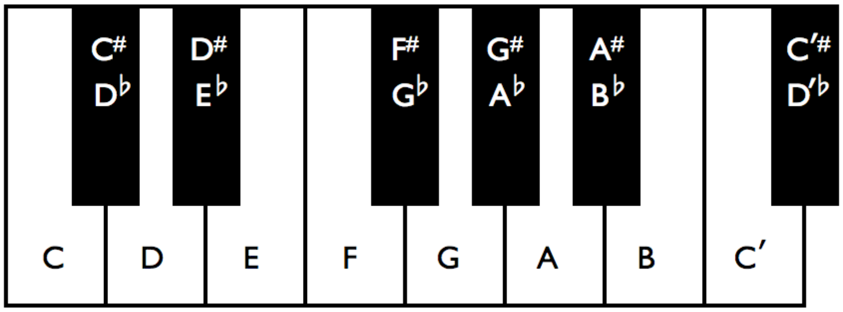


Figure 2: Notes in a chromatic scale on a piano keyboard.

- Fifth: a frequency interval corresponding to 7 half-steps between two notes (e.g., C to G or A to E or B to F$^\sharp$). The approximate frequency ratio of a fifth is $3/2 = 1.5$.
- Fourth: a frequency interval corresponding to 5 half-steps between two notes (e.g., C to F or G to C). The approximate frequency ratio of a fourth is $4/3 = 1.33$.
- Major third: a frequency interval corresponding to 4 half-steps between two notes (e.g., C to E or A to C$^\sharp$). The approximate frequency ratio of a major third is $5/4 = 1.25$.
- Minor third: a frequency interval corresponding to 3 half-steps between two notes (e.g., E to G or A to C). The approximate frequency ratio of a minor third is $6/5 = 1.2$.
- Chord: three notes played simultaneously. An example of a major chord is C-E-G (i.e., the tonic C and the major third and fifth above C). An example of a minor chord is C-E$^\flat$-G (i.e., the tonic C and the minor third and fifth above C).

1.7 Physics terminology

- Position: the location of a point in space, specified by the distance of that point away from a reference point, and a direction.
  Notation: x or y
  Unit: meter, cm, inch, foot, ...
- Displacement: the difference between two positions in space.
  Notation: Δx or Δy (which means *change* in x or change in y)
  Unit: meter, cm, inch, foot, ...
- Time: the reading on a clock.
  Notation: t
  Unit: second, minute, hour, ...
- Duration (or time interval): the difference between two clock readings.
  Notation: Δt (which means *change* in t).
  Unit: second, minute, hour, ...
- Velocity: the rate of change of position with respect to time, which has both a magnitude and direction.
  Notation:
  $$v = \frac{\Delta x}{\Delta t} \tag{1.19}$$
  Unit: m/s, mile/hr (or mph), ...

  In the absence of any outside influences, an object will either remain at rest or keep moving with the same velocity.

- Speed: distance traveled per unit time. *Average* speed is the total distance traveled divided by the total duration.
  Notation: v
  Unit: m/s, mile/hr (or mph), ...

  Example: The speedometer needle in a car measures the instantaneous speed of the car.

  Example: The speed of sound in dry air at a temperature of 25 °C is $v = 346$ m/s.

- Acceleration: the rate of change of velocity with respect to time, which has both a magnitude and direction.
  Notation:
  $$a = \frac{\Delta v}{\Delta t} \tag{1.20}$$
  Unit: (m/s)/s = m/s^2, mph/s, ...

  Example: If my car accelerates from a stop sign at a constant rate of 10 mph/s, then after 6 seconds I'll be traveling at 60 mph.

  Example: A ball freely-falling toward the floor accelerates at a constant rate of approximately 10 m/s^2.

  Example: A weight attached to the end of a spring oscillates vertically up and down with position, velocity, and acceleration all changing with respect to time. At the lowest and highest points, the velocity is zero, while the acceleration has its maximum value (directed either upward or downward). As the weight passes through the equilibrium position, the velocity has it maximum value (directed either upward or downward), while the acceleration is zero.

- Force: that which causes an object to accelerate. A push or a pull. Gravity, electricity, and magnetism are examples of *non-contact* forces.
  Notation: F
  Unit: Newton (N), pound (lb)

- Mass: the resistance that an object offers to changes in its state of motion. The larger the mass, the greater the force required to get the same acceleration.
  Notation: m
  Unit: kilogram, gram, ...

- Newton's 2nd law: force, mass, and acceleration are related by the following formula:
  $$a = F_{\text{net}}/m \quad \text{or, equivalently,} \quad F_{\text{net}} = ma \tag{1.21}$$
  where F_{net} is the *net* force (i.e., the sum of all forces) acting on the object.

  Example: The gravitational force exerted by the Earth on an object with mass m is the object's weight and is given by
  $$F_{\text{gravity}} = mg \tag{1.22}$$
  where $g = 10$ m/s^2 is the free-fall acceleration of an object near the surface of the Earth.

  Example: The restoring force exerted by a spring stretched a distance x from its equilibrium position is given by *Hooke's law*:
  $$F_{\text{spring}} = -kx \tag{1.23}$$
  where k is the spring constant, which tells how stiff the spring is. The minus sign means that the force is directed back toward the equilibrium position.

- Density: the mass of an object divided by the volume occupied by the object.
  Notation: ρ
  Unit: kg/m^3, ...

  Linear mass density of a one-dimensional object like a guitar string is given by the mass of the object divided by its length; it has units of kg/m, ...

  Example: The thicker guitar strings, which play the lower pitch (bass) notes, have a larger linear mass density than the thinner strings.

- Pressure: force per unit area, where the force is perpendicular to the surface.
  Notation:
$$P = \frac{F}{A} \tag{1.24}$$
  where A is the area over which the force is applied.

  Unit: $\mathrm{N/m^2} = 1$ Pa (where Pa stands for Pascal) or $\mathrm{lb/in^2}$ (psi)

  Example: Sound waves are pressure waves in air or in some other medium.

  Example: On Earth, standard atmospheric pressure is approximately 100 kPa, which is the same as 10^5 Pa.

  Exercise: Calculate the pressure exerted by a 120 lb woman standing on the floor, wearing stilettos having an approximately circular heel with radius 0.25 in. Compare that to the pressure exerted by a 10,000 lb elephant, whose feet are approximately circles with radius 10 in.

  Answer:

  Woman:
$$P = \frac{120 \text{ lb}}{2 \cdot \pi \cdot (0.25 \text{ in})^2} \approx 300 \ \frac{\text{lb}}{\text{in}^2} \tag{1.25}$$
  Elephant:
$$P = \frac{10000 \text{ lb}}{4 \cdot \pi \cdot (10 \text{ in})^2} \approx 8 \ \frac{\text{lb}}{\text{in}^2} \tag{1.26}$$
  So the pressure exerted by the woman's heels on the floor is almost 40 times greater than that exerted by the elephant!

Part II
Physics of oscillations and waves

2 Oscillations

- Recall that a musical note typically has a characteristic pitch (fundamental frequency), so its associated pressure wave *repeats* over time.
- Hence, understanding repeated motion is important for understanding both the production and perception of sound.

2.1 Periodic motion

- Oscillation: *any* motion that repeats itself. Also called periodic motion.
- Examples / demos: Mass on a spring, swinging pendulum, a vibrating guitar string, a vibrating air column in a tube, the Earth rotating around its axis, the Earth orbiting the Sun, etc.
- Period: time needed to complete one cycle (T)
- Frequency: number of cycles completed in one second (f)
- Relationship between period and frequency:

$$f = \frac{1}{T} \tag{2.1}$$

- Amplitude: 1/2 peak-to-peak displacement or extent of the motion (usually denoted by A).
- Periodic motion is thus characterised by its fundamental frequency $f = 1/T$, its amplitude, and the shape of the wave as a function of time (called the waveform).
- In the context of sound, the fundamental frequency corresponds to pitch, the amplitude to loudness or sound intensity, and the waveform to the timbre or tone quality of the sound.
- Different waveforms with the same fundamental frequency sound differently due to the presence of higher harmonics (more on this later).
- Examples of different waveforms: sine, triangle, square wave, sawtooth, pulse, random noise, and the sum of two sine waves with slightly different fundamental frequencies (hear beats for this last example).
- <u>Demonstration</u>: Use playsound.m routine to listen to these different sounds.
- <u>Exercise</u>: The frequency range of human hearing is roughly 20 Hz to 20 kHz. This is a factor of 1000 or roughly 10 octaves ($2^{10} = 1024$) in frequency.

  Calculate the associated periods.

  <u>Answer</u>:

$$\text{20 Hz}: \qquad T = \frac{1}{\text{20 Hz}} = 0.05 \text{ s} = 50 \text{ ms} \tag{2.2}$$

$$\text{20 kHz}: \qquad T = \frac{1}{\text{20 kHz}} = 5 \times 10^{-5} \text{ s} = 50\ \mu\text{s} \tag{2.3}$$

- The pressure variation range of human hearing at 1000 Hz is roughly 2×10^{-5} Pa (threshold of hearing) to 20 Pa (threshold of pain). This is a factor of 10^6 in pressure variation or 10^{12} in intensity (Watt/m^2). Compared to standard atmospheric pressure $P_{\text{atm}} = 10^5$ Pa, these values are only 2×10^{-10} and 2×10^{-4} of atmospheric pressure, respectively. So even at the threshold of pain, the pressure change is only a few parts in 10^4 of atmospheric pressure!
- In contrast, the frequency range for human vision is only a factor of 2 (so 1 octave); and the intensity range of human vision is only a factor of 10^5 (versus 10^{12} for human hearing). *Thus, the human ear is a much more sensitive device than the human eye!*

2.2 Simple harmonic motion

- Simple harmonic motion (SHM) is periodic motion with a sinusoidal dependence

$$x(t) = A\,\sin(2\pi f[t - t_0]) = A\,\sin(2\pi f t - \theta_0) \tag{2.4}$$

  Note that the term proportional to t_0 (or θ_0) just shifts the sine curve to the right by t_0.

- Recall from trigonometry that the sine and cosine functions can be defined with respect to a point P on the unit circle:

$$y = \sin\theta\,, \qquad x = \cos\theta \tag{2.5}$$

  and repeat whenever θ changes by 360° or 2π radians (Figure 3).

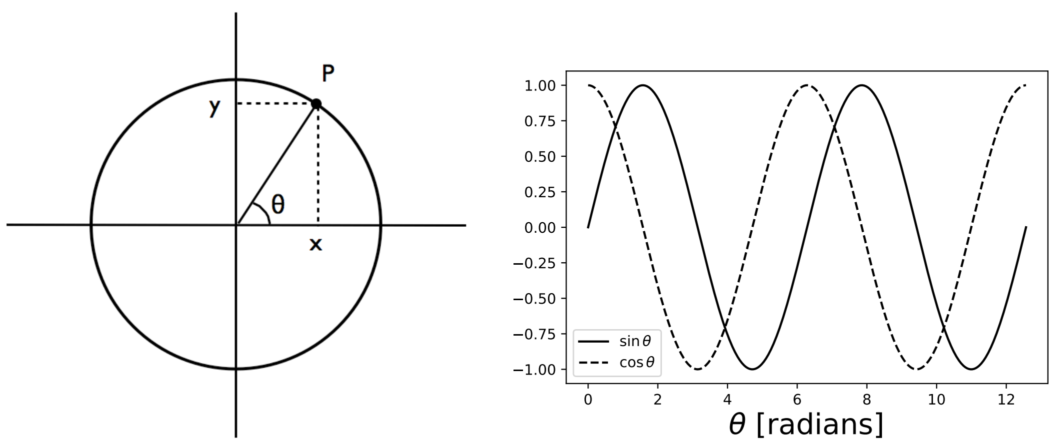


Figure 3: Left panel: Point P on a unit circle (radius equal to one) making an angle θ with respect to the x-axis. Right panel: Plots of $\sin\theta$ and $\cos\theta$.

- Simple harmonic motion is produced whenever there is a *linear restoring force* about a position of stable equilibrium.

  Linear means that the force is directly proportional to the displacement of an object away from its equilibrium position; in other words, if you double the displacement you double the force.

  Restoring means that the force is always directed in such a way as to try to bring the object back to its equilibrium position.

- One example of a linear restoring force is the force exerted by a stretched or compressed spring on a mass:

$$F_{\rm spring} = -kx \tag{2.6}$$

  where k is the spring constant (which tells how stiff the spring is) and x is the diplacement of the mass m away from the equilibrium position. The minus sign indicates that the force is directed back toward the equilibrium position.

- The period and frequency of the associated SHM are given by

$$T = 2\pi\sqrt{\frac{m}{k}}\,, \quad f = \frac{1}{2\pi}\sqrt{\frac{k}{m}} \tag{2.7}$$

- Note that the period is independent of the amplitude of the oscillation. Also, larger masses have longer periods of oscillation, while stiffer springs have shorter periods.

- Demonstration: Illustrate the above by changing the amount of mass attached to a spring.

- Another example of a linear restoring force is the force (or more precisely the torque τ) exerted by gravity on a swinging pendulum bob

$$\tau = -mgL\,\theta \tag{2.8}$$

where L is the length of the pendulum, g is the acceleration due to gravity ($g \approx 10$ m/s^2), and θ is the angular displacement of the pendulum bob away from equilibium.

- We are assuming here that the angular displacements are small, i.e., $\theta \ll 1$. If we allow large angular displacements, then the force becomes non-linear and the motion is no longer simple harmonic (see below).

- The period and frequency of the associated SHM are given by

$$T = 2\pi\sqrt{\frac{L}{g}}\,, \quad f = \frac{1}{2\pi}\sqrt{\frac{g}{L}} \tag{2.9}$$

- Note that the period is again independent of the amplitude of the oscillation. Also, longer pendula have longer periods of oscillation, and the period is *independent* of the mass of the pendulum bob.

- Demonstration: Illustrate the above by changing the length of the pendulum, and then the mass of the pendulum bob. (Parent and child swinging on separate playground swings have the same period, despite their different masses.)

- Demonstration: Use the program periodicmotion.m to illustrate the effect of non-linear restoring forces on the waveforms of the resulting period motions. Listen to the associated sounds and notice the presence of higher harmonics.

2.3 Damped oscillations

- Real-world oscillations eventually damp out due to friction and air resistance.

- A simple form of a damping (or drag) force has the form

$$F_d = -bv \tag{2.10}$$

where v is the velocity of the mass, and b is some (positive) constant that specifies the strength of the damping. An example of such a force is the drag force on a moving car due to air resistance.

2.4 Driven oscillations, resonance

- In order to sustain an oscillation that has damping, it is necessary to supply energy from outside. Such an outside force is called a *driving force*.

- For damped, driven oscillations there are three forces acting:

$$F_s = -kx \quad \text{(restoring force)} \tag{2.11}$$
$$F_d = -bv \quad \text{(damping force)} \tag{2.12}$$
$$F_{\text{driving}} = F_0\cos(2\pi f t) \quad \text{(driving force)} \tag{2.13}$$

where F_0 is the amplitude of the driving force and f is the driving frequency, which might be different from the natural frequency of the oscillation $f_0 = (1/2\pi)\sqrt{k/m}$.

- After an initial transient period corresponding to the damped (undriven) motion, the motion of the mass will be simple harmonic with a frequency equal to the driving frequency f. The phase and amplitude of the motion will depend on the relationship between the driving frequency f and the natural frequency f_0.

  Assuming that the damping is small,

  (i) for $f \ll f_0$, the motion of the mass is in phase with the driving force, and has an amplitude equal to the driving force amplitude F_0/k.

  (ii) for $f \gg f_0$, the motion of the mass is $180°$ out of phase with the driving force, and has an amplitude that tends to zero as the driving frequency tends to infinity.

  (iii) for $f \approx f_0$, the motion of the mass is $90°$ out of phase with the driving force, and has an amplitude that can be much larger than the driving force amplitude F_0/k.

- The case where $f \approx f_0$ is called *resonance*.

- Demonstration: Illustrate these three different cases using a simple pendulum, moving the pivot point by hand with different frequencies.

- Demonstration: Illustrate these three different cases using the program dampeddrivengui.m (choose damping=0.05, amplitude=0.2, frequency=variable).

- Other examples of resonance:

  - pushing a child on a swing or pumping with your legs

  - car bump analogy: the size and spacing of bumps is most effective at slowing a car when they are chosen to be approximately the same size of the car's wheels and the spacing between the front and back wheels of the car

  - blowing over a Coke bottle to produce a note: the bottle has a natural frequency that depends on its specific size and shape; blowing is the driving force having contributions from all frequencies (basically noise); only that frequency of the blowing that matches the natural frequency of the bottle is in resonance

  - a vibrating guitar string doesn't produce much sound on its own, but it acts as the driving force on the body of the guitar, which is coupled to the string via the bridge. If the body of the guitar has natural frequencies close to the frequency of the vibrating string, then the guitar body resonates and produces large amplitude oscillations that are easier to hear.

  - a tuning fork mounted on a sound box is the same as the guitar string example

  - resonance between two pendula having the same length (coupling via the rocking support rod) or between two tuning forks having the same frequency (coupling via the air)

2.5 The not-so-simple simple pendulum

- If you allow large amplitude angular displacements, the motion of a simple pendulum is no longer simple harmonic since the restoring force is no longer linear.

- The damped, driven simple pendulum with a non-linear restoring force exhibits a variety of possible motions ranging from simple harmonic, to periodic, to chaotic.

- Chaotic motion is non-periodic, unpredictable motion associated with deterministic equations. The unpredictability arises from the sensitive dependence on the starting conditions of the motion.

- Demonstration: Run forcedoscillator.m with different amplitudes for the driving force.

  0.2, 0.3: small amplitude SHM oscillations
  0.4: large amplitude, periodic but not simple harmonic oscillations
  0.5: large amplitude, periodic but not simple harmonic oscillations (two complete orbits, but in opposite directions, per period)

0.6, 0.7: non-periodic, chaotic
0.8: large amplitude, periodic but not simple harmonic oscillations (one complete orbit per period)
0.9: large amplitude, periodic but not simple harmonic oscillations (two complete orbits per period)
1.0: non-periodic, chaotic

- *It is amazing that such complex motion can arise from such a simple system!*

3 Waves and sound

- Sound is an example of a pressure wave that propagates from one location to another—e.g., from an instrument to our ears.
- In this section, we learn about general properties of waves, with applications to sound.

3.1 Basic definitions

- Wave: any 'disturbance' that transports energy and momentum from one location to another without the transport of matter.
- Example: surface water waves, waves on a stretched string, pressure waves in a column of air in a tube, electromagnetic waves (light, radio waves, X-rays, ...), sound waves in air, the 'wave' at a football game, ...
- Waves can propagate in one dimension (e.g., along a stretched string or in a column of air), in two dimensions (e.g., along the surface of a pond or the head of a drum), or in three dimensions (e.g., through air or even empty space).
- Wave motion is different from particle motion. The disturbance propagates, not the individual particles.
- The *intensity* of a wave is the rate at which energy passes through a unit area perpendicular to the direction of propagation. (Units: $\mathrm{W/m^2}$ corresponding to energy/(area $\cdot$ time).) For a spherical wave in three dimensions, the intensity is proportional to $1/r^2$, where r is the distance from the source of the wave.
- Although most waves require a medium in which to travel (for example, sound), some waves do not. Electromagnetic waves, like light, can travel through empty space.
- Example: A ringing bell in a vacuum jar. (You can see the bell, but you can't hear the sound.)
- Transverse wave: the disturbance is perpendicular to the direction of wave propagation.
- Longitudinal (or compressional) wave: the distubance is parallel to the direction of wave propagation.
- Electromagnetic waves, like light, are examples of transverse waves.
- Sound is an example of a longitudinal pressure wave consisting of alternating compressions (overdensities) and expansions (underdensities) in a medium. The medium can be air (or any gas), a liquid, or a solid.
- Liquids and gases can only support longitudinal waves, since there is no "sideways" restoring force. (The reason why surface water waves can have a transverse component is gravity is acting as the restoring force.)
- Demonstration: Illustrate transverse and longitudinal waves on a slinky. Illustrate the difference between a wave pulse and a periodic wave on a slinky.

3.2 Wave velocity

- Since the wave disturbance propagates from one location to another, waves have a velocity equal to the distance traveled by the disturbance divided by the corresponding time interval.
- The wave velocity depends only on the properties of the medium and not on the motion of the source relative to the medium.
- Wave velocity on a string:

$$v = \sqrt{F/\mu} \tag{3.1}$$

where F is the tension in the string and $\mu = m/L$ is the mass per unit length of the string.

- The tension in the string is the restoring force that tries to bring the (slightly) deformed string back to its equilibrium position. Note that for a heavier string, the wave velocity is less than that for a light string, assuming the same tension and length of the string.

- More generally, the wave velocity in any medium is given by

$$v = \sqrt{\text{elastic/inertial}} \tag{3.2}$$

  where the elastic property of the medium provides the restoring force and the inertial property provides the resistance that the medium offers to changes in its state of motion.

- Using this more general expression, one can show that the speed of sound in air depends on the temperature of the air:

$$v_{\text{air}} = 331 \text{ m/s} \sqrt{1 + \frac{T_C}{273.15}} \tag{3.3}$$

  where T_C is the temperature in degrees Celsius (centigrade scale).

- Exercise: Show that speed of sound in air at room temperature $T = 25$ °C (77 °F) is given by

$$v_{\text{air}} = 346 \text{ m/s} \tag{3.4}$$

- Exercise: Compare the above answer to the speed of sound in air at $T = 20$ °C (68 °F). Show that the percent difference between the two speeds is approximately 1%, which (as we will show in the next subsection) corresponds to a 1% increase in the frequency (or pitch) of a sound.

- Sound travels in other materials as well. For example,

$$v_{\text{water}} \approx 1440 \text{ m/s}\,, \quad v_{\text{steel}} \approx 5000 \text{ m/s} \tag{3.5}$$

3.3 Periodic waves

- Periodic waves are produced by periodic sources. They are characterized by frequency $f = 1/T$, amplitude A, and wavelength λ (the distance between two like points on the wave). The waves travel with wave speed v given by

$$v = \frac{\lambda}{T} = f\lambda \tag{3.6}$$

- Sound waves can be produced in principle at any frequency, but the human ear is sensitive "only" to frequencies in the approximate range 20 Hz to 20,000 Hz. Dogs can hear higher frequencies; whales and elephants, lower frequencies.

- Exercise: Show that the corresponding wavelengths of sound waves in air traveling with $v = 346$ m/s for $f = 20$ Hz and 20,000 Hz are $\lambda = 17$ m and 1.7 cm, respectively. (We will refer to these numbers later on when we talk about the diffraction of sound waves around obstacles.)

- The intensity for a periodic wave is proportional to the wave velocity, the square of the displacement wave amplitude, and the square of its frequency.

- The human ear is sensitive to intensities ranging from 10^{-12} W/m^2 (threshold of hearing) to 1 W/m^2 (threshold of pain).

- The maximum displacement of the air molecules corresponding to these two intensities at a frequency of 1000 Hz are 1.1×10^{-11} m and 1.1×10^{-5} m, respectively. Note that the maximum displacement of the air molecules at the threshold of hearing is approximately 1/10 the size of a typical atom!!

- The maximum pressure deviations due to the over- and underdensity of the air molecules are 2.8×10^{-5} Pa and 28 Pa, respectively. These numbers should be compared to the ambient atmospheric pressure of 1×10^5 Pa. Thus, even the maximum pressure deviations at the threshold of pain are approximately 10,000 times smaller than the ambient atmospheric pressure.

- The above items illustrate the incredible sensitivity of the human ear to hear sound.
- Simple harmonic waves are produced by SHM sources. In 1-dimension, a plot of a simple harmonic wave has sinusoidal dependence with respect to both: (i) variable position x for fixed t (e.g., a snapshot of the wave), and (ii) variable time t for fixed x (simple harmonic displacement of the particle at x).
- Mathematically, a simple harmonic wave in 1-dimension has the form

$$s(x,t) = s_m \sin\left[2\pi\left(\frac{x}{\lambda} - \frac{t}{T}\right) - \phi_0\right] \tag{3.7}$$

  Note the periodicity in the waveform when x is increased by one wavelength λ and/or t is increased by one period T.
- Note that when we plot air molecule displacement (or pressure deviation from atmospheric pressure) as a function of position x, we are giving a 'transverse-wave representation' of a longitudinal wave.
- Demonstration: Graphical illustration that air molecule displacement and pressure deviations are 90° out of phase with one another—i.e., one is maximum (or minimum) when the other is zero.

3.4 Superposition principle and interference

- Principle of superposition: The sum of two waves is another wave.
- Waves can pass through one another and emerge without changing the identity of the other wave.
- Waves are said to *constructively* or *destructively* interfere when the disturbances in the two waves add or subtract from one another.
- For two waves having the same frequency, whether the two waves add constructively or destructively depends on the *phase difference* between the two waves.
- Demonstration: Illustrate superposition of two sine waves using sumsines.m, keeping the frequency the same, but varying the amplitude and phase of the two waves.
- Example: Noise cancelling headphones use destructive interference to reduce the level of ambient noise. (The headphones sense the ambient noise, invert the corresponding sound waves, and then send the inverted wave to the earphones where they cancel most of the ambient noise.)
- Demonstration: Illustrate interference of waves traveling in two-dimensions by overlaying two transparencies of concentric circles. The clear spaces between the concentric circles represent the crests of the waves; the distance between two neighboring circles is one wavelength. Constructive interference corresponds to clear areas of the overlaid transparencies.

  Note that the number of radial lines corresponding to constructive interference decreases as the spacing between the two point sources decreases.
- Beats are an example of two sound waves with nearly the same frequencies f_1 and f_2 interfering with one another. The beat frequency is given by $f_{\rm beat} = |f_1 - f_2|$.
- Beats are an example of interference in time, as opposed to interference in space.
- Demonstration: Illustrate beats using two tuning forks having frequencies of 440 Hz and 441 Hz.

  We hear a sound with frequency $\bar{f} = (f_1 + f_2)/2 = 440.5$ Hz, whose amplitude is modulated in time with frequency $f_{\rm beat} = |f_1 - f_2| = 1$ Hz.
- Demonstration: Illustrate beats using sumsines.m with same amplitudes, but $f_1 = 9$ Hz and $f_2 = 10$ Hz. Change the amplitude of one of the waves, noting how the modulation envelope changes—e.g., the modulation doesn't dip to zero anymore. Show how changing the phase of the two waves shifts the modulation envelope to the right or to the left.

- Demonstration: Illustrate beats using the Faber Electroacoustics Toolbox using two signal generator and the oscilloscope mode. Set f_1 to 440 Hz and let f_2 be variable. Initially set $f_2 = 441$ Hz to hear beats with a frequency of 1 Hz. Adjust f_2 so that the beat frequency is too large to hear beats. When do we eventually hear two *separate* frequencies? What do we see using iSpectrum in spectrum analyzer mode for this case?

- NOTE: We normally hear beats if the frequency difference is less than about 10 Hz, independent of the center frequency. We can distinguish two separate frequencies (played simultaneously) when the frequency difference is greater than about 10% of the center frequency, which corresponds to about two semitones.

3.5 Standing waves

- Begin by considering waves on a string.

- A right-moving wave becomes a left-moving wave when it is reflected off of a boundary. For a fixed end, the reflected wave is inverted; for a free end, the reflected wave is not inverted. (Note that the string at a fixed end has zero displacement, so the incident and reflected waves on the string must be opposite one another in order to sum to zero at this point.)

  This process also occurs at the left boundary, changing a left-moving wave into a right-moving wave, etc.

- The superposition of right and left-moving waves having just the right frequencies can give rise to standing wave patterns on a string fixed at both ends:

$$\lambda_n = \frac{2L}{n}\,, \quad f_n = \frac{v}{\lambda_n} = n\left(\frac{v}{2L}\right)\,, \quad n = 1, 2, \cdots \tag{3.8}$$

  $n = 1$ is the fundamental frequency; different values of n correspond to different *harmonics*.

- Draw standing wave patterns for a string fixed at both ends.

- Nodes are locations where the displacement is zero. Anti-nodes are locations where the displacement is maximum.

- Note that standing waves are a resonance phenomenon. Only for certain wave frequencies do the right-moving and left-moving waves (which are traveling back-and-forth on the string) add constructively to the new waves that are produced at the source. For other frequencies, the new waves destructively interfere with the waves already traveling along the string.

- Demonstration: Produce standing waves on a long stringlike spring by oscillating one end back-and-forth at the appropriate frequencies.

- Demonstration: Illustrate standing waves using standingwaves.m with arguments 1, 2, 3, 1.7. Compare the amplitude of the summed waves on and off resonance.

- Standing wave vibrations in a column of air can also be set up in a cylindrical tube.

  Tube open at both ends:

$$\lambda_n = \frac{2L}{n}\,, \quad f_n = n\left(\frac{v}{2L}\right)\,, \quad n = 1, 2, 3, \cdots \tag{3.9}$$

  Tube open at one end, closed at the other:

$$\lambda_n = \frac{4L}{n}\,, \quad f_n = n\left(\frac{v}{4L}\right)\,, \quad n = 1, 3, 5, \cdots \tag{3.10}$$

  $n = 1$ is the fundamental frequency.

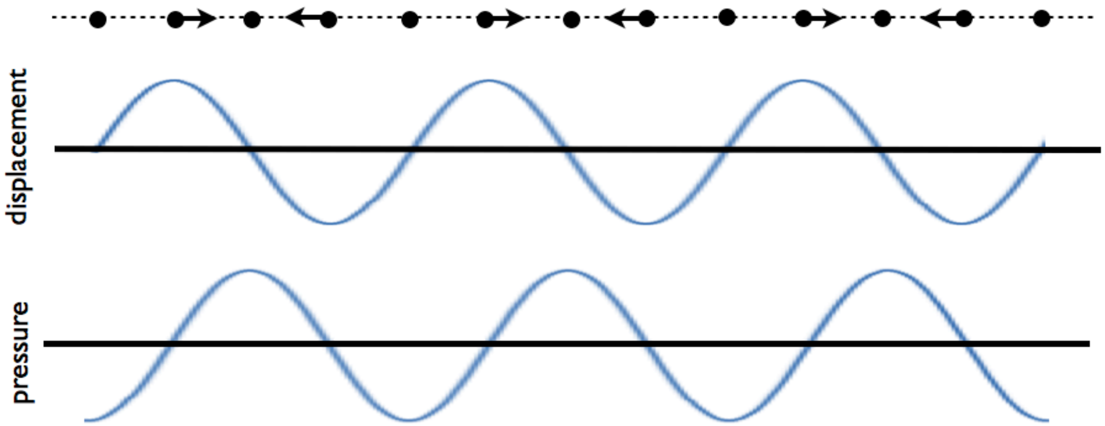


Figure 4: Relationship between air molecule displacement and pressure deviation.

- Draw standing wave patterns for both types of tubes. Compare plots of pressure deviation and air molecule displacement as functions of distance along the tube. (See Figure 4.) Recall that they are 90° out of phase with one another. For example, an open end corresponds to a node for pressure deviation but an anti-node for air molecule displacement.

- Note that a tube open at both ends has both odd and even harmonics, while a tube closed at one end has only odd harmonics.

- Although it is also possible to consider standing waves in a tube closed at both ends (which has the same standing wave patterns as a tube open at both ends), it is of no interest for us in the context of sound, since the waves in the tube have no way to couple their oscillations to the air outside the tube.

- Example: A clarinet is a woodwind instrument that behaves like a tube closed at one end. (The reed corresponds to the closed end.) The notes it produces are dominated by odd harmonics.

- Example: Contrast the clarinet to a flute or recorder, which behave like tubes open at both ends. These instruments produce notes that have both odd and even harmonics.

- The standing wave patterns for a cylindrical tube with one or two open ends actually have their pressure nodes at a slight distance *outside* the tube, due to the two-dimensional nature of the cylinder. If R denotes the radius of the cylinder:

$$L_{\text{eff}} = L + 0.61R \quad \text{(tube open at one end)} \tag{3.11}$$
$$L_{\text{eff}} = L + 1.22R \quad \text{(tube open at both ends)} \tag{3.12}$$

  is the effective length of the tube. To be accurate, the effective length should be used in the expressions for the standing wave wavelengths and frequencies.

- Demonstration: Using a tuning fork with frequency $f = 1024$ Hz and an adjustable length tube that is closed at one end, estimate the speed of sound in air.

  Find two consecutive lengths of the tube that resonate with the tuning fork. The difference of those two lengths, $\Delta L = L_2 - L_1$, must equal half a wavelength λ, so $\lambda = 2\Delta L$. But since $v = f\lambda$, we have

$$v = 2f(L_2 - L_1) \tag{3.13}$$

  as our estimate of the speed of sound in air. (Compare this to $v_{\text{air}} = 346$ m/s at room temperature.)

- One can also talk about standing waves for two-dimensional objects (e.g., a stretched drum head) and for stiff objects like one-dimensional bars and two-dimensional plates.

- For these more complicated objects, the frequencies corresponding to different standing wave patterns do *not* form a simple harmonic series.

- Demonstration: Connect a loudspeaker to a function generator, whose frequency can be varied. Attach a Chladni plate to the speaker, and sprinkle salt on the plate. Adjust the frequency of the function generator to find the resonant frequencies for different standing wave patterns of the plate. (The salt collects along the nodal lines of the standing wave patterns.) Record the resonant frequencies and show that they are not harmonically related to one another.

3.6 Reflection

- When a wave travels from one medium to another, part of the wave is transmitted and part of the wave is reflected.

- For a wave traveling in one-dimension (like a wave on a string, or a sound wave in a column of air), the reflected wave moves in the direction opposite to the incident wave.

- The reflected wave is inverted (i.e., has a 180° phase shift relative to the incident wave) if the incident wave encounters a fixed end. The reflected wave has the same shape (i.e., there is no phase shift relative to the incident wave) if the incident wave encounters a free end.

- Demonstration: Illustrate this by observing incident and reflected waves on a single stretched string, with fixed or free ends.

- For waves traveling in two- or three-dimensions, the wavefronts are approximately lines or planes far from the source. The direction of wave propogation is perpendicular to the wave fronts (it is called a *ray*.)

- The *law of reflection* tells how the direction of the reflected wave is related to the direction of the incident wave:
$$\theta_i = \theta_r \tag{3.14}$$
where θ_i, θ_r are the angles that the incident and reflected waves make with respect to the normal.

- Reflection is responsible for echoes. Reducing the amount of echoes is important in the design of musical performing halls.

- Example: A whispering room in a science museum consists of two parabolic walls facing one another, which focus sound produced at one focal point to the other focal point.

3.7 Refraction

- Refraction is the 'bending' or change in the direction of propagation of a wave as it passes from one medium into another. The amount of bending depends on the wave velocities in the two media.

- Demonstration: A straw in a glass of water appears to be 'bent' when viewed from an angle.

- The *law of refraction* (Snell's law) tells how the direction of the transmitted wave is related to the direction of the incident wave:
$$\frac{\sin\theta_1}{v_1} = \frac{\sin\theta_2}{v_2} \tag{3.15}$$
where v_1, v_2 are the velocities of the waves in the two media. (Fig. 5). The angle with respect to the normal is larger in the medium with the larger wave velocity.

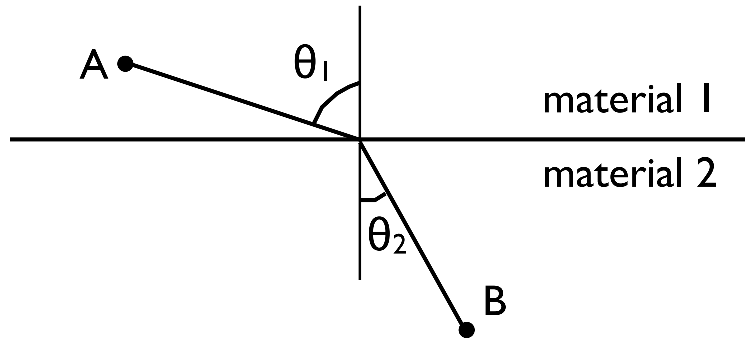


Figure 5: Incident and refracted rays.

- Example: Surface water waves tend to approach the shore at right angles, regardless of the initial direction of propagation far from the shore. This is because the velocity of the waves decreases as the water becomes more shallow closer to the shore. (Treat the water as a succession of media with different wave velocities, and then apply Snell's law to each successive interface to determine the direction of propagation.)

- Example: A similar analysis can be applied to sound waves in air above the surface of the Earth. If the temperature of the air decreases with increasing altitude, so that the velocity of sound in air decreases with increasing altitude, then sound waves will tend to propagate upward to higher altitudes. For this case, it will be difficult to hear sounds at large distances from the source.

- Example: Conversely, if the temperature of the air increases with increasing altitude (called *temperature inversion*), so that the velocity of sound in air increases with increasing altitude, then sound waves will tend to propagate back toward the surface of the Earth. (This condition is called *temperature inversion.*) For this case, one can hear sounds at relatively large distances from the source, since the sound tends to stay close to the surface of the Earth.

- Wind: Suppose the air temperature doesn't change with altitude, but there is a wind blowing (i.e., the air has a velocity) parallel to the surface of the Earth. Due to frictional effects, the air velocity will be smallest near the surface of the Earth; it will increase with increasing altitude.

  Since the air is moving relative to the ground, the relevant sound speed is the *speed of sound with respect to the ground*, not with respect to the air.

- Example: For sound propagating in the same direction as the wind, the velocity of sound with respect to the ground increases with altitude, similar to the temperature inversion case. *Thus, it is easy to hear sound that propagates in the direction of the wind.*

- Example: For sound propagating in the direction opposite to the wind, the velocity of sound with respect to the ground decreases with altitude, similar to the case where the air temperature decreases with altitude. *Thus, it is difficult to hear sound that propagates against the wind.*

3.8 Diffraction

- Diffraction refers to the 'spreading' of a wave as it passes through an opening or around an obstacle in space.

- For a fixed size object, diffraction is most noticeable for waves that have wavelengths comparable to or bigger than the size of the opening or obstacle.

- Visible light has a wavelength (~ 500 nm $= 0.5\ \mu$m) much smaller than normal everyday objects (the thickness of a ordinary piece of copy paper is approximately 100 μm). Hence everday objects cast sharp shadows for visible light, so you cannot see around a corner.
- Sound has much larger wavelengths (a 440 Hz note has a wavelength of about 1 m). Hence sound waves can easily diffract through doorways and objects less than about 100 m in size.
- Similarly, AM radio waves (electromagnetic waves) have a longer wavelength than FM radio waves, and hence can diffract around large buildings more easily than FM waves.
- Exercise: AM radio stations broadcast with frequencies between 530 kHz and 1710 kHz; FM radio stations broadcast with frequencies between 88 MHz to 108 MHz. Using $v = 3 \times 10^8$ m/s for the velocity of electromagnetic waves in air, calculate the wavelength of a (a) 1000 kHz AM radio wave, and (b) 100 MHz FM radio wave.

  Answer: (a) 300 m, (b) 3 m.

3.9 Doppler effect

- Doppler effect: The change in observed frequency due to the motion of the source and/or receiver relative to the medium through which the sound travels.
- Example: The observed frequency of a siren on a police car that is moving toward you is larger than the frequency of the siren when it is at rest. The observed frequency is smaller when the police car is moving away from you.
- Demonstration: Use the function generator and a small speaker to produce a pure tone, e.g., 440 Hz. Then take the speaker, still connected to the function generator by a long cable, and twirl it overhead in a circle. One can hear the frequency of the sound increase and decrease as the speaker approaches and recedes from the listener.

4 Fourier analysis

- *Fourier's Theorem*: Any complex periodic function can be written as a sum of harmonics.
- *Corollary to Fourier's Theorem*: Any complex periodic vibration can be written as a sum of standing wave vibrations.
- Mathematically, Fourier's theorem can be expressed as

$$y(t) = A_0 + A_1 \sin(2\pi f_1 t + \phi_1) + A_2 \sin(2\pi f_2 t + \phi_2) + \cdots \tag{4.1}$$

where $y(t)$ is a periodic function with period T and $f_n = nf_1$ $(n = 1, 2 \cdots)$ are harmonics of the fundamental frequency $f_1 = 1/T$. The term A_0 allows for a constant offset, and A_n and ϕ_n are the amplitudes and phases of the component sinusoids.

- NOTE: The above theorem also applies to periodic functions of position $y(x)$, where L is the period of the function.
- Mathematically, the corollary to Fourier's theorem can be expressed as

$$y(x,t) = A_1 \sin(2\pi x/\lambda_1)\cos(2\pi f_1 t + \phi_1) + A_2 \sin(2\pi x/\lambda_2)\cos(2\pi f_2 t + \phi_2) + \cdots \tag{4.2}$$

where A_n and ϕ_n $(n = 1, 2, \cdots)$ are the amplitudes and phases of the standing wave vibrations, having wavelengths $\lambda_n = 2L/n$ and frequencies $f_n = v/\lambda_n = nv/2L$ (v being the wave velocity).

- Each term above corresponds to a standing wave vibration pattern, with a vibration frequency that is a harmonic of the fundamental frequency $f_1 = v/2L$.
- Decomposing a periodic function into its component sinusoids is called *Fourier analysis*. Constructing a periodic function by adding to together harmonics having different amplitudes and phases is called *Fourier synthesis*.
- If the function $y(t)$ is not periodic, one can still decompose it into sinusoids, but now one has contributions (in principle) from *all* frequencies, and not just the (discrete) harmonics.
- Demonstration: The FFT Analyzer tool from the Faber Electroacoustics Toolbox performs Fourier analysis.

  A pure tone (e.g., a tuning fork) consists of just a single frequency component. A note produced by an instrument (e.g., a guitar or recorder) will have a frequency spectrum dominated by a few harmonics.
- Examples of different waves and their Fourier decompositions:

  (i) square wave: only odd harmonics, with amplitudes 1, 1/3, 1/5, $\cdots$, and phases all 0.

  (ii) triangle wave: only odd harmonics, with amplitudes 1, 1/9, 1/25, $\cdots$, and phases alternating between 0 and 180° (so alternating $\pm$).

  (iii) sawtooth wave: both odd and even harmonics, with amplitudes 1, 1/2, 1/3, $\cdots$, and phases all 0.
- Demonstration: Use the matlab routines fouriersythesize.m, fouriersynthesizeScript.m to synthesize periodic waves given the amplitudes and phases of the harmonics.
- Demonstration: Convert these to periodic functions to audible sound using the matlab routines fouriersythesizesound.m, fouriersynthesizesoundScript.m.
- Demonstration: Synthesize sound waves having the same component amplitudes but different component phases. Can you hear a difference?

  Typically, phase has little effect on the timbre of a sound. This is called *Ohm's law of hearing.* The timbre is only noticeably different when the phases of two Fourier combinations are radically different from one another—e.g., both waves have amplitudes $1, 1/2, 1/3, \cdots$ but phases $0, 0, 0, \ldots$ versus $0, 90°, 0, 90°, \cdots$ (so alternating sines and cosines).

- Demonstration: Synthesize sound waves having a contributions from the 2nd, 3rd, and higher harmonics, but no contribution from the fundamental—e.g., contributions from 440 Hz, 660 Hz, 880 Hz, $\cdots$ but not from 220 Hz. This is an example of a sound have a *missing fundamental*. What pitch do you hear? Do you hear the fundamental frequency even though it has zero amplitude?

  NOTE: Most people hear a pitch corresponding to the fundamental frequency even though it is absent from the spectrum. This is called a *virtual* or *subjective* pitch. Although the ear doesn't physically respond to a frequency that's not present, cognitive processes in the brain infer its presence from the timing of electrical impulses triggered by the periodicity of the sound wave.

Part III
Production of sound

5 String instruments

- This section provides an introduction to the production of sound by string instruments.
- We will consider the harp, guitar, and violin to illustrate the fundamental differences between string instruments.
- We will follow the presentation in the book "How music works," by John Powell. More details can be found in the text by Berg and Stork.

5.1 Sounds from strings

- String instruments come in two basic types: those that are plucked and those that are bowed.
- Examples of plucked string instruments are the harp, guitar, banjo, ... Examples of bowed string instruments are the violin, cello, ...
- Plucking and bowing are the two basic methods for exciting vibrations in a stretched string.
- As discussed in a previous section, the frequency of a standing wave vibration of a stretched string depends on only three things: (i) the length of the string, (ii) what the string is made of, and (iii) the tension in the string.

  (i) The length of the string determines how far the waves have to travel to go from one end to another. The longer the string, the longer the travel time, and hence the lower the frequency of vibration.

  (ii) What the string is made of is important since heavier strings offer more resistance to changes in their state of motion, and hence correspond to lower-frequency waves.

  (iii) The tension in the string is a measure of how tightly the string is stretched. Strings under higher tension are pulled back toward their equilibrium position with greater restoring force, and hence correspond to higher-frequency waves.

- The mathematical formula for the frequency of the nth standing wave vibration is

$$f_n = \frac{n}{2L}\sqrt{\frac{F}{\mu}}\,, \qquad n = 1, 2, 3, \cdots \tag{5.1}$$

  where L is the length of the string, μ is its mass per unit length, and F is the tension in the string. (f_1 is the fundamental frequency and $f_n = nf_1$ is the nth harmonic.)

- A vibrating string by itself doesn't produce much sound. The vibrations of a string need to be amplified by attaching them to some sort of wooden box (called a soundboard), which is able to drive more air and produce a louder sound.
- The interaction between the natural frequencies of the strings and the natural frequencies of the soundboard are responsible for the rich timbre of the different string instruments.

5.2 Physics of a plucked string

- Where you pluck a string determines the timbre of the sound produced by the subsequent vibration.
- As shown in Figure 6, plucking a string in the middle give the purest tones, consisting of just the odd harmonics. Plucking a string near the ends produces "fuller" tones, with contributions from both even and odd harmonics.
- Demonstration: Reproduce some of these plots using the matlab routine `pluckedstringharmonics.m` for different values of the plucking location α.

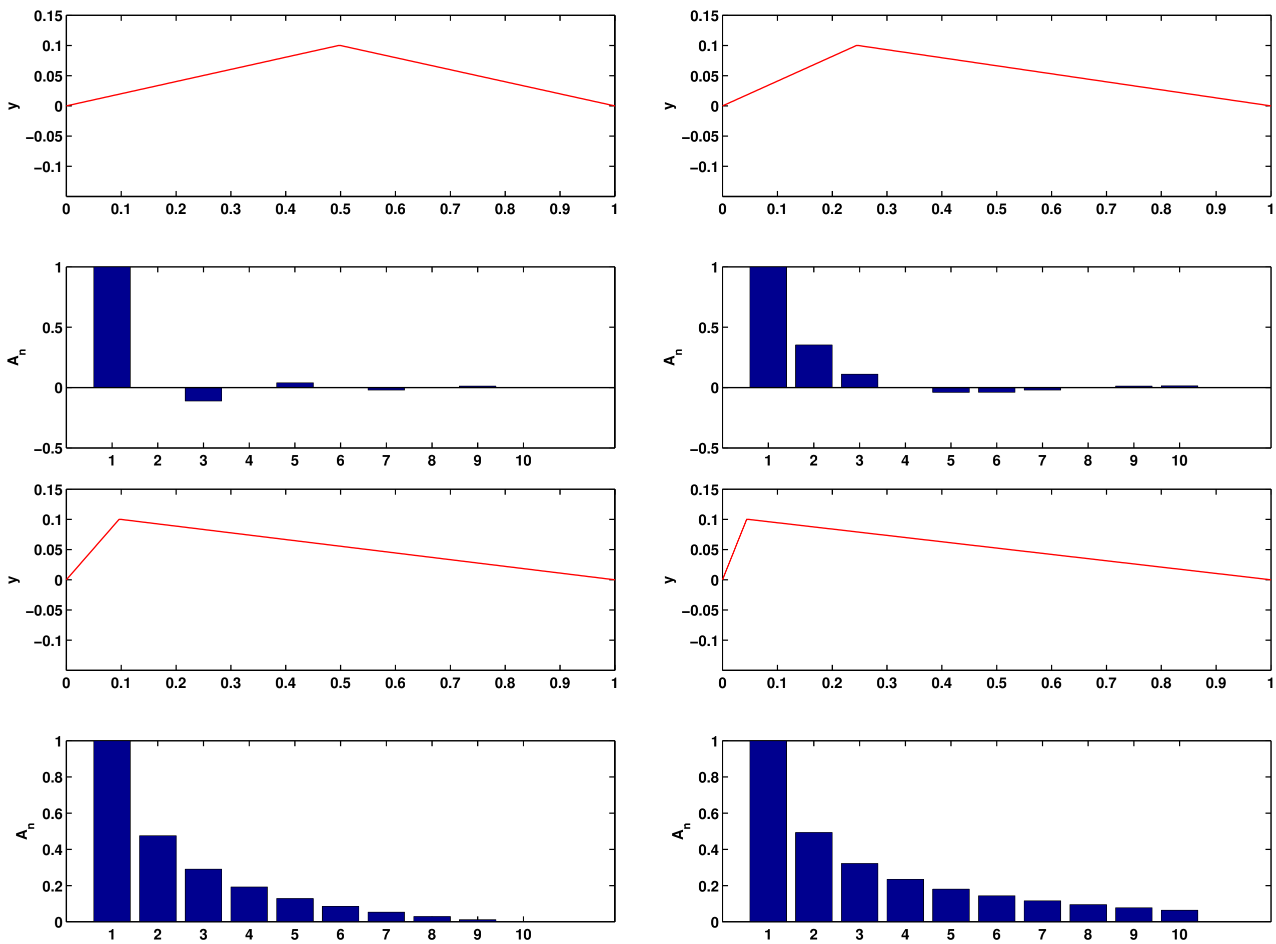


Figure 6: Fourier coefficients corresponding to strings plucked in different locations. Note that a string plucked in the center (top-left plot) corresponds to the purest tone, with small contributions from just the odd harmonics. A string plucked near the end (bottom-right plot) corresponds to a fuller sound, with relatively large contributions from many harmonics.

- Demonstration: Watch the YouTube videos:
"Motion of a plucked string" `http://www.youtube.com/watch?v=_X72on6CSL0` and
"Guitar Oscillations Captured with iPhone 4" `http://www.youtube.com/watch?v=TKF6nFzpHBU`.

- Note that the guitar string oscillations captured with an iPhone show the *same wave* pulse at different locations along the string. (So this is *not* how a real wave looks on a guitar string.) This is due to the *rolling shutter* feature of the iPhone video camera, which opens parallel to the longer side of the iPhone. If the iPhone was positioned such that the longer side of the iPhone was parallel to the guitar strings, then one wouldn't see these wave pulses at all.

- Demonstration: Verify these statements using a guitar and an iPhone. (Note: To see the wave pulses, the strings need to be brightly lit—e.g., use a flood light placed next to the strings if you are doing this indoors. A bright light increases the iPhone frame rate to 30 frames per second.)

- Mathematically, the Fourier coefficients for a plucked string are proportional to $\sin(n\pi\alpha)/n^2$, where $n = 1, 2, \cdots$, and α is the fractional distance from the bridge (i.e., where the string is plucked).

- Note that if $\alpha = 1/N$, then there is no contribution from the Nth harmonic and its multiples.

- For a plucked string, the sideways force on the bridge has the shape of a square wave. The relative durations of the positive and negative parts of the square wave depend on α. The Fourier coefficients for this type of square wave are proportional to $\sin(n\pi\alpha)/n$.

- Demonstration: Illustrate this using the matlab routine `pluckedstring.m` for different values of the plucking location α.

5.3 Physics of a bowed string

- Bowing is another way to excite the vibrations of a stretched string.

- The violin, cello, and bass are examples of bowed string instruments.

- A bow is typically made from hairs from a horse's tail, stretched tight on a stick. The horse hairs are made sticky by rubbing them with rosin (dried tree resin).

- The bow excites vibrations in a string via an alternating *stick-slip* motion. The bow first 'sticks' to a string, dragging it away from its equilibrium position. The string then 'slips' from under the bow, moving back towards, and then overshooting, its equilibrium position (Figure 7).

- This stick-slip motion happens hundreds of times per second, aligning itself with the natural vibrational frequency of the string.

- Demonstration: Wet your finger and rub it across a window pane or rim of a glass. Your finger tip slips and sticks as you press your finger against the glass, producing a squeaking sound as it moves.

- The vibrational motion of a bowed a string is basically a triangular wave that runs around a parabolic arc, as shown in Figure 8.

- Demonstration: Watch the YouTube video "Bowed violin string in slow motion"
`http://www.youtube.com/watch?v=6JeyiM0YNo4`.

- It turns out that *all* the harmonics of the vibrating string are excited by the bowing motion, *regardless of the location of the bowing.* The Fourier coefficients are $(-1)^{n+1}/n^2$ for a bowed string, independent of the location of bowing.

- NOTE: This statement is strictly true only for bowing motion *perpendicular* to the direction of the string, and for a *very narrow* bow. Since real bows have a considerable width, the kink in the triangular wave is smoothed out, eliminating some the harmonics above $N \sim 1/\alpha$. Thus, for real bows, there is a *slight* dependence of the harmonics on the location of bowing.

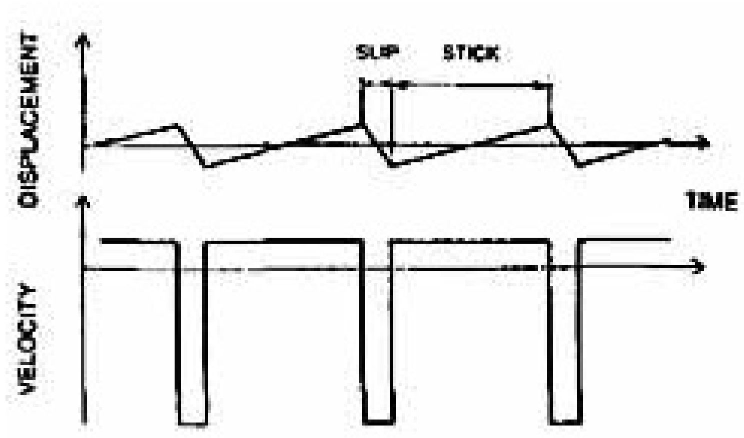


Figure 7: Stick-slip motion of a bowed string. Top panel shows string displacement as a function of time. Bottom panel shows string velocity as a function of time. This example is for the case where the string is bowed one-quarter of the way from the bridge. (Figure taken from `http://www.zainea.com/`.)

- There is a sideways force on the bridge exerted by a component of the tension in the stretched string. In the limit of an infinitesimally narrow bow, this force has the shape of a sawtooth wave, independent of the location of the bowing. The Fourier coefficients of the sawtooth wave are proportional to $(-1)^{n+1}/n$.
- Demonstration: Illustrate these statements using the matlab routine `bowedstring.m` for different values of the bowing location α.

5.4 Harp

- The harp is probably the simplest of all stringed instruments. (See Figure 9.)
- The strings have fixed lengths and they are plucked.
- Each string produces just one note. (That's why a harp needs so many strings!)
- However, by using a foot pedal that changes the tension in the strings, one can raise or lower the note by a semitone. (A harp with such a foot pedal is called a *double-action* harp.)

5.5 Guitar

- The guitar is probably the next simplest string instrument. (See Figure 10.)
- It has six strings all of the same length.
- The strings are made of different materials and are under different tensions, to produce notes with different pitches. (The tension in the strings can be adjusted using the tuning keys.)
- By pressing the strings against frets, which are located on the neck of the guitar (between the nut and the bridge), the effective lengths of the strings can be changed. This increases the number of notes that can be produced.

  i) By pressing a string against the seventh fret, you decrease its length to two-thirds of its full length, producing a sound a perfect fifth above that of the open string.

  ii) By pressing a string against the twelfth fret, you decrease its length to one-half of its full length, producing a sound an octave above that of the open string.

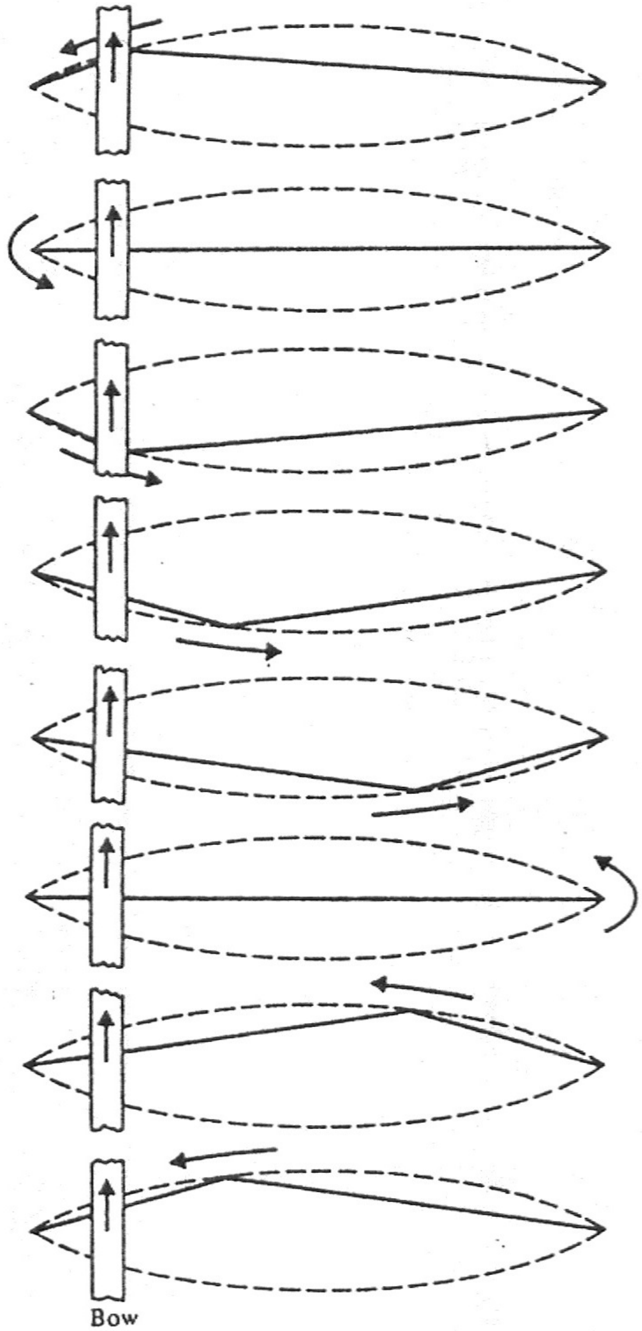


Figure 8: Schematic diagram showing the motion of a bowed string. In the top three panels, the string slips from underneath the bow, with the kink traveling counter-clockwise from bow to bridge to bow again. In the bottom five panels, the string sticks to the bow and is dragged along with it, with the kink traveling counter-clockwise from bow to nut, and eventually back to the bow again. The envelope of the motion consists of two parabolic arcs. (Figure taken from `http://www.colorado.edu/physics/phys1240/`.)

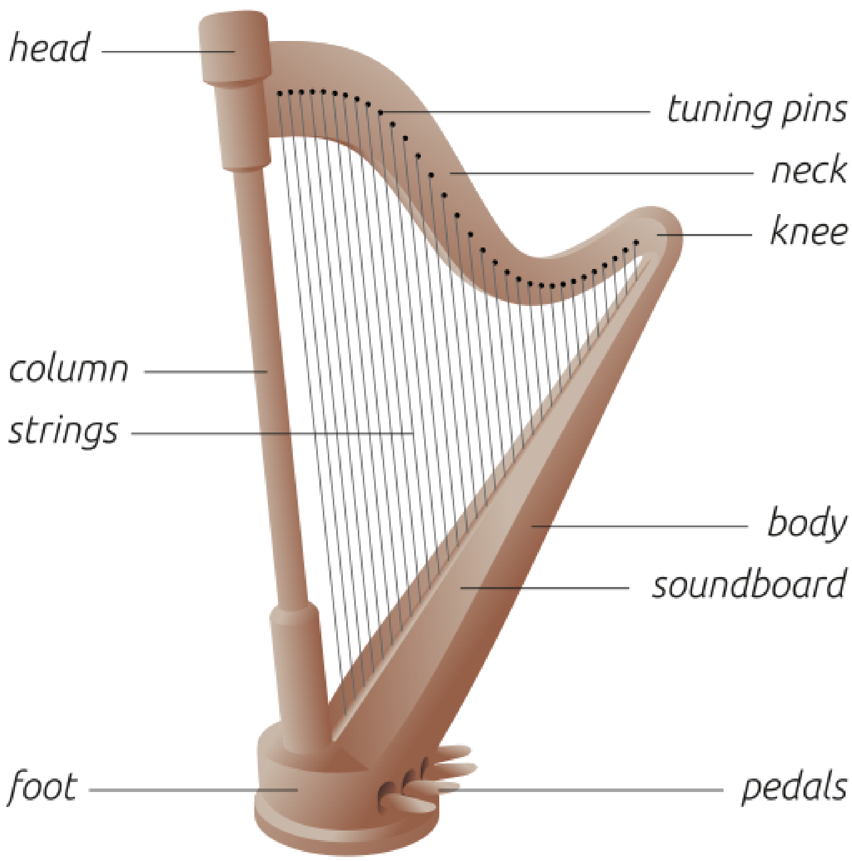


Figure 9: The parts of a harp. (Figure taken from Wikipedia.)

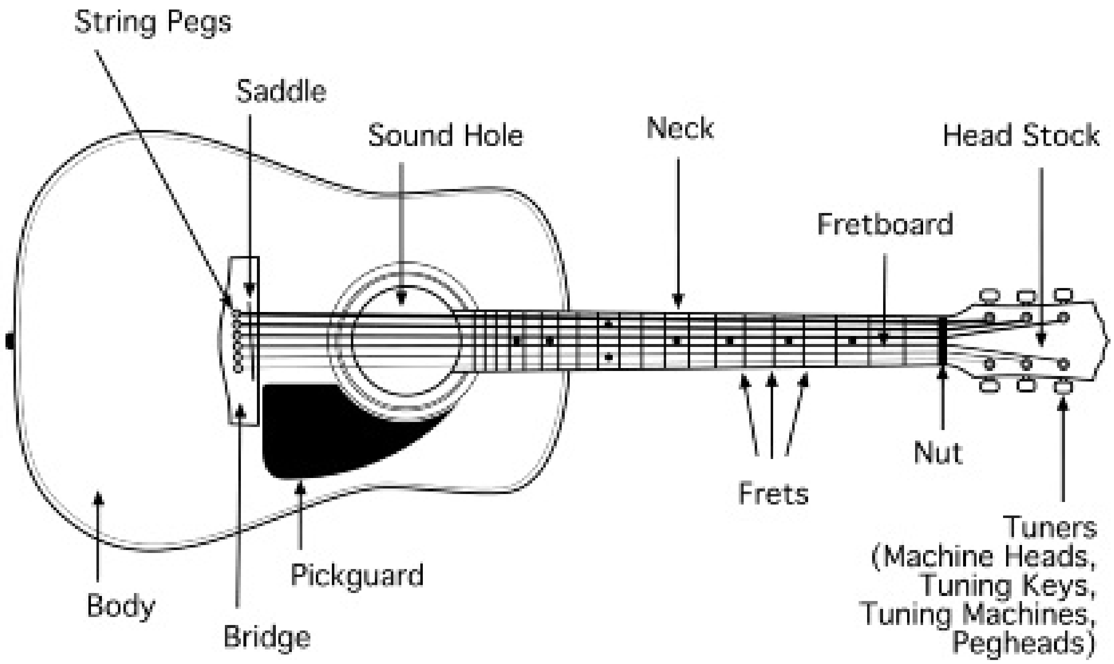


Figure 10: The parts of a classical guitar. (Figure taken from `http://www.betterguitar.com`.)

- The strings are coupled to the body of the guitar at the bridge. The vibrations of the strings are amplified by the body, increasing the loudness of the sound produced.

5.6 Violin

- Although a violin looks superficially similar to a guitar, it differs from a guitar in many ways. (See Figure 11.)
- A violin has only four strings of fixed length and fixed tension. (Similar to a guitar, the tension in the strings can be adjusted by using the tuning keys.)
- Similar to a guitar, the effective length of the violin strings can be shortened by pressing the strings against the neck of the violin. But unlike a guitar, there are no frets on a violin to give fixed, discrete changes in length.
- Thus, a violin doesn't have fixed notes. (This makes it harder for a beginner to learn how to play a violin.)
- The other major difference between the violin and the guitar is that the violin strings are bowed. (A violinist's fingers would quickly damp out the vibrations of a string if it were plucked instead of bowed.)
- By wiggling your finger when you press a string against the neck of the violin, you can produce a note with a varying pitch. This effect is called *vibrato* and is used especially by violinists and sometimes by vocalists.
- The tone quality of a violin note can also be changed by varying the intensity of the bowing.
- This is in contrast to the harp, guitar, or piano, where once a string is plucked or struck, nothing else can be done to change the intensity of the note as the string vibrates.
- The violin strings are coupled to the body of violin at the bridge, just like a guitar. The sound post couples the front and back of the body.

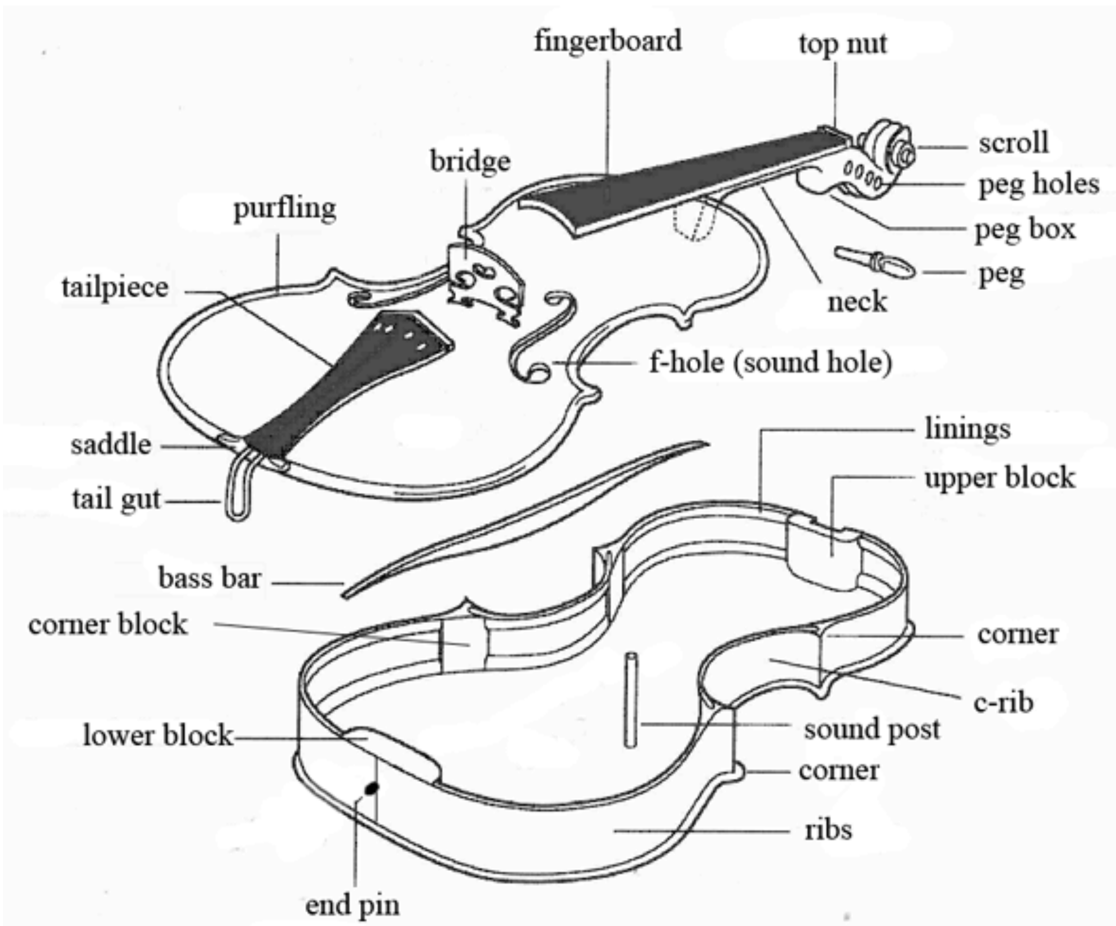


Figure 11: The parts of a violin. (Figure taken from `http://www.violinsonly.info`.)

- The sound hole of a violin (two f-holes) is much smaller than the sound hole of a guitar.
- The complex natural frequencies of the body of the violin and the use of a bow instead of plucking are two main reasons for the rich timbre of notes played on a violin.

6 Wind instruments

- We continue in this section with our discussion of musical instruments.
- Here we will focus on three wind instruments: a flue-organ pipe, a penny whistle, and a clarinet. (We will briefly describe brass instruments at the end of the section.)
- As before, we will follow the presentation of "How music works," by John Powell.

6.1 Sounds from wind in tubes

- Wind instruments come in two basics types: those that have reeds and those that don't.
- Examples of wind instruments that have reeds are the clarinet, oboe, bassoon, ... (Oboes and bassoons actually have double reeds.) Examples of wind instruments that don't have reeds are the flue-organ pipe, penny whistle, flute, recorder, ...
- For both types of wind instruments, the sound is produced by a vibrating air column in a tube.
- The vibrations in the air column are excited by either an *oscillating air stream* directed over an edge (like that from a whistle) or a *vibrating reed.* (For brass wind instruments, a vibrating reed is replaced by the vibrating lips of a brass player.)
- We will discuss these different sources of excitation in the next two subsections.
- For both types of excitation, the oscillation frequency of the air stream or the vibration frequency of the reed (or lips) is determined (via resonance) by the natural frequencies of the air column in the tube.
- Recall that the natural frequencies of an air column in a tube depend on two things, in addition to the speed of sound in air: (i) the length of the tube, and (ii) whether the tube is closed at one or open at both ends.
- Mathematically, the frequency of the nth standing wave vibration of an air column in a cylindrical tube is given by:

  Tube open at both ends:
  $$f_n = n\left(\frac{v}{2L}\right), \qquad n = 1, 2, 3, \cdots \tag{6.1}$$
  Tube open at one end, closed at the other:
  $$f_n = n\left(\frac{v}{4L}\right), \qquad n = 1, 3, 5, \cdots \tag{6.2}$$
  where L is the length of the tube and v is the speed of sound in air ($v = 346$ m/s at room temperature).
- We are assuming that the tubes are sufficiently narrow so that end effects are not important. Otherwise, we would have to increase the effective length of the tube by $0.61R$ for each open end, where R is the radius of the tube.
- Note that for tubes closed at one end, only odd harmonics are present. Tubes open at both ends have both even and odd harmonics.

6.2 Oscillating air streams

- An oscillating air stream is an example of a *flow-controlled* excitation, associated with the displacement of the air molecules.
- Figure 12 shows what happens to a stream of air as it encounters a sharp edge.

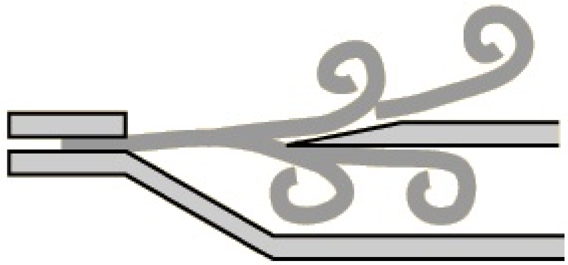


Figure 12: Schematic diagram showing an air stream encountering a sharp edge. The stream splits into vortices that take turns going to one side or the other of the edge. (Figure taken from `hyperphysics.phy-astr.gsu.edu/`.)

- Note that the air stream doesn't divide into two smoothly flowing streams on each side of the edge. Instead, whirlpools (or vortices) are created, which take turns going to one side or the other of the edge. (This is analogous to cars in traffic deciding to go to either one side or the other of a divider in the road, based on the choice of the previous cars.)

- The frequency with which the vortices are formed on opposite sides of the edge is determined by the natural frequencies of the tube to the right of the edge. Resonance occurs when the frequency of vortex formation matches one of the natural frequencies of the air column in the tube.

- At resonance, a vortex enters the tube when the motion of the air molecules in the air column is also directed into the tube.

- Since the vortices are associated with air molecule motion, the location of the sharp edge is an *anti-node* for air molecule displacement (or a node for pressure deviation), and thus corresponds to an *open* end of a tube. If the other end of the tube is open, then Equation 6.1 should be used for the natural frequencies. If the other end of the tube is closed, then Equation 6.2 should be used.

- This is the basic operation of a whistle, or producing a sound by blowing over a bottle.

- Example: For a bottle-shaped object like that in Figure 13, called a *Helmholtz resonator*, there is only one natural frequency given by
$$f = \frac{v}{2\pi}\sqrt{\frac{A}{Vl}} \tag{6.3}$$
where l is the length of the neck, A is its cross-sectional area, V is the volume of the base, and v is the speed of sound in air.

- The air in the neck acts like a mass $m = \rho A l$ attached to a spring with spring constant $k = \rho A^2 v^2/V$, where ρ is the density of the air. (This expression for k follows from Hookes' law $F = -kx$, with $F = -\Delta p\, A$, $\Delta p = B\,\Delta V/V$, $\Delta V = Ax$, and B being the bulk modulus, which is related to the speed of sound in air via $v = \sqrt{B/\rho}$.)

- The body of a guitar or violin with its sound hole acts like a Helmholtz resonator.

- Demonstration: Calculate the resonant frequency of a bottle, approximating the neck length, cross-sectional area of the neck, and volume of the base. Blow over the lip of the bottle to produce a sound, and use iSpectrum to determine its fundamental frequency. How well does the calculated frequency agree with measured frequency?

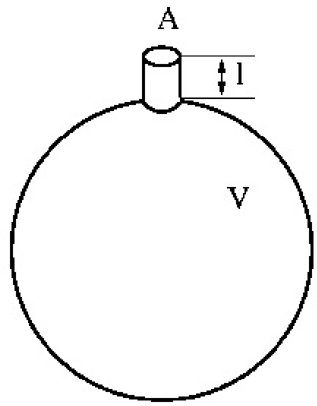


Figure 13: Schematic diagram of a bottle-shaped object, called a Helmholtz resonator. (Figure taken from `http://fisicaondemusica.unimore.it`.)

6.3 Vibrating reeds and tubes

- A vibrating reed (or vibrating lips) is an example of a *pressure-controlled* excitation, associated with pulses of high and low pressure waves.
- Figure 14 shows what happens to a pressure pulse as it propagates in an open tube with a reed at one end.

  (a)-(b): When the reed is open, a positive-pressure pulse enters the tube and propagates to the right.

  (c): At the open end of the tube, the positive-pressure pulse is inverted upon reflection, becoming a negative-pressure pulse. (Recall that an open end of a tube is a node for pressure deviations, or an anti-node for air molecule displacement.)

  (d): The negative-pressure pulse then propagates to the left, reaching the reed-end of the tube, helping to draw the reed shut. No air enters the tube at this time.

  (e): The negative-pressure pulse sees the reed-end of the tube as a closed end, and hence is *not* inverted upon reflection.

  (f)-(g): The negative-pressure pulse propagates to the right and reflects off the open end of the tube.

  (h)-(i): The returning positive-pressure pulse now helps to keep the reed open, allowing an external high-pressure pulse of air to enter the tube.

  The whole process repeats as in (a).
- Summary: At resonance, the reed acts like a pressure-controlled valve, letting in high-pressure pulses of air in phase with the high-pressure vibrations of the air column in the tube.
- Note that the reed is closed for half a cycle, during which time no energy is added to the air column in the tube. This is similar to the slipping stage of a bowed violin string.
- Since a reed acts like a closed end of a tube, the natural frequencies of the vibrating air column consist of only the *odd* harmonics.

6.4 Flue-organ pipe

- A flue-organ pipe is an example of a wind instrument that uses an oscillating air stream as its source of excitation.
- When an organ key is pressed, a constant flow of air is directed toward a sharp edge at one end of the organ pipe, as shown in Figure 15.

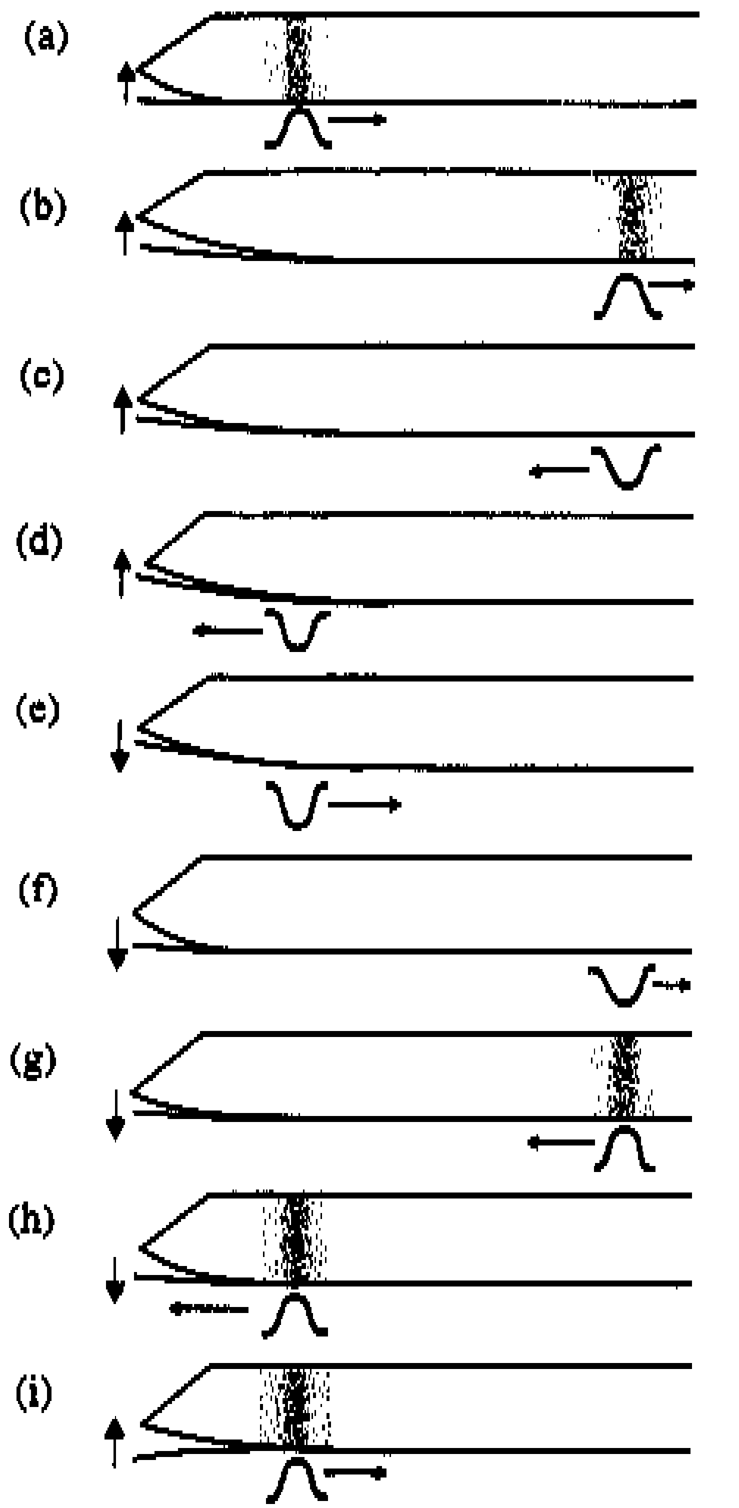


Figure 14: Schematic diagram showing positive and negative-pressure pulses propagating in an open tube with a reed at one end. (Figure taken from “Science of Sound,” by Rossing, Moore, and Wheeler.)

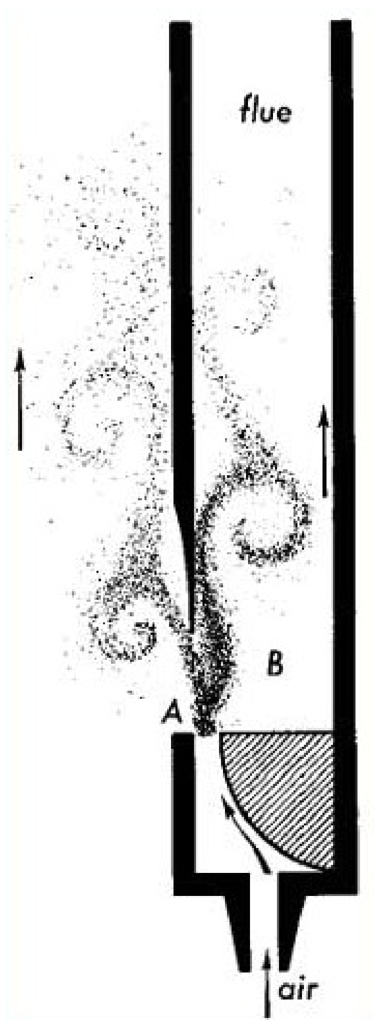


Figure 15: Schematic diagram showing air flow in a flue-organ pipe. (Figure taken from `http://home.comcast.net`.)

- Since the oscillating air stream acts like an open end of a tube, the natural frequencies of the vibrating air column depend on whether the other end of the flue-organ pipe is open or closed.
- If the other end of the pipe is closed, then only odd harmonics are produced.
- There are also reed-organ pipes, where the sharp edge of the pipe is basically replaced by a vibrating reed. (See pipes 4 and 5 in Figure 16.)

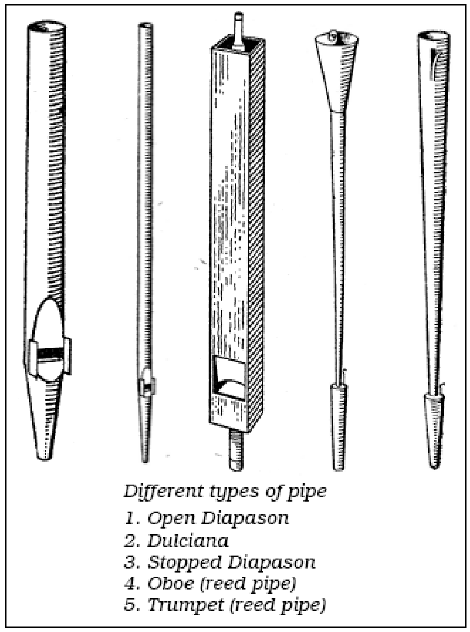


Figure 16: Different types of organ pipes. Flue-organ pipes (1-3) and reed-organ pipes (4, 5). (Figure taken from `http://www.churchmusicdublin.org/`.)

- The timbre of a note produced by an organ pipe is affected by the shape of the bore (e.g., whether it is cylindrical or conical), whether it is open or closed at the other end, and how wide it is. Narrow pipes support higher harmonics better than wide pipes.
- But like a harp, each organ pipe basically produces only one note. To get more than one note per pipe, you need to drill holes in it, like a penny whistle, which we describe next.

6.5 Penny whistle

- A penny whistle is like an organ pipe, but with a series of holes drilled in it.
- These holes are called *tone holes*. They allow you to play different notes, by selectively covering or uncovering one or more holes with your fingers.
- With all the tone holes covered, the effective length of the tube is just the full length of the tube.
- By uncovering a tone hole, you change the effective length of the tube as shown in Figure 18. Note that effective length depends on the *size* of the tone hole. If the radius of the tone hole is the same as the radius of the tube, then the effective length of the tube ends at the center of the tone hole.
- There are 6 tone holes on a penny whistle corresponding to the 7 different notes of a diatonic scale.
- By *partially* covering a tone hole with your finger, you get notes *in between* those defined by the fully-open (or fully-closed) tone holes.

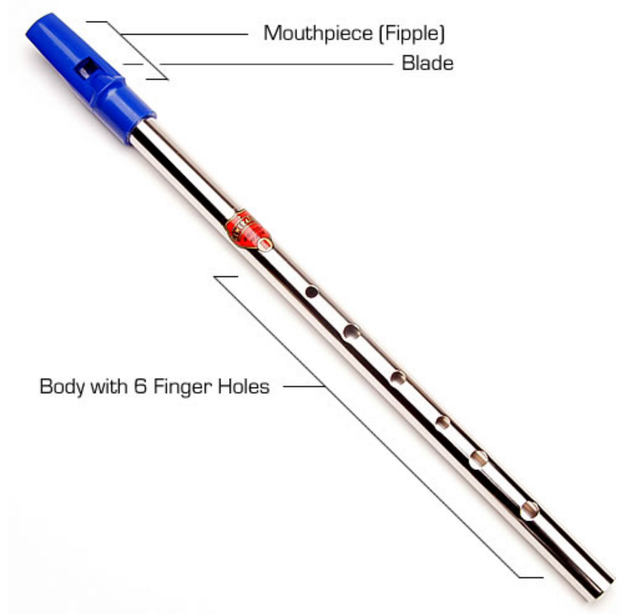


Figure 17: A penny whistle. `http://www.blaynechastain.com/files/webfm/image/course-whistle/`.)

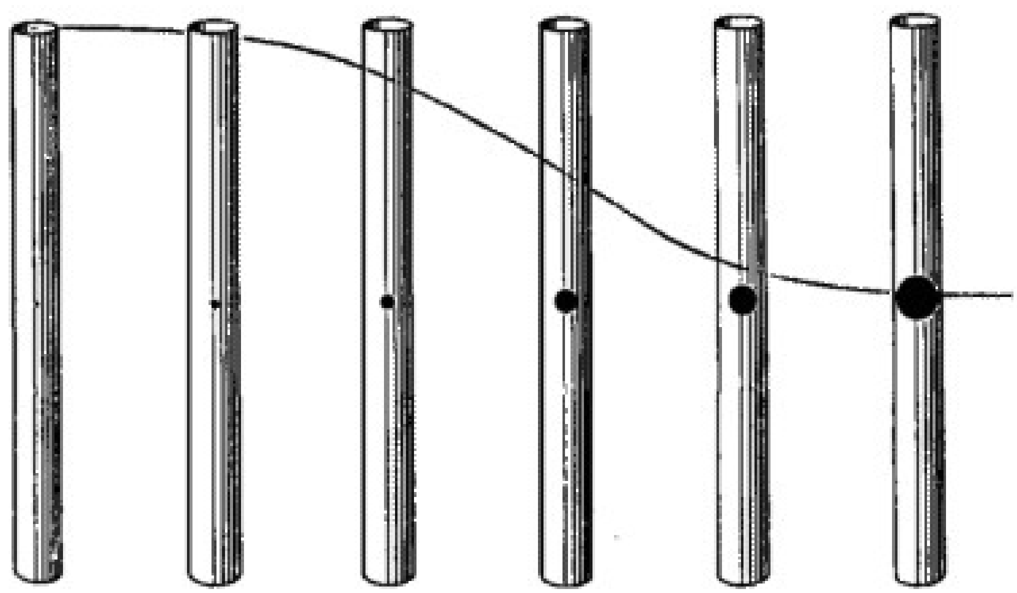


Figure 18: How the effective length of a tube depends on the size of an open tone hole. (Figure taken from `http://www.physics.georgetown.edu/`.)

- By *overblowing* a penny whistle (i.e., blowing with more force than normal), one can excite the second (or even third) harmonic of the vibrating air column, and hence produce notes that are an octave (or a twelfth) above the nominal notes.

6.6 Clarinet

- A clarinet is an example of a reed-pipe system.
- Figure 19 shows the parts of a clarinet and Figure 20 is a schematic drawing of the reed.

Figure 19: Parts of a clarinet. (Figure taken from `http://en.wikibooks.org/wiki/Clarinet`.)

- The vibration frequency of the reed is determined by resonance to match the natural frequencies of vibration of the air column in the tube. (These vibration frequencies are typically much less than the natural frequency of the reed by itself.)
- Since a clarinet has a cylindrical bore and is closed at one end (because of the reed), *odd* harmonics are predominantly produced. (The conical shape of the bell changes the higher harmonics slightly.)
- NOTE: For reed instruments with conical bores, *both* even and odd harmonics are produced.
- Because the reed acts as a closed end for a clarinet, its fundamental frequency is an octave lower than that of a flute, even though they both have the same size. (This is because a flute acts like a tube that is open at *both* ends.)
- Similarly, because of the reed, overblowing a clarinet will produce notes that are a twelfth (three times) above the nominal notes.

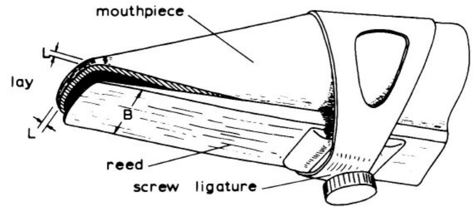


Figure 20: Schematic diagram of a clarinet reed. (Figure taken from `http://www.speech.kth.se/music/publications/leofuks/thesis/`.)

- In addition to tone holes, clarinets have *register keys* that allow you to play notes in a higher register (pitch) than normal.
- By opening the register key located a third of the way from the reed-end of the clarinet, you produce a note that is a musical twelfth (or *three* times the fundamental) higher than with the register key closed, as shown in Figure 21. (One gets the third harmonic, and not the second, since a clarinet produces only the odd harmonics.)

6.7 Brass instruments (in brief)

- Brass instruments are similar to woodwind instruments in that the vibrating lips of a brass player play the same role as a vibrating reed in e.g., a clarinet.
- But since a brass player's lips are *more massive* than a reed on a woodwind instrument, the brass player can more easily play harmonics on his (or her) instrument by simply buzzing his lips at the harmonic frequencies of the vibrating air column in the tube. (No valves or slides are needed to play these harmonic notes.)
- Brass instruments differ from woodwind instruments in that they use valves (for trumpets and tubas) and slides (for trombones) to change the effective length of the air column, instead of using tone holes and register keys.
- Since there are no tone holes for brass instruments, most of the sound is radiated from the *bell* of the instrument.
- Since a trumpet has a cylindrical bore and the lips act like a closed end of a tube, one might think that a trumpet should produce only odd harmonics. But the mouthpiece and the bell change the natural frequencies of the instrument, so that both even and odd harmonics are actually produced.

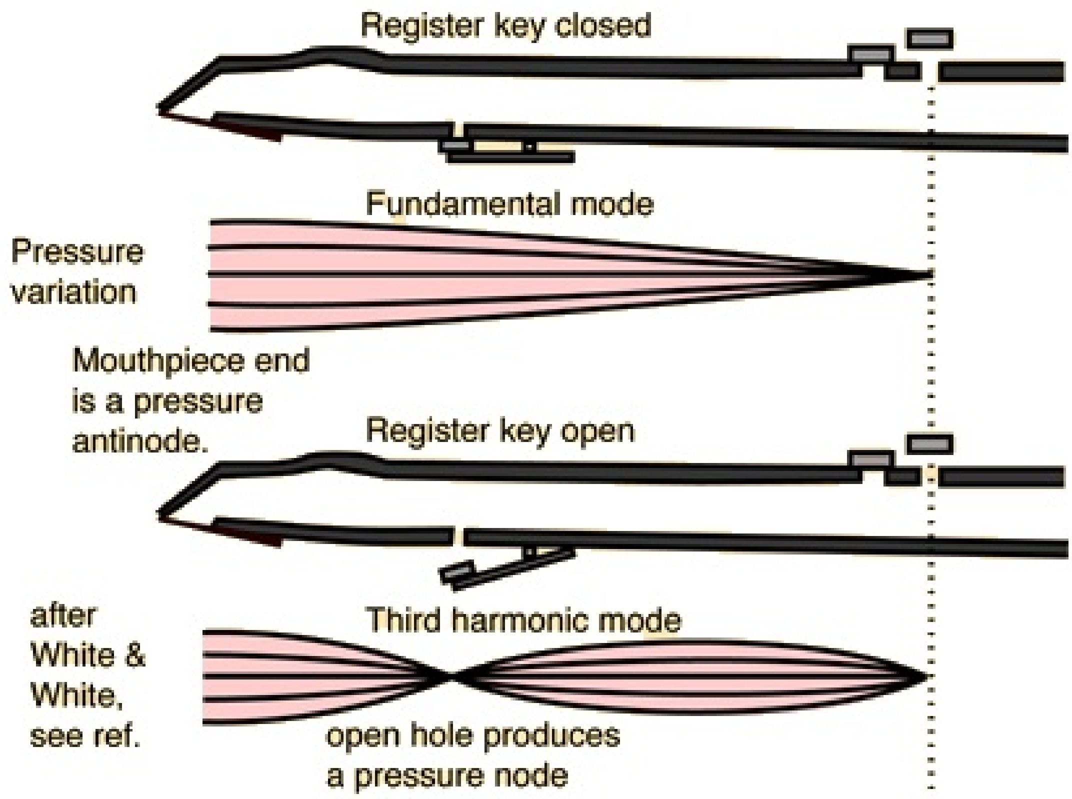


Figure 21: The operation of a clarinet register key. The vertical dotted line shows the location of an open tone hole, which defines the effective length of the clarinet tube. (Figure taken from `http://hyperphysics.phy-astr.gsu.edu`.)

7 Percussion instruments

7.1 Classification of percussion instruments

- Percussion instruments are usually classified in terms of the main vibrating element:

  (i) *bars or rods*, e.g., glockenspiel, xylophone, marimba, vibraphone, chimes, and triangle.

  (ii) *membranes*, e.g., timpani (kettle drum), bass drum, tomtom, snare drum, etc.

  (iii) *plates*, e.g., cymbals, gongs, and bells (either large church bells or smaller hand bells).

- Percussion instruments can also be classified in terms of whether they produce a definite pitch:

  (i) *tuned*, e.g., glockenspiel, xylophone, marimba, vibraphone, chimes, triangle, timpani, gongs, and bells.

  (ii) *untuned*, e.g., bass drum, tomtom, snare drum, and cymbals.

- The piano is another example of a tuned percussion instrument, whose main vibrating elements are struck strings. It will be discussed in its own section later in these notes.

- Percussion instruments are characterized by a vast number of overtones that, in general, are *not harmonically related*. (This is different from string and wind instruments, which produce overtones that are harmonics of a fundamental frequency.)

- Tuned percussion instruments are designed so as to emphasize a particular frequency or are tuned so that the first few overtones are (close to) harmonics of the fundamental vibration mode.

7.2 Vibrating bars and rods

- The transverse vibrations of a bar or rod are similar to the vibrations of a stretched string, with the restoring force provided by the inherent stiffness of the bar and not to any externally-applied tension.

- Figure 22 shows the first three transverse (bending) vibrational modes for a thin bar with free ends. Note that the location of the nodes closest to the ends of the bar move further out as n increases, and that the frequencies f_n of the partials are not harmonics of the fundamental f_1.

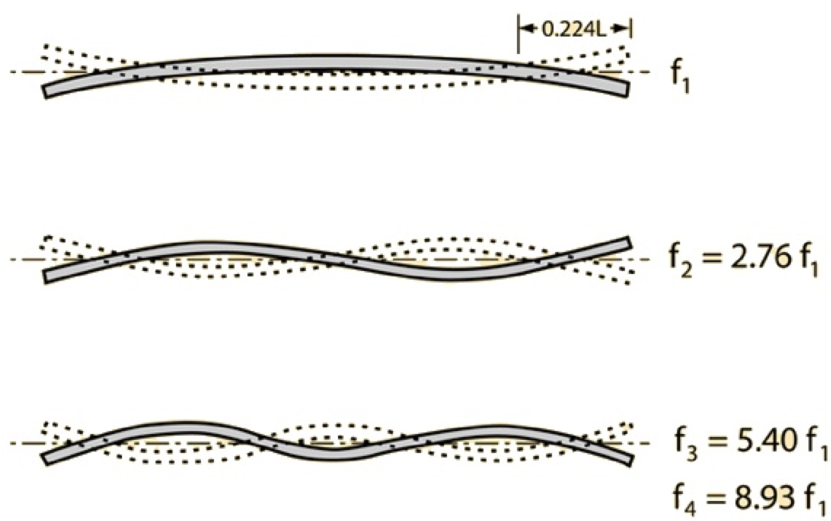


Figure 22: The first three transverse vibrational modes for a thin bar with free ends, and their corresponding frequencies. (Figure taken from `hyperphysics.phy-astr.gsu.edu/`.)

- The frequencies of the different transverse vibrational modes for a thin bar (length L, thickness t) with free ends are given by

$$f_n = 0.1134\, m^2 \frac{vt}{L^2}\,, \quad \text{where} \quad m = 3.011,\ 5,\ 7,\ \cdots,\ (2n+1),\ \cdots \tag{7.1}$$

  and $v = \sqrt{E/\rho}$ is the wave velocity in the bar. (E is called *Young's modulus* and ρ is the mass density (mass/volume) of the bar, both of which depend on the material. E is basically a spring constant for a bar under tension, defined by $F/A = E\Delta l/l$ for small displacements Δl.)

- In terms of the fundamental frequency f_1, the frequencies of the partials are given by

$$f_1\,, \quad f_2 = 2.76\, f_1\,, \quad f_3 = 5.40\, f_1\,, \quad f_4 = 8.93\, f_1\,, \quad \cdots \quad f_n = \left(\frac{2n+1}{3.011}\right)^2 f_1\,, \quad \cdots \tag{7.2}$$

  Note that these frequencies are *not* harmonically related to the fundamental frequency.

- One can also set up *longitudinal* vibrations in a bar or rod, whose modes have a simple harmominc relation:

$$f_n = n\frac{v}{2L}\,, \quad \text{where} \quad n = 1,\ 2,\ \cdots \tag{7.3}$$

  This formula is identical in form to that for standing waves on a stretched string, with $v = \sqrt{\tau/\mu}$ replaced by $v = \sqrt{E/\rho}$.

- Demonstration: Strike a 440 Hz thin bar supported over a resonator box. Where do you think the supports are located?

- Demonstration: Set up longitudinal vibrations in a long thin rod by holding the middle of the rod between your thumb and forefinger, while stroking the rod with your other hand. (You might need to wear a rosin-covered glove to get enough friction to excite the longitudinal vibrations.)

7.3 Glockenspeil, xylophone, marimba, vibraphone

- Glockenspiel bars are rectangular. Since the bars are supported at the nodes of the fundamental vibration mode, the fundamental frequency is the dominant pitch for a glockenspiel bar.

- Xylophone, marimba and vibraphone bars are "scooped out" near the center of the bar, forming an arch. In this narrower section of the bar, the sound waves travel slower, decreasing the frequency of the transverse vibrational modes.

- The arch serves two purposes: (i) since the vibrational frequencies are lower, one can get the lower pitches without needing excessively long bars; (ii) the arch allows one to tune the overtone frequencies to harmonics of the fundamental.

- For xylophones, the first overtone is tuned to the 3rd harmonic of the fundamental. For marimbas and vibraphones, the arch is deeper and tunes the first overtone to the 4th harmonic.

- The xylophone, marimba, and vibraphone also have *resonator tubes* below each bar, whose lengths are chosen to match the fundamental frequencies of the different bars.

- The resonator tubes increase the loudness of the sound at the expense of a shorter decay time for the vibrations of the bars. (The tubes efficiently couple the vibration of the bars to motion of the surrounding air.)

- Since the natural frequencies of the resonator tubes are harmonics of the fundamental frequency of the tube, the resonator tubes select out the fundamental and any *harmonically-related* frequencies associated with the vibrations of the bars.

- The resonator tubes of a vibraphone have motor-driven discs that alternately open and close, producing a vibrato effect associated with the changing intensity of the sound. The speed of the motor can be adjusted in order to produce a slow vibrato or fast vibrato. A vibraphone is sometimes played without vibrato by turning the motor off.

7.4 Chimes and triangles

- Chimes are long vertical pipes (rods) that are typically struck near the top.
- The perceived *strike tone* for a chime is the missing fundamental for modes $n = 4,\ 5,\ 6$. These modes have relative frequencies $9^2 : 11^2 : 13^2$, which are approximately equal to $2 : 3 : 4$.
- A triangle is a thin solid metal rod that is bent in the shape of a triangle. When struck, the tone produced is similar to that of a straight rod of the same length.

7.5 Vibrating membranes

- A vibrating membrane is the 2-dimensional analogue of a (1-dimensional) vibrating string.
- Since a membrane is a 2-dimensional object, the modes of oscillation are labeled by two integers (m, n). For circular membranes, $m = 0, 1, 2, \cdots$ specifies the number of nodal diameters and $n = 0, 1, 2, \cdots$ specifies the number of nodal circles. (If the membrane is fixed at the circumference, n starts at 1.)
- Figures 23 and 24 show several vibrational modes of an ideal circular membrane, fixed at the circumference like a drumhead. Adjacent parts of the membrane move in opposite directions.

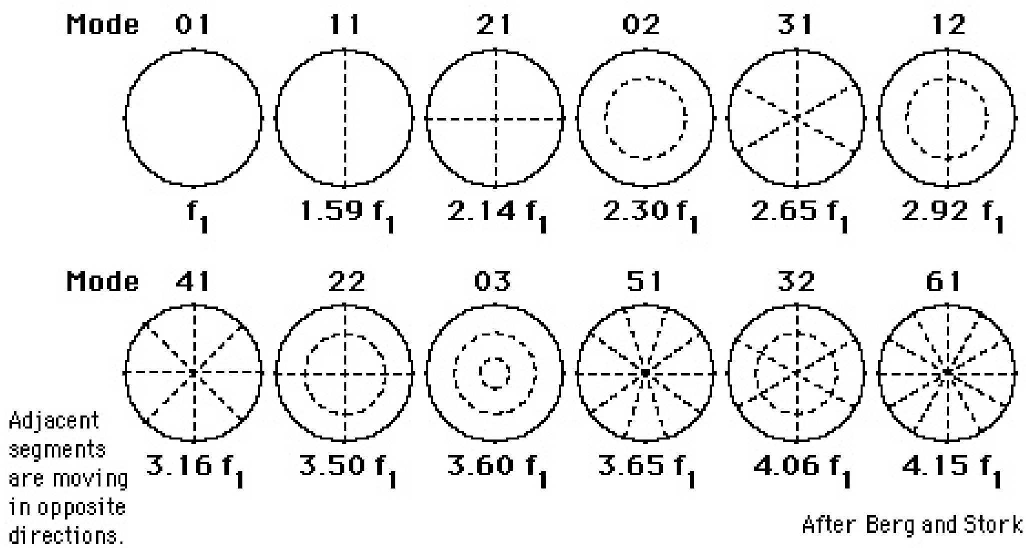


Figure 23: Several vibrational modes of an ideal circular membrane, fixed at the circumference. The modes are labelled by two integers (m, n), where m is the number of nodal diameters and n is the number of nodal circles. Also given are the vibrational frequencies of each mode in terms of the fundamental frequency f_1. (Figure taken from `http://www.mwit.ac.th/ physicslab/`, adapted from "Physics of sound," by Berg and Stork.)

- The frequencies of the different vibrational modes for an ideal circular membrane of radius r fixed at the circumference are given by

$$f_{mn} = \frac{v}{2\pi r}\, x_{mn}\,, \quad m = 0, 1, 2, \cdots \quad n = 1, 2, 3, \cdots \tag{7.4}$$

where $v = \sqrt{\tau/\sigma}$ is the wave velocity on the membrane. (τ is the tension and σ is the 2-dimensional mass density (mass/area) of the membrane.)

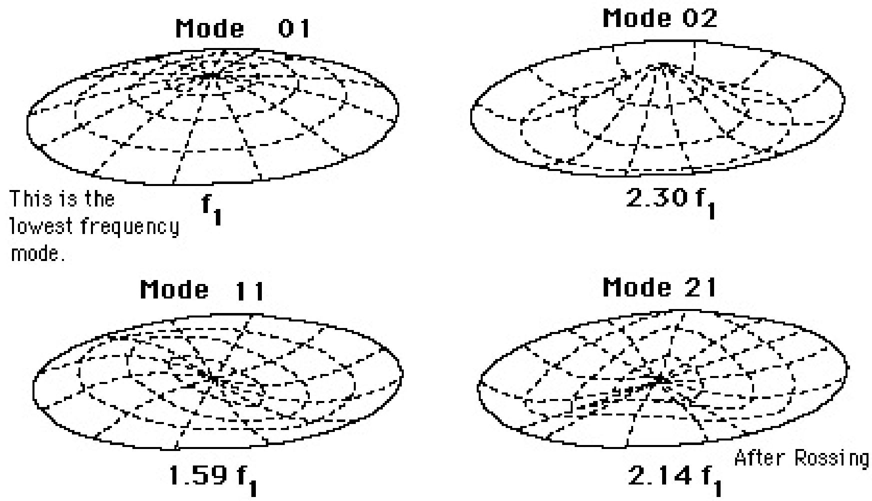


Figure 24: A perspective view of the first four vibrational modes of an ideal circular membrane, fixed at the circumference. See also Figure 23. (Figure taken from `http://www.mwit.ac.th/ physicslab/`.)

- Technical note: x_{mn} denotes the nth zero of a special function called a *Bessel function* $J_m(x)$, which is similar to a damped sinusoid. The values of x_{mn} determine the relative frequencies of the vibrational modes, some of which are given in Figures 23 and 24.

- In terms of the fundamental frequency f_1, the frequencies of the first few partials are given by

$$f_{01} = f_1\,, \quad f_{11} = 1.59 f_1\,, \quad f_{21} = 2.14 f_1\,, \quad f_{02} = 2.30 f_1\,, \quad f_{31} = 2.65 f_1\,, \quad \cdots \tag{7.5}$$

  Note that these frequencies are *not* harmonically related, just as we saw for the transverse vibrational modes of a thin bar or rod.

- Demonstration: `http://en.wikipedia.org/wiki/Vibrations_of_a_circular_membrane` has an animation of the various vibrational modes of a circular membrane:

- Drums are the prime example of percussion instruments that use vibrating membranes as their source of sound.

- The membrane is usually made out of animal skin or some type of synthetic material, like mylar.

7.6 Timpani

- Timpani (or kettle drums) consist of a circular membrane stretched across the top of kettle-shaped enclosure. There are tension screws along the circumference of the drum head and a foot pedal, both of which can be used to change the tension in the drum head.

- A timpani drum head does not behave like an ideal membrane due to: (i) some inherent stiffness associated with the material of the membrane, and (ii) its interaction with the air enclosed in the kettle. The inherent stiffness tends to raise the frequency of the higher vibrational modes, while the interaction with the air tends to lower the frequency of the lower vibrational modes.

- These two features tend to bring the overtones into a near harmonic relationship. The frequencies of the dominant modes (1,1), (2,1), (3,1), (4,1) for a real kettle drum have the approximate ratio 1 : 1.5 : 2 : 2.5 or, equivalently, 2 : 3 : 4 : 5, which is part of a harmonic series. (The (0,1) mode is not relevant, since it dies down quickly because it motion tries to change the total air volume in the kettle.)

- These (nearly harmonic) modes give the timpani its strong sense of pitch.
- NOTE: Some people perceive the missing fundamental ($0.5f_{11}$) as the pitch associated with these modes. Others perceive the pitch to equal just f_{11}. It depends on the strength of the partials f_{21}, f_{31}, and f_{41} relative to f_{11}.

7.7 Bass drum, tomtom, snare drum

- A bass drum is a large two-headed drum.
- The tension in the two heads (the *batter head*, which is struck, and the *carry head*, on the other side) is adjusted differently in order to produce an indefinite pitch.
- The vibrations of the two drum heads are coupled by the enclosed air.
- Tomtoms are either single-headed or two-headed drums of various sizes. Like the bass drum, tomtoms have an indefinite pitch.
- A snare drum is basically a tomtom with *snares* (wires) stretched across the drum head opposite the batter head (called the *snare head*).
- The vibrations of the batter head are coupled to the snare head and the snares via the enclosed air between the two drum heads.

7.8 Vibrating plates

- A vibrating plate is similar to a vibrating membrane, with the restoring force coming from the inherent stiffness of the plate and not from an external tension.
- Figure 25 shows several vibrational modes of a 15-in cymbal, which is behaves like a thin circular plate that is supported at the center and is free at the circumference. The vibrational modes are labelled by integers (m, n), similar to that for circular membranes.
- Cymbals, gongs, and bells are all examples of vibrating plates.
- E.F.F. Chladni (in 1827) noticed that one can illustrate various vibrational modes of a thin plate by bowing the edge of the plate, supported at the center, with fine sand (or salt) sprinkled on top of the plate. As the plate vibrates, the sand accumulates along the nodal lines of the vibrational mode.
- <u>Demonstration</u>: The YouTube video `http://www.youtube.com/watch?v=wMIvAsZvBiw` shows various vibrational patterns for a thin square metal plate. These vibrational patterns are sometimes called *Chladni patterns* after E.F.F. Chladni.
- <u>Demonstration</u>: Repeat the YouTube video demonstration in class for both a square plate and circular plate.

7.9 Cymbals and gongs

- A cymbal is a thin disc of metal with a slight bulge in the center where it is supported.
- The high-energy crash of two cymbals excites many high and low frequency vibrational modes. The sound that is produced has an indefinite pitch.
- After crashing two cymbals together, the musician holds the cymbals with their faces pointing outward toward the audience. This is because the direction of the radiated sound is predominantly *perpendicular* to the face of the cymbals, in the direction of the vibrations of the metal disc.
- Gongs are flat discs of metal with the edges turned over to form a ridge (sort of like a metal trash can cover). They produce sound with a fairly definite pitch.
- Non-linear effects in both cymbals and gongs transfer energy from one vibrational mode to another, making the pitch change with time (usually from low to high).

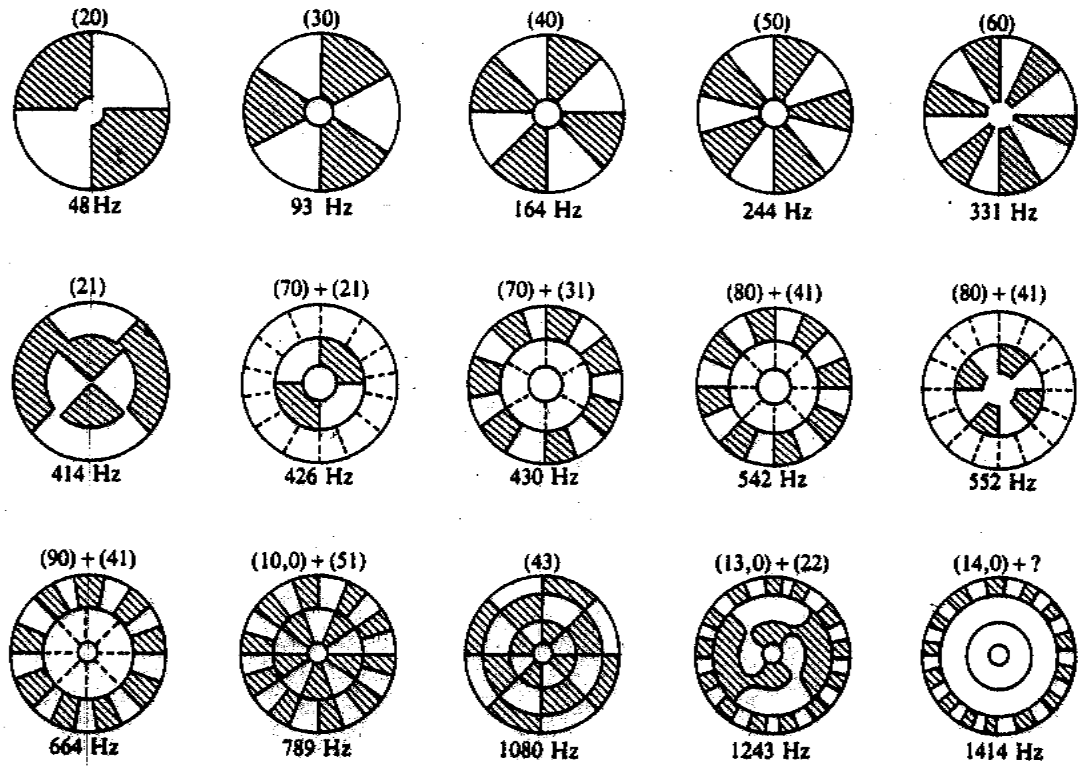


Figure 25: Several vibrational modes of a 15-in cymbal that is supported at the center and is free at the circumference. The first six vibrational patterns are for single modes; the remaining patterns are for combinations of two or more modes. Also given are the frequencies of the various vibrational modes. (Figure taken from "Science of Sound," by Rossing, Moore, and Wheeler.)

7.10 Bells

- Bells are circular plates whose edges have been bent so that several of the overtones are in near harmonic relation with one another.
- Figure 26 show several vibrational modes of a large church bell. The lower circles show the vibrational modes at the rim of the bell. The marks show the location of nodal points on the rim.

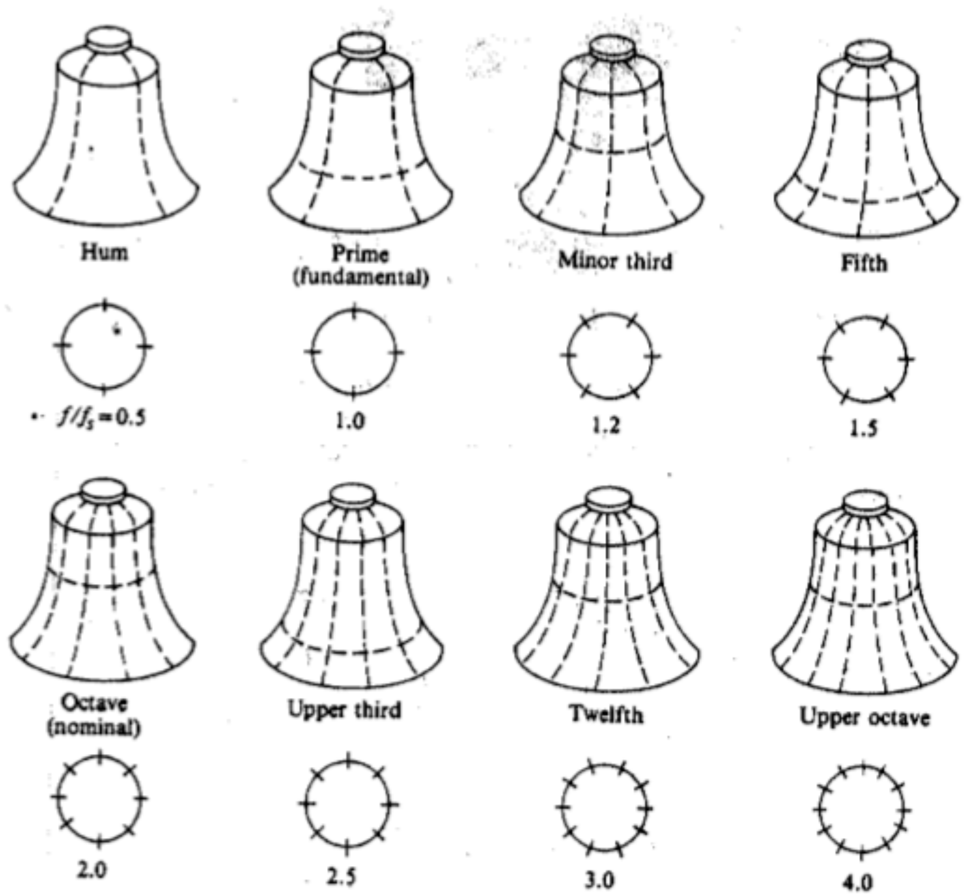


Figure 26: Several vibrational modes of a large church bell. The lower circles show the vibrational modes at the rim of the bell. Also given are the frequencies of the various vibrational modes in terms of the prime (fundamental) mode. (Figure taken from "Science of Sound," by Rossing, Moore, and Wheeler.)

- Similar to chimes, the strike note of a bell is the missing fundamental for the harmonically-related octave, twelfth, and upper octave overtones having frequency ratios 2 : 3 : 4. Although there is an actual vibration mode (called the *prime* mode) whose frequency coincides with the missing fundamental, the perceived pitch is predominantly due to the other frequencies.

8 Piano

8.1 Brief historical overview

- The modern piano is one of the most popular musical instruments today.
- It was invented in 1709 by Bartolomeo Cristofori in Florence.
- It is a marvel of engineering, with a very sophisticated "action", which converts the pressing of a key into the striking of a hammer against a stretched string.
- The force of the hammer strike is controlled by the force with which the key is depressed, giving rise to a large dynamic range of the loudness of the produced sound.
- The original name of the piano, "piano-forte", literally means "soft-loud", corresponding to the large dynamic range of the produced sound.
- Precursors of the piano are the *clavichord* and the *harpsichord*, both of which are keyboard instruments but lack the dynamic range of the modern piano.
- The strings of a clavichord are struck by a small piece of metal (called a *tangent*), which is attached to the end of each key. The tangent stays in contact with the struck string, which vibrates between the location of the tangent and the bridge.
- The strings of a harpsichord are plucked by a a *plectrum* (originally made from a birds quill), which is attached to a vertical piece of wood (called a *jack*) at the end of each key. Since the force of the plucking is *independent* of the force used to depress the key, the loudness of a note produced on a harpsichord is fixed—i.e., it cannot be changed like a piano.

8.2 Construction of a grand piano

- The main parts of a modern grand piano are the keyboard, action, strings, soundboard and frame.
- Figure 27 is an exploded view of a grand piano, and Figure 28 shows a simplified view of a cross-section of the piano.
- There are 88 keys corresponding to the notes A_0 to C_8, tuned to equal-temperament (but see details in a later section).
- There are 243 strings, varying in length from 2 m at the bass end to 5 cm at the treble end.

  - 8 single strings, wrapped (lowest notes)
  - 5 pairs of strings, wrapped
  - 7 triples of strings, wrapped
  - 68 triples of strings, unwrapped (highest notes)
- Wrapping increases the mass density of the string needed for the bass notes, without increasing the *inharmonicity* of the strings too much. (The inharmonicity of the strings is due to the inherent stiffness of the wire used for piano strings. See below for more details.)
- The individual strings are typically subject to tensions of order 1000 N (or about 220 lb). Since there are over 200 piano strings, this means that the piano must support over 40000 lb (or 20 tons) of tension.
- A cast iron frame is needed to support such large forces.
- The soundboard is generally made of spruce, approximately 1 cm thick. The vibrating strings are coupled to the soundboard via the bridge. The soundboard is what produces the large volume of sound produced by a piano.

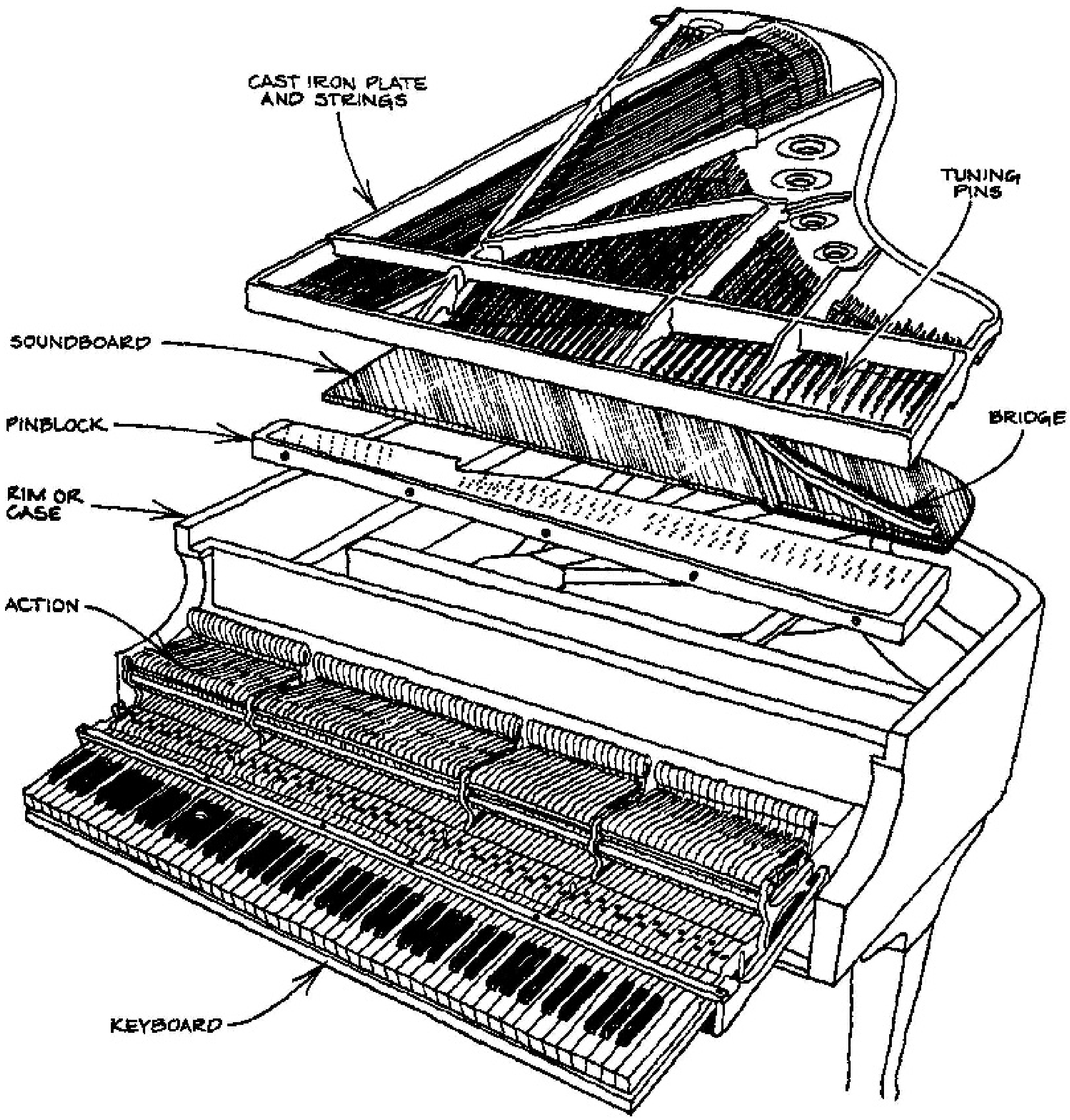


Figure 27: Exploded view of a grand piano. (Figure taken from `http://www.motspheres.com`.)

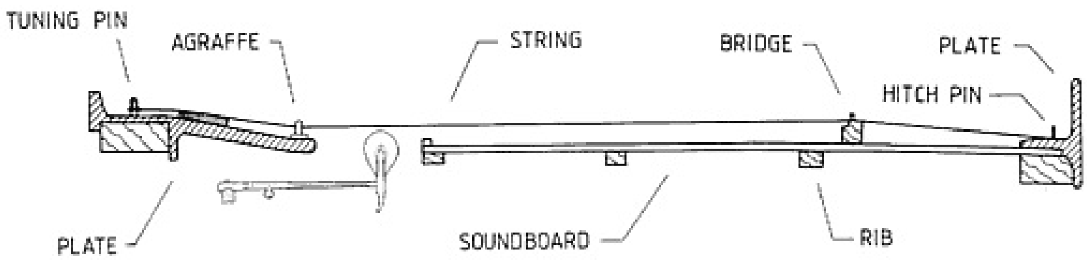


Figure 28: Simplified cross-section of a grand piano, showing a hammer, a string, and its connection to the soundboard and frame via the bridge and pins. (Figure taken from `http://www.speech.kth.se/music/`.)

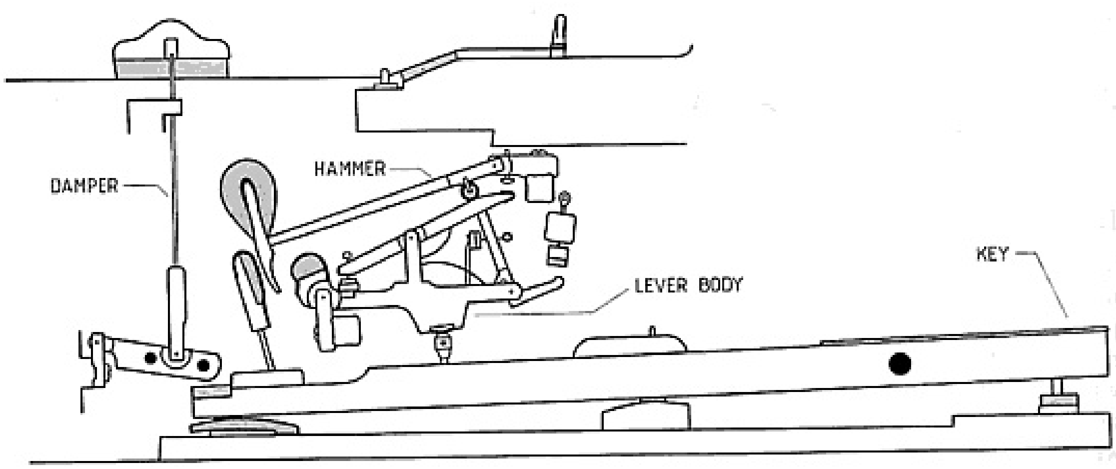


Figure 29: Detailed view of the piano action, which connects a key to its hammer. (Figure taken from `http://www.speech.kth.se/music/`.)

- Figure 29 shows a detailed view of the piano "action", which converts the pressing of a key into the striking of a hammer against a piano string.

- Demonstration: Watch the YouTube video "Grand Piano Action Model" (`https://www.youtube.com/watch?v=95hnnb7KLAg`) to see a piano action in action.

8.3 Attack and decay transients of piano notes

- The notes produced by a piano have a very short attack transient and a long decay transient.

- The decay transient actually has two phases: a period of rapid decay followed by a period of a slower decay, sometimes called the *aftersound*.

- The different decay rates are due to slight mistunings (on purpose) of the three strings per note. The tensions in the neighboring strings are adjusted so that the corresponding frequencies differ by about 1 or 2 cents.

- When a note is played, the three strings initially vibrate in phase and efficiently transfer their energy to the soundboard, each pulling on the bridge in the same direction at the same time. This leads to a loud sound with a rapid decay.

- But as time progresses, the strings begin to vibrate *out of phase* with one another, leading to less efficient transfer of energy to the soundboard, since the vibrations pull on the bridge in opposite directions at the same time. This leads to a softer sound with a slower decay.

- Demonstration: Play the matlab file `pianoC4.mat` both forward and backward using the routine `playrecordedsound.m` to illustrate the difference between the attack and decay transients.

- Demonstration: Do the same with `happybirthday-backwards.mat`.

- A musician can use the *pedals* on a piano to change how the notes sound.

  - Right pedal: sustain pedal; lifts the dampers so that all notes are sustained

  - Left pedal: expression pedal (or *una corda* pedal); shifts the action slightly so that the hammers strike only 2 of the 3 strings

  - Center pedal: (*sostenuto* pedal); sustains only those notes that were depressed *prior* to depressing the pedal, and not the subsequent notes

8.4 Inharmonicity of piano strings

- Due to the inherent stiffness of a real piano wire, the frequencies of the vibrational modes are not exact harmonics of the fundamental frequency.

- There is a bending force intrinsic to the wire, in addition to the applied tension, which tends to increase the frequency of the vibrational modes.

- The frequencies of the vibrational modes are given by

$$f_n = nf_0\left(1 + \frac{1}{2}n^2 B\right) \tag{8.1}$$

  where

$$f_0 = \frac{1}{2L}\sqrt{\frac{\tau}{\mu}}\,, \qquad B = \frac{\pi^3 r^4 E \mu f_0^2}{\tau^2} = \frac{\pi^3 r^4 E}{4L^2\tau} \tag{8.2}$$

- In the above formulae, f_0 is the frequency that an ideal string would have, E is Young's modulus (which specifies the stiffness of the material), τ is the tension in the string, r is its radius, L is its length, and μ is its linear mass-density (mass/length). The formula for f_n assumes that $n^2 B \ll 1$, which is typically the case.

- Note that the fundamental frequency f_1 for a real string is not the same as f_0 for an ideal string, but equals

$$f_1 = f_0\left(1+\frac{1}{2}B\right) \quad \text{which implies} \quad f_n = nf_1\left[1+\frac{1}{2}(n^2-1)B\right] \tag{8.3}$$

  for the vibrational frequencies in terms of the fundamental frequency f_1 for the real string.

- The deviation of these frequencies from pure harmonics is proportional to (n^2-1). For small values of n the difference of f_n from its pure harmonic nf_1 is small, but as n increases the difference can become large.

- Example: For a steel piano string which has $E = 2\times10^{11}$ N/m^2, $\rho = 7840$ kg/m^3, $r = 0.5$ mm, $\tau = 150$ lb, and a length L corresponding to the fundamental frequency $f_0 = 231.63$ Hz (C_4) for an ideal string:

$$\frac{f_2}{2f_1} = 1\ \text{cent}\,,\quad \frac{f_3}{3f_1} = 3\ \text{cent}\,,\quad \frac{f_4}{4f_1} = 5\ \text{cent}\,,\quad \frac{f_8}{8f_1} = 20\ \text{cent}\,,\quad \frac{f_{16}}{16f_1} = 79\ \text{cent} \tag{8.4}$$

  so the deviations for the second, third, and fourth octaves are definitely noticeable.

- The deviation is smaller if one uses *wrapped* strings, which are more flexible than solid strings with the same mass density. This reduces the deviation to about 30 cents for the fourth octave instead of 79 cent, as calculated above for solid steel wire strings.

- As we shall see in the next subsection, pianos are actually tuned to have *stretched* octaves corresponding to these higher frequency ratios, as first systematically studied by O.L. Railsback around 1960.

8.5 Stretched tuning

- A piano is tuned to equal temperament within an octave, and has stretched tuning (due to the inharmonicity of the strings) from one octave to another.

- Stretched tuning means that the higher octaves are tuned sharp and lower octaves tuned flat relative to equal-tempered tuning, as we shall see below.

- The chromatic notes in a octave are equally spaced by a semitone interval, which corresponds to the frequency ratio $2^{1/12} = 1.0596$.

- Step 1: The piano tuner uses a tuning fork (or an electronic tuning device) to set one note in the octave, e.g., C_4, to its correct equal-tempered value. In this case, C_4 should have the absolute frequency 261.63 Hz.

- Step 2: The tuner then "lays the temperament" by tuning all of the chromatic notes in an octave interval (e.g., C_4 to C_5) by working around the circle of fifths: C to G, G to D, etc.

- Since the frequency ratio of an equal-tempered fifth

$$2^{7/12} = 1.4983 \tag{8.5}$$

  is less than that of a perfect fifth ($3/2 = 1.5$), there will be beats between the second harmonic of G_4 and the 3rd harmonic of C_4.

- The frequency difference between an equal-tempered fifth and a perfect fifth is given by

$$\delta = 1.4983f_1 - 1.5f_1 = -0.0017f_1 \tag{8.6}$$

  and the corresponding beat frequency is

$$f_{\text{beat}} = 2|\delta| = 0.0034f_1 \tag{8.7}$$

Interval	ET freq of first note (Hz)	Beat frequency (Hz)
C_4 to G_4	261.63	0.89
G_4 to D_5	392.00	1.33
D_4 to A_4	293.66	1.00
A_4 to E_5	440.00	1.50
E_4 to B_4	329.63	1.12
B_3 to $F_4^\sharp$	246.94	0.84
$F_4^\sharp$ to $C_5^\sharp$	369.99	1.26
$C_4^\sharp$ to $G_4^\sharp$	277.18	0.94
$G_4^\sharp$ to $D_5^\sharp$	415.30	1.41
$D_4^\sharp$ to $A_4^\sharp$	311.12	1.06
$A_4^\sharp$ to F_5	466.16	1.58
F_4 to C_5	349.23	1.19

Table 1: Frequencies and beat frequencies for equal-tempered tuning of the chromatic scale from C_4 to C_5 using the method of fifths. (After a similar table in "Science of Sound," by Rossing, Moore, and Wheeler.)

- Since the beat frequency depends on f_1, it will be different when tuning the equal-tempered fifths G_4 to D_5, D_4 to A_4, etc., than it is for tuning the equal-tempered fifth C_4 to G_4.
- Table 1 lists the beat frequencies for equal-tempered tuning of the chromatic scale from C_4 to C_5 using the above method of fifths.
- Step 3: The piano tuner listens for these beat frequencies, adjusting the tension in the G_4 string, D_5 string, etc. to give the beat frequencies in this table.
- NOTE: Rather than try to listen for e.g., 0.89 beats per second, the tuner listens for 8.9 beats in a 10 second interval. The counting of beats is easier to do over this longer time interval.
- Step 4: Once the chromatic notes in an octave have been set, the piano tuner moves from that octave to other octaves by matching e.g., the second harmonic of G_4 with the fundamental of G_5, etc., and then the fourth harmonic of C_4 with the fundamental of C_6, etc. (This is done by adjusting the tension of the G_5 string to eliminate any beats with the second harmonic of G_4, etc.)
- But since the second harmonic of any note has a frequency that is greater than $2\times$ its fundamental frequency (due to the inharmonicity of the piano strings), the frequency of the new note will be *sharper* than the equal-tempered frequency for that new note.
- And as one tunes the higher octaves the notes become even sharper, since the deviation of the frequencies from true harmonics increases with n as we saw above.
- Using similar reasoning, one can show that for octaves below C_4 (e.g., starting with C_3, C_2, etc.), the notes will be tuned *flatter* than equal-tempered tuning.
- Around 1960 O.L. Railsback measured the tunings of many pianos. Figure 30, called a Railsback curve, shows the deviation of the frequencies of the notes of an actual piano (in green) from that of equal-tempered tuning (which corresponds to a horizontal line at 0 cents). The jagged line is an average of tunings measured from several pianos.

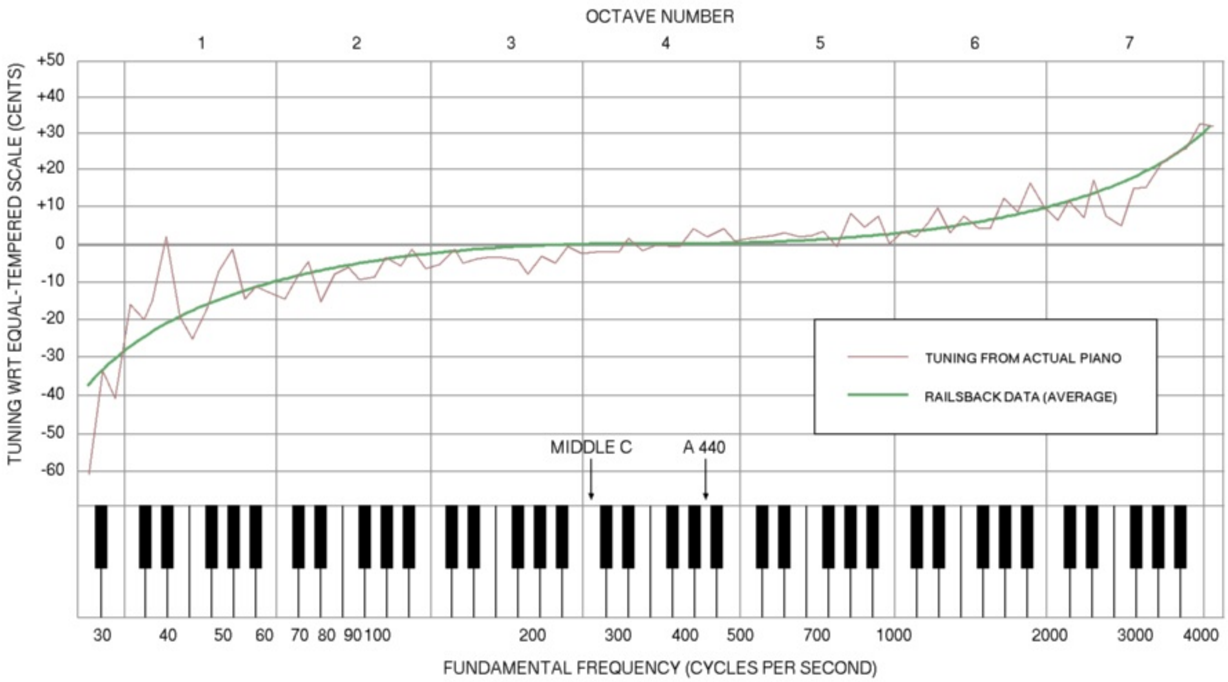


Figure 30: A Railsback curve showing the deviation of the frequencies of the notes of a single piano (and an average for many piano) from that of equal-tempered tuning. (Figure taken from `http://en.wikipedia.org`.)

- This sloping curve, which indicates flatter notes in the lower octaves and sharper notes in the upper octaves, is an illustration of stretched tuning.

9 Voice

9.1 Vocal organs

- The main organs responsible for the human voice are the lungs, larynx, pharynx, mouth, and nose. (See Figures 31 and 32.)

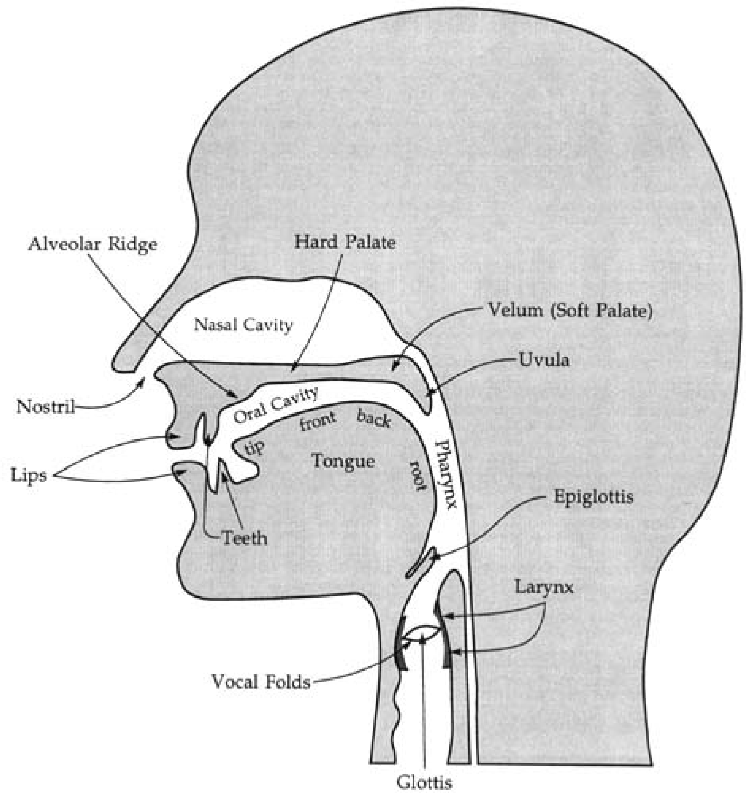


Figure 31: The main vocal organs. (Figure taken from `http://emedia.leeward.hawaii.edu/hurley/Ling102web/`.)

- Schematically, the human voice organs can be thought of as an "instrument" with a *power supply*, *oscillator*, and *resonator*.
- The lungs are the power supply. It is a source of excess air pressure, producing an air stream that passes through the *glottis* (a V-shaped opening between the vocal folds in the larynx).
- The *vocal folds* (sometimes incorrectly called vocal *chords*) are the oscillator for the human voice, playing a role similar to the buzzing lips of a brass player.
- The resonator for the human voice instrument is the *vocal tract*, consisting of the larynx, pharynx, and oral and nasal cavities. It plays a role similar to the tube of trumpet or the body of a violin.

9.2 Vocal folds

- The vocal folds are folds of muscle in the larynx. They are roughly 9-13 mm long for women and about 15-20 mm long for men.
- The frequency of vibration of the vocal folds for voiced sounds is determined primarily by the mass and tension of the folds. For normal speech, the vibration rate may vary by about an octave (a factor of 2 in frequency).

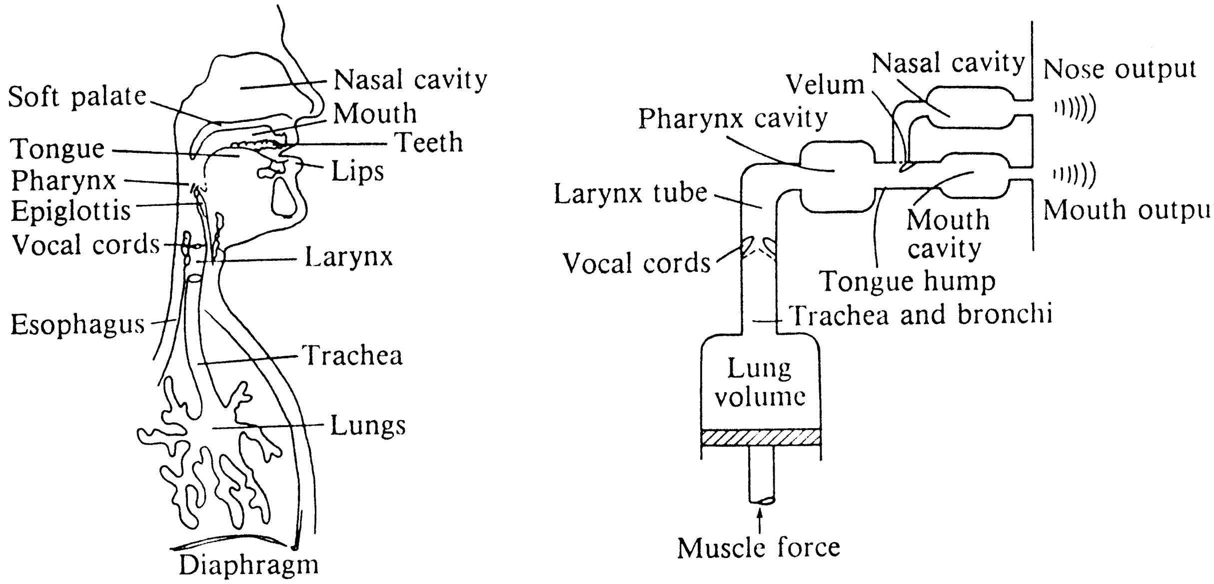


Figure 32: Similar to Figure 31. The right-hand panel is a schematic representation of the key components of the human voice. (Figure taken from "Science of Sound," by Rossing, Moore, and Wheeler.)

- For males, a typical vibration frequency is 110 Hz (corresponding to A_2); for females, 220 Hz; and for children, 300 Hz.
- The forces responsible for the vibration of the vocal folds are a combination of: (i) excess air pressure in the trachea, which tends to push the vocal folds apart, and (ii) the *Bernoulli effect*, which tends to bring the vocal folds together, due to the reduced pressure in the moving air stream. The opening and closing of the vocal folds repeats hundreds of times per second.
- Demonstration: Illustrate the Bernoulli effect by blowing over a horizontal piece of paper. The paper moves upward, since the stream of air above the paper has reduced pressure relative to the still air below it.
- The frequency spectrum of vibrating vocal folds is similar to that of a triangular wave. The components have amplitudes that decrease with harmonic number like $1/n^2$.
- Exercise: Show that this frequency spectrum corresponds to a sound pressure level that decreases by 12 dB per octave.
- Answer: An octave is a factor of 2 in frequency. This leads to a decrease in amplitude by $1/n^2 = 1/2^2 = 1/4$, and a decrease in intensity by $(1/4)^2 = 1/16$. Since each factor of 1/2 in intensity corresponds to a 3 dB decrease in sound pressure level, a factor of $1/16 = (1/2)^4$ in intensity corresponds to a 4×3 dB = 12 dB decrease in sound pressure level.
- It is also possible to produce sounds that are *unvoiced*—e.g., that do not require the vocal folds to vibrate.

  (i) For consonants like 's', 'sh', and 'f', the vocal folds are completely open. The sound is produced by turbulent flow of air through a constriction in the vocal tract, e.g., through clenched teeth or through a small opening between the teeth and the lips.

  (ii) For the consonant 'h', the vocal folds are completely closed, but then suddenly open letting a burst air pass through the glottis.

9.3 Formants

- The radiated sound from the human voice depends on both the spectrum of the vocal fold vibrations and the resonant (or natural) frequencies of the vocal tract.
- The vocal tract acts as a *filter*, selecting only those frequencies of the vocal fold spectrum that coincide with the resonant peaks. This is similar to how the body of a violin resonates with only certain frequencies of the violin strings.
- Figure 33 shows: (a) the spectrum for the vocal fold vibrations, (b) the resonant frequencies of the vocal tract for a typical vowel sound, and (c) the resultant spectrum of the radiated sound.

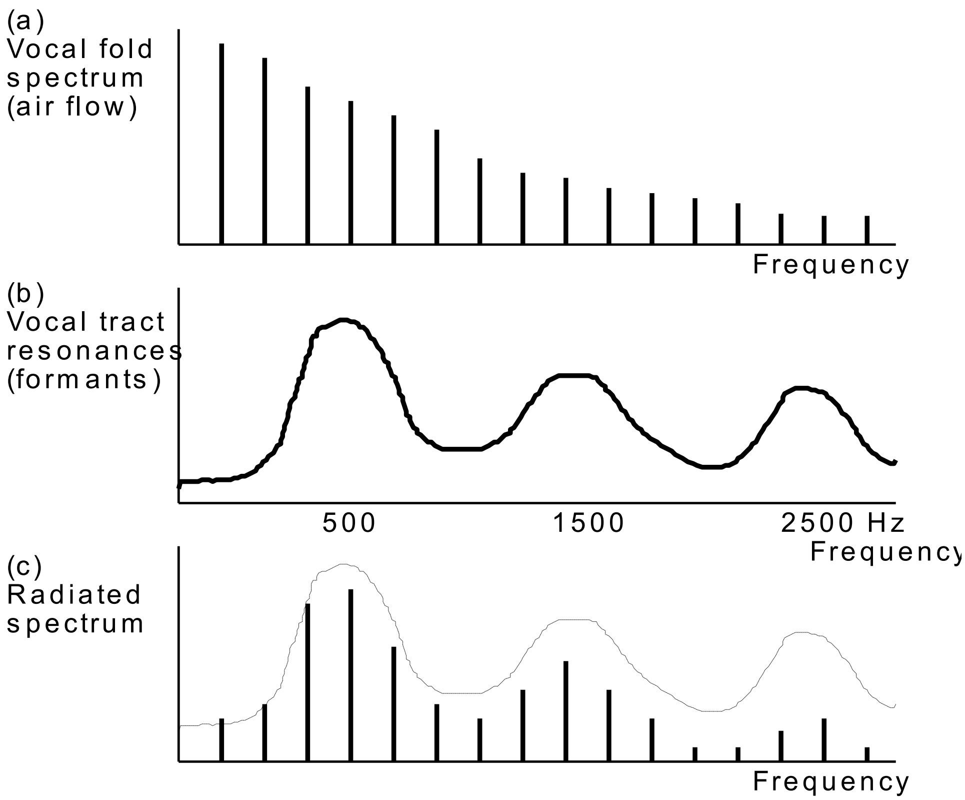


Figure 33: (a) Spectrum for vocal fold vibrations, (b) resonant frequencies for the vocal tract, and (c) resultant spectrum of the radiated sound. (Figure taken from "MU1217 Lecture Notes," Cardiff University by Dr. Bernard Richardson.)

- The resonant peaks for the vocal tract are called *formants*, or *formant regions*.
- The peak frequencies and shapes of the formant regions depend on the geometry of the vocal tract:

  (i) By opening the jaw wider, one can increase the peak frequency of the first formant.

  (ii) By changing the position of the body of the tongue, one can shift the peak frequency of the second formant.

  (iii) By changing the position of the tip of the tongue, one can shift the peak frequency of the third formant.

  (iv) By lowering the larynx, one can produce an extra formant (called the *singer's formant*) between the third and fourth formants. (This will be discussed further below.)

- Demonstration: By breathing in helium, one sounds like Donald Duck.
- The explanation for this change in pitch and tone quality is a shift of the formant regions to higher frequencies, since the speed of sound in helium is about 3 times greater than the speed of sound in air. So even though the vocal folds are oscillating at the same frequencies as they were in air, the radiated sound has higher frequency components since the resonant peaks of the formant regions are about 1.5 times higher than before. (Note: The factor is 1.5 and not 3 since the air in the vocal tract is replaced by a *mixture* of air and helium.)
- If one breathes in sulfur hexafluoride instead of helium, the effect is opposite, since the speed of sound in sulfur hexafluoride is less than that in air. (Sulfur hexafluoride is denser than air.)
- NOTE: It is *very* dangerous to breathe in sulfur hexafluoride, since—being more dense than air—it is harder to flush out of your lungs after you breathe it in.

9.4 Models of the vocal tract

- The simplest model of the vocal tract is to treat it as a cylindrical tube, approximately 17 cm long, which is closed at one end (at the location of the vocal folds).
- Exercise: Show that the first 4 natural frequencies for such a tube, closed at one end, are approximately 500 Hz, 1500 Hz, 2500 Hz, and 3500 Hz.
- Demonstration: Verify the above experimentally using a tuning fork having a frequency of approximately 500 Hz and an adjustable length tube that is closed at one end.
- More sophisticated models of the vocal tract consist of two cylindrical tubes with different cross-sectional areas and different lengths, which add up to 17 cm. For modeling consonant sounds, the two cylindrical tubes are separated by a short and very narrow constriction.
- Figure 34, based on data taken from X-ray photographs, shows the shape of the vocal tract for different vowel sounds and the corresponding formant regions.
- Note the shape of the vocal tract for the vowel sounds 'hod', 'heed', and 'who'd' in Figure 34. The corresponding models for these sounds are: (i) a small cross-sectional area tube (near the vocal folds) followed by a large cross-sectional area tube; (ii) a large cross-sectional area tube followed by a small cross-sectional area tube; (iii) two nearly-equal cross-sectional area tubes separated by a short, narrow tube in between.
- Demonstration: Use the matlab routine `sound_spectrogram.m` to record and calculate sound spectrograms for different vowel sounds. Verify, for example, that 'who'd' and 'heed' have formant regions as shown in Figure 34.

9.5 Speech

- The basic unit of speech is called a *phoneme*. Phonemes are either *vowel* sounds or *consonant* sounds.
- All vowels are *voiced*—i.e., the sound involves vibrations of the vocal folds.
- Different vowel sounds are distinguished by different peak frequencies and shapes for the formant regions as we saw in Figure 34.
- Consonants are either voiced or unvoiced. For example, 'z' is voiced while 's' is unvoiced.
- Demonstration: Hold your hand against your adam's apple and make the sounds for 'z' and 's'. You should be able to feel a vibration for 'z' but not for 's'.
- Table 2 is a list of vowel sounds for American Engish (both pure vowels and dipthongs), with representative words.

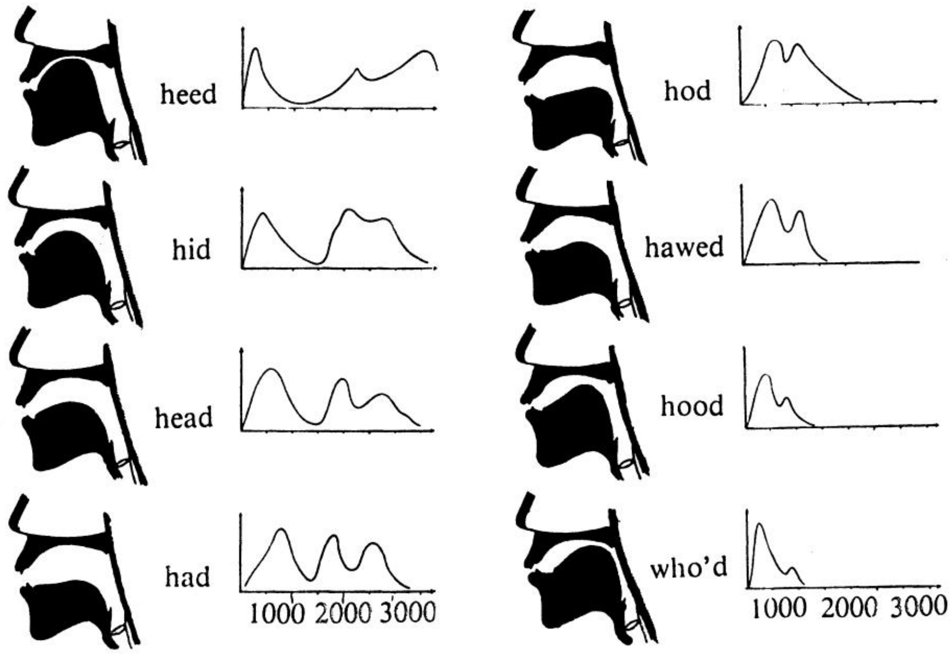


Figure 34: Images based on X-ray photographs, showing the shape of the vocal tract for different vowel sounds. To the right of each image is a graph of the corresponding formant regions. (Figure taken from "Science of Sound," by Rossing, Moore, and Wheeler.)

Pure vowels		Diphthongs	
ee	heat	ou	tone
i	hit	ei	take
e	head	ai	might
ae	had	au	shout
uh	the	oi	toil
ah	father	ju	fuse
aw	call		
u̇	put		
oo	cool		
ŭ	ton		
er	bird		

Table 2: A list of pure vowels and diphthongs for American English. (Based on a similar table in "Science of Sound," by Rossing, Moore, and Wheeler.)

	Plosive		Fricative		Nasal	Semivowel	Liquids
Place of articulation	Unvoiced	Voiced	Unvoiced	Voiced			
Lips	p	b			m	w	
Lips and teeth			f	v			
Teeth			th (thin)	th (then)			
Gums	t	d	s	z	n	y	l, r
Palate			sh	zh			
Soft palate	k	g			ng		
Glottis			h				

Table 3: Classification of consonants in American English. (Based on a similar table in "Science of Sound," by Rossing, Moore, and Wheeler.)

- Table 3 is a classification of consonants for American English, indicating the type of consonant (e.g., *plosive* or *fricative*), whether it is voiced or unvoiced, and where the sound is articulated.
- Demonstration: Use the matlab routine `sound_spectrogram.m` to record and calculate sound spectrograms for different words like "science", and sentences like "I can see you" and "This is a sound spectrogram". Note in particular the high-frequencies associated with the 's' sounds in these examples.
- Demonstration: Illustrate the effect of filtering by using the matlab routine `filter_recorder.m`. Apply low-pass, high-pass, and band-pass filters to a recorded sentence like those in the previous demo, and see what happens to the filtered output.
- Demonstration: Repeat the above for a 'chirp'—i.e., a whistle with increasing frequency.

9.6 Singing

- Although sung vowels and spoken vowels are fundamentally the same, singers will modify the vowels somewhat in order to improve the musical tone of the sounds.
- The main differences between singing and speaking are, in singing:

  (i) the larynx is lowered

  (ii) the jaw opening is larger

  (iii) the tip of the tongue is advanced for back vowels—e.g., oo, ah, ...

  (iv) the lips are protruded for front vowels—e.g., ee, i, ...
- Lowering the larynx leads to 'darker' vowels sounds (i.e., fewer high harmonics).
- Lowering the larynx and opening the pharynx also leads to the formation of an *additional* formant region that peaks between 2500 and 3000 Hz.
- This additional formant is called the *singer's formant*. It allows an opera singer (usually a male) to project above the sound of the musical instruments. (See Figure 35.)
- By opening his/her jaw wider, a singer can reduce the effective length of his/her vocal tract. This increases the peak frequency of the first formant region.
- Sopranos use this technique to 'tune' the first formant region to match the fundamental frequencies of high-pitched notes. (See Figure 36.)
- It is not clear if techniques like diaphragm breathing are important for good singing, provided the subglottal pressure at the larynx is sufficient to produce loud sounds. Formant tuning and being able to produce a singer's formant are probably more important for successful singing.

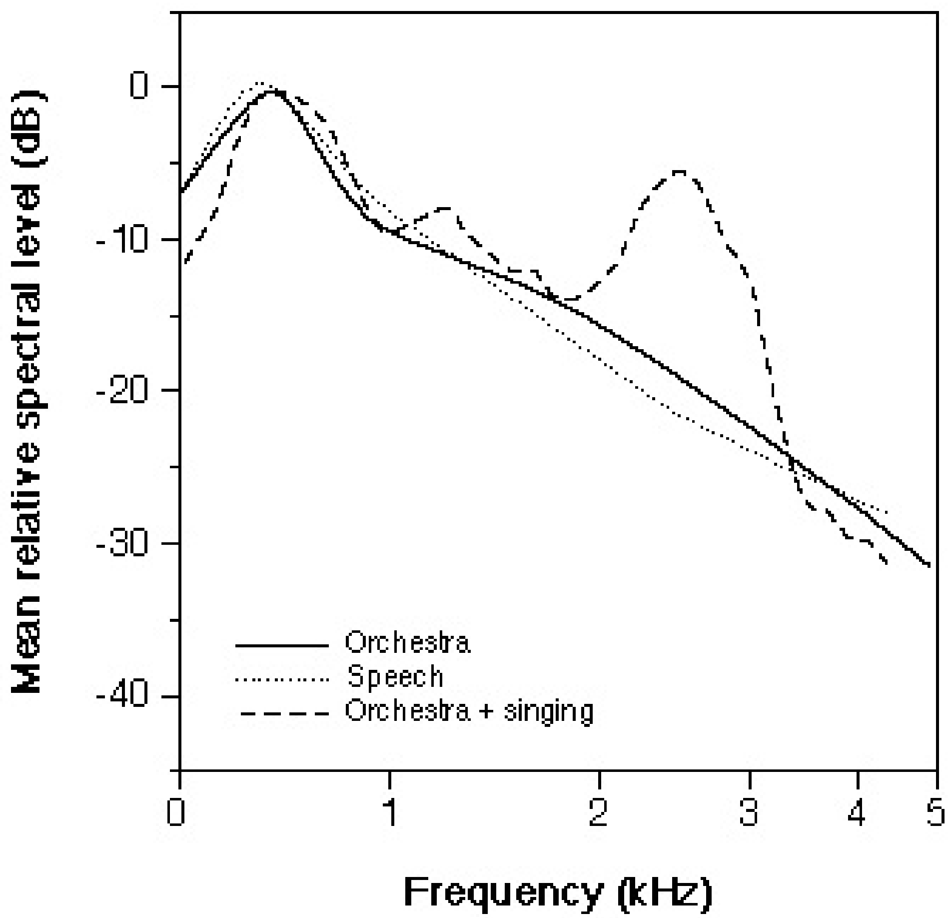


Figure 35: Sound pressure level for sound produced by an orchestra, speech, and an orchestra plus singing. The peak between 2 and 3 kHz for an orchestra plus singing is due to the presence of a singer's formant. (Figure taken from `http://www.ncvs.org/ncvs/tutorials/voiceprod/`.)

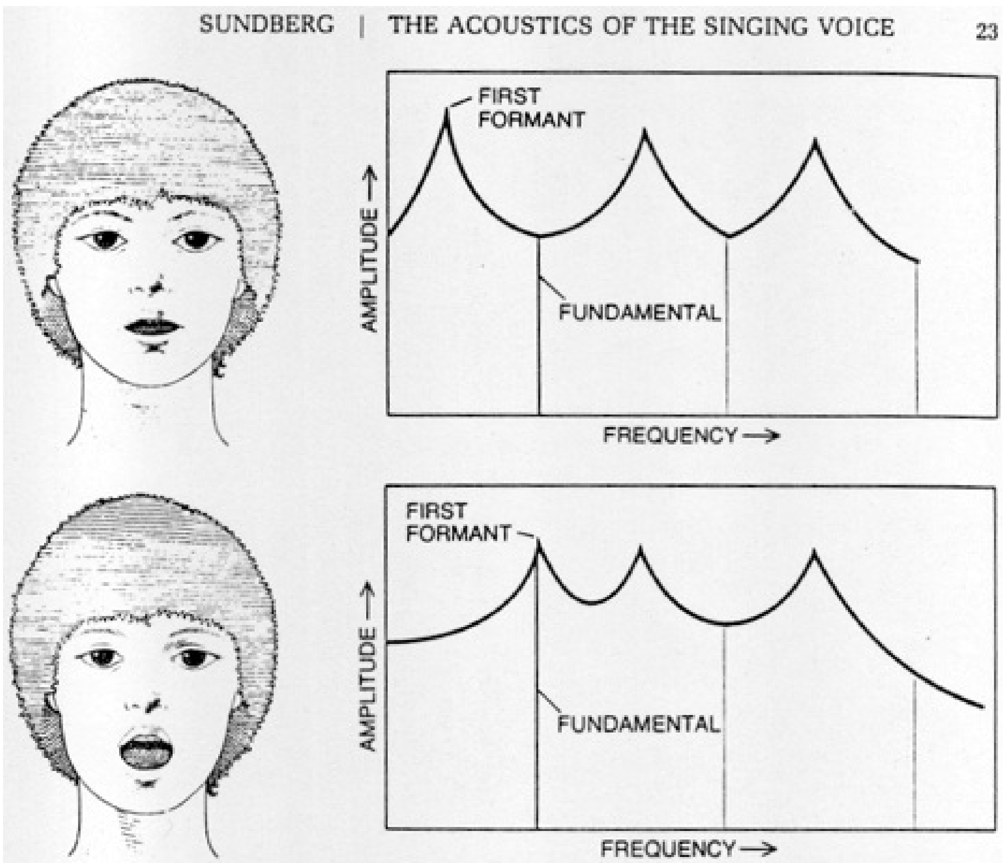


Figure 36: Illustration of the shift in the peak frequency of the first formant region by having a wider jaw opening. The peak frequency of the first formant is tuned to have the same frequency as the fundamental of the sung note. (Figure taken from "The Acoustics of the Singing Voice," by Johan Sundberg.)

Part IV
Perception of sound

10 Hearing

10.1 Range of human hearing

- The frequency range of human hearing is roughly 20 Hz to 20 kHz. This is a factor of 1000 or roughly 10 octaves ($2^{10} = 1024$) in frequency.
- The pressure variation range of human hearing at 1000 Hz is roughly 2×10^{-5} Pa (threshold of hearing) to 20 Pa (threshold of pain). This is a factor of 10^6 in pressure variation or 10^{12} in intensity (Watt/m^2).
- Compared to standard atmospheric pressure $P_{\rm atm} = 100$ kPa, these values are only 2×10^{-10} and 2×10^{-4} of atmospheric pressure, respectively. So even at the threshold of pain, the pressure change is only a few parts in 10^4 of atmospheric pressure!
- In contrast, the frequency range for human vision is only a factor of 2 (so 1 octave); and the intensity range of human vision is only a factor of 10^5 (versus 10^{12} for human hearing). Thus, the human ear is a much more sensitive device than the human eye!

10.2 Fechner's law

- G. T. Fechner in "Elements of psychophysics" ($\sim$ 1860): "As stimuli are increased by multiplication, sensation increases by addition."
- In later sections we'll see that Fechner's law applies *approximately* to our perceptions of both the pitch and loudness of a sound. (Fechner's law applies to other sensations as well, such as sight and smell.)
- Mathematically, Fechner's law means that sensation is proportional to the *logarithm* of a stimulus.
- Recall the defintion of the logarithm:

$$y = \log x \quad \Leftrightarrow \quad x = 10^y \tag{10.1}$$

- Some useful values to remember:

$$\log 2 \approx 0.3\,, \quad \log 3 \approx 0.5\,, \quad \log 4 \approx 0.6\,, \quad \log 5 \approx 0.7\,, \quad \log 10 = 1 \tag{10.2}$$

- Logarithms allow us to compress a large dynamic range of a stimulus to a much smaller (i.e., manageable) range. This was probably advantageous to us from an evolutionary perspective.

10.3 The human ear

- The human auditory system consists of two parts: (i) the peripheral auditory system made up of the ear, and (ii) the central auditory system made of the brain and the auditory nervous system.
- The human ear can be divided functionally into three basic parts (the outer, middle, and inner ear) as shown in Figure 37.
- The outer ear consists of the pinna and auditory canal, which collects the sound (i.e., pressure waves in the air) and funnels it to the middle ear.
- The middle ear consists of the ear drum and ossicles (three small bones called the hammer, anvil, and stirrup, based on their shapes). The pressure wave at the ear drum is amplified ($\sim 30\times$) by the ossicles as it is transmitted to the oval window.
- The inner ear consists of the semi-circular canals, which are responsible for balance, and the cochlea, which is responsible for converting the mechanical sound waves to electrical impulses that get sent to the brain via the auditory nerve.
- A schematic representation of the ear is shown in Figure 38.

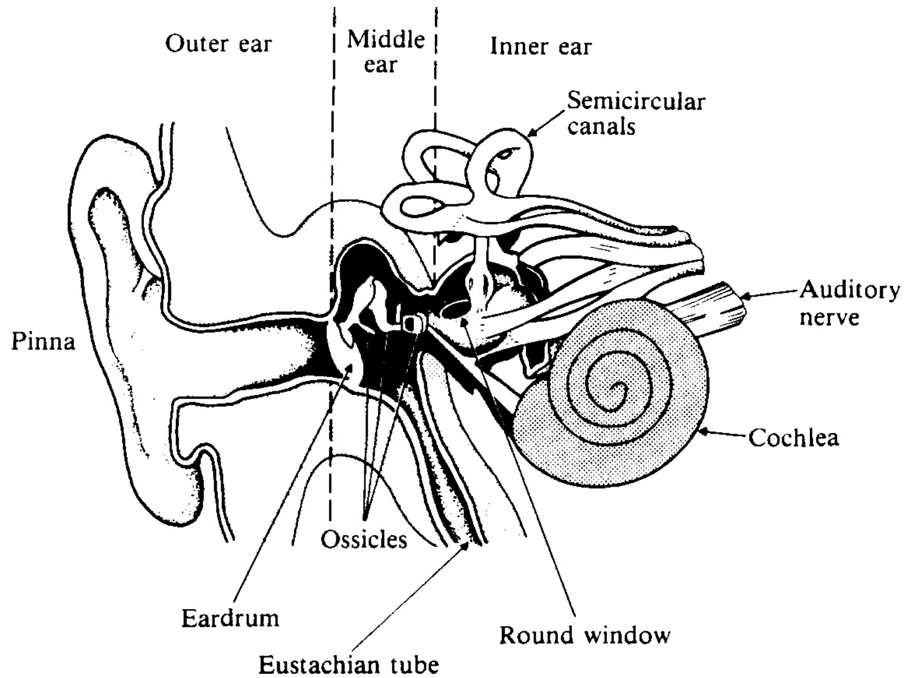


Figure 37: A drawing of the human ear, showing the division into the outer, middle, and inner ear. Note that this drawing is not to scale. (From "Science of Sound" by Rossing, Moore, and Wheeler.)

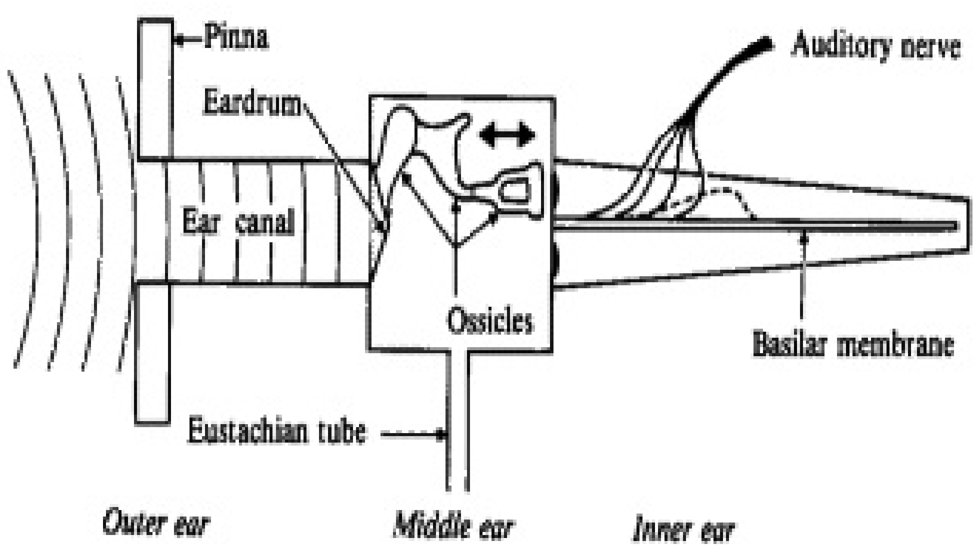


Figure 38: A schematic representation of the human ear, with an uncoiled cochlea. (From "Science of Sound" by Rossing, Moore, and Wheeler.)

- In Figure 38, the cochlea is unwound. The length of an unwound cochlea is approximately 3.5 cm (so a little less than 1.5 inches). The basilar membrane divides the cochlear tube into two sections called the *scala vestibuli* at the top of the diagram, and the *scala tympani* at the bottom. The ossicle bones are, from left to right, the hammer, anvil, and stirrup.
- Figure 39 is a schematic representation of an unwound cochlea.

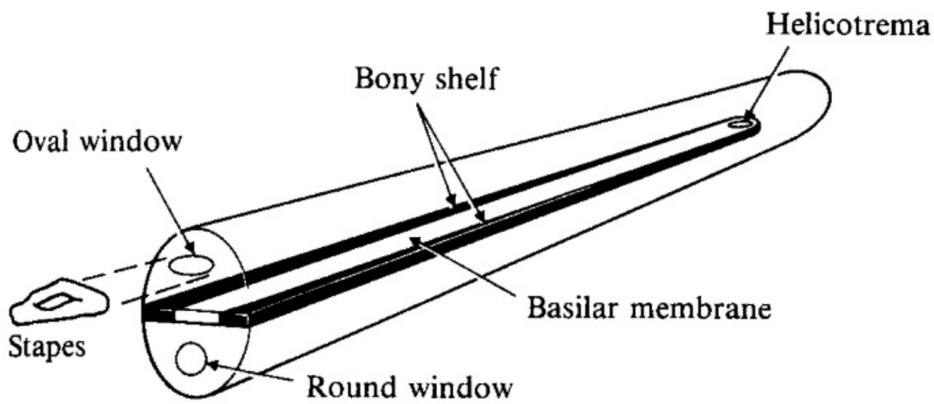


Figure 39: A schematic representation of an unwound cochlea. The region above the basilar membrane is called the *scala vestibuli*; the region below the basilar membrane is called the *scala tympani*. The helicotrema is a hole in the basilar membrane connecting the two sections of the cochlea. (From "Science of Sound" by Rossing, Moore, and Wheeler.)

- The organ of Corti rests on the basilar membrane. It contains rows of tiny hair cells that are connected to the auditory nerve fibers.
- Each hair cells has many hairs, called *stereocilia*, which bend in response to motions of the basilar membrane. The bending of the stereocilia stimulate the hair cells, which trigger the auditory nerve fibers to send electrical impulses to the brain.
- When the stapes (stirrups) press against the oval window, a pressure wave is set up in the cochlear fluid, causing the basilar membrane to have ripples as shown in Figure 40.

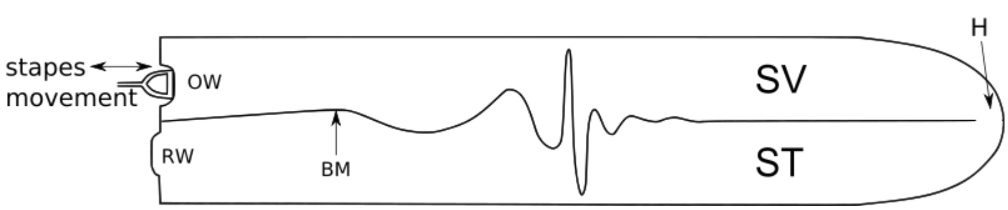


Figure 40: A schematic representation showing the effects of a pressure wave propagating through the cochlear fluid. OW stands for oval window; RW for round window; SV for scala vestibuli; ST for scala tympani; and H for helicotrema. The location of maximum displacement of the basilar membrane (BM) depends on the frequency of the sound. (From `commons.wikimedia.org`.)

- The location of maximum displacement of the basilar membrane depends on the frequency of the incident sound wave: *High-frequency* sound waves produce maximum displacements of the basilar membrane *closer* to the stirrups; low-frequency sound waves produce maximum displacements of the basilar membrane *further* from the stirrups.

- A pure tone (i.e., a single frequency sound wave) excites a relatively wide region (1.3 mm, having approximately 1300 neurons) of the basilar membrane. This region is called a *critical band*.
- There are 24 critical bands on the basilar membrane that span the audible frequency range from 20 to $\approx$ 20 kHz.
- The center frequencies of the critical bands are spaced logarithmically along the basilar membrane, similar to the frequencies of the keys on a piano keyboard. In other words, equal distances along the basilar membrane correspond to equal ratios of the central frequencies as shown in Figure 41.

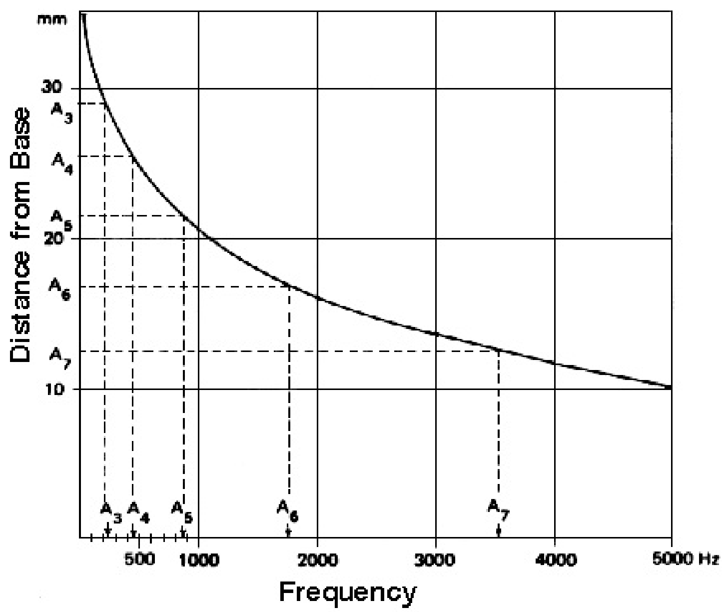


Figure 41: Position of the maximum displacement on the basilar membrane as a function of the pure tone frequency. (From "Science of Sound" by Rossing, Moore, and Wheeler.)

- Thus, the location of maximum displacement of the basilar membrane has a *logarithmic frequency response*, in agreement with Fechner's law.
- The bandwidths of the critical bands are approximately constant ($\approx$ 100 Hz) for frequencies below about 500 Hz; at higher frequencies, the bandwidths increase in proportion to the center frequencies, having values approximately equal to one-fourth of an octave (so 3 semitones or approximately 15%) about each center frequency.
- The fact that the human ear acts like a spectrum analyser for sound waves (with different parts of the basilar membrane repsonding to different frequency sound waves) is called the *place theory of pitch*.
- Sound waves can enter the inner ear via *air conduction* as discussed above. They can also enter the inner ear via *bone conduction* through the skull.
- Humming is an example of a sound that is heard mostly via bone conduction. By plugging your ears with your fingers when humming (thus blocking the air path), the humming can actually sound louder.
- Hearing through bone conduction is the reason why our voices sound different when we hear ourselves speak on a tape recorder. The microphone of a tape recorder picks up our voice as it is heard via air conduction. But we hear our own voice partly via air conduction and partly via bone conduction.

10.4 Binaural hearing and sound localization

- The fact that we have two ears allows us to better locate the source of a sound (similar to depth perception using two eyes—i.e., binocular vision).
- For frequencies above about 4000 Hz, localization is accomplished via an *intensity* difference of the sound reaching the two ears. At these higher frequencies, sound does not easily diffract around the head, so one ear is effectively in the "shadow" of the head and receives less intense sound. The brain interprets the ear that hears the less intense sound as being further from the source of sound.
- For frequencies below about 1000 Hz, localization is accomplished via a difference in the *time of arrival* (or, equivalently, a phase difference) of the sound reaching the two ears. The brain interprets the ear that hears the time-delayed sound as being further from the source of the sound.

11 Loudness

- Loudness is our perception of the relative strength of a sound. It depends on the amplitude of the sound wave, its frequency, and its duration. (Here we will concentrate on the loudness of pure tones; the perceived loudness of complex tones depends on the frequency distribution of the sound.)
- Loudness is a *relative* quantity in the sense that it involves a comparison between two different sounds—e.g., we say that this sound is twice as loud as that one. In many of the formulae below, we will take our reference sound to be at the threshold of hearing.
- As we shall describe in more detail below, the human ear is most sensitive to sounds with frequencies around 4000 Hz. For example, a sound wave with a frequency of 100 Hz is perceived to be quieter than a sound wave with the same amplitude, but with a frequency of 4000 Hz.
- Sounds with durations less than about a half a second are perceived to be quieter than the same sound having a duration of about a second.
- The loudness of a sound that has a duration greater than several tens of seconds is perceived to *decrease* slightly over time.
- This is due to *sensory adaptation* of the brain. A long-duration stimulus is increasingly ignored by the brain as it realizes that there is no associated danger. The firing rate of neurons decreases as the duration of the stimulus increases; the firing rate increases when there is a *change* in the stimulus.
- Loudness depends on the *intensity* of a sound wave but is not equal to it. As we shall see below, loudness is more closely related to the *logarithm* of the intensity, in line with Fechner's law relating stimulus to sensation.

11.1 Intensity, sound intensity level (W/m^2, dB)

- Recall that the intensity of a wave is the rate at which energy passes through a unit area perpendicular to the direction of wave propagation. The units of intensity are W/m^2, corresponding to energy/(area$\cdot$time).
- The intensity I of a sound wave is related to the pressure deviation p via

$$I = \frac{p^2}{\rho v} \tag{11.1}$$

  where ρ is the density of air and v is the speed of sound. Thus, the intensity is proportional to the *squared amplitude* of the sound wave.
- The *sound intensity level* SIL is defined as

$$\mathrm{SIL} = \log(I/I_0) \text{ bels} = 10\log(I/I_0) \text{ dB} \tag{11.2}$$

  where $I_0 = 10^{-12}$ W/m^2 is the intensity for the threshold of hearing at $f = 1000$ Hz.
- Note that 10 dB (decibels) equals 1 bel. (A bel is named after Alexander Graham Bell, the inventor of the telephone.)
- Example: Doubling the intensity of a sound wave corresponds to an increase in the sound intensity level by $10\log 2 \approx 3$ dB.

  Example: Increasing the intensity of a sound wave by a factor of 10 corresponds to an increase in the sound intensity level by $10\log 10 = 10$ dB.
- The *just noticeable difference* (JND) for intensity is about 1 dB. This is the increase in sound intensity level needed in order for the change in sound intensity (for two notes played sequentially) to be just noticeable.
- In terms of a percent, a 1 dB increase in SIL corresponds to a 26% increase in intensity since

$$10^{1/10} = 1.26 \tag{11.3}$$

11.2 Sound loudness level (phon)

- Although sound intensity level SIL is a quantative measure of the strength of a sound wave, it doesn't take into account the frequency-dependence of the response of the human ear.

- In 1933, H. Fletcher and W.A. Munson experimentally determined *equal-loudness* curves for pure tones by carrying out tests on many people. Their results were very similar to those shown in Figure 42, which have been recommended by the International Standards Organization.

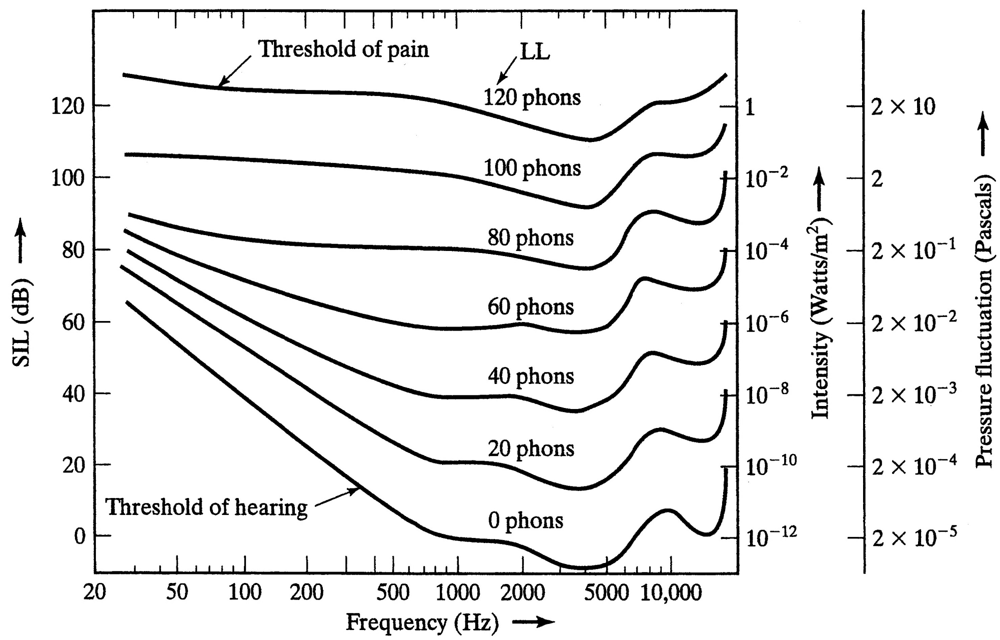


Figure 42: Equal-loudness curves (labeled in units of phons) for the human ear. Note that the ear is most sensitive to frequencies around 4000 Hz. (From "Physics of Sound" by Berg and Stork.)

- The equal-loudness curves are labeled by values of constant *sound loudness level* L_L. The units of L_L are *phons*. Numerically,

$$L_L \text{ (phon)} = \text{SIL (dB)} \quad \text{at } f = 1000 \text{ Hz} \tag{11.4}$$

  In general, the value of L_L depends on both SIL and the frequency f.

- Curves of constant phon are perceived to be equally-loud at different frequencies. This means that at low frequencies (e.g., 100 Hz), the sound pressure level must be substantially higher than at $f = 1000$ Hz for a sound to be perceived as being equally loud. From the figure, we see that our ears are most sensitive to sounds having frequencies near $f = 4000$ Hz.

- Exercise: What is the sound intensity level at 100 Hz for a sound with a sound loudness level of 40 phon?

- Answer: Using Figure 42, we see that the $L_L = 40$ phon equal-loudness curve has SIL = 60 dB at $f = 100$ Hz.

- A standard sound-level meter measures the sound intensity level with different frequency responses: (i) C-weighting has a flat frequency response, corresponding to a measurement of SIL in dB; (ii) A-weighting has a frequency response that mimics the sensitivity of the human ear, corresponding to a measurement of L_L in phon (sometimes denoted dBA).

11.3 Subjective loudness (sone)

- If the sound loudness level L_L of one sound is 10 phon larger than another sound, then most people perceive the first sound *to be twice as loud* as the second sound. (This corresponds to ΔSIL = 10 dB or a factor of 10 increase in intensity at a frequency of 1000 Hz.)

- To quantify this "twice as loud" dependence, one defines *subjective loudness* S via

$$S = 2^{(L_L - 40)/10} \text{ sone} \tag{11.5}$$

so that $S = 1$ sone is equivalent to $L_L = 40$ phon, and S increases by a factor of 2 whenever L_L increases by 10 phon.

- Table 4 compares different measures of loudness.

	SIL (dB) at 1000 Hz	L_L (phon)	S (sone)
Threshold of hearing	0	0	1/16
Recording studio	20	20	1/4
Quiet office	40	40	1
Ordinary conversation	60	60	4
Normal piano practice	80	80	16
Piano fortissimo	100	100	64
Threshold of pain	120	120	256

Table 4: Comparison of different measures of loudness.

- Figure 43 is a table with the Occupational Safety and Health Administration (OSHA) guidelines for noise exposure.

- The range of sound level in a musical performance is called its *dynamic range*.

- The dynamic range of a performance is typically between 10-20 dB, although it may as high as 40 dB. (Remember that a 10 dB increase in SIL at 1000 Hz corresponds to a 2× increase in perceived loudness.)

11.4 Sound from multiple sources

- Consider two instruments playing a note at the same time. The sound waves from each instrument combine with one another to produce a resultant wave with amplitude p_{tot} that can have any value between $p_1 + p_2$ and $|p_1 - p_2|$, which depends on the phase relation of the waves, where p_1, p_2 are the pressure amplitudes for the two component waves.

- On average, for random phases, the resultant amplitude will have a value

$$p_{\text{tot}} = \sqrt{p_1^2 + p_2^2} \tag{11.6}$$

This is called an *incoherent* superposition of two waves.

- Since intensity is proportional to the square of the pressure amplitude, it follows that the intensity of the incoherent sum of the two sound waves is given, on average, by

$$I_{\text{tot}} = I_1 + I_2 \tag{11.7}$$

Thus, *intensities add*.

TABLE G-16 - PERMISSIBLE NOISE EXPOSURES (1)

Duration per day, hours	Sound level dBA slow response
8	90
6	92
4	95
3	97
2	100
1 1/2	102
1	105
1/2	110
1/4 or less	115

Footnote(1) When the daily noise exposure is composed of two or more periods of noise exposure of different levels, their combined effect should be considered, rather than the individual effect of each. If the sum of the following fractions: C(1)/T(1) + C(2)/T(2) C(n)/T(n) exceeds unity, then, the mixed exposure should be considered to exceed the limit value. Cn indicates the total time of exposure at a specified noise level, and Tn indicates the total time of exposure permitted at that level. Exposure to impulsive or impact noise should not exceed 140 dB peak sound pressure level.

Figure 43: OSHA permissible noise exposure levels. (From `www.osha.gov`, Standard Number: 1910.95.)

- Hence, 2 violins playing simultaneously produce an intensity $2\times$ greater than that of a single violin; 10 violins playing simultaneously produce an intensity $10\times$ greater. The corresponding increase in the sound intensity level SIL will be

$$10 \log 2 \approx 3 \text{ dB} \quad \text{(two violins)} \tag{11.8}$$

$$10 \log 10 = 10 \text{ dB} \quad \text{(ten violins)} \tag{11.9}$$

- Note that to double the subjective loudness level S requires L_L to increase by 10 phon (or SIL to increase by 10 dB at 1000 Hz). This corresponds to a $10\times$ increase in the intensity I of the sound at 1000 Hz. *Hence, we would need 10 violins playing simultaneously to sound twice as loud as a single violin.*

12 Pitch and timbre

- Demonstrations: http://www.personal.psu.edu/meb26/INART50/psychoacoustics.html
- Question: What distinguishes one musical tone from another?
- Answer: duration, pitch, timbre, loudness or intensity, attack and decay transients, others (??)
- We talked about loudness in a previous section. In this section we focus on pitch and timbre.
- We shall see that the attack and decay transients affect the timbre of a tone. Also one needs a duration of about 5 cycles to be able to determine the pitch of a tone.
- Pitch is simply how high or how low (bass) a tone is. It is a one-dimensional attribute (the frequency of the fundamental).
- To describe timbre or tone quality, one uses words like "bright" or "dark" $\cdots$, which are much more vague. It is multi-dimensional attribute.

12.1 Perfect pitch

- Perfect (or absolute) pitch: Ability to determine the absolute pitch of a sound without regard to a reference tone. Only about 1 out of 10,000 people in the population has perfect pitch.
- It is much more common for a person to have *relative* pitch, which is the ability to determine the pitch of a sound relative to some reference tone.
- There is controversy regarding the origin of perfect pitch. Possible theories: (i) it is hereditary, (ii) it can be acquired through training, even in old age, (iii) everyone has perfect pitch when they are born, but it is lost through neglect, i.e., without training, (iv) anyone can be taught to have absolute pitch if it is done at the right stage in their development (i.e., imprinting as a child).

12.2 Pitch discrimination

- *Just noticeable difference* (JND) in frequency: the minimum difference in the frequencies of two pure tones played sequentially that can be distinguished from one another.
- At high frequencies (i.e., above about 1000 Hz), the JND is approximately 0.5% of the center frequency, corresponding to roughly 1/10th of a semitone. At low frequencies (i.e., below about 500 Hz), the JND is approximately 3 Hz, independent of the center frequency. (See Figure 44.)
- Example: From the figure, we see that at 2000 Hz the JND is approximately 10 Hz. At 200 Hz the JND is approximately 3 Hz.
- The JND has the same general dependence on frequency as does the critical bandwidth for the cochlea, but the JND is about a factor of 30 times smaller than the critical bandwidth. (Recall: At high frequencies, the critical bandwidth is about 15% of the center frequency, corresponding to one-fourth of an octave, or 3 semitones. For example, at 2000 Hz, the critical bandwidth is approximately 300 Hz.)
- *Limit of frequency discrimination* (LFD): the minimum difference between two pure tones played simultaneously that can be distinguished from one antother.
- The limit of frequency discrimination is about 10% of the center frequency, corresponding to about 2 semitones.
- Analogy: The sense of touch on the skin of your forearm using two pencil erasers separated by some distance.
- Demonstration: Do the pencil eraser experiment. Note that it is harder for you to tell if the pencil erasers are touching different parts of your forearm if the erasers touch your skin at the same time.

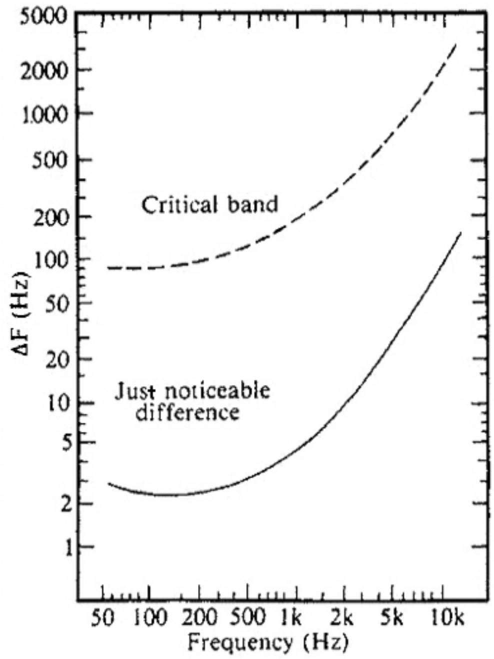


Figure 44: Just noticeable difference in frequency and comparison to the critical bands on the cochlea. (Figure taken from "Science of Sound," by Rossing, Moore, and Wheeler.)

12.3 Missing fundamental and second-order beats

- If a complex tone containing many harmonics—*but not the fundamental*—is played, a person will identify the pitch of the sound to be that of the missing fundamental.

- The missing fundamental is inferred by the the brain from the timing of electrical impulses triggered by the periodicity of the complex sound wave. (Recall that a complex sound wave containing many harmonics—but not the fundamental—still has a period equal to the inverse of the fundamental frequency. This is true even though the sound wave *does not have any power at the fundamental frequency.*)

- Demonstration: Illustrate this with the matlab routine `fouriersynthesizeScript.m` for wave type 'missing'.

- The missing fundamental does not correspond to a physical wave in the ear. The missing fundamental is an example of the *periodicity theory* of pitch.

- Contrast this to the *place theory* of pitch, which says that the pitch of a sound is determined by the location on the basilar membrane that is excited by the sound wave. But if the sound wave doesn't contain power at a particular frequency, then that part of the basilar membrane won't be excited.

- Exercise: A complex tone consists of harmonics 200 Hz, 300 Hz, 400 Hz, etc. What pitch will be heard?

- Exercise: A complex tone consists of the harmonics 300 Hz, 500 Hz, 700 Hz, etc.. What pitch will be heard? (Hint: These are odd harmonics.)

- Answer: 100 Hz for both of the exercises.

- Ordinary (*first-order*) beats are heard when two pure tones with nearly identical frequencies f_1 and f_2 are played simultaneously. We hear beats at a frequency given by

$$f_b = |f_1 - f_2| \tag{12.1}$$

corresponding to alternating periods of constructive and destructive interference. We are sensitive to the low frequency (< 10 Hz) amplitude modulation of a tone with frequency $\bar{f} = (f_1 + f_2)/2$.

- *Second-order* beats arise when two pure tones with frequencies that are nearly one octave apart, i.e., $f_2 \approx 2f_1$, are played simultaneously. We hear beats at a frequency given by

$$f_b = |2f_1 - f_2| \tag{12.2}$$

The beats are a result of periodicity processing within the brain just like for the missing fundamental.

- <u>Demonstration</u>: Using the matlab routine `playintervalFreq.m`, produce sounds with frequencies: (i) $f_1 = 440$ Hz and $f_2 = 442$ Hz. Listen for beats. (ii) Repeat for $f_1 = 220$ Hz and $f_2 = 442$ Hz. For both cases, we hear beats with a frequency of 2 Hz. The brain processes the periodicity of the combined waveform.

12.4 Aural harmonics and aural combination tones

- Asymmetries in the amplitude response of the human ear cause a pure tone to be represented by a distorted wave having harmonics of the original sine wave. These are called *aural harmonics.*

- Aural harmonics are not part of the original pure tone. They are created by the non-linear amplitude response of the ear to sound pressure:

$$x(t) = a_0 + a_1 p(t) + a_2 p^2(t) + a_3 p^3(t) + \cdots \tag{12.3}$$

where a_n are constants.

- Figure 45 shows a pure tone $p(t) = \sin(2\pi f t)$ with $f = 1$ Hz, and the non-linear output $x(t)$ for the special case $a_0 = 0$, $a_1 = 1$, $a_2 = 1/2$, $a_3 = 1/3$, with all higher-order $a_n = 0$. Note that $x(t)$ is not a simple sinusoid.

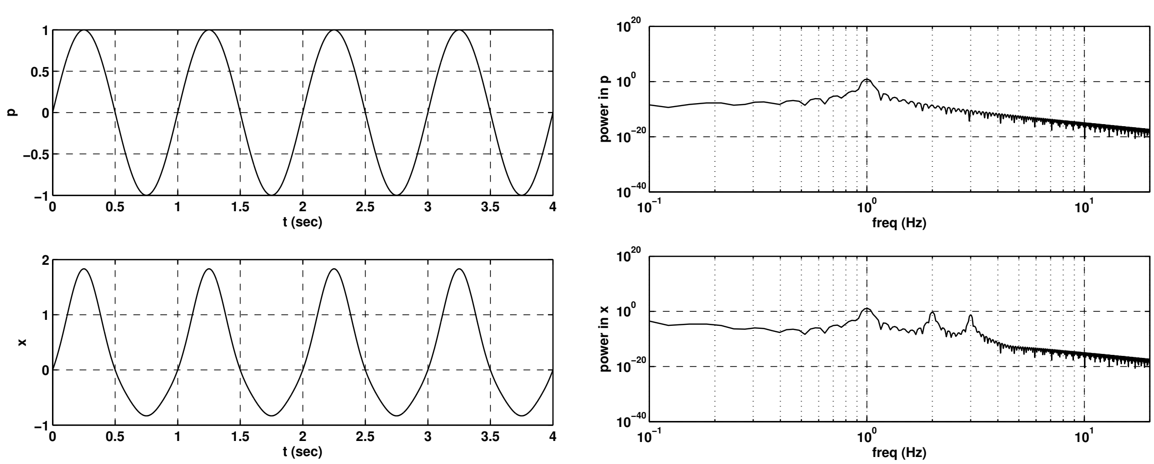


Figure 45: Illustration of aural harmonics produced by the non-linear response of the ear. Top two panels: pure sinusoidal input and its corresponding frequency spectrum. Bottom panels: non-linear output and its corresponding frequency spectrum containing linear, quadratic, and cubic contributions.

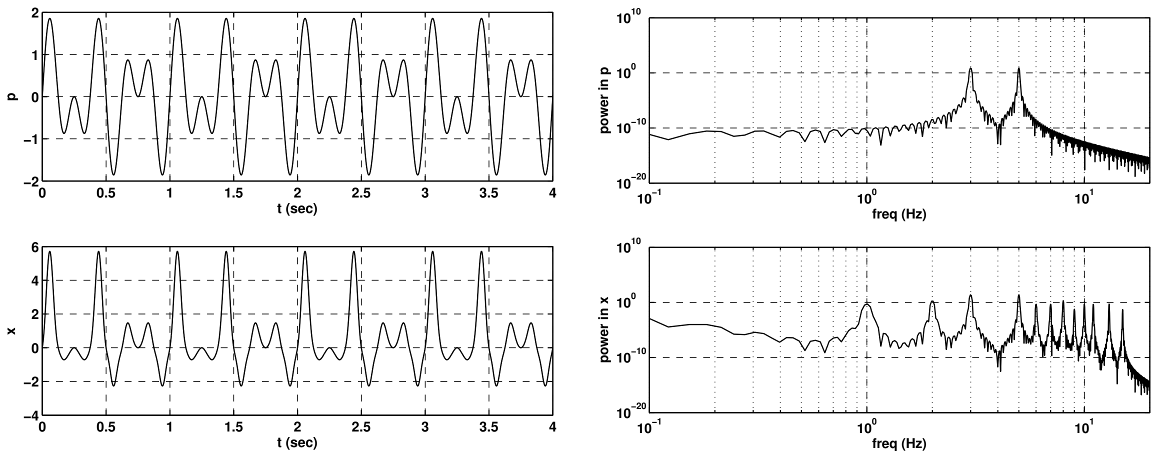


Figure 46: Same as the previous figure, but for a complex tone consisting of a sum of two sine waves with frequencies $f_1 = 3$ Hz and $f_2 = 5$ Hz. From the bottom-right-hand plot we see that the non-linear output has contributions from several different combination tones $f_c = |mf_1 \pm nf_2|$, ranging from 1 Hz to 15 Hz. There is no contribution from 4, 12, or 14 Hz for this particular example, where $x(t)$ contains only linear, quadratic, and cubic terms.

- If two pure tones (with frequencies f_1 and f_2) are played sufficiently loudly, then the non-linear response of the ear will create aural *combination tones* having frequencies equal to the sum and difference of the two original frequencies:

$$f_c = |mf_1 \pm nf_2| \tag{12.4}$$

  where m and n are integers.

- Figure 46 illustrates this for a complex tone made up of two sinusoids with frequencies $f_1 = 3$ Hz and $f_2 = 5$ Hz.

- Combination tones exist as physical waves having power at the sum and difference frequencies. The difference tones are more easily heard than the sum tones. The effect is greatest when the two tones are loud. The most easily heard difference tones have frequencies given by

$$|f_1 - f_2|\,, \quad |2f_1 - f_2|\,, \quad |3f_1 - f_2| \tag{12.5}$$

- Example: Play two notes simultaneously on a piano a fifth apart, e.g., C_4 and G_4. The frequencies differ by (approximately) a factor of 3/2. If the two notes are played sufficiently loudly, one hears the difference tone C_3, which is an octave below C_4.

- The fact that one can hear bass notes from a small loudspeaker is due to a combination of (i) the aural difference tone of adjacent harmonics, and (ii) our perception of the missing fundamental.

- Application: This is an advantage for speaker manufacturers, as well as for manufacturers of pipe organs, since to physically produce the lowest bass notes would require extremely long organ pipes. We let the ear and the brain create the lowest bass notes from the higher-order harmonics.

12.5 Timbre

- Timbre (or tone quality) can be defined as "Any attribute that allows a listener to judge that two sounds are disimilar using any criteria other than pitch, loudness, or duration" (Pratt and Doak, 1976).

- Timbre depends on the relative contribution of the harmonics or overtones that make up the sound wave.
- Aural harmonics introduced by non-linearities in the ear change the timbre of a pure tone.
- Ohm and Helmholtz did some experiments (in the 1800's) showing that the timbre of a complex tone depends primarily on the amplitude of the harmonic components and *not* on their relative phases. This is called *Ohm's law of hearing*.
- Demonstration: Illustrate this using the matlab routine `fouriersynthesizesoundScript.m` with inputs `'sawtooth'` and `'sawtoothphase'`.
- Ohm's law of hearing is important for sound reproduction systems. Although the process of sound recording and playback (which uses electromagnetic induction) may change the relative phases of the harmonics, the played-back sound sounds virtually identical to that of the original sound.
- Timbre also depends on the *attack* and *decay transients* of the sound.
- A note from an instrument sounds different if it is played forward or backward in time.
- Demonstration: Illustrate this using the matlab routine `playrecordedsound('pianoC4',0)` and `playrecordedsound('pianoC4',1)`.
- Demonstration: Also play
  `playrecordedsound('happybirthday', 0)` followed by
  `playrecordedsound('happybirthday-backwards',0)`, and
  `playrecordedsound('happybirthday-backwards',1)`.

12.6 Consonant and dissonant tones

- Two complex tones are consonant (sound pleasing) when several of the harmonics or overtones coincide.
- Consonance / dissonance depends on the frequency difference bewteen the two tones and not on the frequency ratio. If the frequency difference is greater than the critical bandwidth, then the notes sound consonant. If the frequency difference is slightly less than the critical bandwidth, then the notes sound rough. If the frequency difference is roughly equal to a quarter of the critical bandwidth then the notes have maximum dissonance.
- Example: C to G is a fifth, which always has a frequency ratio of (approximately) 3/2.

  C_4 to G_4 has $\Delta f = 130$ Hz which is greater than the critical bandwidth $\Delta f_{\rm cb} \approx 100$ Hz at the center frequency $\bar{f} = 328$ Hz, and so sounds consonant.

  C_3 to G_3 has $\Delta f = 65$ Hz which is slightly less than the critical bandwidth $\Delta f_{\rm cb} \approx 90$ Hz at the center frequency $\bar{f} = 164$ Hz, and so sounds rough.

  C_2 to G_2 has $\Delta f = 32$ Hz which is about one-third of the critical bandwidth $\Delta f_{\rm cb} \approx 90$ Hz at the center frequency $\bar{f} = 82$ Hz, and so sounds dissonant.
- Demonstration: Illustrate this using a dual function generator, first for the frequencies $f_1 = 262$ Hz (for C_4) and $f_2 = 392$ Hz (for G_4); then for $f_1 = 131$ Hz (for C_3) and $f_2 = 196$ Hz (for G_3); and finally for $f_1 = 65$ Hz (for C_2) and $f_2 = 98$ Hz (for G_2).

12.7 A pitch paradox

- In 1964, Roger Shepard constructed a repeating set of complex tones (called *Shepard tones*), whose pitch appears to increase (or decrease) forever.
- Demonstration: Illustrate this with the matlab routine `shepardscale('ascending',1)`.
- Demonstration: Also watch the YouTube videos by David Huron (http://vimeo.com/34749558).

- The Shepard scale is the musical analog of the optical illusion of a *never-ending staircase*, originally developed by the mathematician Lionel Penrose in 1958, and subsequently used by M.C. Escher in his 1960 print 'Ascending and Descending.' (See Figure 47 of a never-ending staircase.)

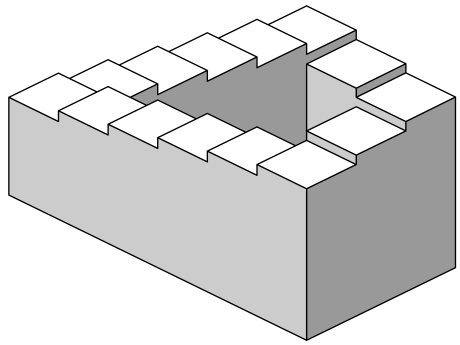


Figure 47: Never-ending staircase.

- Demonstration: Watch the YouTube vide called the Shepard-Penrose mix (http://www.youtube.com/watch?v=PCs1lckF5vI)
- Question: How does the Shepard scale work?
- Answer: EXTRA CREDIT, 2 points for an explanation.

Part V

Room acoustics, reproduction of sound

13 Auditorium and room acoustics

- This section briefly discusses issues related to the acoustical properties of rooms and auditoria.
- We will pay particular attention to the distinction between direct, reflected, and reverberant sound. We will also learn how to calculate the reverberation time of a room.
- Demonstration: Watch the YouTube video `https://www.youtube.com/watch?v=BYBSA9v8IRE` about anechoic (i.e., echo-free) chambers.
- Demonstration: For a general introduction to architectural and environmental acoustics, listen to "One man's quest to find the sonic wonders of the world," on NPR's 'Fresh Air' program with Terry Gross.

13.1 Direct sound

- *Direct sound* is the sound that is received from a source in the absence of any reflections. See Figure 48.

Figure 48: Direct sound only. No reflections.

- In an anechoic chamber (or in empty space) one hears only the direct sound from a source.
- For an omni-directional source, the power is spread out uniformly over the area of a sphere. The intensity at a distance r from a source with power P is given by

$$I = \frac{P}{4\pi r^2} \tag{13.1}$$

(Recall: intensity is the amount of energy per unit time that passes through a unit area directed perpendicular to the flow.)

- Exercise: Calculate the power received by the ear of a listener standing a distance of 10 m from an omni-directional 1 watt source. Approximate the human ear as a square with side length of 5 cm.
- Answer: The area of the ear is given by $A_{\rm ear} = (.05\ {\rm m})^2 = 0.0025\ {\rm m}^2$. Since the absorbed power is equal to the intensity times the area,

$$P_{\rm ear} = IA_{\rm ear} = \frac{P}{4\pi r^2}\,A_{\rm ear} = \frac{1}{4\pi 10^2}0.0025 = 2\times 10^{-6}\ {\rm watts} \tag{13.2}$$

Note that the ratio of the received power to the source power is 2×10^{-6} or 1/500,000, which is a very small fraction of the radiated power.

- For a directional source, which radiates only over a portion of the sphere,

$$I = \frac{QP}{4\pi r^2} \tag{13.3}$$

where Q is the *directivity factor*. For a directional source, both I and Q depend on the direction from the source.

- Example: For a source that radiates only in a hemisphere, e.g., a loudspeaker mounted to the ceiling, $Q = 2$. For a source that radiates only in an octant, e.g., a loudspeaker in a corner, $Q = 8$.

13.2 Reflected sound

- *Reflected sound* is the sound that is received from a source after it has been reflected from one or more surfaces.
- Recall that sound reflects according to the formula that the angle of incidence equals the angle of reflection. The reflected sound appears to come from an image source, as shown in Figure 49.

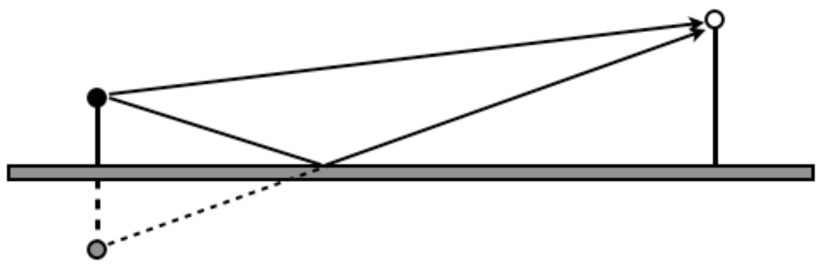


Figure 49: Direct sound plus reflection off a floor. The reflected sound appears to come from an image source on the other side of the floor.

- *Specular reflection* occurs when the wavelength of sound is much larger than the length scale of the roughness of the surface. Parallel rays in become parallel rays out.
- *Diffuse reflection* occurs when the wavelength of sound is of order or less than the length scale of the roughness of the surface. Parallel rays in become non-parallel rays out.
- Since the reflected sound travels farther than the direct sound, it will arrive later than the direct sound. (The speed of sound is 346 m/s in air at room temperature. This corresponds to roughly 1000 ft/s or 1 foot per millisecond.)
- The brain interprets the reflected sound and the direct sound to be the same if the time between the reception of the reflected and direct sound is less than $\sim$35 msec. Longer delays give rise to an echo.
- A related phenomenon is the *precedence effect*: The brain interprets the source of sound to be in the direction from which the first sound is heard.
- Since the reflected sound travels farther than the direct sound, the sound intensity level of the reflected sound will be less than that for the direct sound. Non-zero absorption by the reflecting wall will also reduce the sound intensity level.
- Exercise: A listener stands 4 m in front of a 1 watt omni-directional loudspeaker. It is 1.5 m from a reflecting wall.

  (a) Calculate the time of arrival for both the direct and reflected sound.

  (b) Calculate the decrease in sound intensity level for the reflected sound relative to the direct sound, assuming perfect reflection.

  (c) Calculate the decrease in sound intensity level due to a non-zero absorption coefficient, e.g., $a = 0.2$.
- Answer:

  (a) Using geometry, one can show that the distance that the reflected sound travels is $2\sqrt{2^2 + 1.5^2} = 5$ m:

$$t_{\text{direct}} = \frac{4\ \text{m}}{346\ \text{m/s}} = 12\ \text{msec}\,, \qquad t_{\text{reflected}} = \frac{5\ \text{m}}{346\ \text{m/s}} = 15\ \text{msec} \tag{13.4}$$

(b) The decrease in SIL comes from the increased distance from the source for the reflected sound:

$$\Delta\text{SIL} = 10\log[1/(r_{\text{reflected}}/r_{\text{direct}})^2]\text{ dB} = 10\log[(4/5)^2]\text{ dB} = -2\text{ dB} \tag{13.5}$$

(c) For a non-zero absorption coefficient $a = 0.2$:

$$\Delta\text{SIL} = 10\log(1-a)\text{ dB} = 10\log 0.8\text{ dB} = -1\text{ dB} \tag{13.6}$$

13.3 Multiple reflections

- Since a room has multiple walls, a ceiling, and a floor, sound can undergo multiple reflections when traveling from the source to a listener.
- Figures 50, 51, 52 show paths for multiple reflections. The situation gets complicated very quickly!!

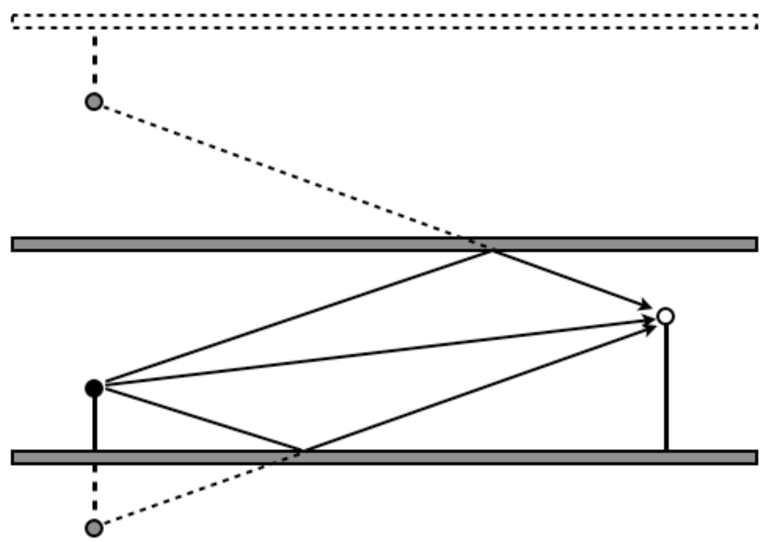


Figure 50: Direct sound plus reflections off a floor and a ceiling. The reflected sound appears to come from image sources on the other side of the floor and the other side of the ceiling.

13.4 Reverberant sound

- *Reverberant sound* is the sound that is formed from multiple reflections, coming from many different directions, and overlapping in time.
- Figure 53 is a graph showing the direct sound, early reflected sound, and reverberant sound as a function of time. Recall that sounds separated by less than 35 msec are perceived as the same sound.
- The sound intensity level decays as function of time since the energy in the initial sound pulse is lost by absorption to the walls, etc.
- Figure 54 shows the build-up and decay of the sound pressure level ($L_p \approx$ SIL) of the reverberant sound for a sustained source of sound.
- The *reverberation time* is the time for the reverberant sound intensity level to decrease by 60 dB. This is equivalent to a drop in intensity of the reverberant sound by a factor 10^{-6} of its original value.

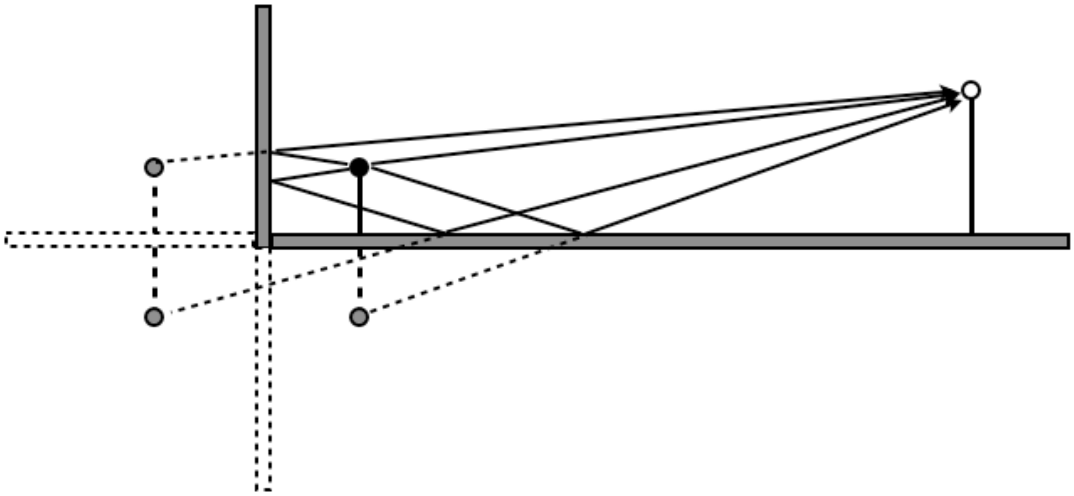


Figure 51: Direct sound plus reflections off a floor and a back wall. The reflected sound appears to come from image sources on the other side of the floor and the other side of the back wall.

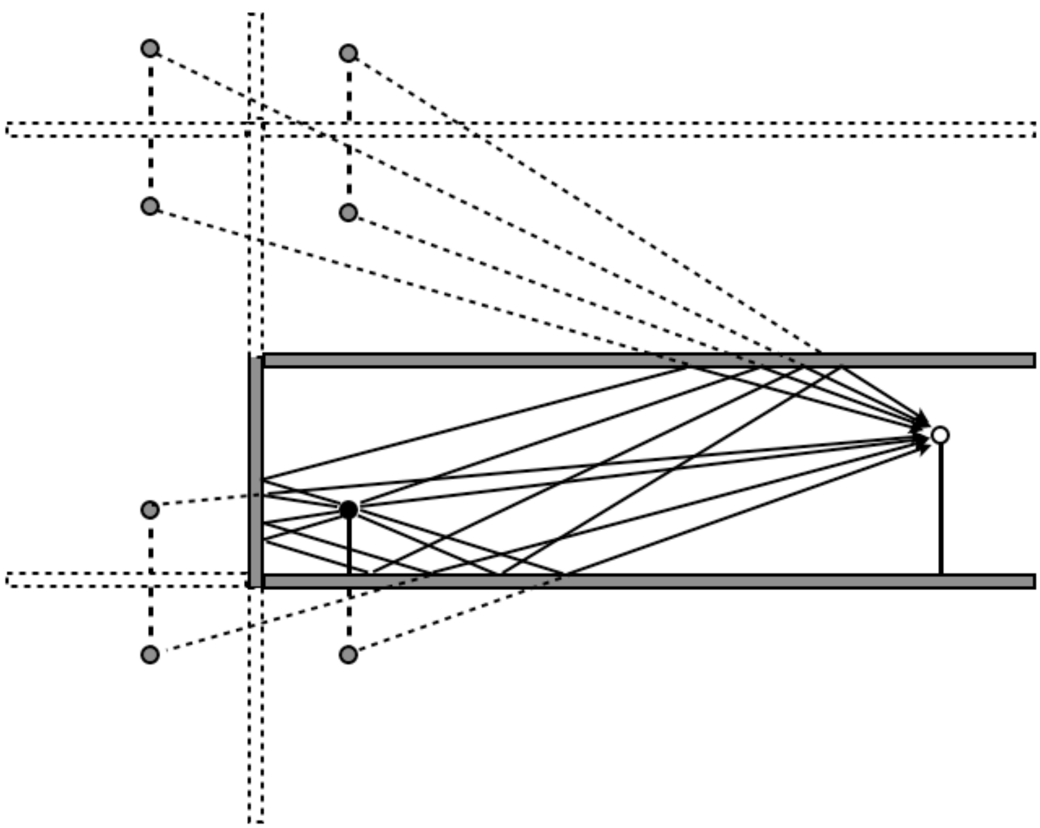


Figure 52: Direct sound plus reflections off a floor, a ceiling, and a back wall. The reflected sound appears to come from image sources on the other sides of the floor, ceiling, and back wall.

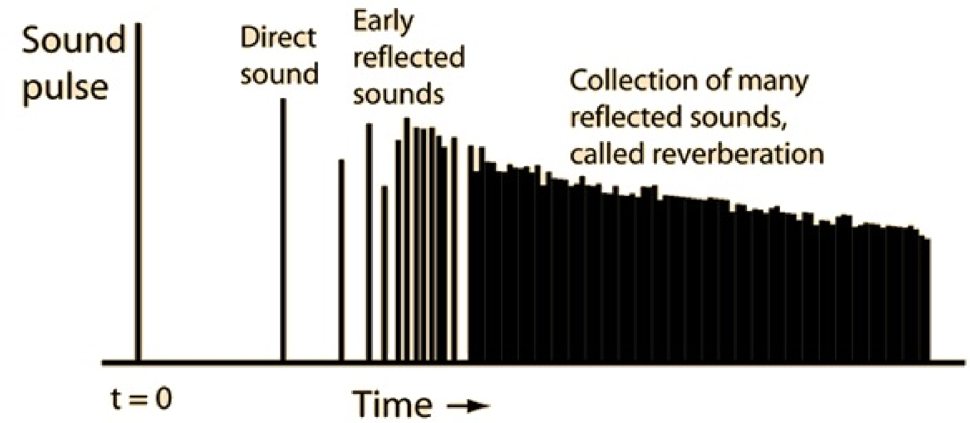


Figure 53: A graph showing the distinction between direct, early reflected, and reverberant sound. The sound pulse is produced at $t = 0$. The horizontal axis give the arrival time of the direct and reflected pulse. The vertical axis is a measure of the sound intensity level in the direct and reflected sounds. (Figure taken from http://hyperphysics.phy-astr.gsu.edu/.)

- The reverberation time depends on the frequency of the sound. High frequency sounds typically have reverberation times that are less than those for low frequency sounds.
- Figure 55 is a graph of ideal reverberation times for rooms of different sizes and for various functions.
- Table 5 contains acoustical characteristics of several different concert halls.

	Year built	Volume (m^3)	Number of seats	Reverberation time (sec) Frequency (Hz)		
				125	500	2000
Teatro alla Scala, Milan	1778	11,245	2289		1.2	
Royal Opera House	1858	12,240	2180		1.1	
Royal Albert Hall	1871	86,600	6080	3.4	2.6	2.2
Carnegie Hall, New York	1891	24,250	2760	1.8	1.8	1.6
Symphony Hall, Boston	1900	18,740	2630	2.2	1.8	1.7
Royal Festival Hall	1951	22,000	3000	1.4	1.5	1.4
Philharmonic Hall, Berlin	1963	36,030	2200		2.0	
St. David's Hall, Cardiff	1983	22,000	2200	1.8	1.9	1.8

Table 5: Acoustical characteristics of various concert halls. (From a table in "MU1217 Lecture Notes," Cardiff University by Dr. Bernard Richardson.)

13.5 Calculating the reverberation time

- The reverberation time T_R can be calculated from the following formula:

$$T_R = 0.05 \, \frac{V}{A_{\text{eff}}} \text{ s} \tag{13.7}$$

where V is the total volume of the room (in ft^3) and A_{eff} is the total absorption (in units of *sabin*).

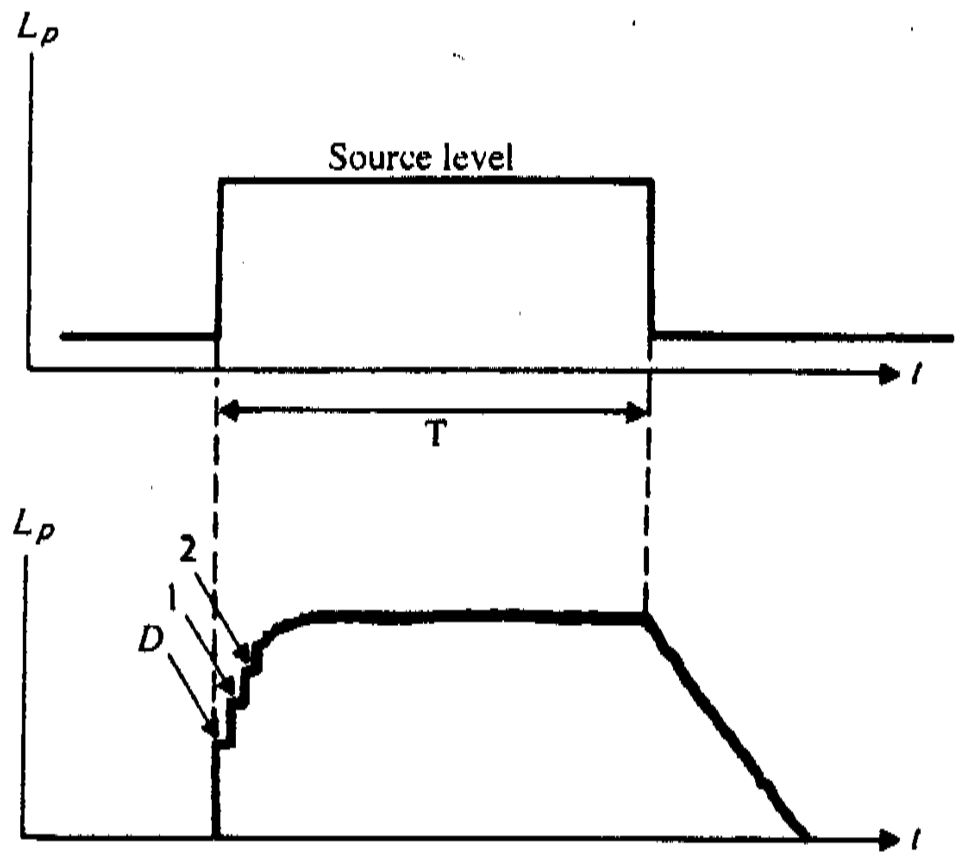


Figure 54: Build-up and decay of the sound pressure level ($L_p \approx$ SIL) of the reverberant sound for a sustained source of sound. D indicates the increase in power due to the arrival of the direct sound; 1 and 2 are the same for the arrival of the first and second reflections. (Figure taken from "Science of Sound," by Rossing, Moore, and Wheeler.)

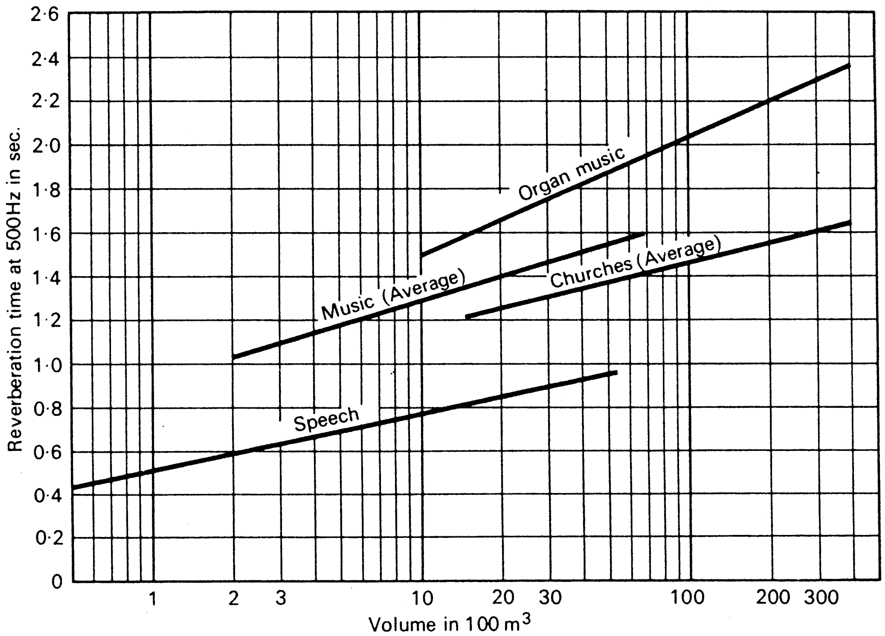


Figure 55: Ideal reverberation times for rooms of different sizes and for various functions. (Figure taken from "MU1217 Lecture Notes," Cardiff University by Dr. Bernard Richardson.)

- Total absorption A_{eff} is given by

$$A_{\text{eff}} = A_1 a_1 + A_2 a_2 + \cdots + B_1 + B_2 + \cdots \tag{13.8}$$

  where A_i and a_i are the surface area (in ft^2) and absorption coefficient (dimensionless) of surface $i = 1, 2, \ldots$ (e.g., the walls, floor, ceiling) in ft^2, and B_i is the total absorption for chairs, people in the room, etc.

- One sabin is equivalent to one square-foot of a perfectly absorbing surface—e.g., a window one square-foot in area.

- The sabin is named after Wallace C. Sabine (1869-1937), who was the first person to systematically study room acoustics.

- Note that an absorption coefficient $a = 0$ corresponds to no absorption or total reflection. An absorption coefficient $a = 1$ corresponds to complete absorption—e.g., an open window or no reflection.

- The reverberation time depends on the frequency of the sound, since the absorption coefficients of different materials depend on frequency.

- Table 6 gives the absorption coefficients for several different materials evaluated at different octave intervals.

- Table 7 gives the absorption (in units of m^2) for seats and people in the audience. To convert to absorption in units of sabin, multiply the values in the Table by 10.8.

- Exercise: Calculate the reverberation time at 500 Hz for a room with dimensions 20 m × 15 m × 8m (high). The walls are painted concrete, the ceiling is plaster, and the floor is carpet on pad. Also, assume that there are 200 upholstered seats, and that they are half-filled with people.

Material	Frequency (Hz) 125	250	500	1000	2000	4000
Concrete (painted)	0.10	0.05	0.06	0.07	0.09	0.08
Plywood panel	0.28	0.22	0.17	0.09	0.10	0.11
Plaster on lath	0.14	0.10	0.06	0.05	0.04	0.03
Gypsum board, 1/2 in.	0.29	0.10	0.05	0.04	0.07	0.09
Glass window	0.35	0.25	0.18	0.12	0.07	0.04
Curtains	0.14	0.35	0.55	0.72	0.70	0.65
Carpet (on concrete)	0.02	0.06	0.14	0.37	0.60	0.65
Carpet (on pad)	0.08	0.24	0.57	0.69	0.71	0.73
Acoustical tile, suspended	0.76	0.93	0.83	0.99	0.99	0.94

Table 6: Absorption coefficients (dimensionless) for different materials evaluated at different octave intervals. (Based on a similar table in "Science of Sound," by Rossing, Moore, and Wheeler.)

Material	Frequency (Hz) 125	250	500	1000	2000	4000
Wood or metal seat, unoccupied	0.014	0.018	0.020	0.036	0.035	0.028
Upholstered seat, unoccupied	0.13	0.26	0.39	0.46	0.43	0.41
Adult	0.23	0.32	0.39	0.43	0.46	—
Adult in an upholstered seat	0.27	0.40	0.56	0.65	0.64	0.56

Table 7: Absorption (in units of m^2) for different types of seats with and without people evaluated at different octave intervals. To convert to absorption in units of sabin, multiply the values in the Table by 10.8. (Based on a similar table in "Science of Sound," by Rossing, Moore, and Wheeler.)

- Answer: First converting to units of feet, and using data from Tables 6 and 7 at $f = 500$ Hz:

$$
\begin{aligned}
L &= 20 \text{ m} \times 3.28 \text{ ft/m} = 65.6 \text{ ft} && (13.9)\\
W &= 15 \text{ m} \times 3.28 \text{ ft/m} = 49.2 \text{ ft} && (13.10)\\
H &= 8 \text{ m} \times 3.28 \text{ ft/m} = 26.24 \text{ ft} && (13.11)\\
V &= L \times W \times H = 2400 \text{ m}^3 = 8.47 \times 10^4 \text{ ft}^3 && (13.12)\\
A_{\text{eff}} &= 0.06\,[2(L \times H) + 2(W \times H)] + 0.06(L \times W) + 0.57(L \times W)\\
&\quad + 10.8(100 \times 0.39 + 100 \times 0.56)\\
&= 3.42 \times 10^3 \text{ sabin} && (13.13)\\
T_R &= 1.2 \text{ sec} && (13.14)
\end{aligned}
$$

- Comparing with Figure 55, we see that this value for T_R is close to ideal for music (average) for a room of this volume (2400 m^3).

13.6 Criteria for good acoustical design

- Adequate loudness: Everybody in the audience should be able to hear the speaker or the performers. The room shouldn't be too large or be too absorptive.
- Uniformity: The sound should be spread uniformly throughout the room. There shouldn't be "live" spots or "dead" spots in the room.
- Reverberance or liveness: The reverberation time should be sufficiently large so that the listener feels bathed in sound from all directions.
- Clarity: The reverberation time shouldn't be so large that the revererberant sound masks subsequent sounds. (Clarity and reverberance are competing qualities.)
- Freedom from echoes: The room should be free of echoes.
- Freedom from background noise: The room should be free of excessive levels of background or external noise.

13.7 Problems to avoid

- Echoes: Reflected sounds should arrive early enough (< 35 msec) in order to avoid echoes.
- Since curved surfaces tend to focus sound to certain areas in a room, these should be avoided.
- Figure 56 shows how sound produced at one focal point of a room with elliptical walls is focused at the other focal point.
- Parallel walls can produce *flutter echoes* since there are multiple paths for the reflected sound to reach the listener, corresponding to an infinite number of image sources. See Figure 57.
- Shadows: Sound shadows produced by overhanging balconies or columns in the room should be avoided as much as possible.
- Resonances: Room resonances should be avoided as much as possible, since they would distort (i.e., filter) the sound produced by the source. One can avoid room resonances by *not* having parallel walls and ceilings.
- Note that a rectangular box-shaped room with dimensions $L \times W \times H$ would have resonances at the discrete frequencies

$$f_{lmn} = \frac{v}{2}\sqrt{\left(\frac{l}{L}\right)^2 + \left(\frac{m}{W}\right)^2 + \left(\frac{n}{H}\right)^2}\,, \qquad l, m, n = 0, 1, 2, \cdots \tag{13.15}$$

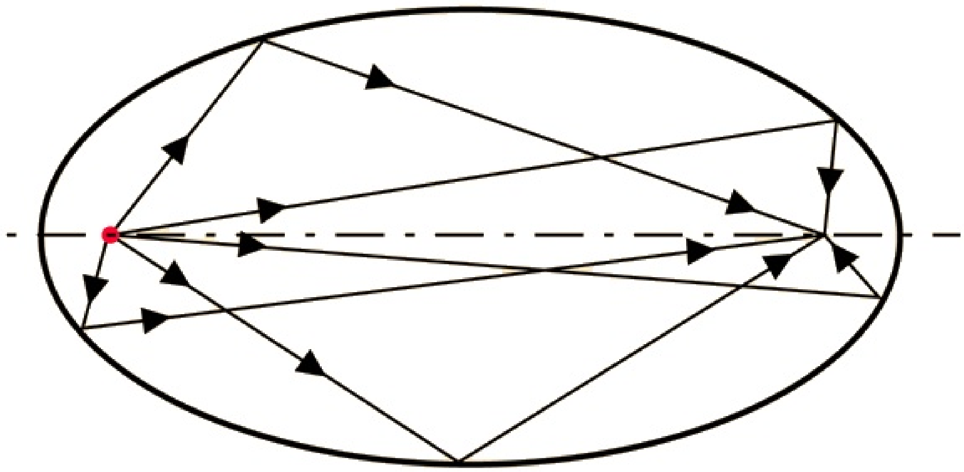


Figure 56: Illustration of the so-called "whispering chamber effect" for a room shaped like an ellipse (top view). The whisperer sits at the focal point indicated by the red dot, and listener sit at the other focal point. (Figure taken from http://hyperphysics.phy-astr.gsu.edu/.)

where v is the speed of sound in air. (This is just a generalization of standing waves in a 1-dimensional tube closed a both ends to 3 dimensions.)

- Exercise: Calculate the first few resonant frequencies for a rectangular-shaped shower stall having dimensions 1 m $\times$ 1 m $\times$ 2m.

- Answer:

$$f_{001} = 85 \text{ Hz}\,, \quad f_{010} = 170 \text{ Hz}\,, \quad f_{100} = 170 \text{ Hz}\,, \tag{13.16}$$

$$f_{002} = 170 \text{ Hz}\,, \quad f_{110} = 240 \text{ Hz}\,, \quad f_{111} = 255 \text{ Hz}\,, \cdots \tag{13.17}$$

- Although resonances might be good for singing in the shower, they are not good for concert halls in general!

- Background noise: External noise due to heating, ventillation or air conditioning systems should be kept to a minimum.

- For example, one should try to isolate the air conditioning unit from the ceiling joists, which would strongly couple the vibrations of the air conditioning unit into the building.

13.8 Terminology used for concert halls

- Intimacy or presence: The impression of being in a small concert hall.

- Spaciousness: The sound appears to come from all directions and from a source wider than the visual width of the source.

- Reverberance or liveness: (As before.) The feeling that the listener is bathed in sound.

- Clarity: (As before.) Being able to distinguish between discrete sounds in a musical performance.

- Warmth: Liveness of the bass notes (75 Hz–350 Hz) relative to the treble notes.

- Brilliance: Liveness of the treble notes (> 350 Hz) relative to the bass notes.

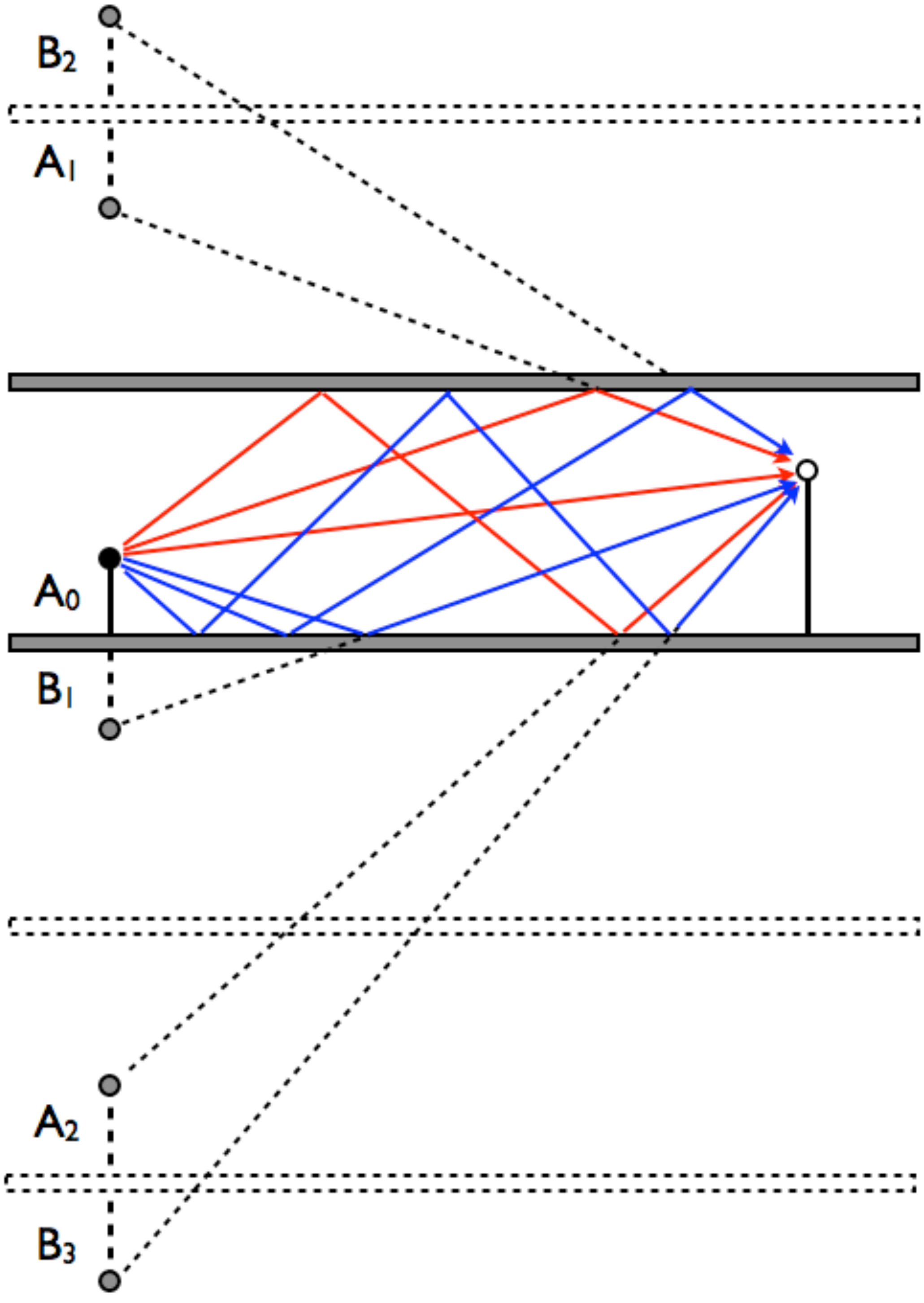


Figure 57: Multiple reflections and image sources due to parallel floor and ceiling. Multiple reflections from closely-spaced parallel surfaces can produce so-called flutter echoes.

- Blend: Mixing of sounds from different instruments.
- Ensemble: The ability of the performers to hear and respond to music that they hear from the other performers. A time delay of < 50 msec between notes requires the performing stage to be < 50 ft.

14 Electrical reproduction of sound

- A basic understanding of electricity and magnetism is needed to describe the functioning of musical recording and reproduction equipment.
- Here, we briefly describe electrical circuits (voltage, current, resistance, power, $\cdots$) and also Faraday's law of induction, which underlies the operation of microphones and loudspeakers.

14.1 Basic electricity

- When a voltage source (such as a battery) is connected to a load (such as a flashlight bulb) via a closed path (called a circuit), an electric current flows (electric charges in motion).
- Voltage is denoted by V and is measured in volts (V). The load has a *resistance*, denoted by R, which is measured in ohms (Ω). Current is denoted by I and is measured in ampères or amps (A).
- For certain materials, V, I, and R are related by

$$R = V/I \quad \text{or, equivalently,} \quad V = IR \tag{14.1}$$

  This is called *Ohm's law of electricity*, not to be confused with *Ohm's law of hearing*, even though it's by the same guy.
- For a battery, which has a *polarity* (+ and − terminals), the current I flows in only one direction, which we take to be from the + terminal of the battery to the − terminal. This is called a *direct current* (DC) circuit. (See the left-hand panel of Figure 58.)

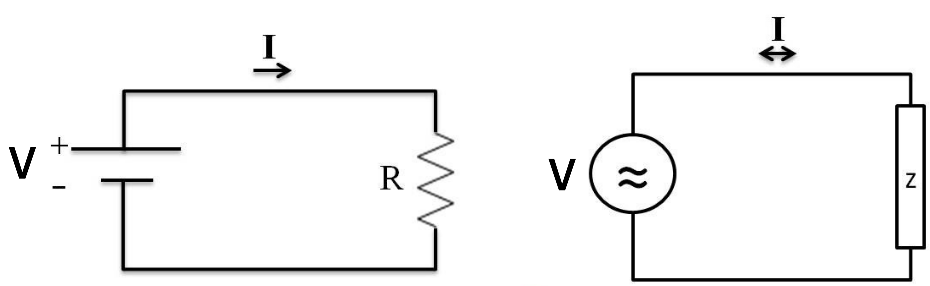


Figure 58: Left: Direct current (DC) circuit, consisting of a voltage source V and a load having resistance R. Right: Alternating current (AC) circuit, consisting of a voltage source V and a load having impedance Z. For a DC circuit, the current I flows in only one direction. For an AC circuit, the current I flows alternately clockwise and counterclockwise around the circuit. (Figures from "PHYS 1406: Physics of Sound & Music" Course Guide by Prof. Borst.)

- For a household wall outlet, the voltage V alternates sinusoidally with time (60 cycle/sec or 60 Hz), producing a current I that also alternates sinusoidally, traveling both clockwise and counterclockwise around the circuit. Such a circuit is called an *alternating current* (AC) circuit. (See the right-hand panel of Figure 58.)
- For AC circuits, there is a relation

$$V = IZ \tag{14.2}$$

  which is similar to Ohm's law, but it involves the *impedance* Z of the load.
- Impedance is basically a generalization of resistance, which holds for electrical devices such as *capacitors* (closely-spaced metal plates, which store electric charge) and *inductors* (coils of wire, which store electric current), as well as for ordinary resistors.

- The product
$$P = VI \tag{14.3}$$
is the electrical *power* associated with a circuit. (For AC circuits, the formula is slightly more complicated than this, due to the difference between resistance and impedance.) Power is measured in Watts (W), such as that for a 100-W light bulb.

- Exercise: Using Ohm's law $V = IR$, show that
$$P = VI = I^2 R \tag{14.4}$$

- In general, power is the rate at which *work* is being done or *energy* is being produced. If we denote the work done or energy produced in time interval Δt by W, then
$$P = W/\Delta t \quad \text{or, equivalently,} \quad W = P\,\Delta t \tag{14.5}$$
Work (or energy) is measured in Joules (J), where 1 Joule $\equiv$ 1 Watt $\cdot$ sec.

- Application: A kilowatt-hour (kWh) is a convenient unit of energy that you will see on electric bills. In terms of Joules
$$1 \text{ kWh} = 1 \text{ kW} \times 1 \text{ hr} = 1000 \text{ W} \times 3600 \text{ s} = 3.6 \times 10^6 \text{ J} \tag{14.6}$$

- Exercise: Suppose that you paid \$100 for last month's electric bill at a cost of \$0.13 per kWh.

  (a) How much energy (in kWh) did you use?

  (b) What was the average power consumption (in Watt) over the month (assume 30 days).

- Answer:

  (a)
$$W = \$100 \div \$0.13/\text{kWh} = 769 \text{ kWh} \tag{14.7}$$
  (b)
$$P = \frac{W}{\Delta t} = \frac{769 \text{ kWh}}{30 \times 24 \text{ h}} = 1.1 \text{ kW} = 1,100 \text{ W} \tag{14.8}$$
This is equivalent to having eleven 100-W light bulbs on continuously for a month.

14.2 Basic magnetism

- Permanent magnets (such as refrigerator magnets) have North (N) and South (S) poles that attract pieces of iron.
- Like poles repel and unlike poles attract, similar to positive (+) and negative (−) electric charges. Isolated magnetic poles do not exist in nature (unlike isolated electric charges).
- The Earth has an intrinsic magnetic field with its South magnetic pole located near Earth's North geographic pole. A compass needle is a small magnet that aligns its North pole with Earth's South magnetic pole (and hence points toward geographic North).
- In 1820, Hans Christian Oersted discovered that an *electric current produces a magnetic field.*
- He noticed that a current-carrying wire causes a compass needle to deflect. The magnetic field *circles around* the current-carrying wire in a plane perpendicular to the direction of the current.
- A consequence of this fact is that it's possible to create an *electromagnet* by sending an electric current through a coil of wire. An electromagnet acts just like a permanent bar magnet, with N and S magnetic poles.
- Demonstration: Make an electromagnet from a screwdriver, battery, and some wire wrapped around the screwdriver.

14.3 Faraday's law

- In 1831, British scientist Michael Faraday discovered that a change in the magnetic flux through a coil of wire induces a voltage in the coil. This is called *Faraday's law of electromagnetic induction.*

- Mathematically, the induced voltage is given by

$$V = -N\,\frac{\Delta\Phi}{\Delta t} \tag{14.9}$$

  where N is the number of loops of the coil and $\Delta\Phi$ is the change in the magnetic flux passing through a loop in time interval Δt.

- Demonstration: Connect a coil of wire to a galvanometer (a sensitive current meter). Move a permanent magnet toward or away from the coil and watch how the galvanometer needle deflects. Repeat, but now keep the magnet stationary and move the coil instead. Note that only the *relative* motion of the magnet and coil is important. (See Figure 59.)

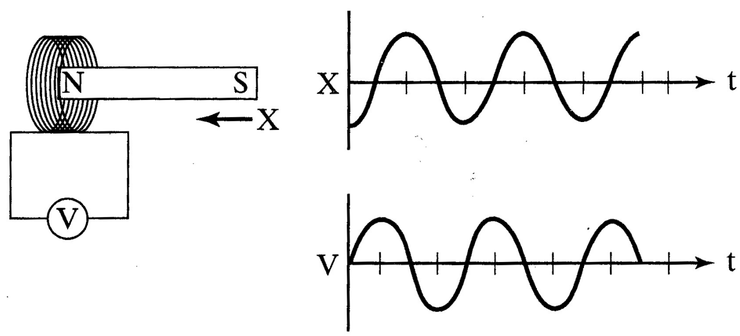


Figure 59: Illustration of Faraday's law. As a magnetic moves back and forth (X vs t) in the vicinity of a coil of wire, an alternating voltage (V vs t) is induced in the coil. (Figure from "Physics of Sound," by Berg and Stork.)

- Faraday's law of induction has had a profound influence of technology, as it underlies the operation of electric motors and generators. It also underlies the operation of certain types of microphones and loudspeakers (these will be described in the next subsection).

- Demonstrations:

  (i) Illustrate how a hand-powered AC generator works.

  (ii) Illustrate the operation of a simple DC motor constructed from electromagnets.

  (iii) Show a "do-it-yourself" DC motor constructed from a D-cell flashlight battery, a small magnet, paper clips, and a coil of (magnet) wire stripped on one side.

- Note that electric motors and electric generators are basically "inverses" of one another:

  Electric generator: mechanical energy is converted to electrical energy by physically rotating a coil in an external magnetic field. An voltage is induced in the coil by Faraday's law.

  Electric motor: electrical energy is converted to mechanical energy by sending a current through a coil of wire. This current creates an electromagnet, which interacts with the external magnetic field, causing the coil to rotate.

14.4 Microphones and loudspeakers

- Faraday's law of electromagnetic induction also underlies the operation of so-called "dynamic" microphones and loudspeakers.

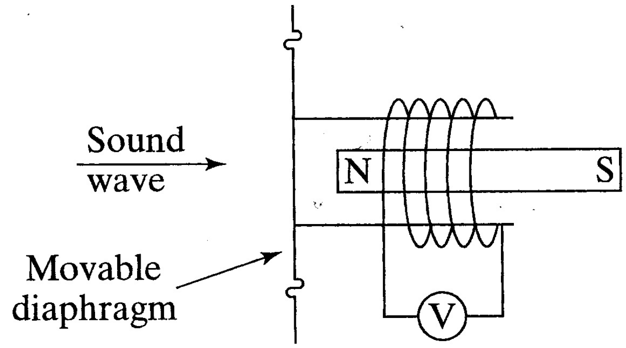


Figure 60: Schematic diagram of a dynamics microphone. The pressure deviations associated with a sound wave push back and forth on the diaphragm of the microphone. A coil of wire, which is attached to the diaphragm, thus moves in the presence of a magnetic field. The changing magnetic flux through the coil induces a voltage V in the coil which follows the fluctuations of the sound wave. (Figure from "Physics of Sound," by Berg and Stork.)

- A schematic diagram of a dynamic microphone is shown in Figure 60.
- When a sound wave impinges on the movable diaphragm, the pressure deviations in the wave cause the diaphragm to move back-and-forth in response to the sound. A coil of wire, which is attached to the diaphragm, thus moves with respect to a fixed magnetic field, inducing a time-varying voltage in the coil (according to Faraday's law) that follows the pressure deviations in the sound wave. This voltage can then be amplified and used as input to other electronics that record or transmit the sound.
- A loudspeaker works like a microphone in reverse, see Figure 61.

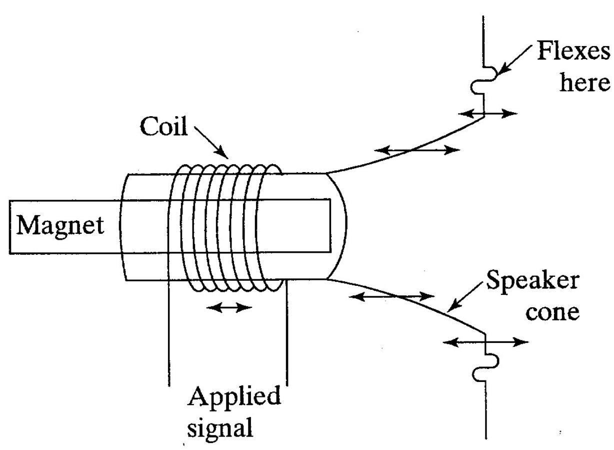


Figure 61: An time-varying electrical signal (e.g., a voltage) applied to the coil creates a magnetic field that interacts with that of the permanent magnet. The speaker cone, which is attached to the coil, moves back-and-forth in response to this interaction, thus producing a sound wave. (Figure from "Physics of Sound," by Berg and Stork.)

- A time-varying voltage, which is an electrical representation of the sound, is applied to a coil of wire that is attached to the loudspeaker cone. A current then flows in this coil creating an electromagnet that interacts with an external magnetic field, causing the loudspeaker cone to move back-and-forth. This motion creates pressure deviations in the air, which is the sound wave that we then hear.

- Figure 62 shows three different types of loudspeakers: (i) an "infinite-baffle" loudspeaker, which is mounted in either a wall or a ceiling; (i) an "acoustic suspension" loudspeaker, which is a loudspeaker cone and coil suspended inside an air-tight box; and (ii) a "tuned-port" loudspeaker, which is an acoustic suspension loudspeaker whose box has a hole in it, which emphasizes the lower (bass) frequencies.

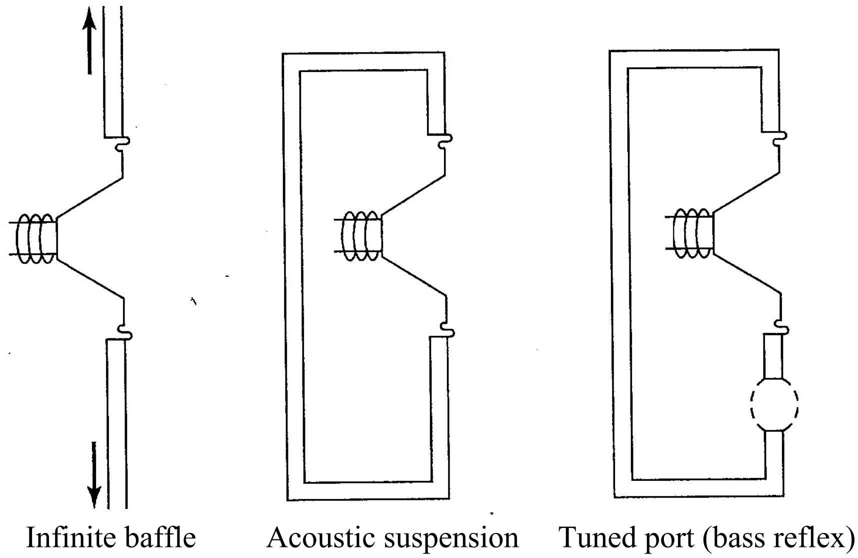


Figure 62: Three different types of loudspeakers: Infinite-baffle loudspeaker (left), acoustic-suspension loudspeaker (middle), and tuned-port loudspeaker (right). The cone of an infinite-baffle loudspeaker can be mounted in either a wall or a ceiling. (Figures from "Physics of Sound," by Berg and Stork.)

- Figure 63 shows the frequency response curves for: (a) an acoustic-suspension loudspeaker; (b) an empty box with a hole in it, which acts like a Helmholtz resonator; and (c) a tuned-port speaker, which is a combination of the frequency response curves shown in (a) and (b).
- Compared to the acoustic-suspension loudspeaker, the tuned-port loudspeaker is better at producing lower (bass) frequency sounds. Without the hole, the suspension box would need to be larger for the loudspeaker to produce these lower frequencies.

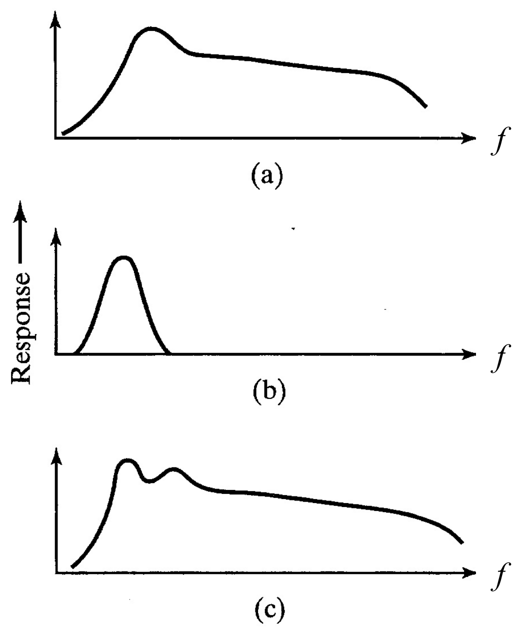


Figure 63: (a) Frequency response for an acoustic-suspension loudspeaker. (b) Frequency response of an empty box with a hole (tuned port), which acts like a Helmholtz resonator. (c) Combined frequency response for a tuned-port loudspeaker. (Figures from "Physics of Sound," by Berg and Stork.)

Part VI
Music theory

15 Musical scales and intervals

- A scale is basically a division of the octave into a succession of notes, in ascending or descending order.
- The most common scales in Western music are the *chromatic*, *diatonic*, and *pentatonic* scales, consisting of 12, 7, and 5 intervals per octave, respectively.
- In this section, we shall describe these scales and the most common musical intervals.

15.1 Chromatic scale

- A chromatic scale divides the octave into 12 intervals called *semitones* or *half-steps*. In *equal temperament* tuning (more about this in a later section), these intervals all have the same size, corresponding to a frequency ratio of
$$2^{1/12} = 1.05946 \tag{15.1}$$
- Oftentimes it is more convenient to work in term of *cents*, where a cent is the frequency interval corresponding to 1/100 of a semitone:
$$2^{1/1200} = 1.000578 \tag{15.2}$$
- The human ear can distinguish frequencies that differ by about 10 cents, which corresponds to a frequency ratio of $2^{10/1200} = 1.00579$, or about a 0.5% difference in frequency. (This percentage difference in frequency is the *just noticeable difference* for pitch discrimination.)
- Figure 64 shows the corresponding notes in a chromatic scale on a piano keyboard:
$$\mathrm{C} - \mathrm{C}^\sharp - \mathrm{D} - \mathrm{E}^\flat - \mathrm{E} - \mathrm{F} - \mathrm{F}^\sharp - \mathrm{G} - \mathrm{A}^\flat - \mathrm{A} - \mathrm{B}^\flat - \mathrm{B} - \mathrm{C}'$$

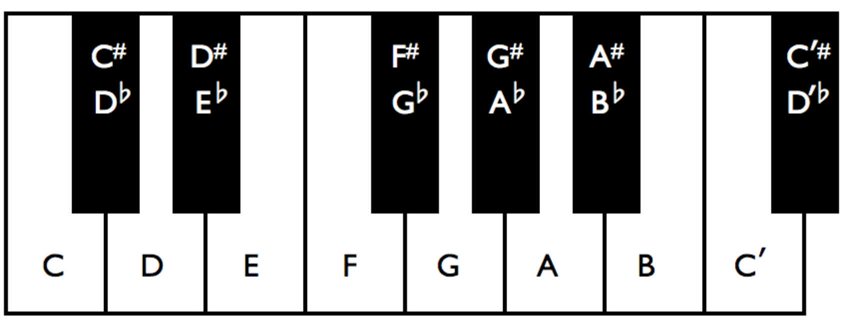


Figure 64: Notes in a chromatic scale on a piano keyboard.

- On a piano, or any instrument tuned to equal temperament, the sharps and flats are equal to one another—e.g., $\mathrm{C}^\sharp$ and $\mathrm{D}^\flat$ are tuned to the same frequency. These are called *enharmonic* notes.
- A full piano keyboard, which has 88 keys ranging from A_0 to C_8, is a basically a logarithmic frequency scale, with neighboring keys (white-black or white-white) corresponding to a frequency interval of a semitone.
- Figure 65 shows the correspondence between the treble and bass staves and their equal-tempered frequencies. This is another example of a logarithmic frequency scale.
- The staff lines are not equally-spaced in this graph, since some divisions correspond to a minor third (three semitones), while other divisions correspond to a major third (four semitones).

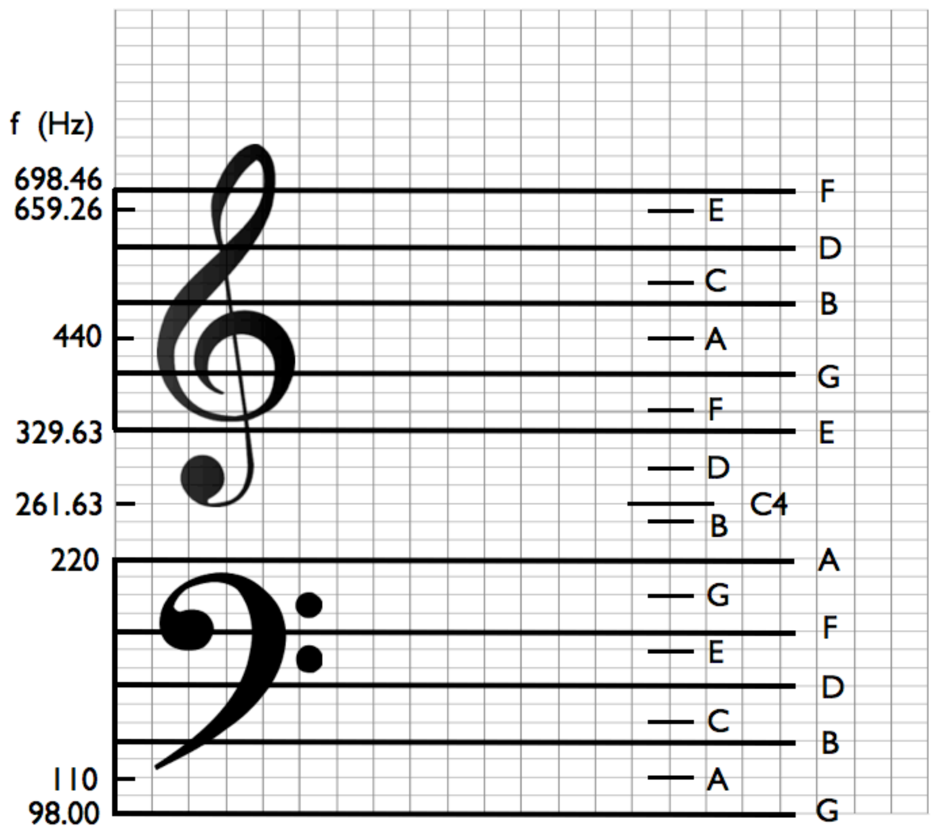


Figure 65: Correspondence between the treble and bass staves and their equal-tempered frequencies.

15.2 Diatonic scale

- A diatonic scale divides the octave into 7 intervals consisting of both tones and semitones. The order of tones and semitones defines the *major* and *minor* interval orders.

- The diatonic major interval order is T-T-S-T-T-T-S (2-2-1-2-2-2-1). This is the standard

$$\text{do} - \text{re} - \text{mi} - \text{fa} - \text{sol} - \text{la} - \text{ti} - \text{do}$$

  interval order.

- The diatonic minor interval order is T-S-T-T-S-T-T (2-1-2-2-1-2-2).

15.3 Musical intervals

- The most important musical interval is the octave, corresponding to a frequency ratio of exactly 2.

- Other common musical intervals are the fifth (C-G), fourth (C-F), major third (C-E), and minor third (C-$E^\flat$).

- Table 8 is a list of these and other musical intervals and their corresponding *just* and equal-tempered frequency ratios. (We will describe the just temperament tuning system in a later section.)

Interval	# semitones	Just freq ratio	ET freq ratio	Difference (cents)	Example
Octave	12	$2:1 = 2.000$	2.000	0	C-C′
Fifth	7	$3:2 = 1.500$	1.498	2	C-G
Fourth	5	$4:3 = 1.333$	1.335	-2	C-F, G-C′
Major third	4	$5:4 = 1.250$	1.260	-14	C-E
Minor third	3	$6:5 = 1.200$	1.189	16	C-$E^\flat$, A-C′

Table 8: Common musical intervals and their corresponding just and equal-tempered frequency ratios.

- Mathematically, a musical interval corresponds to a *ratio* of frequencies. Multiplication of frequency ratios corresponds to addition of frequency intervals.

- For example, an octave equals a fifth plus a fourth, and a fifth equals a major third plus a minor third:

$$\frac{3}{2} \cdot \frac{4}{3} = \frac{2}{1}, \qquad \frac{5}{4} \cdot \frac{6}{5} = \frac{3}{2} \tag{15.3}$$

15.4 Harmonic series

- Notes in a harmonic series have frequencies that are integer multiples of some fundamental frequency: $f_n = nf_1$, where $n = 1, 2, \cdots$.

- Successive harmonics can be related to the musical intervals as shown in Figure 66, where the fundamental frequency corresponds to A_2.

- A comparison of the harmonic frequencies and equal-tempered frequenices are given in Table 9.

- Note that the largest discrepancy in equal temperament tuning is for the 7th harmonic.

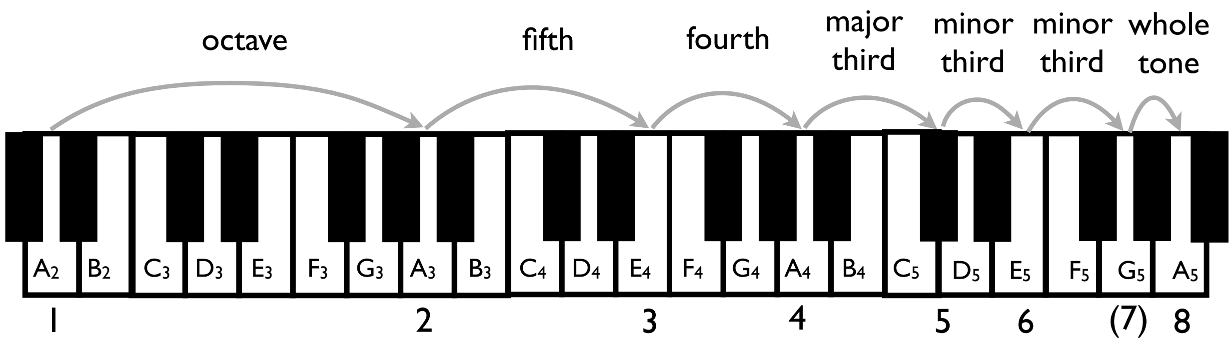


Figure 66: Notes on a piano keyboard corresponding to the first eight harmonics of A_2.

Harmonic	Exact freq (Hz)	Equal-tempered freq (Hz)	Difference (cents)	Piano note
1	110	110.00	0	A_2
2	220	220.00	0	A_3
3	330	329.63	2	E_4
4	440	440.00	0	A_4
5	550	554.37	−14	$C_5^\sharp$
6	660	659.26	2	E_5
7	770	783.99	−31	G_5
8	880	880.00	0	A_5

Table 9: Relation between harmonic frequencies and equal-tempered frequencies.

15.5 Chords

- Chords: There are three major chords (triads) having just frequency ratios 4:5:6:

$$\mathrm{C}-\mathrm{E}-\mathrm{G}\,,\quad \mathrm{F}-\mathrm{A}-\mathrm{C}\,,\quad \mathrm{G}-\mathrm{B}-\mathrm{D}$$

  These three notes correspond to a major third followed by a minor third.

- There are two minor chords (triads) having just frequency ratios 10:12:15:

$$\mathrm{E}-\mathrm{G}-\mathrm{B}\,,\quad \mathrm{A}-\mathrm{C}-\mathrm{E}$$

  These three notes correspond to a minor third followed by a major third.

15.6 Circle of fifths

- A useful construct is the so-called *circle of fifths* shown in Figure 67.
- Successive notes on the circle are separated by a musical fifth.
- In equal temperament, the notes next to one another have the same frequency; this is not the case for other tuning systems such as Pythagorean or just temperament, as we shall see in a later section. For these other tuning systems, the circle of fifths does *not* close.

15.7 Major and minor key signatures

- For example, the C-major and C-minor diatonic scales are:

$$\mathrm{C}-\mathrm{D}-\mathrm{E}-\mathrm{F}-\mathrm{G}-\mathrm{A}-\mathrm{B}$$

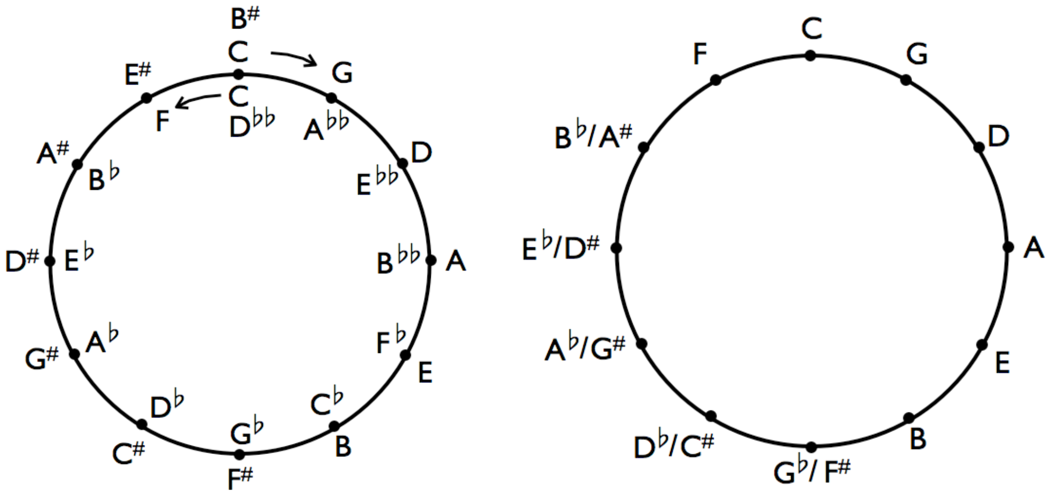


Figure 67: Circle of fifths. Left panel: General circle of fifths. Right panel: Simplified circle of fifths for equal temperament tuning.

and

$$\mathrm{C}-\mathrm{D}-\mathrm{E}^{\flat}-\mathrm{F}-\mathrm{G}-\mathrm{G}^{\sharp}-\mathrm{A}^{\sharp}$$

- The notes in the G-major scale are the same as for C-major except for F$^\sharp$:

$$\mathrm{G}-\mathrm{A}-\mathrm{B}-\mathrm{C}-\mathrm{D}-\mathrm{E}-\mathrm{F}^{\sharp}$$

- The notes in the F-major scale are the same as for C-major except for B$^\flat$:

$$\mathrm{F}-\mathrm{G}-\mathrm{A}-\mathrm{B}^{\flat}-\mathrm{C}-\mathrm{D}-\mathrm{E}$$

- Note that the A-minor scale is

$$\mathrm{A}-\mathrm{B}-\mathrm{C}-\mathrm{D}-\mathrm{E}-\mathrm{F}-\mathrm{G}$$

so the notes are *exactly* the same as for C-major. But because the tonic (and tonal interval order) are different for these two scales, they sound differently.

15.8 Pentatonic scale

- A pentatonic scale is probably the oldest division of the octave. It was developed independently in many different cultures.
- A pentatonic scale divides the octave into 5 intervals (3 intervals are whole tones and 2 intervals are three semitones wide).
- Examples of pentatonic scales in major interval order are

$$\mathrm{C}-\mathrm{D}-\mathrm{E}-\mathrm{G}-\mathrm{A} \tag{15.4}$$

and

$$\mathrm{F}^{\sharp}-\mathrm{G}^{\sharp}-\mathrm{A}^{\sharp}-\mathrm{C}^{\sharp}-\mathrm{D}^{\sharp} \tag{15.5}$$

which are just the black keys on a piano.

- The pentatonic scale also has minor and various modal interval orders. The major interval order is: 2-2-3-2-3. The minor interval order is: 3-2-2-3-2.

- A pentatonic scale can be constructed from just fifths and the octave (or, equivalently, from just fifths and fourths) as a subset of Pythagorean temperament (see next section).

- A guitar (or any 6-string instrument) can be tuned to a pentatonic scale *using only the guitar to do the tuning* as described below:

  [To play a note *an octave above* an open note, touch the string at its half-way point (12th fret on a guitar) as you pluck it with your other hand. Similarly, to play a note *a fifth above* an open note, touch the string at its one-third point (7th fret on a guitar) as you pluck it with your other hand.]

  Instructions (tuning order: string 6 (top), 1 (bottom), 3, 5, 2, 4):

  - Adjust the tension of string 6 to any value that gives it a nice sound (let us assume it's a C)

  - Tune the open note of string 1 to the note an octave above string 6 (then string 1 will be C′)

  - Tune the open note of string 3 to the note a fifth above string 6 (then string 3 will be G)

  - Tune the note an octave above string 5 to the note a fifth above string 3 (then string 5 will be D)

  - Tune the open note of string 2 to the note a fifth above string 5 (then string 2 will be A)

  - Tune the note an octave above string 4 to the note a fifth above string 2 (then string 4 will be E)

  The result is the guitar tuned to the pentatonic scale in major interval order 2-2-3-2-3; for this example, it is the key of C corresponding to the notes C-D-E-G-A-C′.

- Demonstration: Examples of music played using a pentatonic scale are "Amazing Grace," "Auld Lang Syne," "My Girl" (by the Temptations), and Etude Op.10 No.5 (Black Keys) by Chopin.

  Watch YouTube video of pianist Vladimir Horowitz playing Chopin's etude (http://www.youtube.com/watch?v=jaMA8LWW3C0).

15.9 Normal tuning of a violin and a guitar

- A violin has four strings. They are normally tuned to

$$\mathrm{G}_3 - \mathrm{D}_4 - \mathrm{A}_4 - \mathrm{E}_5 \tag{15.6}$$

  which are all separated by fifths.

- A guitar has six strings. They are normally tuned to

$$\mathrm{E}_2 - \mathrm{A}_2 - \mathrm{D}_3 - \mathrm{G}_3 - \mathrm{B}_3 - \mathrm{E}_4 \tag{15.7}$$

  which are all separated by fourths except for G_3 to B_3, which is separated by a major third.

- Note that using perfect fourths (ratio 4:3) and a perfect major third (ratio 5:4), the notes E_2 and E_4 are not *exactly* two octaves apart, since

$$\frac{4}{3}\cdot\frac{4}{3}\cdot\frac{4}{3}\cdot\frac{5}{4}\cdot\frac{4}{3} = \frac{80}{81}\cdot 4 \tag{15.8}$$

  which is slightly less than 4.

- Slight discrepancies, such as this, were ultimately the reason for introducing the equal temperament (ET) tuning system (more about this later). The product of four equal-tempered fourths and an equal-tempered major third equals two octaves exactly.

16 Tuning systems

- An International Conference in London in 1939 established $A_4 = 440$ Hz as the standard reference frequency. Before that time, A_4 was tuned to different frequencies ranging from 422.5 Hz (Handel) to 444 Hz, depending on composer, orchestra, ...
- A tuning system is basically an assignment of precise frequencies to all the other musical notes.
- Since notes an octave apart have a frequency ratio of exactly 2 in all tuning systems, it suffices to assign frequencies to all the notes within a single octave scale. The frequencies of notes outside this reference octave scale are simply related to the frequencies of the notes within the reference octave.
- Example: The note A_3 has frequency $440/2 = 220$ Hz in all tuning systems.
- In this section, we describe three different tuning systems: equal temperament, Pythagorean temperament, and just temperament.
- As we shall see below, there is no such thing as a perfect tuning system. Each tuning system has its own advantages and disadvantages.

16.1 Equal temperament

- Equal temperament is the most common tuning system in use today.
- In equal temperament tuning, all semitone intervals have the same size, corresponding to a frequency ratio of

$$2^{1/12} = 1.05946 \tag{16.1}$$

  A *whole tone* is exactly equal to two semitones.
- An equal-tempered chromatic scale was originally proposed independently by Vincenzo Galilei (father of Galileo Galilei) and Chu Tsai-Yu (a Chinese scholar) around 1580.
- Equal temperament didn't become popular until the 1700's. "The Well-Tempered Clavier" by Johannes Sebastian Bach demonstrated the usefulness of equal-temperament. Using all 24 major and minor keys, this piece sounded good only in equal-tempered tuning, and not in the other tuning systems that we will describe below.
- Table 10 gives the frequency ratios for a chromatic scale in equal-temperament tuning.

Note	ET freq ratio
C	$2^{0/12} = 1.000$
$C^\sharp/D^\flat$	$2^{1/12} = 1.059$
D	$2^{2/12} = 1.122$
$D^\sharp/E^\flat$	$2^{3/12} = 1.189$
E	$2^{4/12} = 1.260$
F	$2^{5/12} = 1.335$
$F^\sharp/G^\flat$	$2^{6/12} = 1.414$
G	$2^{7/12} = 1.498$
$G^\sharp/A^\flat$	$2^{8/12} = 1.587$
A	$2^{9/12} = 1.682$
$A^\sharp/B^\flat$	$2^{10/12} = 1.782$
B	$2^{11/12} = 1.888$
C'	$2^{12/12} = 2.000$

Table 10: Notes in a chromatic scale and the corresponding equal-temperament frequency ratios.

- On a piano, or any instrument tuned to equal temperament, the sharps and flats are equal to one another—e.g., $C^\sharp$ and $D^\flat$ are tuned to the same frequency. These are called *enharmonic* notes.
- NOTE!! This equality of sharps and flats is *not* true, in general, for other tuning systems such as Pythagorean or just temperament (see below). For these systems, one must first decide to tune, for example, to either $C^\sharp$ or $D^\flat$, etc.
- It is common to choose:

$$C - C^\sharp - D - E^\flat - E - F - F^\sharp - G - A^\flat - A - B^\flat - B - C'$$

  as the fundamental notes in a chromatic scale.
- <u>Exercise</u>: Show that middle C in equal temperament has a frequency of 261.63 Hz:

$$C_4: \quad A_3 \cdot 2^{3/12} = \frac{440 \text{ Hz}}{2} \cdot 2^{3/12} = 261.63 \text{ Hz} \tag{16.2}$$

16.2 Pythagorean temperament

- Pythagorean temperament is a tuning system constructed from just fifths and the the octave (or, equivalently, from just fifths and fourths).
- Using the circle of fifths (cf. Fig. 67), we can tune a chromatic scale to Pythagorean temperament as illustrated in Figure 68.

$$A^\flat \xleftarrow{2/3} E^\flat \xleftarrow{2/3} B^\flat \xleftarrow{2/3} F \xleftarrow{2/3} C \xrightarrow{3/2} G \xrightarrow{3/2} D \xrightarrow{3/2} A \xrightarrow{3/2} E \xrightarrow{3/2} B \xrightarrow{3/2} F^\# \xrightarrow{3/2} C^\#$$

Figure 68: A portion of the circle of fifths for calculating a chromatic scale of frequencies in Pythagorean temperament.

- Applying the appropriate number of factors of 3/2 (or 2/3) and then scaling by 1/2 (or 2) to bring the resulting frequency to within an octave of C, we obtain the numbers in Table 11.

Note	Pyth freq ratio	ET freq ratio	Difference (cents)
C	$1:1 = 1.000$	1.000	0
$C^\sharp$	$2187:2048 = 1.068$	1.059	14
D	$9:8 = 1.125$	1.122	4
$E^\flat$	$32:27 = 1.185$	1.189	−6
E	$81:64 = 1.266$	1.260	8
F	$4:3 = 1.333$	1.335	−2
$F^\sharp$	$729:512 = 1.424$	1.414	12
G	$3:2 = 1.500$	1.498	2
$A^\flat$	$128:81 = 1.580$	1.587	−8
A	$27:16 = 1.688$	1.682	6
$B^\flat$	$16:9 = 1.778$	1.782	−4
B	$243:128 = 1.898$	1.888	10
C'	$2:1 = 2.000$	2.000	0

Table 11: Pythagorean chromatic scale and comparison to equal-tempered frequency ratios. The Pythagorean diatonic whole-tone interval is 9:8. The Pythagorean chromatic semitone interval is either 256:243 or 2187:2048.

- Table 12 compares Pythagorean and equal-tempered frequencies for all notes on the circle of fifths traversed once in the clockwise direction and once in the counter-clockwise direction.

Note	Pyth freq ratio	ET freq ratio	Difference (cents)
$D^{\flat\flat}$	0.987	1.000	-23
C	1.000	1.000	0
$D^{\flat}$	1.053	1.059	-10
$C^{\sharp}$	1.068	1.059	14
$E^{\flat\flat}$	1.110	1.122	-20
D	1.125	1.122	4
$E^{\flat}$	1.185	1.189	-6
$D^{\sharp}$	1.201	1.189	18
$F^{\flat}$	1.249	1.260	-16
E	1.266	1.260	8
F	1.333	1.335	-2
$E^{\sharp}$	1.352	1.335	22
$G^{\flat}$	1.405	1.414	-12
$F^{\sharp}$	1.424	1.414	12
$A^{\flat\flat}$	1.480	1.498	-22
G	1.500	1.498	2
$A^{\flat}$	1.580	1.587	-8
$G^{\sharp}$	1.602	1.587	16
$B^{\flat\flat}$	1.665	1.682	-18
A	1.688	1.682	6
$B^{\flat}$	1.778	1.782	-4
$A^{\sharp}$	1.802	1.782	20
$C^{\flat}$	1.873	1.888	-14
B	1.898	1.888	10
C'	2.000	2.000	0
$B^{\sharp}$	2.027	2.000	23

Table 12: A comparison of notes in the Pythagorean and equal-temperament tuning systems going once around the circle of fifths in both the clockwise and counter-clockwise directions.

- Note that the circle of fifths doesn't close since $B^{\sharp}$ and C' are not equal to one another. In addition, $C^{\sharp}$ doesn't equal $D^{\flat}$, $C^{\flat}$ doesn't equal B, $B^{\flat\flat}$ doesn't equal A, etc.
- Pythagorean comma: ratio of frequencies for $B^{\sharp}$ and C':

$$\frac{B^{\sharp}}{C'} = \frac{(3/2)^{12}(1/2)^{6}}{2/1} = \frac{(3/2)^{12}}{2^{7}} = \frac{531441}{524288} = 1.0136 \tag{16.3}$$

- The above is equivalent to the statement that 12 fifths doesn't exactly equal 7 octaves. They differ by a ratio of 1.0136, which corresponds to an interval of 23 cents.
- Pythagorean temperament (by construction) has perfect fourths and fifths in the key of C, but in the key of $C^{\sharp}$, the fifth from $C^{\sharp}$ to $A^{\flat}$ is too flat by a factor of the Pythagorean comma:

$$\frac{A^{\flat}/C^{\sharp}}{3/2} = \frac{(128/81)/(2187/2048)}{3/2} = \frac{1}{1.0136} \tag{16.4}$$

This is called a *wolf fifth* due to its 'howling' sound (which I can't hear!).

- Demonstration: Use the matlab routine playinterval.m to compare the fifths C to G and $C^{\sharp}$ to $A^{\flat}$ in Pythagorean temperament.

- In addition, a Pythagorean third (e.g., from C to E) differs signicantly from a just major third that has a frequency ratio of 5/4. This difference is called the *syntonic comma*:

$$\delta = \frac{81/64}{5/4} = \frac{81}{80} = 1.0125 \tag{16.5}$$

  Thus, a Pythagorean third is too sharp by just over 1% or approximately 22 cents.

16.3 Just temperament

- Just temperament is a tuning system constructed from perfect fifths, major thirds, and the octave (or, equivalently, from perfect fifths, perfect fourths, major thirds, and minor sixths). In the key of C-major, it has three beatless major chords and two beatless minor chords.
- Using Figure 69 as a schematic for going up and down in frequency by perfect fifths and major thirds, we obtain the frequency ratios given in Table 13.

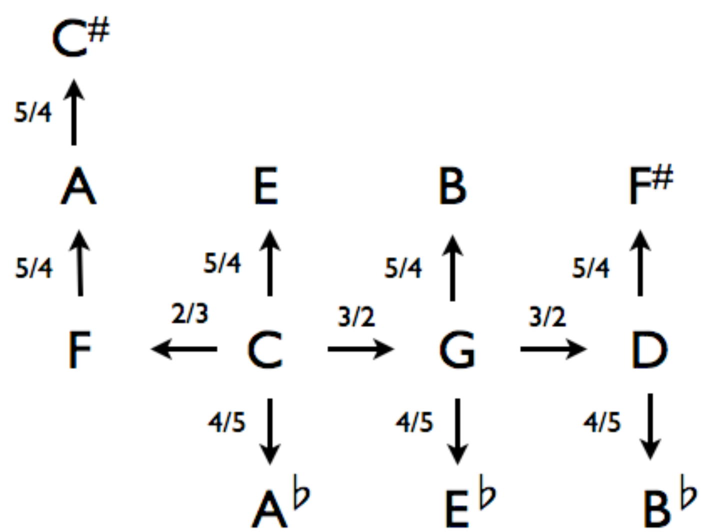


Figure 69: Schematic for calculating a chromatic scale of frequencies in just temperament.

- Just temperament is arguably the best possible tuning system, provided one stays within the key in which you tuned to, e.g., the key of C for the above tuning.
- But just as we saw for Pythagorean temperament, just temperament becomes increasingly out of tune as one move away from the key of C. For example, the fifth $C^\sharp$ to $A^\flat$ in the key of $C^\sharp$ differs from a perfect fifth by 41 cents.

16.4 Comparison of different tuning systems

- Figure 70 compares the frequency ratios for a chromatic scale for the three different tuning systems that we considered in the previous subsections. One can easily see from this figure that the fifth $C^\sharp$ to $A^\flat$ is too flat in Pythagorean tuning and too sharp in just temperament tuning.
- Table 14 quantifies this comparison, giving the frequency ratios of two different fifths (C to G and $C^\sharp$ to $A^\flat$) in the three different tuning systems that we considered. The last column gives the difference in cents between these frequency ratios and 3/2 for a just perfect fifth.

Note	Just freq ratio	ET freq ratio	Difference (cents)
C	$1:1=1.000$	1.000	0
$C^\sharp$	$25:24=1.042$	1.059	-29
D	$9:8=1.125$	1.122	4
$E^\flat$	$6:5=1.200$	1.189	16
E	$5:4=1.250$	1.260	-14
F	$4:3=1.333$	1.335	-2
$F^\sharp$	$45:32=1.406$	1.414	-10
G	$3:2=1.500$	1.498	2
$A^\flat$	$8:5=1.600$	1.587	14
A	$5:3=1.667$	1.682	-16
$B^\flat$	$9:5=1.800$	1.782	18
B	$15:8=1.875$	1.888	-12
C'	$2:1=2.000$	2.000	0

Table 13: Just chromatic scale and comparison to equal-tempered frequency ratios. The just diatonic whole-tone interval is either 9:8 or 10:9. The just chromatic semitone interval is either 16:15, 25:24, or 27:25.

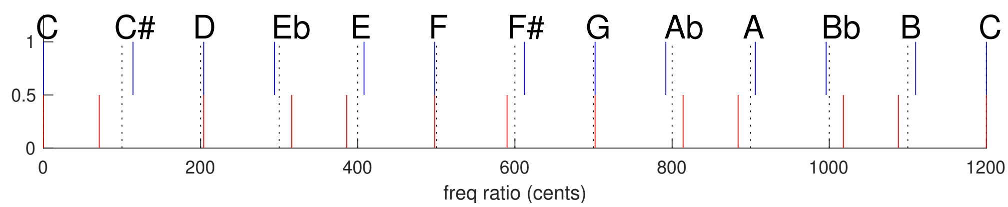


Figure 70: Frequency ratios for a chromatic scale for the three different tuning systems that we considered. The horizontal axis is the frequency ratio expressed in cents (1200 cents equals one octave). The dotted vertical lines correspond to equal-tempered tuning (equally spaced at 0, 100, 200, $\cdots$ 1200 cents); the blue and red vertical lines correspond to the Pythagorean and just temperament tuning systems, respectively.

Fifth	Temperament	Freq ratio	Difference (cents)
C-G	equal	1.498	-2
C-G	pyth	1.500	0
C-G	just	1.500	0
$C^\sharp$-$A^\flat$	equal	1.498	-2
$C^\sharp$-$A^\flat$	pyth	1.480	-23
$C^\sharp$-$A^\flat$	just	1.536	41

Table 14: Frequency ratios for two different fifths for the three different tuning systems that we considered. The last column is the difference in cents from a perfect fifth, which has a frequency ratio of 3/2.

- Demonstration: Use the matlab routine fifths.m to compare the two different fifths, C to G and $C^\sharp$ to $A^\flat$, in the three different tuning systems that we considered.

- There is no perfect tuning system. Each system makes some compromise:

  i) Equal temperament: Music in equal temperament can be played equally-well in all keys at the expense of imperfect intervals in all keys—e.g., a fifth doesn't have a frequency ratio of 3/2, a major third doesn't have a frequency ratio of 5/4, etc. The circle of fifths closes.

  ii) Pythagorean tuning has perfect fifths in the key of C, but major thirds that are sharper than just major thirds by the syntonic comma $\delta = 1.0125$ (22 cents) even in the key of C. This means that major chords like C-E-G sound out of tune. Also, even fifths are no longer perfect fifths when one moves away from the key of C—e.g., the wolf fifth $C^\sharp$ to $A^\flat$ is 23 cents too flat.

  iii) Just temperament is the ideal tuning system in the key of C having perfect fifths and major thirds with frequency ratios of 3/2 and 5/4, respectively. But just like Pythagorean tuning, just temperament gets progressively worse as one moves away from the key of C—e.g., the fifth $C^\sharp$ to $A^\flat$ is now 41 cents too sharp.

A Mathematical details

- Logarithm identities:

$$\log(ab) = \log a + \log b \tag{A.1}$$
$$\log(a/b) = \log a - \log b \tag{A.2}$$
$$\log(a^b) = b \log a \tag{A.3}$$
$$a^{\log b} = b^{\log a} \tag{A.4}$$

  These are true for logarithms with respect to any base.

- Trigonometric identities:

$$\sin A \, \sin B = \frac{1}{2}\left[\cos(A - B) - \cos(A + B)\right] \tag{A.5}$$
$$\cos A \, \cos B = \frac{1}{2}\left[\cos(A - B) + \cos(A + B)\right] \tag{A.6}$$
$$\sin A \, \cos B = \frac{1}{2}\left[\sin(A - B) + \sin(A + B)\right] \tag{A.7}$$

- Adding two sinusoids (same frequency):

  The sum of two sine functions having the same frequency (but possibly different amplitudes and phases) yields another sinusoid with the *same* frequency:

$$A_1 \sin(2\pi f t + \phi_1) + A_2 \sin(2\pi f t + \phi_2) = A \sin(2\pi f t + \phi) \tag{A.8}$$

  where

$$A = \sqrt{A_1^2 + A_2^2 + 2A_1A_2 \cos(\phi_1 - \phi_2)}\,, \quad \tan\phi = \frac{A_1 \sin\phi_1 + A_2 \sin\phi_2}{A_1 \cos\phi_1 + A_2 \cos\phi_2} \tag{A.9}$$

- Adding two sinusoids (different frequencies):

  The sum of two sinusoids with the same amplitude but *different* frequencies and phases is given by

$$A \sin(2\pi f_1 t + \phi_1) + A \sin(2\pi f_2 t + \phi_2) = 2A \cos\left(\pi\Delta f t + \frac{(\phi_2 - \phi_1)}{2}\right) \sin\left(2\pi \bar{f} t + \frac{(\phi_1 + \phi_2)}{2}\right) \tag{A.10}$$

  where

$$\bar{f} = \frac{1}{2}(f_1 + f_2)\,, \quad \Delta f = f_2 - f_1 \tag{A.11}$$

  If $\bar{f}$ is in the audio range (between about 20 Hz and 20 kHz) and $|\Delta f| \lesssim 15$ Hz, we hear *beats* with frequency

$$f_{\text{beat}} = |\Delta f| = |f_2 - f_1| \tag{A.12}$$

  The beats are amplitude modulations of a tone with average frequency $\bar{f}$.

  Note that the modulated signal is periodic only if $f_1/f_2 = m/n$ for some positive integers m and n (*comensurate* frequencies). For such a case, the period of the modulated signal is given by $T = mT_1 = nT_2$, where T_1 and T_2 are the periods of the two component sinusoids—i.e., $T_1 = 1/f_1$ and $T_2 = 1/f_2$.

- Amplitude modulation:

  The product of two sine functions having different frequencies and phases is given by:

$$\begin{aligned}\sin(2\pi f_1 t + \phi_1) \sin(2\pi f_2 t + \phi_2) = &\frac{1}{2}\big(\cos\left[2\pi(f_2 - f_1)t + (\phi_2 - \phi_1)\right] \\ &- \cos\left[2\pi(f_1 + f_2)t + (\phi_1 + \phi_2)\right] \big)\end{aligned} \tag{A.13}$$

  Note the presence of *sidebands* at the sum and difference frequencies $f_2 - f_1$ and $f_2 + f_1$. This is a non-linear combination of sinusoids, as opposed to a linear combination such as addition of two sinusoids.

- Displacement of air molecules:

  For a sinusoidal wave propagating in the x-direction, the displacement of an air molecule away from its equilibrium position at x is given by:

$$s(x,t) = s_m \sin(kx - \omega t) \tag{A.14}$$

  where $k = 2\pi/\lambda$ and $\omega = 2\pi/T$. (Note: $\omega/k = v$.) It follows that

$$\frac{\Delta s}{\Delta x} = k s_m \cos(kx - \omega t) \,, \quad u \equiv \frac{\Delta s}{\Delta t} = -\omega s_m \cos(kx - \omega t) \tag{A.15}$$

  Thus,

$$\frac{\Delta s}{\Delta x} = -\frac{k}{\omega}\frac{\Delta s}{\Delta t} = -\frac{1}{v}\frac{\Delta s}{\Delta t} = -\frac{u}{v} \tag{A.16}$$

- Relationship between $\Delta V/V$ and $\Delta s/\Delta x$:

$$V = A\Delta x \,, \quad \Delta V = A\Delta s \quad \Rightarrow \quad \frac{\Delta V}{V} = \frac{\Delta s}{\Delta x} \tag{A.17}$$

- Relationship between pressure p and displacement s:

$$p = -B\frac{\Delta V}{V} \,, \quad \frac{\Delta V}{V} = \frac{\Delta s}{\Delta x} = -\frac{u}{v} \,, \quad v = \sqrt{\frac{B}{\rho}} \quad \Rightarrow \quad p = \rho v u = \rho v \frac{\Delta s}{\Delta t} \tag{A.18}$$

  Thus, p is proportional to u, which means that p and s are 90 degrees out of phase with one another.

- Relationship between intensity I and pressure p:

$$I = (\text{energy density}) \cdot v = 2 \cdot \frac{\frac{1}{2} m u^2}{V} \cdot v = \rho u^2 v = \frac{p^2}{\rho v} \tag{A.19}$$

  Thus, I is proportional to p^2.

- Fourier series:

  Any periodic function $y(t)$ with period T can be written as a sum of sines and cosines:

$$y(t) = A_0 + [a_1 \sin(2\pi f_1 t) + b_1 \cos(2\pi f_1 t)] + [a_2 \sin(2\pi f_2 t) + b_2 \cos(2\pi f_2 t)] + \cdots \tag{A.20}$$

  where $f_n = n f_1$ ($n = 1, 2 \cdots$) are harmonic frequencies of the fundamental $f_1 = 1/T$.

  Alternatively, one can write

$$y(t) = A_0 + A_1 \sin(2\pi f_1 t + \phi_1) + A_2 \sin(2\pi f_2 t + \phi_2) + \cdots \tag{A.21}$$

  where A_n and ϕ_n are the amplitudes and phases of the sinusoids.

  In terms of a_n and b_n,

$$A_n = \sqrt{a_n^2 + b_n^2} \,, \quad \tan\phi_n = b_n/a_n \tag{A.22}$$

- Plucked guitar string:

  The vibrations of a plucked guitar string (length L) can be written as a sum of standing wave vibrations

$$y(x,t) = A_1 \sin(2\pi x/\lambda_1)\cos(2\pi f_1 t) + A_2 \sin(2\pi x/\lambda_2)\cos(2\pi f_2 t) + \cdots \tag{A.23}$$

  where A_n ($n = 1, 2, \cdots$) are the amplitudes of the standing wave vibrations, having wavelengths $\lambda_n = 2L/n$ and frequencies $f_n = v/\lambda_n = nv/2L = n f_1$ (v being the wave velocity on the string).

  Note that each term in the sum can be written as the sum of a right-moving and left-moving wave:

$$A_n \sin(2\pi x/\lambda_n)\cos(2\pi f_n t) = \frac{1}{2}\left[\sin(k_n(x - vt)) + \sin(k_n(x + vt))\right] \tag{A.24}$$

  where $k_n = 2\pi/\lambda_n = n\pi/L$.

B Matlab computer demonstrations

The following is a list of matlab routines that can be used to illustrate a particular music or physics concept.

- addsines.m: add two sine waves with the same amplitude, frequency, but with a phase offset ϕ (in degrees)
- aliasing.m: script to illustrate aliasing for a sine wave sampled too slowly
- amplitudemodulation.m: combine carrier and signal waves using amplitude modulation
- auralharmonics.m: illustrate non-linear response of the human ear
- bandpass.m: apply bandpass filter (routine by fjsimons-at-alum.mit.edu)
- bar_free_notension.m: find discrete frequencies for free bar (no tension)
- beats.m: add two sine waves with the same amplitude and phase, but different frequencies
- bowedstring.m: illustrate the physics of a bowed violin string
- comparetemperaments.m: compare equal, pythagorean, mean-tone and just temperaments
- dampeddrivengui.m: integrate $F = ma$ for damped, driven harmonic motion
- doublependulum.m: small oscillations of coplanar double pendulum
- fifths.m: compare two different fifths (C–G) and (C#–Ab) in different temperaments
- filter_recorder.m: record sound, filter, and calculate spectrogram
- forcedoscillator.m: illustrate chaotic motion of driven oscillator
- fourieranalyze.m: fourier analyze a time series
- fourierdecompose.m: calculate fourier coefficients for plucked guitar string fixed at both ends
- fouriersynthesize.m: fourier synthesizer using amplitude and phase information from first N harmonics
- fouriersynthesizeScript.m: script for fouriersynthesize
- fouriersynthesizegui.m: fourier synthesizer using first eight harmonics
- fouriersynthesizesound.m: fourier synthesize sound using amplitude and phase information from first N harmonics
- fouriersynthesizesoundScript.m: script for fouriersynthesizesound
- frequencymodulation.m: combine carrier and signal waves using frequency modulation
- harmonics.m: calculate nearest equal-tempered frequencies for first 8 harmonics of f_0
- hipass.m: apply high-pass filter (routine by fjsimons-at-alum.mit.edu)
- inharmonicity.m: calculates the deviation of the nth partial from the exact nth harmonic for a real piano string
- intervals.m: calculate frequency ratios for musical intervals and compare to equal temperament
- just.m: calculate frequency ratios in just temperament and compare to equal temperament
- loudness.m: plot loudness versus intensity curves
- lowpass.m: apply low-pass filter (routine by fjsimons-at-alum.mit.edu)

- meantone.m: calculate frequency ratios in mean-tone temperament and compare to equal temperament
- multiplesources.m: plot various quantities versus number of sources
- note2freq.m: convert note to frequency
- periodicmotion.m: numerically integrate 1-D equation for periodic motion where $F(x) = -\text{sign}(x)k|x|^p$
- phasemodulation.m: combine carrier and signal waves using phase modulation
- playchord.m: play three notes in unison
- playchromaticscale.m: play chromatic scale
- playdiatonicscale.m: play diatonic scale in major interval order
- playinterval.m: play two notes in unison (notes specified by name)
- playintervalFreq.m: play two notes in unison (notes specified by frequency)
- playnote.m: play note (note specified by name)
- playnoteFreq.m: play note (note specified by frequency)
- playrecordedsound.m: ply recorded sound (if reverse=1, reverse sound in time domain)
- playshepardtone.m: play notes an octave apart in equal temperament with intensities weighted by a gaussian centered at C4
- playsound.m: play sound (if reverse=1, reverse sound in time domain)
- playrecordedsound.m: play recorded sound that was saved to a file (if reverse=1, reverse sound in time domain)
- pluckedstring.m: illustrate physics of a plucked string
- pluckedstringharmonics.m: determine fourier coefficients for a plucked string
- pythagorean.m: calculate frequency ratios in pythagorean temperament and compare to equal temperament
- ratio2cents.m: convert ratio of frequencies to an interval in cents
- recordsound.m: record sound and fourier analyse (if reverse=1, reverse sound in time-domain)
- recordsoundandsave.m: record sound and write to .mat file
- shepardintensity.m: plot gaussian intensity distribution for shepard tones
- shepardscale.m: play shepard tone scale
- sound_spectrogram.m: record sound and calculate spectrogram
- standingwaves.m: demonstrate standing waves on a string fixed at both ends (variable freq)
- sumsines.m: gui for adding two sine waves with variable amplitude, frequency, and phase
- tape_recorder.m: record sound and fourier analyse (allow forward/reverse and speed adjustment)
- vibratingstring.m: small transverse vibrations of a stretched string fixed at both ends (approximate string as N discrete mass points)
- vibratingstringScript.m: script for running vibratingstring

C Homework 1

Show all work.

1. Write the following numbers in scientfic notation: (a) 2 gigabytes, (b) 5 msec.
2. Evaluate the following quantity both as simple fraction and as a decimal number.

$$\left(\frac{3}{2}\right)^4 \cdot \left(\frac{1}{2}\right)^2$$

   Approximately, what frequency interval does this correspond to—an octave, a perfect fifth, a major third, a minor third, $\cdots$?
3. What physical quantity corresponds to pitch?
4. One note is 10 octaves higher in pitch than another note. What frequency ratio does this correspond to?
5. Sound in dry air at 25°C travels at 346 m/s. (a) Convert this to ft/s. (b) Determine how long it would take sound to make a roundtrip to a wall 50 ft away.
6. Light travels at approximately 3×10^8 m/s. Given that the distance between the Earth and Moon is approximately $d = 384400$ km, determine how long it takes light to make a round trip from the Earth to moon. (Hint: Don't forget to convert km to m.)
7. The musical note A_2 has a fundamental frequency of 110 Hz. (a) Calculate the first eight harmonics of A_2. (b) How many octaves above A_2 is the eigth harmonic?
8. Concert A_4 has a fundamental frequency of 440 Hz. Calculate the period in (a) seconds, (b) milliseconds.
9. What's the main feature that distinguishes a musical note from noise?
10. Why does the same musical note, for example C_4, sound differently when it is played on two different instruments—e.g., a flute and a clarinet?

D Homework 2

Show all work.

1. Draw a graph of displacement versus time for general periodic motion. Show three complete cycles of the motion. Label the axes, and clearly indicate the period T and amplitude A of the motion.

2. Draw a graph of displacement versus time for simple harmonic motion. Show three complete cycles of the motion. Label the axes, and clearly indicate the period T and amplitude A of the motion.

3. What type of force produces simple harmonic motion.

4. Using the relation 2π rad $= 360°$, determine the number of radians in an angle of (a) $90°$, (b) $270°$, (c) $30°$, and (d) $60°$.

5. A 1 kg mass attached to a spring oscillates up and down with a period of $T = 2$ s. (a) What is the corresponding frequency of oscillation? (b) Is this a frequency that we can hear with our ears?

6. Suppose we replace the 1 kg mass from the previous problem with a 2 kg mass. What is the new period of oscillation?

7. A child and an adult go to a park and swing on neighboring swings, which have the same length. Suppose the adult weighs twice as much as the child. Does the child swing back-and-forth faster than the adult? If not, explain why.

8. Give two examples of a damped oscillation.

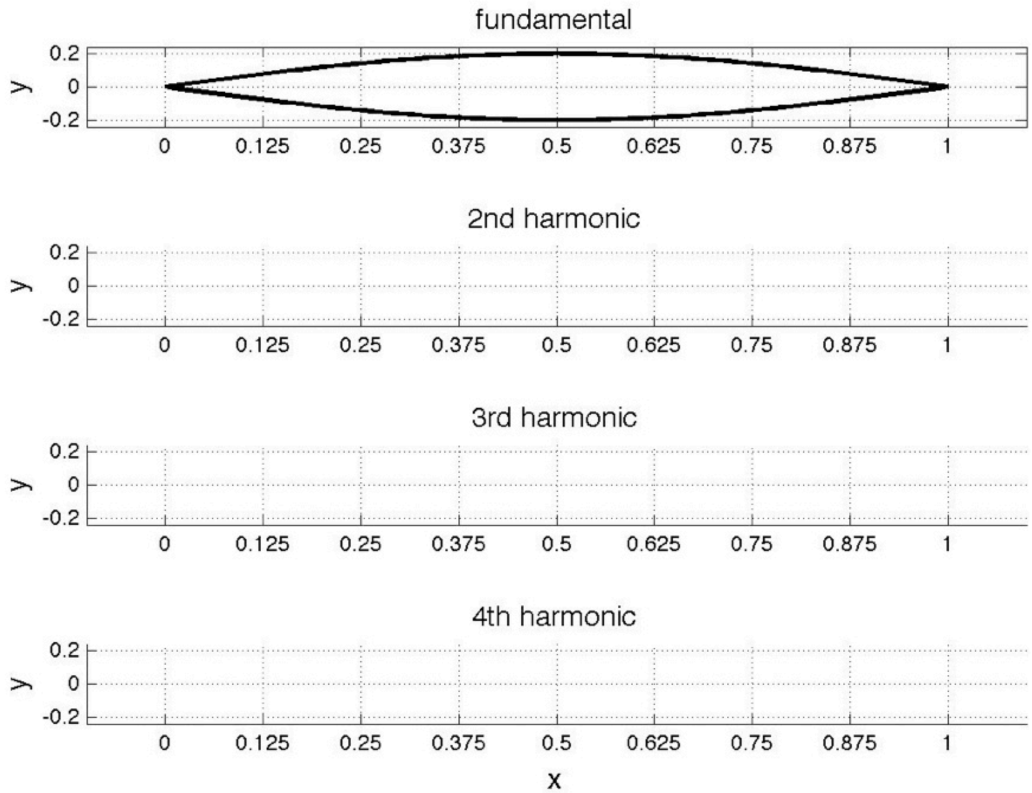
9. In order to drive a damped oscillation so that the amplitude of the resulting motion is as large as possible, how should the driving frequency f and natural frequency of the oscillation f_0 be related to one another?

10. Given two examples of resonance.

E Homework 3

Show all work.

1. Give examples of (a) a transverse wave and (b) a longitudinal wave.
2. Give examples of (a) a wave that requires a medium in which to travel and (b) a wave that travels in empty space.
3. Give an example that shows that waves transport energy.
4. A guitar string is stretched with a tension $\tau = 50$ lb. It has a length $L = 60$ cm and a mass $m = 2$ g. Calculate the wave velocity for waves on the string, $v = \sqrt{\tau/\mu}$ where $\mu = m/L$. (Note: Don't forget to convert lb to N, g to kg, and cm to m.)
5. Microwaves are examples of electromagnetic waves which travel at $v = 3 \times 10^8$ m/s in vacuum. If the frequency of the waves in a microwave oven is $f = 2.45 \times 10^9$ Hz, what is the corresponding wavelength λ? Is it short enough to fit inside the microwave oven? It should be.
6. Calculate the velocity of sound in air at a temperature of (a) 0 °C, (b) 20 °C, (c) 25 °C, and (d) 30 °C.
7. Calculate the wavelength of sound in air at room temperature for frequencies (a) $f = 20$ Hz and (b) $f = 20$ kHz. (Recall: $v_{\text{air}} = 346$ m/s at room temperature.)
8. Two notes with frequencies $f_1 = 440$ Hz and $f_2 = 441$ Hz are played simultaneously. (a) What is the frequency of the sound that you hear? (b) What is the beat frequency?
9. The standing wave pattern for the fundamental frequency for a string fixed at both ends is shown below. Sketch the 2nd, 3rd, and 4th harmonics on the following graphs. If $L = 1$ m, what are the wavelengths of the the fundamental frequency and these harmonics?



10. Standing waves are set up in the guitar string from the Problem 4. Calculate the value of the fundamental frequency and the 2nd and 3rd harmonics.

F Homework 4

Show all work.

1. (a) What type of tube (open at both ends or closed at one end) produces sound with only odd harmonics? (b) Give an example of a wind instrument that produces only odd harmonics.
2. What is the difference between harmonics and overtones?
3. Standing waves are set up in a column of air in a tube that is closed at one end. Is the closed end a node or anti-node for (a) air molecule displacement, (b) pressure deviation from atmospheric pressure?
4. You clap your hands and hear an echo from a nearby canyon wall 0.50 sec later. How far away is the wall? Assume $v = 346$ m/s for the speed of sound in air. (Hint: Don't forget that sound must travel to the canyon wall and back to be heard as an echo.)
5. Does sound tend to carry when there is a temperature inversion (i.e., when the temperature increases with height above the ground) or when the temperature decreases with increasing height above the ground? Explain why it carries.
6. What happens to a wave pulse as it propagates through a dispersive material?
7. Glass is a dispersive material for light with the wave speed for red light being greater than the wave speed for violet light. Which color (red or violet) will be bent more (i.e., refracted through a bigger angle) as it passes from air into glass.
8. Explain why sound diffracts through a doorway but visible light does not?
9. (a) What must the wavelength of an electromagnetic wave be in order for it to diffract through a doorway like sound? (b) What frequency does this correspond to? (Recall that $v = 3 \times 10^8$ m/s for the speed of light in air.)
10. A stationary source of sound has a (fundamental) frequency $f = 440$ Hz. What is the observed frequency if an observer moves (a) away from the source at a speed v_o equal to the speed of sound in air $v = 346$ m/s, and (b) toward the source at a speed equal to the speed of sound?

G Homework 5

Show all work.

1. Given that $A_4 = 440$ Hz, calculate the frequency of middle C (i.e., C_4) in (a) equal temperament, (b) Pythagorean tuning, and (c) just temperament.
2. Do semitone intervals correspond to the same frequency ratio in all tuning systems? If not, give an example of a tuning system that has different semitone intervals.
3. How many cents does a fifth correspond to in (a) equal temperament, and (b) just temperament?
4. What is a Pythagorean comma? Explain in words and give its numeric value.
5. What is a syntonic comma? Explain in words and give its numeric value.
6. Do all fifths in equal temperament correspond to the same frequency ratio? If not, give an example of two fifths that have different frequency ratios.
7. Do all fifths in Pythagorean tuning correspond to the same frequency ratio? If not, give an example of two fifths that have different frequency ratios.
8. What is the main difference between just temperament and Pythagorean tuning?
9. What is the main advantage of equal temperament over other tuning systems like Pythagorean or just temperament?

H Homework 6

Show all work.

1. The effective area of an ear drum is approximately 5.5×10^{-5} m^2. The pressure variation for ordinary conversation is approximately 10^{-2} N/m^2. Calculate the force on an eardrum during ordinary conversation.
2. Calculate the logarithms of the following numbers: 2, 10, 1000, 1/2, 0.01.
3. In what sense is the basilar membrane like a piano keyboard?
4. What is the sound intensity level at 100 Hz of a pure tone having a sound loudness level of 50 phon?
5. What weighting (A or C) should you use for a sound level meter if you want it to measure the sound loudness level in phon?
6. What is the relationship between dB, phon, and sone?
7. Two violins playing simultaneously each produce sound waves with amplitude $p = 10^{-2}$ N/m^2. Calculate the amplitude $p_{\rm tot}$ of the resultant wave assuming that the waves add: (a) in phase with one another, and (b) incoherently (which is the normal case).
8. Suppose one violin has a sound intensity level $L_p = 70$ dB. Calculate the sound intensity level for: (a) two violins playing simultaneously, (b) ten violins playing simulatenously.
9. Suppose one violin has a subjective loudness level $S = 8$ sone. Calculate the subjective loudness level for: (a) two violins playing simultaneously, (b) ten violins playing simulatenously.
10. According to OSHA, (a) what is the maximum duration per day for noise exposure to a piano being played fortissimo? (b) What is the maximum duration per day if you listen to two pianos simultaneously, each being played fortissimo?

I Homework 7

Show all work.

1. (a) Explain the difference between just noticeable difference for pitch and the limit of frequency discrimination. (b) Calculate both of these values at a center frequency of 1000 Hz.
2. Compare and contrast the place theory of pitch perception and the periodicity theory of pitch perception.
3. (a) What is the missing fundamental frequency for the harmonics 220 Hz, 330 Hz, 440 Hz, etc. (b) Does the missing fundamental frequency correspond to a real physical wave in the ear?
4. What is responsible for the production of aural harmonics? Do they correspond to real physical waves in the ear?
5. (a) What is the frequency of the difference tone associated with C_4 and G_4? (b) What note does this correspond to?
6. Describe Ohm's law of hearing and why it is important for sound reproduction systems.
7. Give an example of two frequencies that would produce (a) first-order beats and (b) second-order beats. What are the beat frequencies for each?
8. What is an attack transient and how does it affect the timbre of a sound? Give an example.
9. Do two notes a perfect fifth apart always sound consonant? If not, explain in what circumstances the notes would sound dissonant.
10. What is the difference between pitch height and pitch class?

J Homework 8

Show all work.

1. What is the main difference between a harp and a guitar?
2. The natural frequencies of the standing wave vibrations on a string depend on three things. What are they?
3. Calculate the fundamental frequency for a guitar string having a length of 60 cm, a mass of 2 gm, subject to a tension of 50 pounds. (Hint: Convert cm to m, gm to kg and pounds to Newtons before substituting into the formula for the frequency.)
4. Do the harmonics of a plucked string depend on the location of plucking?
5. Suppose you pluck a guitar string 1/5th of the way from the bridge. What harmonics will be absent?
6. Why are the YouTube videos showing oscillations of a guitar string captured with an iPhone somewhat misleading about wave propagation on a vibrating string?
7. Explain how a violin bow excites vibrations in a violin string.
8. Do the harmonics of a bowed string depend on the location of bowing? (Assume that the bow is very narrow.)
9. Why are the strings of a string instrument usually connected to a box-shaped object?
10. Why wouldn't a violin finger board work if it was on a guitar?

K Homework 9

Show all work.

1. What two things (in addition to the speed of sound in air) do the vibration frequencies of an air column in a tube depend on?
2. What type of tube produces only odd harmonics?
3. Is an open end of a tube a node for air-molecule motion or pressure deviation?
4. What happens to a positive-pressure pulse when it reflects off (a) an open end of a tube, (b) a closed end of a tube.
5. Give an example of an instrument that uses (a) a flow-controlled excitation, (b) a pressure-controlled excitation.
6. A flute and a clarinet both have a length of 66 cm. Calculate the fundamental frequency of (a) the flute, (b) the clarinet when all the tone holes are closed.
7. How big must a tone hole be in order for the effective length of the tube to end at the tone hole?
8. What's the purpose of a register key on a clarinet? How does it work?
9. What does it mean for a brass player "to play harmonics"?
10. Why doesn't a trumpet produce only odd harmonics.

L Homework 10

Show all work.

1. (a) Write down the frequencies of first few transverse vibrational modes of a thin bar with free ends. (b) Are these frequencies harmonically related?
2. Why do xylophone, marimba, and vibraphone bars have arches?
3. (a) What is the purpose of the resonator tubes under xylophone, marimba, and vibraphone bars. (b) How does the presence of the tubes affect the decay time of the vibrating bars?
4. What is responsible for the perceived strike tone of chimes?
5. What do the integers (m, n) for the vibrating modes of a membrane stand for?
6. What two factors make the vibrations of a timpani drum head non-ideal?
7. Why does a bass drum make a “thud” sound while timpani produce a definite pitch?
8. What is the main difference between a vibrating membrane and a vibrating plate?
9. Why do cymbal players hold the cymbals with their faces pointing toward the audience after clashing them together?
10. What are Chladni patterns?

M Homework 11

Show all work.

1. List the main parts of a piano.
2. Why doesn't a harpsichord have a large dynamic range?
3. Why do pianos have cast iron frames?
4. What is the purpose of having wrapped strings instead of solid strings with the same linear mass density?
5. (a) How many piano strings are there for most keys? (b) Are the strings for the same key tuned to the same frequency?
6. Draw a picture showing the attack and decay transient of a piano note.
7. (a) Are the vibrational frequencies of a piano string equal to harmonics of the fundamental frequency? (b) If not, are the corresponding vibrational frequencies larger or smaller than those of pure harmonics?
8. What is the frequency difference between an equal-tempered fifth and a perfect fifth?
9. (a) Explain what is meant by stretched tuning. (b) Are all of the notes tuned to equal temperament?
10. Draw a Railsback curve, labeling the horizontal and vertical axes.

N Homework 12

Show all work.

1. List the main vocal organs.
2. Explain the Bernoulli effect.
3. Is it possible to produce sounds that do not require the vocal folds to vibrate? If so, give an example.
4. What are formants and hope do they shape the spectrum of the radiated sound?
5. What distinguishes different vowel sounds from one another?
6. Describe a simple model of the human vocal tract that predicts formant peak frequencies at 500 Hz, 1500 Hz, 2500 Hz, etc.
7. What is a sound spectrogram? Why is it more useful than a single spectrum?
8. If I record myself whistling at 1000 Hz, what type of filter will reproduce the recorded sound:
   (a) a low pass filter with cutoff frequency at 500 Hz.
   (b) a high pass filter with cutoff frequency at 1500 Hz.
   (c) a band pass filter with a low cutoff frequency at 500 Hz and a high cutoff frequency at 1500 Hz.
9. What is a singer's formant, and why is it useful for an opera singer to be able to produce one?
10. How can a soprano tune the peak of her first formant region to match the fundamental frequency of a high-pitched note?

O Homework 13

Show all work.

1. Calculate the sound power level L_W of a 1-watt omni-directional loudspeaker.
2. Calculate the sound pressure levels L_p of the loudspeaker from the previous problem at distances of 1 m, 2 m, and 4 m from the loudspeaker.
3. You are standing 50 ft in front of a high cliff. If you clap your hands, how long does it take for the reflected sound to reach your ears? Will you hear an echo?
4. Explain the difference between direct, early reflected, and reverberant sound.
5. You need to convert a concert hall into a room that will be used for large public speaking events. What acoustical changes would you make?
6. Calculate the reverberation time at 500 Hz for a room having dimensions 8 m $\times$ 10 m $\times$ 3 m (high), concrete walls (0.06), plywood floor (0.17), and suspended acoustical tile ceiling (0.83). Also, assume that there are 50 wooden chairs (0.02 m^2/chair) and 30 students (0.39 m^2/student).
7. Calculate the sound pressure level of the reverberant sound at 500 Hz for a 1-watt omni-directional loudspeaker in a room that has a surface area $S = 1000\ \text{m}^2$ and a total absorption $A = 400\ \text{m}^2$ at 500 Hz. (Hint: You will first need to calculate the average absorption $\bar{a}$ and the room constant R.)
8. What is the critical distance from the loudspeaker for the above problem?
9. Why do you sound better when you sing in the shower?
10. Explain in what sense 'liveness' and 'clarity' of sound are competing acoustical properties.