Acoustic

Signal Processing

Explorers Computer Technology Post 631

April 3, 2008

by Don Braun

(last modified on 4/14/2008)

1

1. Title: This lesson will involve learning about acoustic signals and how computers can record, playback, and manipulate them.
2. "Acoustic" signals (from the Greek word for something audible) are otherwise known as "sound" (from the Latin word for hearing).

Goals of this lesson

1. Use the free audio editor and recorder program "Audacity" (downloaded from http://audacity.sourceforge.net) to experiment with sound.
2. Understand basic concepts and terminology of acoustics.
3. Learn or review a little math and use it to describe sound.

Download this PowerPoint presentation & notes from Web

http://explorersposts.grc.nasa.gov/post631/2007-2008/

What is sound?

A sound is the compression and expansion of a medium such as a gas, a liquid, or a solid, that propagates (that is, travels) through that medium.

The medium is the substance or material that carries the sound wave.

2

1. "Goals of this lesson – 1. Audacity": Hands-on demos and exercises at this Explorers meeting will use the free software "Audacity" version 1.2.6.
2. "Goals of this lesson – 2. Concepts and terminology": The basic concepts and terminology of acoustics explained in this presentation include sound, medium, traveling or standing waves, longitudinal or transverse or surface waves, energy, power, radians, sine, cosine, amplitude, frequency, period, wavelength, speed, phase, exponentials, logarithms, sound intensity, decibels, octave, semitone, superposition of waves and beats, interference nodal lines and antinodal lines, spectrum, fast Fourier transform (FFT), anatomy of the human ear, anatomy of the human vocal tract, and phonetic speech synthesis and recognition.
3. "Goals of this lesson – 3. Math": On your applications for this Explorer program, a number of you said you hoped the program would improve your math skills. Also, on your weekly evaluations of the meetings some of you said certain presentations spent too much time explaining things you already knew; however, those topics were not well known by some of the other Explorers. Some Explorers asked for more variety in the topics. I hope this lesson will address those concerns. This presentation is an overview of a variety of subjects, and each of you will probably find some topics in this presentation to be unfamiliar. We will not have time to investigate any of the subjects in great detail. If some unfamiliar subjects seem hard to follow, don't be too concerned; it is generally sufficient for you to understand the basic principles of each topic. I will present some mathematical details for those who want to understand them. If you're not that interested in the math, please be patient.
4. "Download this presentation": Because there is more information on the slides of this presentation than we will have time to cover thoroughly, printed copies of the slides will be handed out so you can refer to them later if you want. The PowerPoint presentation file "AcousticSignalProc.ppt", which includes these detailed notes for the slides, will also be posted on the Explorers Web site https://explorersposts.grc.nasa.gov/post631/2007-2008/.
5. I'd like this discussion to be interactive, so please ask questions or make helpful comments at any time. I'll try to present a lot of useful information in a fairly short time and to make it as interesting as I can, so please give me your undivided attention. Don't worry, there will be no exam at the end, so just keep your mind alert and enjoy the experience.
6. "What is sound?": I want an explanatory physical definition, not just something like "an acoustic signal", "something to listen to", or "something to put on an iPod". The "medium" is the "stuff" that sound passes through.

How fast does sound travel?

741 mph ≈ 1,087 feet/s ≈ 331 m/s in dry air at sea level at 32°F;
770 mph ≈ 1,130 feet/s ≈ 344 m/s in dry air at sea level at 70°F;
3,349 mph ≈ 4,911 feet/s ≈ 1,497 m/s in pure water at 77°F;
11,400 mph ≈ 16,700 feet/s ≈ 5,100 m/s in steel.

Other things being equal, the speed of sound (called "Mach 1") increases as the density of the medium decreases and as the stiffness of the chemical bonds between particles in the medium increases. Pressure in the medium hardly influences the speed of sound.

Why do you see lightning before you hear its thunder?

The speed of light is c = 299,792,458 m/s ≈ 186,282 miles/second, so you see a lightning bolt from one mile away after only (1 mile) / c ≈ 5.7 μs but you hear it after (5280 ft) / (1130 ft/s) ≈ 4.7 s . Light travels almost one million times as fast as sound in air.

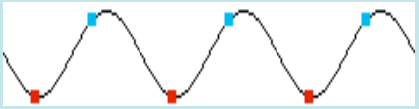
Would an astronaut on a space walk hear a bell if she or he hit it with a hammer?

No! Although the bell would vibrate, there is no medium like air in the vacuum of space to propagate a sound wave to the astronaut's ears.

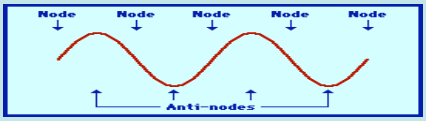
3

1. "How fast does sound travel": The transmission of sound can be explained using a toy model consisting of an array of balls interconnected by springs as in the upper left figure on slide #5. For a real material the balls represent molecules and the springs represent the chemical bonds between them. Sound passes through the model by compressing and expanding the springs, transmitting energy to neighboring balls, which transmit energy to *their* springs, and so on. The speed of sound through the model depends on the stiffness of the springs: stiffer springs transmit energy more quickly. All other things being equal, sound will travel more slowly in denser materials, and faster in stiffer ones. (The preceding text of this note is from the Web page http://en.wikipedia.org/wiki/Speed_of_sound .) For example, sound moves slower through the denser (higher pressure) cold air at 32°F than it does through the less dense (lower pressure) warm air at 70°F . Although material density increases from air to water to steel, neither the density rule nor the stiffness rule stated on this slide should be applied because different states of matter are involved (i.e., a gas, a liquid, and a solid), which does not satisfy the requirement of "other things being equal."
2. "Why do you see lightning before you hear its thunder?": Light travels about $c / v \approx (300 \cdot 10^6 \text{ m/s}) / (344 \text{ m/s}) \approx 10^6$ times as fast as sound in air.
3. "Astronaut on a space walk hits a bell": A sound wave cannot even exist, let alone travel, if there is no medium to compress and expand.

There are 2 categories of waves based on motion of the waves: traveling and standing



<http://www.ncat.edu/~gpii/>



<http://www.glenbrook.k12.il.us/gbssci/phys/mmedia/waves/harm4.html>

<http://www.glenbrook.k12.il.us/gbssci/phys/Class/sound/u11l1a.html>

<http://www.ncat.edu/~gpii/>

Try the following Java applet for a standing wave in a wind instrument like a pipe organ: http://www.walter-fendt.de/ph11e/stlwaves.htm

A traveling wave (like sound in an open area) is a repeating disturbance that moves through a medium from one location to another.

A standing wave (like a vibrating string fixed at both ends) is a disturbance that repeats without propagating through the medium.

4

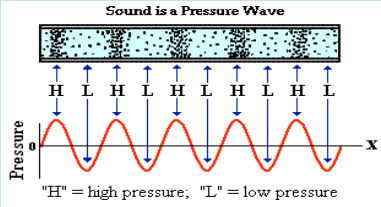
1. The upper left animated gif figure shows a traveling transverse sinusoidal wave (like a slinky or heavy stretchable cord under tension, whose left end is moved up and down, and whose right end is free). Traveling transverse waves will be described on slide #6. Click on this Web hyperlink to also see animated gif pictures of a traveling longitudinal wave of particles, a standing transverse sinusoidal wave (similar to the traveling transverse sinusoidal wave shown here in the upper left), and the standing longitudinal wave of particles (as shown in the lower right figure on this slide).
2. The lower left animated gif figure shows a traveling longitudinal wave (like sound emanating from a tuning fork in an open area so there are no echoes) showing compression and expansion of the medium (like air). This diagram clearly shows the wave motion from left to right, but it does not clearly show that individual particles of the medium only move right and then back left the same very small distance rather than moving all the way to the right along with the wave. See the Web hyperlink in the upper left quadrant as mentioned in item #1 just above to more clearly see the motion of particles in the medium carrying a traveling longitudinal wave.
3. The upper right animated gif shows a standing transverse sinusoidal wave (like a violin string attached at both ends). The medium particle at any "node" does not move at all. The medium particle at an "antinode" moves up and down a maximum amount.
4. The lower right animated gif shows a standing longitudinal air pressure wave (like sound resonating in a tube that is closed at both ends). The air (medium) particles at a "node" do not move at all, and this example shows 7 equally-spaced nodes (including a node at each end of the tube where the air particles cannot move against the end cap). The air particles at an "antinode" move left and right a maximum amount, and this example shows 6 equally-spaced antinodes, with one antinode halfway between each pair of adjacent nodes. Every 10th short vertical line is drawn bold just to help the viewer see that individual lines move left and right only a short distance, except at a node where there is no motion at all. This is a standing longitudinal wave (in contrast to a traveling longitudinal wave as shown in the similar animated gif in the on lower left figure on the next slide) because each compression seems to move a fairly short distance left and right rather than moving all the way along the tube.
5. Click on the Web hyperlink near the lower right to see an excellent interactive demo of a standing longitudinal wave (like sound in an organ pipe with both ends open, with one end closed, or even a pipe with both ends closed)!

What actually moves through the medium as sound travels?

Particles of a medium do not travel long distances along with a sound wave, but energy does. Particles move forward (parallel to the direction of propagation) a tiny distance, pushing on neighboring particles of the medium and transferring energy to them, and then return close to their original position, so sound waves are called "longitudinal". Push the end of a slinky to make a longitudinal wave.

Traveling longitudinal wave on a slinky
<http://www.glenbrook.k12.il.us/gbssci/phys/mmedia/waves/lw.html>

Traveling longitudinal sound pressure wave
<http://www.ncat.edu/~gpii/>



What is energy?

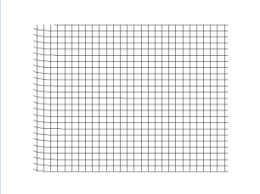
In physics, a force is an action that can cause an object to accelerate (speed up),
work is a force acting over a distance,
energy is the ability to do work, and
power is the rate at which energy is delivered per unit time.

5

1. "What actually moves through the medium as sound travels?": The particles of the medium do not travel along with the wave. Instead, the wave is a disturbance that passes through the medium always moving away from the source, but that wave merely moves the particles a little and then lets them go back to their "rest" positions. In the case of sound, the particles move forward (in the direction of wave propagation) and backward, so sound waves are called "longitudinal", which means "running lengthwise". Only energy (not any individual particle) actually travels with the wave from starting point to ending point. "Sound" relates to energy traveling through a medium.
2. In the lower left animated gif figure, short vertical lines close together represent compressed (high pressure) air. Lines far apart represent rarified (low pressure) air. Every 10th line is drawn bold just to help the viewer see that individual lines move left and right a short distance, but do not follow the wave all the way toward the right. All lines move the same amount, but slightly "out of phase" with adjacent lines. However, high density concentrations of lines (representing compressed air) appear to travel all the way from left to right. Every 10th line is drawn bold just to help the viewer see that individual lines move left and right only a short distance, but those bold lines do not follow the wave all the way toward the right. All lines move the same amount, but slightly "out of phase" with adjacent lines. However, it appears that each high density concentration of lines (representing compressed air) travels all the way from left to right as required for a traveling wave, in contrast to the similar animated figure of a standing longitudinal wave in the lower right on the previous slide.
3. "Force, work, energy, power": As an example of the effect of a force acting on an object, consider pushing constantly with a force of $F = 1$ pound on a hockey puck that is initially at rest and weighs about 1/3 pound (use $m = 0.322$ pound $/ g$, where g is the acceleration of gravity). At the end of $t = 1$ second, the hockey puck will have moved across the (frictionless) ice rink a distance of $x = 50$ feet, the final speed of the puck will be $v = 100$ feet/second, its constant acceleration will have been $a = 100$ (feet/second)/second throughout the $t = 1$ second interval, the work (W) that your pushing did on the puck will be equal to its final energy (E) where $W = E = 50$ foot pounds, and the power you imparted to the puck will have been $P =$ 50 foot pounds / second throughout the $t = 1$ second interval. The following calculations show how to find those results, using standard formulas from elementary physics.
   $F = 1$ lb;
   $g = 32.2$ ft / s^2 = (acceleration of gravity);
   $m = (0.322 \text{ lb}) / g = 0.322 \text{ lb} / (32.2 \text{ ft} / \text{s}^2) = 0.01 \text{ lb s}^2 / \text{ft}$;
   $t = 1$ s;
   $a = F / m = (1 \text{ lb}) / (0.01 \text{ lb s}^2 / \text{ft}) = 100 \text{ ft} / \text{s}^2$;
   $v = a\,t = (100 \text{ ft} / \text{s}^2)(1 \text{ s}) = 100$ ft/s;
   $x = a\,t^2 / 2 = (100 \text{ ft/s}^2)(1 \text{ s})^2 / 2 = 50$ ft;
   $W = F\,x = (1 \text{ lb})(50 \text{ ft}) = 50$ ft lb;
   $E = m\,v^2 / 2 = (0.01 \text{ lb s}^2 / \text{ft})(100 \text{ ft/s})^2 / 2 = 50 \text{ ft lb} = W$;
   $P = E / t = (50 \text{ ft lb}) / (1 \text{ s}) = 50 \text{ ft lb / s}$.

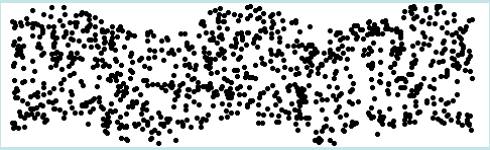
Other categories of traveling waves based on particle motion

For a "traveling transverse wave", particles of the medium move (back and forth) in a direction perpendicular to the direction in which the wave travels, like people in a stadium "doing the wave", where the crowd is the medium.



<http://en.wikipedia.org/wiki/Transverse_wave>

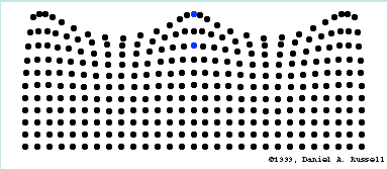
Traveling transverse wave



<http://www.kettering.edu/~drussell/Demos/waves/wavemotion.html>

Try the interactive demo of a traveling transverse wave at http://www2.biglobe.ne.jp/~norimari/science/JavaApp/nami1/e-nami.html .

A wave on the ocean surface is called a "traveling surface wave" in which particles of the medium move in vertical circles, but the wave propagates horizontally.



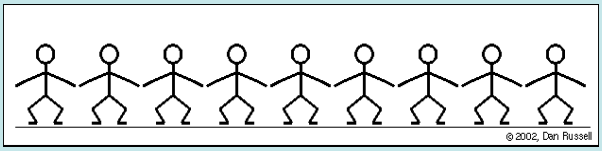
Traveling surface wave
(also called a circular wave)

<http://www.kettering.edu/~drussell/Demos/waves/wavemotion.html>

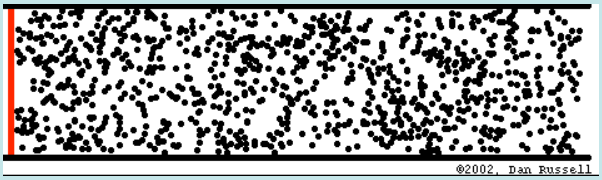
6

1. The upper left animated file TransversePlaneWave.gif is the file http://en.wikipedia.org/wiki/Image:Onde_cisaillement_impulsion_1d_30_petit.gif from the Web page http://en.wikipedia.org/wiki/Transverse_wave .
2. The upper right animated file with dots moving up and down to create a transverse wave that travels from left to right is from the Web page http://www.kettering.edu/~drussell/Demos/waves/wavemotion.html .
3. Another example of a traveling transverse wave would be to stretch a slinky horizontally, and then to quickly jerk the left end up and down to send a pulse from left to right along the slinky, as shown in the excellent interactive demo at the Web hyperlink http://www2.biglobe.ne.jp/~norimari/science/JavaApp/nami1/e-nami.html . That demo allows interactively changing the stiffness of springs (simulating bonds) between balls (simulating particles of the medium) with varying mass per ball (thus varying the density of the medium).
4. The lower left animated file with dots near the top moving clockwise in small circles to create a surface wave that travels from left to right is also from the Web page http://www.kettering.edu/~drussell/Demos/waves/wavemotion.html . Dots near the top move in larger circles than dots lower down, as shown by the two blue dots.

Identify each wave as longitudinal, transverse, or surface



Traveling transverse wave



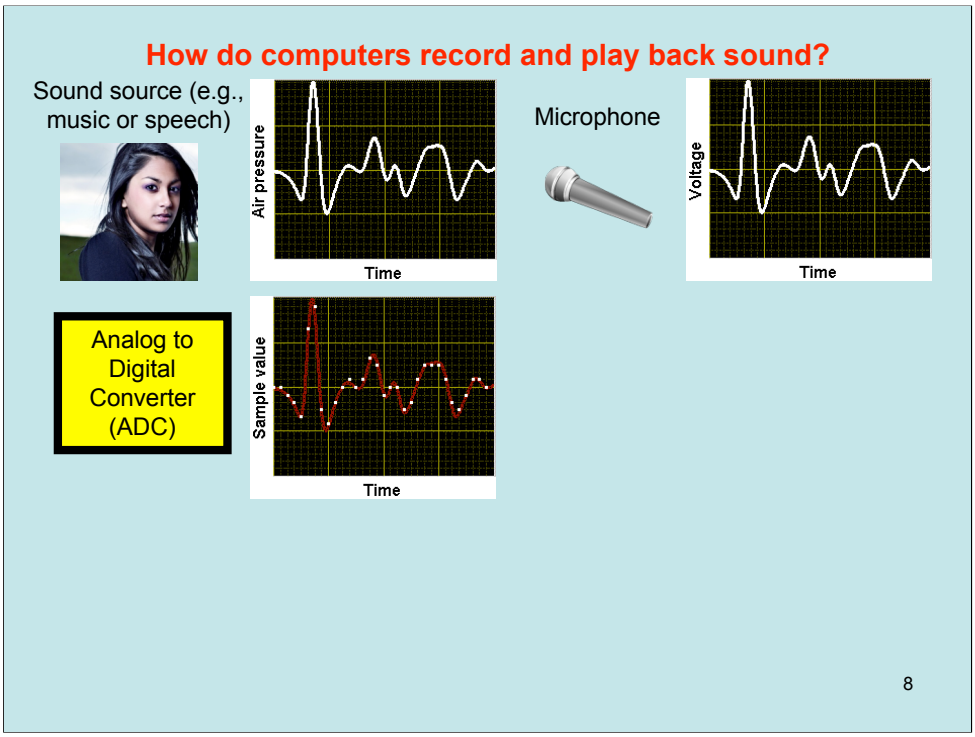
Traveling longitudinal wave (like sound)

Traveling transverse wave

<http://www.kettering.edu/~drussell/Demos/waves-intro/waves-intro.html>

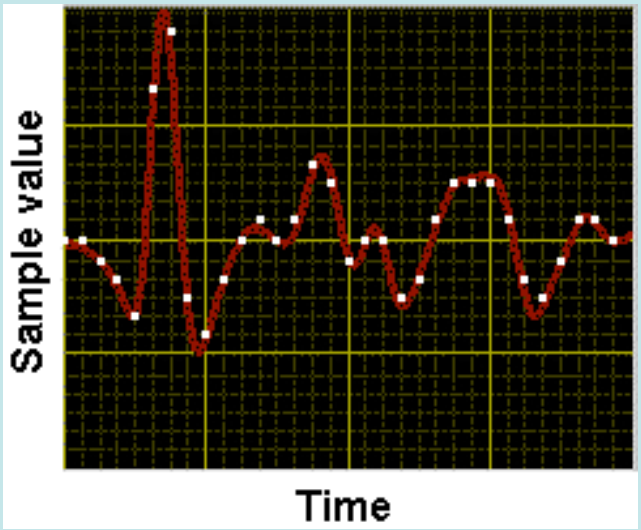
7

1. For the traveling <u>transverse</u> wave at the top, the particles (individuals "doing the wave" in a stadium) move perpendicular (up and down) to the direction of wave propagation (left to right) and then those particles return to their initial (rest) positions.
2. For the traveling <u>longitudinal</u> wave in the middle, the particles (air molecules) move parallel to the direction of wave propagation (left to right) and then those particles return to their initial (rest) positions.
3. For the traveling <u>transverse</u> wave at the bottom, the particles (bits of a stretched piece of elastic or slinky) move perpendicular (up and down) to the direction of wave propagation (left to right) and then those particles return to their initial (rest) positions.



1. Girl and her "Air pressure" plot: The girl's speaking or talking (with her vocal tract including the vocal cords, throat, mouth, and nasal cavity; or else perhaps a musical instrument, a clap, or another sound source) causes the air pressure to fluctuate with time. For example, when either one of the two tines of a tuning fork moves toward the surrounding air, it locally compresses that air and causes a pulse of increased air pressure (above the mean ambient air pressure) to propagate away from that tine. In contrast, when the tine of a tuning fork moves away from the surrounding air, it locally expands the air and causes a pulse of decreased air pressure (below the mean ambient air pressure) to propagate away from the tine. The air pressure waveform is a continuous function of time; that is, at every instant of time the air pressure at any given location has some value. The picture of the girl is from Microsoft Office 2003 Clip Art.
2. Microphone and its "Voltage" plot: The microphone converts the pressure at any instant in time to a voltage proportional to that pressure, thus producing a continuous voltage waveform that is analogous to the continuous air pressure waveform measured at the location of the microphone.
3. ADC and its waveform: The ADC digitizes the voltage waveform by sampling it at equally-spaced discrete time slices and quantizing the value of the waveform at each time slice to an integer value as explained on the next slide.

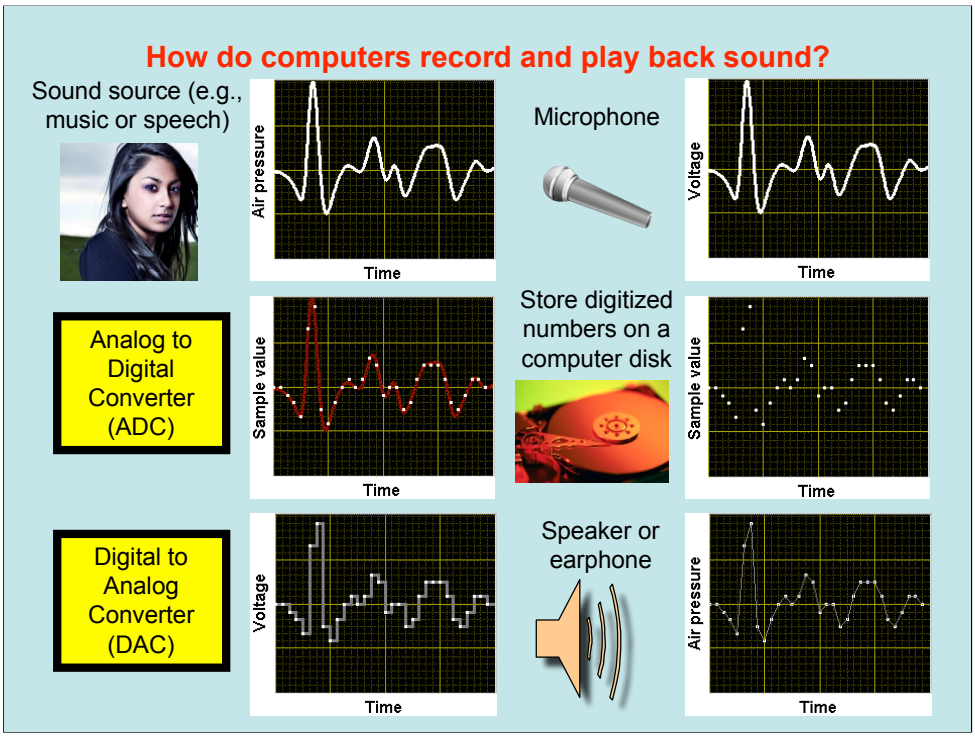
Quantization errors introduced by digitizing an analog signal




```
 0,  0, -1, -2, -4,  8, 11, -3,    -5, -2,  0,  1,  0,  1,  4,  3,
-1,  0,  0, -4, -2,  1,  3,  3,     3,  1, -2, -3, -1,  1,  1,  0
```

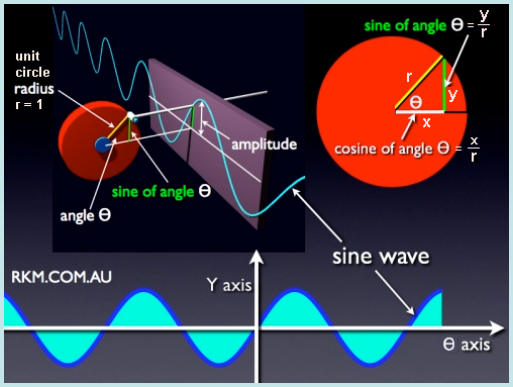
9

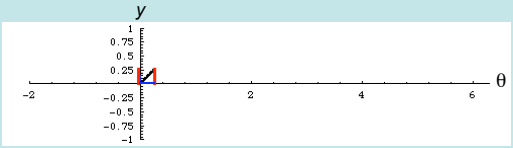
1. Observe the quantization errors as the white square samples shown on this graph approximate points on the red continuous waveform. Those small errors in the vertical positions of the white squares compared with the red curve are caused because each digitized sample is taken at a time instant corresponding to a vertical grid line, and its value is rounded by the ADC to the quantized sample value corresponding to the nearest horizontal grid line crossing the red continuous waveform at that time.
2. The 32 sample values are represented by the integers
   "0, 0, –1, –2, –4, 8, 11, –3, –5, –2, …, 1, 1, 0"
   shown in two rows at the bottom of this slide, where each of the four groups of 8 samples is between two major vertical grid lines on the graph. Those 32 samples would be produced by a signed 5-bit ADC whose sample values must be in the range
   $-16 = -2^4 \le \text{sample} \le 2^4 - 1 = 15$.
   By comparison, audio compact disks (CDs) typically use 16 bits/sample, for a maximum range of
   $-32{,}768 = -2^{15} \le (\text{ADC sample}) \le 2^{15} - 1 = 32{,}767$.
   Professional audio recording equipment may reduce quantization errors by using an ADC that digitizes 24 bits/sample, for a maximum range of
   $-2{,}147{,}483{,}648 = -2^{31} \le (\text{ADC sample}) \le 2^{31} - 1 = 2{,}147{,}483{,}647$.



1. Disk and its "Sample value" plot: The 32 discrete samples (i.e., the integers) produced by the ADC can be stored in a file on a computer disk. Reading that file later recovers those 32 samples, but the exact values of the original waveform at the sample times have been rounded to integers (probably introducing quantization errors), and no information is available for what the original continuous waveform did between the samples (possibly introducing what are called "aliasing errors" if the original continuous signal contained energy at frequencies greater than the Nyquist frequency, which is half the sampling frequency). For brevity, no technical explanation of aliasing will be given here.
2. If the digitized numbers from the ADC are encoded or compressed before recording the data on disk (for example, to reduce required disk space), the data read back from the disk should be decoded or uncompressed to produce sampled data representing a signal similar to the original sample values from the ADC.
3. DAC and its output "Voltage" plot: The DAC converts the numbers read from the disk into voltages. Typically, the output voltage is held constant by the DAC during the brief time interval between one sample and the next one as shown on this graph.
4. Speaker/earphone and its "Air pressure" plot: The DAC output voltages (perhaps after some electronic filtering to produce a smoother waveform and a more pleasant sound without clicks) drives a speaker or earphone. That produces an air pressure waveform: compressing air to increase pressure when the speaker cone moves outward, and expanding air to decrease pressure when the speaker cone moves inward.

Define sine & cosine functions of angles measured in radians





<http://www.wku.edu/~tom.richmond/Sine.html>

Animation movie of a sine wave (sinusoidal wave) y = s.mov

The radian measure of an angle is the length of the circular arc subtended (included) by that angle divided by the radius of the circle.

The *y*-coordinate of the point on the unit circle determined by a central angle (in standard position) of θ radians is the sine, denoted sin(θ) .

The *x*-coordinate of that point is the cosine, denoted by cos(θ) .

<http://www.rkm.com.au/ANIMATIONS/animation-sine-wave.html>

To the left of the *y*-axis, you see a unit circle being swept out, with the radian measure of the angle (arc length) shown in blue, and the sine of that angle (the *y*-coordinate) shown in red. To the right of the *y*-axis, you see the points of the graph of
y = sin(θ) drawn in black. The angle θ is shown in blue and the value of sin(θ) is shown in red.

11

This slide explains two new subjects that should be understood to use Audacity:
(a) how to measure angles in radians (where 2π radians is a full circular arc) instead of degrees (where 360 degrees is a full circular arc);
(b) definitions of the sine function and cosine function, which are very important in trigonometry.

To activate the excellent upper right animation (which was copied from http://www.rkm.com/au/ANIMATIONS/animation-sine-wave.html) during the presentation of this slide in PowerPoint:

1. press the "ESCape" key to exit "slide show mode" and to enter PowerPoint "edit" mode for this slide;
2. double-click the filename "Animation movie of a sine wave (sinusoidal wave) y = s.mov" near the upper right corner of this slide, just under the title "Define sine & … radians" to start the QuickTime movie;
3. after the movie plays, exit the QuickTime Player to return to the PowerPoint "edit' mode for this slide;
4. finally, continue presenting the slide show from this slide by typing Shift+F5 (that is, while holding down the Shift key, press the function key F5).

That method will use QuickTime to show the movie in a window that can be maximized to fill the screen.

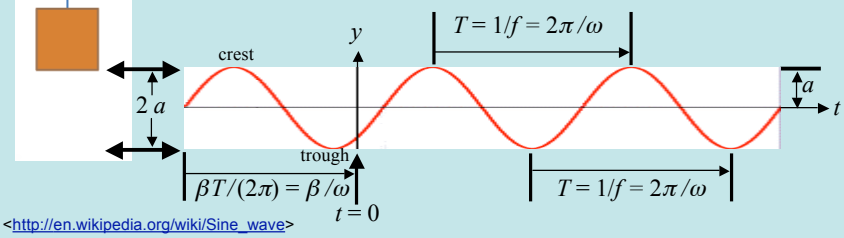
Alternatively, click on the hyperlink "http://www.rkm.com.au/ANIMATIONS/animation-sine-wave.html" while in PowerPoint's presentation mode to see the entire Web page (including that movie) in your Internet browser.

Parameters of a temporal sine wave

Parameters of the temporal sine wave (or sinusoidal function of time)

$y = a \sin(\omega t + \beta) = a \sin(2\pi f t + \beta) = a \sin(2\pi t / T + \beta)$ are

a = amplitude = peak deviation from the center at zero,
ω = angular frequency typically measured in radians/second,
$f = \omega / (2\pi)$ = angular frequency typically in cycles/second,
$T = 1/f$ = period typically measured in seconds, and
β = initial (at time $t = 0$) phase angle in radians, with the variable
t = time typically measured in seconds.



<http://en.wikipedia.org/wiki/Sine_wave>

The oscillating height of an undamped spring-mass system around the equilibrium is a sine wave.

See a helpful interactive demo of frequency, wavelength, and phase of a sinusoidal wave at http://id.mind.net/~zona/mstm/physics/waves/introduction/introductionWaves.html .

12

To see an excellent interactive demo of frequency, wavelength, and phase of a sinusoidal wave, click on the Web hyperlink at the bottom of this slide!

What is a "common logarithm" (which is also called a "base 10 log" and is denoted by "$\log_{10}$") ?

$a = \log_{10} x$ means $10^a = x$

where $0 < x < \infty$ and $-\infty < a < \infty$.

The notation "$\log_{10} x$" is read as "log base 10 of x" or "common log of x" .

Therefore, $a = \log_{10}(10^a)$ and $10^{(\log_{10} x)} = x$;

that is, 10 must be raised to the exponent $\log_{10} x$ to calculate x .

To remember this, memorize the mnemonic: "A logarithm is an exponent."

The common logarithm is the inverse of the power function raising 10 to an exponent:

$a = 3 \rightarrow \boxed{10^a} \rightarrow 10^3 = 1000$ $\qquad$ $\log_{10} 1000 = 3 \leftarrow \boxed{\log_{10} x} \leftarrow x = 1000$

13

The Audacity program uses logarithms to scale the axes of graphs and to compute units called decibels that are commonly used to express the intensity of sound.

1. For any given positive value of x, the base "10" must be raised to the exponent "$\log_{10}(x)$" (which is called a on this slide) to calculate x .
2. The statement "$a = \log_{10}(x)$ means $10^a = x$" defines a "common logarithm" or "base 10 log".
3. In the sentence beginning with the word "Therefore", the equation "$a = \log_{10}(10^a)$" follows by using "$10^a = x$" from the sentence just above it to substitute for x in the equation "$a = \log_{10}(x)$" in that preceding sentence. Likewise, the equation "$10\text{^}(\log_{10}(x)) = x$" follows by using "$a = \log_{10}(x)$" from the sentence just above it to substitute for a in the equation "$10^a = x$" in that preceding sentence.
4. In the diagram at the bottom of this slide, the "10^a" box at the lower left takes the input value $a = 3$ and calculates the output value $10^3 = 1000$. However, the "$\log_{10}(x)$" box at the lower right takes the input value $x = 1000$ and calculates the output value $\log_{10}(1000) = 3$. Thus, those two boxes compute inverse functions; that is, each box undoes the calculation done by the other box.

Example numeric evaluations of the common logarithm

Use column headings "10^a" and "a" for a table of exponentials, or use column headings "x" and "$\log_{10} x$" for a table of logarithms.

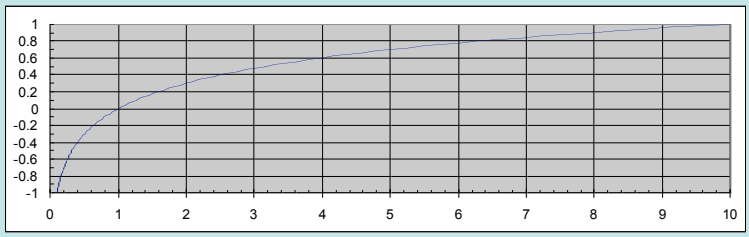
$10^a = x$	$a = \log_{10} x$
1	0
2	0.30103
5	0.69897
10	1
20	1.30103
50	1.69897
100	2
200	2.30103
500	2.69897
1000	3

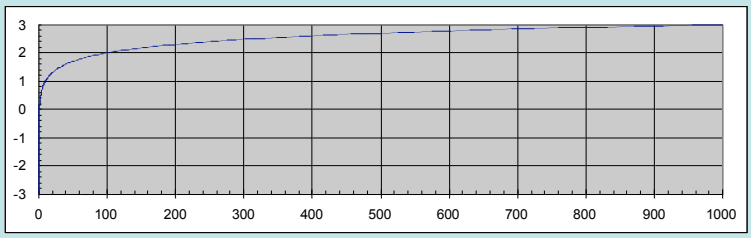
$10^a = x$	$a = \log_{10} x$
1	0
1/2 = 0.500	−0.30103
1/5 = 0.200	−0.69897
1/10 = 0.100	−1
1/20 = 0.050	−1.30103
1/50 = 0.020	−1.69897
1/100 = 0.010	−2
1/200 = 0.050	−2.30103
1/500 = 0.020	−2.69897
1/1000 = 0.001	−3

14

1. Comparing the two columns of this table shows that $\log_{10}(1/x) = -\log_{10}(x)$ for these examples. This is true in general for any positive value of $x > 0$.
2. The left side of this slide shows that while the values of $10^a = x$ increase over the large range from 1 to 1000, the corresponding values of $a = \log_{10} x$ increase over the much smaller range from 0 to 3.
3. The right side of this slide shows that while the values of $10^a = x$ decrease over the small range from 1 to 1/1000, the corresponding values of $a = \log_{10} x$ decrease over the much larger range from 0 to −3.

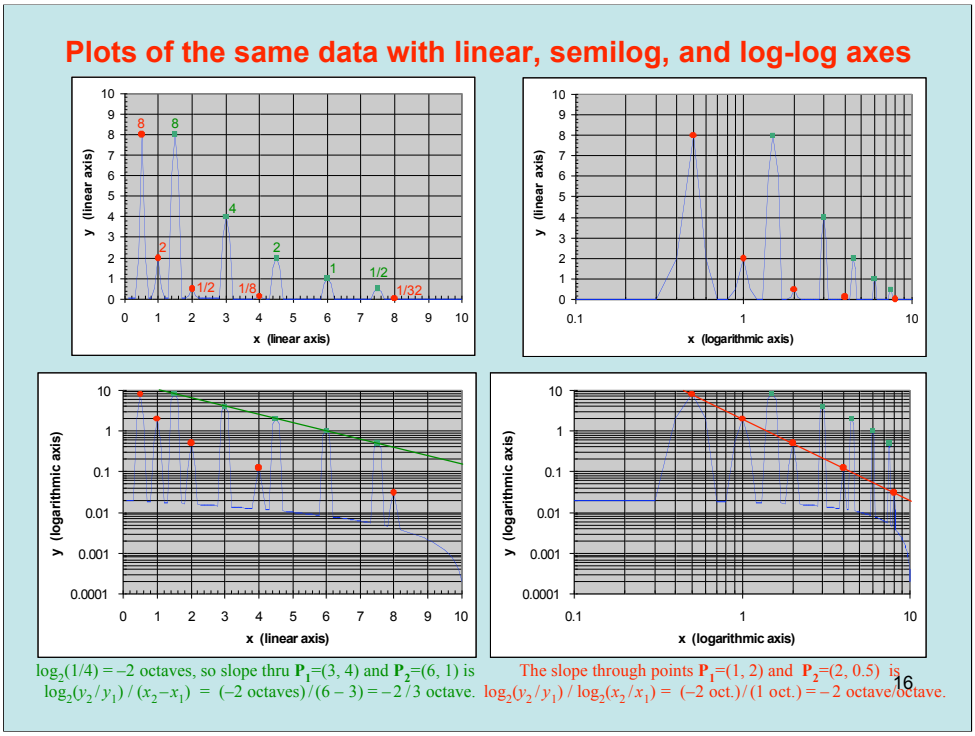
Plots of $\log_{10} x$ versus x with linear axes





15

1. The part of the function $\log_{10}(x)$ shown in the upper graph for $0.1 < x < 10$ is also shown at the leftmost 1/10 of the lower graph, with rescaled axes.
2. The lower graph shows the part of the function $\log_{10}(x)$ for $0.001 < x < 1000$.
3. The blue curve representing the function $\log_{10}(x)$ approaches negative infinity as x decreases toward zero.
4. The values in the tables of the previous slide are included on these two graphs.



1. The same data values are used on each of the 4 plots.
2. The 5 peaks marked by green squares are at coordinates (1.5, 8), (3, 4), (4.5, 2), (6, 1), (7.5, 1/2). Those x coordinates are linearly spaced at identical intervals of 1.5 between consecutive peaks. Those y coordinates are geometrically spaced, with the y coordinate of each peak being 1/2 of the y value of the preceding peak. Therefore, the 5 green square peaks lie along a straight line in the lower-left (log y axis versus linear x axis) plot.
3. The other 5 peaks marked by red circles are at coordinates (1/2, 8), (1, 2), (2, 1/2), (4, 1/8), (8, 1/32). Those x coordinates are geometrically spaced, where the x coordinate of each peak is 2 times the x coordinate of the preceding peak. Those y coordinates are also geometrically spaced, with the y coordinate of each peak being 1/4 of the y coordinate of the preceding peak. Therefore, the 5 red circle peaks lie along a straight line in the lower-right (log y axis versus log x axis) plot.
4. The upper-left (linear y versus linear x) plot is called a "linear" graph.
5. The lower-left (log y versus linear x) plot is the most common kind of "semilog" graph, with "log-linear" axes.
6. Increasing any positive number by "n octaves" for any integer n means multiplying that number by 2^n. For example, increasing the frequency of a musical note from 110 Hz (which is called the note A_2) by 3 octaves yields the frequency 2^3 (110 Hz) = 880 Hz (which is called the note A_5). Similar to the base 10 logarithms already discussed, the base 2 logarithm of any positive number x is denoted by $\log_2(x) = a$ such that $x = 2^a$. Using a theorem about logarithms similar to one on the next slide, $\log_2(x_2 / x_1) = \log_2(x_2) - \log_2(x_1) = b$ = (the number of octaves that x_1 must be raised to equal x_2) because $x_2 / x_1 = 2^b$ and so $x_2 = 2^b x_1$.
7. The upper-right (linear y versus log x) plot is an alternate kind of "semilog" graph, with "linear-log" axes.
8. The lower-right (log y versus log x) plot is called a "log" graph, with "log-log" axes.

Basic properties of powers of 10 and of common logarithms

Basic properties of powers of 10 from algebra include

$10^a\, 10^b = 10^{a+b}$ and $10^a / 10^b = 10^{a-b}$ and $(10^a)^b = 10^{(a\,b)}$

for any real values $-\infty < a < \infty$ and $-\infty < b < \infty$.

Examples with $a = 5$ and $b = 2$:

$$10^5\, 10^2 = (10 \cdot 10 \cdot 10 \cdot 10 \cdot 10)\ (10 \cdot 10) = 10^7 = 10^{5+2};$$

$$10^5 / 10^2 = (10 \cdot 10 \cdot 10 \cdot 10 \cdot 10) \ / \ (10 \cdot 10) = 10^3 = 10^{5-2};$$

$$(10^5)^2 = (10 \cdot 10 \cdot 10 \cdot 10 \cdot 10)\ (10 \cdot 10 \cdot 10 \cdot 10 \cdot 10) = 10^{10} = 10^{(5 \cdot 2)}.$$

The 3 corresponding properties of common logarithms follow, using the definitions

$x = 10^a$ and $y = 10^b$, so $\log_{10} x = a$ and $\log_{10} y = b$

where $0 < x < \infty$, $0 < y < \infty$, $-\infty < a < \infty$, and $-\infty < b < \infty$.

Theorem: $\log_{10}(x\,y) = \log_{10} x + \log_{10} y$.

Proof: $\log_{10}(x\,y) = \log_{10}(10^a\, 10^b) = \log_{10}(10^{a+b}) = a + b = \log_{10} x + \log_{10} y$.

Theorem: $\log_{10}(x / y) = \log_{10} x - \log_{10} y$.

Proof: $\log_{10}(x / y) = \log_{10}(10^a / 10^b) = \log_{10}(10^{a-b}) = a - b = \log_{10} x - \log_{10} y$.

Theorem: $\log_{10}(x^b) = b \ \log_{10} x$.

Proof: $\log_{10}(x^b) = \log_{10}((10^a)^b) = \log_{10}(10^{(a\,b)}) = a\,b = b\,a = b \ \log_{10} x$.

17

1. The 1st theorem "$\log_{10}(x\,y) = \log_{10}(x) + \log_{10}(y)$" can be remembered as "the log of a product equals the sum of the logs". Essentially, logs convert multiplication into the "simpler" operation of addition.
2. The 2nd theorem "$\log_{10}(x / y) = \log_{10}(x) - \log_{10}(y)$" can be remembered as "the log of a quotient equals the difference of the logs". Essentially, logs convert division into the "simpler" operation of subtraction.
3. The 3rd theorem "$\log_{10}(x^b) = b \ \log_{10}(x)$" can be remembered as "the log of a power equals the exponent times the log". Essentially, logs convert a power (that is, raising a base to an exponent) into the "simpler" operation of multiplication.

Measure sound pressure amplitude or intensity in decibels

1. P = (power of sound) = $\dfrac{\text{energy of the sound}}{\text{time interval being considered}}$, typically measured in watts = W.

2. $4\pi r^2$ = (area of spherical sound wavefront at distance $r = v\,t$ from source), in meters^2 = m^2.

3. $I(r) = \dfrac{P_{\text{src}}}{4\pi r^2} = \dfrac{\text{power}}{\text{unit area}}$ = (sound intensity at distance r from source of power P_{src}),

   in $\dfrac{\text{watts}}{\text{meter}^2} = \dfrac{\text{W}}{\text{m}^2}$, because energy (and hence power) is conserved if no obstacle absorbs the sound.

4. $\dfrac{P_{\text{src}}}{4\pi r^2} = I(r) = \dfrac{P(r)}{A} \propto \dfrac{a^2(r)}{A}$, where a microphone (with a given area A of its diaphragm at distance r from sound source generating total power P_{src}) senses $I(r)$ = intensity, $P(r)$ = power, and $a^2(r)$ = squared max pressure wave amplitude, which are each inversely proportional to distance^2 = r^2.

5. Perceived volume or loudness depends on $P(r)$, frequency, and sensitivity of the listener's ear.

6. $I_{\text{TOH}} = 10^{-12}\,\dfrac{\text{W}}{\text{m}^2} = \left(\text{one trillionth of a } \dfrac{\text{watt}}{\text{meter}^2}\right)$ = (the threshold of "normal" human hearing).

7. The decibel (dB) is a logarithmic unit of measurement that is typically used to express power, intensity, or amplitude relative to a reference level. A decibel is a tenth of a bell (B). Using the reference I_{TOH},

   $I_{\text{dB}}(r) = 10\,\log_{10}\dfrac{I(r)}{I_{\text{TOH}}} = 10\,\log_{10}\dfrac{P(r)}{P_{\text{TOH}}} = 20\,\log_{10}\dfrac{a(r)}{a_{\text{TOH}}}$ in (dimensionless) decibels.

18

1. Equations #1, #2, and #3: A sound generated in an open space expands as a spherical wavefront whose total power is constant over time as its area $4\pi r^2 = 4\pi(v\,t)^2$ increases.
2. Equations #3, #4, and #7: The sound power $P(r)$ and the maximum pressure amplitude squared $a^2(r)$ sensed by the microphone are both inversely proportional to the square of its distance (r^2) from the source, and both are directly proportional to the given cross sectional area A of the microphone's diaphragm (or the area of the eardrum if a human ear senses the sound). The intensity $I(r)$ of the sound waveform is inversely proportional to the square of its distance (r^2) from the source, but $I(r)$ is a property of the sound itself at that location independent of what sensor is used to measure the sound. That is, for given values of P_{src} and r, intensity $I(r)$ does not depend on the cross sectional area A of the microphone's diaphragm, whereas both $P(r)$ and $a^2(r)$ are proportional to the area A. The symbol "$\propto$" in equation #4 means "is proportional to", so there is a constant b (independent of P_{src} and r and A) such that $P(r) = b\,a^2(r)$.
3. Equation #6: The reference "threshold of hearing (TOH)" intensity level $I_{\text{TOH}} = 10^{-12}$ W/m^2 is considered the lowest intensity sound a person with "normal" hearing can hear.
4. Equation #7: Throughout this slide, assume the microphone's diaphragm of area A (measured in meter^2) is located at any particular given distance r (in meters) from the sound source, which emanates any particular given power P_{src} (in watts). The intensity $I(r)$ (in watts / meter^2) and $I_{\text{dB}}(r)$ (in dimensionless decibels) at distance r are constant for any given sound source, independent of area A. However, the sensed sound power $P(r) = A\,I(r)$ (in watts) and the sensed sound pressure maximum amplitude squared $a^2(r)$ (in SI units of (newtons / meter^2)^2 = (kg / (m s^2))^2, or else in American units of (pounds / inch^2)^2 = psi^2) for the sound wave at distance r are both inversely proportional to the diaphragm area A, and both $P(r)$ and $a^2(r)$ are inversely proportional to the squared distance (r^2) from the sound source. The reference power $P_{\text{TOH}} = A\,I_{\text{TOH}}$ and the corresponding reference amplitude squared a_{TOH}^2 are each proportional to the diaphragm area A, and they represent the threshold of hearing (TOH) for a "normal" human. The last "=" of equation #7 uses the fact that
   $P(r) / P_{\text{TOH}} = a^2(r) / a_{\text{TOH}}^2 = (a(r) / a_{\text{TOH}})^2$
   because power is proportional to pressure amplitude squared, and so
   $10\,\log_{10}(P(r) / P_{\text{TOH}}) = 10\,\log_{10}(a(r) / a_{\text{TOH}})^2 = 20\,\log_{10}(a(r) / a_{\text{TOH}})$
   where the last "=" just above applies a property of logarithms from the previous slide.

Common sounds with estimates of intensity and decibel level

$$\text{Decibel intensity level} = 10 \log_{10}\left(\frac{I}{I_{\text{TOH}}}\right) \quad \text{where intensity} \quad I = \frac{\text{power}}{\text{unit area}} \quad \text{is measured in} \quad \frac{\text{watts}}{\text{meter}^2} = \frac{\text{W}}{\text{m}^2}$$

Source	Intensity (I)	Decibel intensity level	Multiple of TOH intensity
Threshold of Hearing (TOH)	$I_{\text{TOH}} = 10^{-12}$ W/m^2	0 dB	10^0
Rustling leaves	10^{-11} W/m^2	10 dB	10^1
Whisper	10^{-10} W/m^2	20 dB	10^2
Normal conversation	10^{-6} W/m^2	60 dB	10^6
Busy street traffic	10^{-5} W/m^2	70 dB	10^7
Vacuum cleaner	10^{-4} W/m^2	80 dB	10^8
Large orchestra	$6.3*10^{-3}$ W/m^2	98 dB	$10^{9.8}$
iPod at maximum volume level	10^{-2} W/m^2	100 dB	10^{10}
Front rows of a rock concert	10^{-1} W/m^2	110 dB	10^{11}
Threshold of pain	10^1 W/m^2	130 dB	10^{13}
Military jet takeoff	10^2 W/m^2	140 dB	10^{14}
Instant perforation of eardrum	10^4 W/m^2	160 dB	10^{16}

<http://www.glenbrook.k12.il.us/gbssci/Phys/Class/sound/u11l2b.html>

19

1. This table is based on the Web page http://www.glenbrook.k12.il.us/gbssci/Phys/Class/sound/u11l2b.html .
2. The units of sound intensity (I) here are W / m^2 = watts / meter^2 .
3. (Decibel intensity level) = $10 \log_{10}(I / I_{\text{TOH}})$ where $I_{\text{TOH}} = 10^{-12}$ W/m^2 = (sound intensity at the human threshold of hearing).

Power, intensity, and distance from a sound source

1. Consider a sound source emanating power P (watts) and two people at distances r_1 and r_2 who hear intensity levels I_1 and I_2 (W/m^2) corresponding to decibel levels I_{dB1} and I_{dB2}. Then

$$I_1 = \frac{P}{4\pi r_1^2} \text{ and } I_2 = \frac{P}{4\pi r_2^2}; \text{ so } I_1 r_1^2 = \frac{P}{4\pi} = I_2 r_2^2; \text{ so } \frac{I_1}{I_2} = \left(\frac{r_2}{r_1}\right)^2; \text{ so}$$

$$20 \log_{10}\left(\frac{r_2}{r_1}\right) = 10 \log_{10}\left(\frac{r_2}{r_1}\right)^2 = 10 \log_{10}\left(\frac{I_1}{I_2}\right) = 10 \log_{10}\left(\frac{I_1 / I_{\mathrm{TOH}}}{I_2 / I_{\mathrm{TOH}}}\right)$$

$$= 10 \log_{10}\left(\frac{I_1}{I_{\mathrm{TOH}}}\right) - 10 \log_{10}\left(\frac{I_2}{I_{\mathrm{TOH}}}\right) = I_{\mathrm{dB1}} - I_{\mathrm{dB2}}; \text{ so}$$

$$\frac{r_2}{r_1} = 10^{(I_{\mathrm{dB1}} - I_{\mathrm{dB2}})/20}.$$ This is an important result.

2. For example, to reduce the sound intensity of a rock concert heard in the front row ($I_{\mathrm{dB1}} = 110$ dB at distance r_1 from the music source) to the sound intensity of normal conversation ($I_{\mathrm{dB2}} = 60$ dB at distance r_2), move away from r_1 to r_2 where

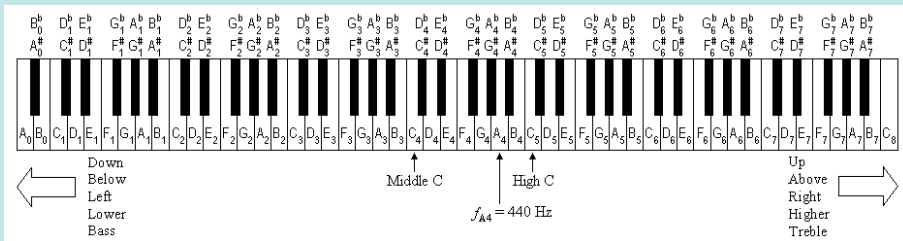
$$\frac{r_2}{r_1} = 10^{(I_{\mathrm{dB1}} - I_{\mathrm{dB2}})/20} = 10^{(110-60)/20} = 10^{2.5} \approx 320.$$

If $r_1 = 15$ feet then $r_2 \approx 320$ (15 feet) $= 4800$ feet ≈ 0.9 mile.

20

1. For the 6-sided equation in item #1 stating $20 \log_{10}(r_2/r_1) = \ldots = I_{\mathrm{dB1}} - I_{\mathrm{dB2}}$,
   - the 1st "=" applies the 3rd theorem at the bottom of slide #17,
   - the 2nd "=" applies the equation $I_1/I_2 = (r_2/r_1)^2$ derived on this slide's line just above this equation,
   - the 3rd "=" applies simple algebra,
   - the 4th "=" applies the 2nd theorem on slide #17, and
   - the 5th "=" applies the definition of expressing sound intensity in decibels from item #7 at the bottom of the slide #18.
2. Based on the table in the previous slide, the example in item #2 of this slide assumes the estimate that the decibel intensity of sound is $I_{\mathrm{dB1}} = 110$ dB at a front row distance of $r_1 = 15$ feet from the source of the music (such as public address loudspeakers).

Names of musical notes on a piano keyboard



Frequency ratio of semitones for the equal-tempered scale

The frequency of any note (e.g., High C = C_5) is twice the frequency of the note one octave below it (e.g., Middle C = C_4), so $f_{C5} = 2\, f_{C4}$.
The ratio of the frequency of any note divided by the frequency of the note one semitone below it (corresponding to the next lower piano key) is always the same for the equal-tempered scale. Denote that ratio by r, where

$$r = \frac{f_{C5}}{f_{B4}} = \frac{f_{B4}}{f_{A\#4}} = \frac{f_{A\#4}}{f_{A4}} = \frac{f_{A4}}{f_{G\#4}} = \cdots = \frac{f_{D4}}{f_{C\#4}} = \frac{f_{C\#4}}{f_{C4}}.$$ So,

$$2\, f_{C4} = f_{C5} = r\, f_{B4} = r^2\, f_{A\#4} = r^3\, f_{A4} = r^4\, f_{G\#4} = \cdots = r^{11}\, f_{C\#4} = r^{12}\, f_{C4}.$$

Therefore, $r^{12} = 2$, so $r = 2^{1/12} \approx 1.059463094$, and hence raising a note by one semitone increases its frequency by 5.9463094 %. 21

1. The Audacity program identifies frequencies by the names of the nearest musical notes (such as A_4 for 440 Hz).
2. As shown on this slide, the piano keyboard includes 88 notes that cover more than 7 octaves. The naming pattern shown above can be extended to notes lower or higher than those a piano can play.
3. The multisided equation
   $2 f_{C4} = f_{C5} = r f_{B4} = r^2 f_{A\#4} = r^3 f_{A4} = r^4 f_{G4\#} = \ldots = r^{11} f_{C4\#} = r^{12} f_{C4}$
   can be derived as follows.
   - The 1st "=" states that note C5 is one octave above note C4 as explained previously on this slide.
   - The 2nd "=" follows by multiplying the 1st and 2nd sides the previous equation ($r = f_{C5} / f_{B4}$) by f_{B4}.
   - The 3rd "=" follows by multiplying the 1st and 3rd sides the previous equation ($r = f_{B4} / f_{A4\#}$) by $r f_{A4\#}$.
   - The 4th "=" follows by multiplying the 1st and 4th sides the previous equation ($r = f_{A4\#} / f_{A4}$) by $r^2 f_{A4}$.
   - The 5th "=" follows by multiplying the 1st and 5th sides the previous equation ($r = f_{A4} / f_{G4\#}$) by $r^3 f_{G4\#}$.
   - Continue applying this iterative method 8 more times (explicitly including the 6 sides represented by "…" in each equation), finally yielding the last "=" in this multisided equation which states $r^{11} f_{C\#4} = r^{12} f_{C4}$.

Compute frequencies & wavelengths of equal-tempered notes

1. The frequency values in the equal-tempered scale commonly assume that the frequency of the note A_4 is exactly $f_{A4} = 440$ Hz (in units of Hz = Hertz = cycles / second).
2. The note A in any octave corresponds to note index $n = 9$ because it is 9 semitones above the note C at the bottom of the octave. The note A_4 has the octave index $u = 4$.
3. The speed v of a wave that has wavelength λ and frequency f is
   $v = \lambda f$, typically measured in units of $\frac{\text{meters}}{\text{second}} = \left(\frac{\text{meters}}{\text{cycle}}\right)\left(\frac{\text{cycles}}{\text{second}}\right)$.
4. Wavelength values in the table on the next slide assume the speed of sound is
   v = 344 meters/second = 34,400 cm/s, which is approximate for dry air at 70 degrees Fahrenheit.
5. The note C_0 (which corresponds to note index $n = 0$ and octave index $u = 0$) is
   (12 semitones/octave) (4 octaves) + (9 semitones) = 57 semitones below the reference note A_4.
6. In general, the number of semitones from the note A_4 to the note with any note index n and any octave index u is $12(u-4) + (n-9) = n - 57 + 12u$, so the number of octaves (with 12 semitone intervals per octave) from note A_4 of frequency $f_{A4} = 440$ Hz to the note with those indexes is $(n-57)/12 + u$.
7. For any note index n and any octave index u, the frequency f and the wavelength λ of the corresponding note are related by
   $(344 \text{ m/s}) / \lambda \approx v / \lambda = f = f_{A4}\, 2^{(n-57)/12 + u} \approx (16.35159783 \text{ Hz})\, 2^{(n/12 + u)}$, so that
   $\lambda = v / f \approx (34{,}400 \text{ cm/s}) / \left((440 / \text{s})\, 2^{(n-57)/12 + u}\right) \approx (2103.769941 \text{ cm})\, 2^{-(n/12 + u)}$.
8. The lowest (leftmost) note on a piano keyboard is A_0, which corresponds to
   note index $n_{A0} = 9$, octave $u_{A0} = 0$, frequency $f_{A0} = 27.50$ Hz, and wavelength $\lambda_{A0} \approx 1250$ cm.
9. The highest (rightmost) note on a piano keyboard is C_8, which corresponds to
   note index $n_{C8} = 0$, octave $u_{C8} = 8$, frequency $f_{C8} \approx 4186.01$ Hz, and wavelength $\lambda_{C8} \approx 8.22$ cm.
10. Based on the previous two sentences, the number of keys on a piano is
    $12(u_{C8} - u_{A0}) + n_{C8} - n_{A0} + 1 = 12(8-0) + 0 - 9 + 1 = 88$ keys, covering more than 7 octaves.

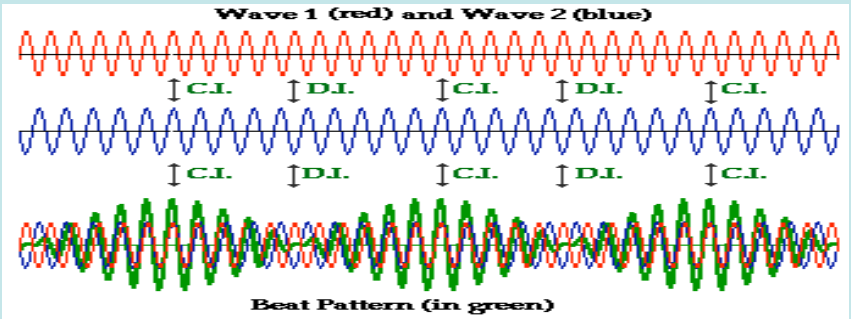
22

Frequencies and wavelengths for the equal-tempered scale

n	Note	Octave u=0	Octave u=1	Octave u=2	Octave u=3	Octave u=4	Octave u=5	Octave u=6	Octave u=7
0	C	16.35 Hz 2100 cm	32.70 Hz 1050 cm	65.41 Hz 526 cm	130.81 Hz 263 cm	261.63 Hz 131 cm	523.25 Hz 65.7 cm	1046.50 Hz 32.9 cm	2093.00 Hz 16.4 cm
1	$C^{\#}$ $= D^{b}$	17.32 Hz 1990 cm	34.65 Hz 993 cm	69.30 Hz 496 cm	138.59 Hz 248 cm	277.18 Hz 124 cm	554.37 Hz 62.1 cm	1108.73 Hz 31.0 cm	2217.46 Hz 15.5 cm
2	D	18.35 Hz 1870 cm	36.71 Hz 937 cm	73.42 Hz 469 cm	146.83 Hz 234 cm	293.66 Hz 117 cm	587.33Hz 58.6 cm	1174.66 Hz 29.3 cm	2349.32 Hz 14.6 cm
3	$D^{\#}$ $= E^{b}$	19.45 Hz 1770 cm	38.89 Hz 885 cm	77.78 Hz 442 cm	155.56 Hz 221 cm	311.13 Hz 111 cm	622.25 Hz 55.3 cm	1244.51 Hz 27.6 cm	2489.02 Hz 13.8 cm
4	E	20.60 Hz 1670 cm	41.20 Hz 835 cm	82.41 Hz 417 cm	164.81 Hz 209 cm	329.63 Hz 104 cm	659.26 Hz 52.2 cm	1318.51 Hz 26.1 cm	2637.02 Hz 13.0 cm
5	F	21.83 Hz 1580 cm	43.65 Hz 788 cm	87.31 Hz 394 cm	174.61 Hz 197 cm	349.23 Hz 98.5 cm	698.46 Hz 49.3 cm	1396.91 Hz 24.6 cm	2793.83 Hz 12.3 cm
6	$F^{\#}$ $= G^{b}$	23.12 Hz 1490 cm	46.25 Hz 744 cm	92.50 Hz 372 cm	185.00 Hz 186 cm	369.99 Hz 93.0 cm	739.99 Hz 46.5 cm	1479.98 Hz 23.2 cm	2959.96 Hz 11.6 cm
7	G	24.50 Hz 1400 cm	49.00 Hz 702 cm	98.00 Hz 351 cm	196.00 Hz 176 cm	392.00 Hz 87.8 cm	783.99 Hz 43.9 cm	1567.98 Hz 21.9 cm	3135.96 Hz 11.0 cm
8	$G^{\#}$ $= A^{b}$	25.96 Hz 1330 cm	51.91 Hz 663 cm	103.83 Hz 331 cm	207.65 Hz 166 cm	415.30 Hz 82.8 cm	830.61 Hz 41.4 cm	1661.22 Hz 20.7 cm	3322.44 Hz 10.4 cm
9	A	27.50 Hz 1250 cm	55.00 Hz 625 cm	110.00 Hz 313 cm	220.00 Hz 156 cm	440.00 Hz 78.2 cm	880.00 Hz 39.1 cm	1760.00 Hz 19.5 cm	3520.00 Hz 9.77 cm
10	$A^{\#}$ $= B^{b}$	29.14 Hz 1180 cm	58.27 Hz 590 cm	116.54 Hz 295 cm	233.08 Hz 148 cm	466.16 Hz 73.8 cm	932.33Hz 36.9 cm	1864.66 Hz 18.4 cm	3729.31 Hz 9.22 cm
11	B	30.87 Hz 1110 cm	61.74 Hz 557 cm	123.47 Hz 279 cm	246.94 Hz 139 cm	493.88 Hz 69.7 cm	987.77 Hz 34.8 cm	1975.53 Hz 17.4 cm	3951.07 Hz 8.71 cm

1. The frequency values in this table use the common assumption that the frequency of the note A_4 is exactly $f_{A4} = 440$ Hertz. The note A_4 corresponds to note index $n = 9$ and octave index $u = 4$.
2. The wavelength values in this table also assume that the speed of sound is $v = 344$ meters/second = 34,400 cm/s, which is approximately true in dry air at 70 degrees Fahrenheit.
3. The note C_0 (which corresponds to indexes $n = 0$ and $u = 0$) is $12\,(4) + 9 = 57$ semitones below the reference note A_4.
4. The note C_4 is called "middle C" (the first C to the left of the center of a piano keyboard); note C_5 is called "high C" (the first C to the right of the center of a piano keyboard).
5. In general, the number of semitones from the note A_4 to the note with any note index n and any octave index u is
   $12\,(u - 4) + (n - 9) = n - 57 + 12\,u$,
   so the number of 12-semitone octaves from note A_4 to the note with those indexes is
   $(n - 57 + 12\,u) / 12 = (n - 57)/12 + u$.
6. For any note index n and any octave index u, the frequency f and the wavelength λ of the corresponding note are related by
   $(344 \text{ m/s}) / \lambda \approx v / \lambda = f = f_{A4}\ 2^{((n-57)/12+u)} \approx (16.35159783 \text{ Hz})\ 2^{(n/12+u)}$,
   so that
   $\lambda = v / f \approx ((34{,}400 \text{ cm/s}) / ((440 / \text{s})\ 2^{((n-57)/12+u)}) \approx (2103.769941 \text{ cm})\ 2^{(n/12+u)}$.
7. The lowest (leftmost) note on a piano keyboard is A_0, which corresponds to $n = 9$, $u = 0$, $f = 27.50$ Hz, and $\lambda \approx 1250$ cm.
8. The highest (rightmost) note on a piano keyboard is C_8, which corresponds to $n = 0$, $u = 8$, $f \approx 4186.01$ Hz, and $\lambda \approx 8.22$ cm.
9. By items #8 and #7 just above, the number of keys on a piano is
   $12\,(u_{C8} - u_{A0}) + n_{C8} - n_{A0} + 1 = 12\,(8 - 0) + 0 - 9 + 1 = 88$.

Superposition (that is, adding) of waves can form beats



Beats are the periodic fluctuations heard in the intensity of a sound when two sound waves of very similar frequencies interfere with one another. In the figure, the red sine wave at the top is added to the slightly lower frequency blue sine wave in the middle, forming the green beating waveform at the bottom.

C.I. = Constructive Interference: When two compressions come together, they combine to make a region of very compressed air (very high pressure). When two rarefactions come together, they combine to make a region of very spread-out air (very low pressure). Both of these are examples of constructive interference. A location of continually constructive interference is called an antinode.

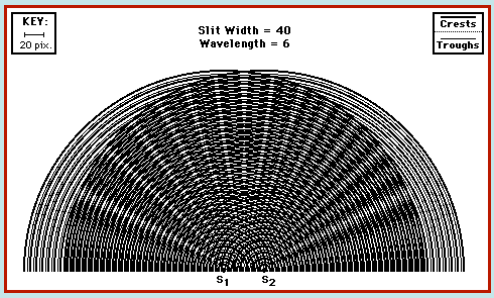
D.I. = Destructive Interference: When a compression from one wave meets a rarefaction from another wave with equal amplitude, they cancel out, so that there's no disturbance in air pressure at that point. A location of continually destructive interference is called a node.

The beat frequency (that is, the rate in repetitions/second at which the loud beats repeat) is equal to the difference in frequency of the two notes that interfere to produce the beats.

24

1. If you listen to two tones that have the same intensity but have two very slightly different frequencies, you hear an oscillating pattern of loud sound and then silence, loud sound and then silence…called **beats**. The frequency of that oscillation (*how often* the beats occur) is called the **beat frequency**.
2. For example, if two sound waves with frequencies of 300 Hz and 302 Hz are heard simultaneously, a beat frequency of 302 Hz – 300 Hz = 2 Hz will be detected; that is, the resulting waveform will reach its maximum intensity 2 times per second.
3. In the plots shown on this slide, if the red wave 1 has the frequency 110 Hz (which will contain exactly 11 cycles in 0.1 second) then the blue wave 2 has the frequency 100 Hz (which will contain exactly 10 cycles in that same 0.1 second). The resulting beat frequency of the green sum of those waves is then 110 Hz – 100 Hz = 10 Hz, which reaches its maximum intensity 10 times per second, which means it repeats after 1 / (10 cycles/second) = 0.1 seconds/cycle .
4. A piano tuner can use beats to tune a piano string. She will pluck the string and tap a tuning fork for the proper note at the same time. If the two sound sources -- the piano string and the tuning fork -- produce detectable beats then their frequencies are not identical. She will then adjust the tension of the piano string and repeat the process until the beats can no longer be heard because the time between beats gets too long.

Two sources of sound with the same frequency interfere



Animated interference pattern formed by two sound wave sources having the same wavelengths and frequencies

<http://www.glenbrook.k12.il.us/gbssci/phys/mmedia/waves/ipl.html>

The number of antinodal lines in front of the two sources is about twice the number of wavelengths between the sources.

Java applets for interference:

http://webphysics.davidson.edu/Applets/Ripple/Ripple_JS.html

http://id.mind.net/~zona/mstm/physics/waves/interference/twoSource/TwoSourceInterference1.html

25

1. The waveforms are like ripples caused on the surface of a pond when it is repeatedly tapped (at the same rate) by two sources s_1 and s_2. The circular ripples expanding from source s_1 interfere with (that is, add to the instantaneous height of) the circular ripples expanding from source s_2 at each location and time instant. This produces "antinodal lines" with high amplitude oscillation where the interference is constructive, and it produces "nodal lines" where destructive interference causes the ripples to continuously cancel each other.
2. This animation shows that the number of antinodal lines and nodal lines decreases as the wavelength of both sets of ripples increases (that is, as their frequency decreases) for the same spacing between the two sources (that is, for a constant "slit width").

Mathematical basis of spectrum analysis and the FFT

1. Any real function of continuous time equals the sum of a possibly infinite number of sine waves of various frequencies. The (continuous) Fourier transform computes the amplitude and initial phase angle for those sine waves at each frequency, where the frequencies may need to be infinitesimally closely spaced for some functions. It was named after Jean Baptist Joseph Fourier, 1768 – 1830.

2. In particular, if the function of continuous time $g(t)$ is periodic with a repetition rate of f_P (typically in units of cycles/second = Hertz = Hz) so its period is $1/f_P$ (seconds/cycle), which means $g(t) = g(t + u/f_P)$ for all times t and for all integers $u = 0, \pm1, \pm2, \pm3, \ldots,$ then that function equals the sum of sine waves (called a Fourier series) whose frequencies are only integer multiples of the fundamental frequency f_P, where $m f_P$ is called the mth harmonic frequency for any positive integer $m = 1, 2, 3, \ldots$.

3. If an analog to digital converter (ADC) digitizes an even number N of samples of the function of time $g(t)$ at a uniform sampling frequency of F_s (samples/second) spanning a duration of N/F_s (seconds), then each of those samples (at times $t_k = k/F_s$ for each sample index $k = 0, 1, 2, \ldots, N–1$) equals a sum of $N/2 + 1$ sine waves at uniformly spaced frequencies $f_n = n F_s / N$ for frequency indexes $n = 0, 1, 2, \ldots, N/2$. That is,

$$g\left(\frac{k}{F_s}\right) = g(t_k) = \sum_{n=0}^{N/2} a_n \sin\left(\frac{2\pi nk}{N} + \theta_n\right)$$

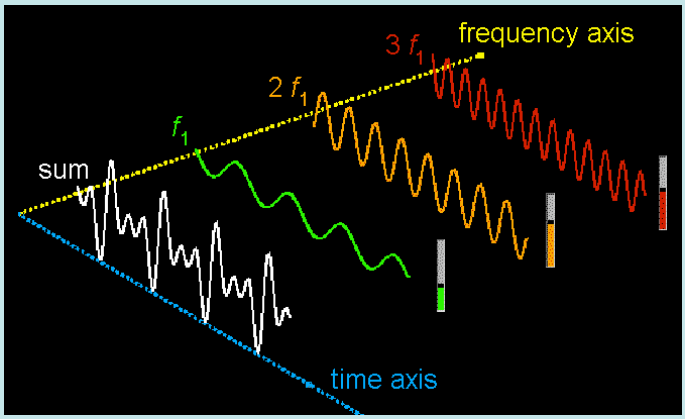
$$= a_0 \sin\theta_0 + a_1 \sin\left(\frac{2\pi k}{N} + \theta_1\right) + a_2 \sin\left(\frac{4\pi k}{N} + \theta_2\right) + a_3 \sin\left(\frac{6\pi k}{N} + \theta_3\right) + \ldots + a_{N/2} \sin\left(\pi k + \theta_{N/2}\right)$$

where the nonnegative amplitudes $a_n \geq 0$ and the phase angles θ_n of those $N/2 + 1$ sine waves can be computed from the N given samples by the fast Fourier transform (FFT) algorithm.

26

1. The definition of the discrete Fourier transform will be explained in these notes for this slide, although this involves some mathematical concepts (like a few trigonometric identities, complex number theory, and Euler's formula) that are beyond the level assumed by the rest of this presentation.
2. The implicit assumption is made throughout this slide that the function $g(t)$ is "purely real" in the sense of complex number theory; that is, the imaginary part of $g(t)$ is assumed to equal zero for all time $-\infty < t < \infty$. This assumption gives the Fourier transforms "complex conjugate symmetry". In terms of data sampled by an ADC at a uniform rate of F_s (measured in Hz = samples/second), complex conjugate symmetry means that $G(F_s - f_n) = G(f_{N-n}) = \text{conjugate}(G(f_n))$ for all $n = 0, 1, 2, \ldots, N-1$, where $G(f_n) = G(n F_s / N)$ is the FFT complex spectrum of time-domain real samples $g(t_k) = g(k / F_s)$ as defined below in the notes for this slide.
3. The requirement that the total number of samples N should be an *even* integer is also assumed for the particular representation shown on this slide with the sum of only $N/2 + 1$ sines.
4. For any frequency index $n = 0, 1, 2, 3, \ldots, N - 1$ corresponding to the frequency $f_n = n F_s / N$, define
$c_n = (1 / N)$ (sum of $g(k / F_s) \cos(2\pi k n / N)$ for all $k = 0, 1, 2, \ldots, N - 1$) and
$d_n = (1 / N)$ (sum of $g(k / F_s) \sin(2\pi k n / N)$ for all $k = 0, 1, 2, \ldots, N - 1$) . Also define
j = (one of the two imaginary unit constants of complex number theory, such that $j^2 = -1$).
5. The discrete Fourier transform (the DFT, which can be computed by the much more efficient FFT algorithm) at any discrete frequency $f_n = n F_s / N$ (measured in Hz) is
$G(f_n) = G(n F_s / N) = c_n + j\, d_n = (1 / N)$ (sum of $g(k / F_s) \exp(j\, 2\pi k n / N)$ for all $k = 0, 1, 2, \ldots, N - 1$) ,
using Euler's formula that $e^{(j\,x)} = \exp(j\,x) = \cos(x) + j \sin(x)$ for any real number x where $e \approx 2.718281828495$ is the base of natural (Naperian) logarithms.
6. Although that DFT function $G(f_n)$ can be evaluated for any integer n, those values of the complex spectrum repeat with a period of N frequency-domain samples; that is, $G(f_n) = G(f_{n+N})$ for any integer n .
7. Let $\mu = G(0) = c_0$ = (the average of all N digitized samples)
= $(1 / N)$ (the sum of all N samples $g(k/F_s)$ for $k = 0, 1, 2, \ldots, N - 1$) .
If $\mu \geq 0$ then $a_0 = \mu \geq 0$ and $\theta_0 = \pi / 2$, so the first summation term (at frequency $f_0 = 0$ Hertz) is $a_0 \sin\theta_0 = \mu$;
otherwise,
if $\mu < 0$ then $a_0 = -\mu > 0$ and $\theta_0 = -\pi / 2$, so the first summation term (at frequency $f_0 = 0$ Hertz) is again $a_0 \sin\theta_0 = \mu$.
8. Corresponding to (half the sampling frequency) $= F_s / 2 = f_{(N/2)} = (n F_s / N$ where $n = N / 2$),
let $h = G(F_s / 2) = c_{(N/2)}$ = (the sum of $(-1)^k g(k/F_s)$ for all $k = 0, 1, 2, \ldots, N - 1$) .
If $h \geq 0$ then $a_{N/2} = h \geq 0$ and $\theta_{N/2} = \pi / 2$, so the last summation term (at frequency $f_{N/2} = F_s / 2$) is $a_{N/2} \sin(\pi k + \theta_{N/2}) = (-1)^k h$; otherwise,
if $h < 0$ then $a_{N/2} = -h > 0$ and $\theta_{N/2} = -\pi / 2$, so the last summation term (at frequency $f_{N/2} = F_s / 2$) is again $a_{N/2} \sin(\pi k + \theta_{N/2}) = (-1)^k h$.
9. For any $n = 1, 2, 3, \ldots, N / 2 - 1$, the amplitude of the sine term at frequency $n F_s / N$ is
$a_n = 2\ |G(f_n)| = 2\ \text{sqrt}(c_n^2 + d_n^2) \geq 0$,
and the phase angle θ_n of that sine term $a_n \sin(2\pi k n / N + \theta_n)$ is
$\theta_n = (\arg(G(f_n)) + \pi / 2) \bmod (2\pi) = (\text{atan2}(d_n, c_n) + \pi / 2) \bmod (2\pi) = \text{atan2}(c_n, d_n) \bmod (2\pi)$, so
$\sin(\theta_n) = 2 c_n / a_n$ and $\cos(\theta_n) = 2 d_n / a_n$, where
$\arg(z) = z / |z|$ is the argument function of any nonzero complex number z , where
$\text{atan2}(y, x) = \beta$ is the arctangent function such that $\sin(\beta) = y / r$ and $\cos(\beta) = x / r$ assuming $r = \text{sqrt}(x^2 + y^2)$, where
$x \bmod y = x - y\ \text{floor}(x / y)$ is the "modulo" function which calculates the remainder of any real x divided by any nonzero real y , and where
$\text{floor}(x)$ = (the greatest integer that is $\leq x$) .

Graphical example and demos of spectrum analysis



Sum of 3 harmonics (based on http://zone.ni.com/cms/images/devzone/tut/a/8c34be30580.gif)

Fourier series and waves animated applet: http://www.kettering.edu/~drussell/Demos/Fourier/Fourier.html

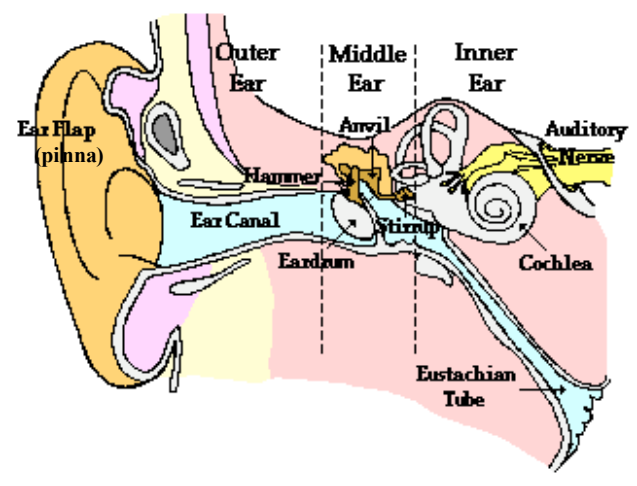
Applet to adjust amplitudes of first 8 harmonics: http://www.earlevel.com/Digital%20Audio/harmonigraf.html

Nice demo to listen to Fourier series harmonics: http://www.jhu.edu/~signals/listen-new/listen-newindex.htm

27

1. This graphical example shows that adding the green ("f_1"), yellow ("$2f_1$"), and red ("$3f_1$") sine waves at each instant of time t produces the white ("sum") waveform. The Fourier transform (FFT) of the sum waveform calculates the frequencies, amplitudes, and initial phase angles of the sine waves that will add up to that sum waveform.
2. Because the 3 sine waves shown here have harmonic frequencies which are the first 3 multiples of the fundamental frequency f_1, the sum waveform must repeat with a period of $1/f_1$ (in seconds).
3. The three Web links at the bottom of this slide help to illustrate the FTT. The last two are interactive, and the last one demonstrates the effect that changing the FFT spectrum has on the sound of its corresponding time-domain waveform.

Anatomy of the human ear



<http://www.glenbrook.k12.il.us/gbssci/phys/Class/sound/u11l2d.html>

The tiniest hair cells are near the entrance of the cochlea and are sensitive to the highest frequencies of sound (up to about 20 KHz). Somewhat larger hair cells are near the apex of the cochlea and respond to low frequencies (down to about 20 Hz).

The hammer (malleaus), anvil (incus), and stirrup (stapes) bones and coiled cochlea are unique to mammals.

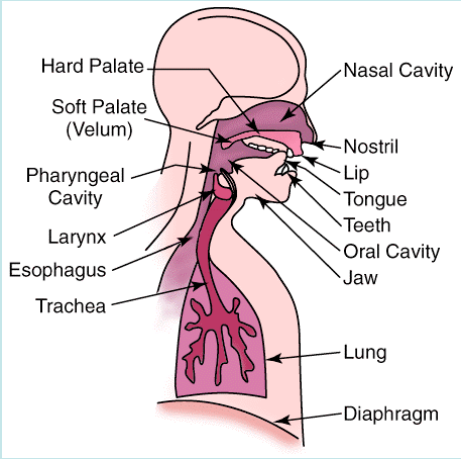
<http://biology.clc.uc.edu/fankhauser/Labs/Anatomy_&_Physiology/A&P202/Special_Senses/Histology_Ear.htm>

A traveling wave in the liquid-filled cochlea causes thousands of tiny, frequency-sensitive hair cells (cilia) to vibrate, which sends electrical impulses along the auditory nerve corresponding to the hairs that are moving. Higher amplitude sounds at any particular pitch produce more rapidly repeating nerve impulses from the corresponding hair.

28

1. The repetition rates of electrical impulses sent along the auditory nerve bundle to the brain convey information about the amplitudes of sounds at each of thousands of frequencies from about 20 Hz to about 20,000 Hz = 20 KHz.
2. Essentially, the frequency and magnitude information that could be computed by an FFT is sent to the brain, but the phase angle information is discarded as being irrelevant for the brain to discern information from the sound.
3. The fluid-filled "semicircular canals" in the inner ear above entrance to the cochlea provide a person's sense of balance.
4. The Eustachian tube connects to the throat to allow air pressure to be balanced on both sides of the eardrum while swallowing.
5. The outer ear flap on the side of the head is also called the pinna or auricle. It modifies the volume of sounds from in front or behind a person to help distinguish the direction the sound is coming from. The tiny time delay (less than about 0.0005 second = 500 μs) between when a sound arrives at a person's left and right ears also helps the brain discern the direction the sound is coming from.

Anatomy of the human vocal tract



<http://cobweb.ecn.purdue.edu/~ee649/notes/figures/vocal_apparatus.gif>

29

1. The "larynx" is also called the "voicebox" and contains the vocal cords (also called vocal folds).
2. When muscles pull on the vocal cords putting them in tension, they are caused to vibrate by air pressure rising from the lungs through the trachea (airway), producing periodic pulses of air that resonate through the upper parts of the vocal tract including the upper throat, mouth, and nasal cavity. When a person sings, increasing the pull of the muscles that put vocal cords in tension makes them vibrate faster and produces a higher pitch.
3. The pharynx located just above the larynx splits the trachea (i.e., the airway or wind pipe to the lungs) and the esophagus (i.e., the food pipe to the stomach).

English phonemes (elemental sounds)

Pure vowels (unchanging sound)			
Notation	Example	Notation	Example
ee	heat	ah	father
I	hit	aw	call
e	head	U	put
ae	had	oo	cool
uh	the	ʌ	ton
		er	bird

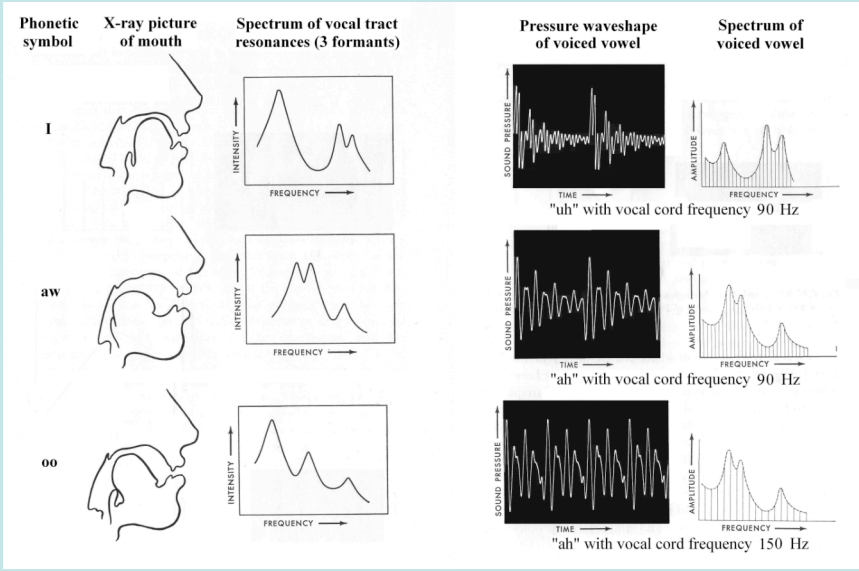
Diphthong vowels (changing sound)	
Notation	Example
ou	tone
ei	take
ai	might
au	shout
oi	toil

Place of articulating the consonant	Manner of articulating the consonant; (unvoiced) or voiced				
	Plosive	Fricative	Semi-vowel	Liquids ("l" = lateral)	Nasal
Labial (lips)	(p) b		w		m
Labio-dental (lips-teeth)		(f) v			
Dental (teeth)		(θ) th			
Alveolar (gums)	(t) d	(s) z	y	l r	n
Palatal (hard front roof)		(sh) zh			
Velar (soft back roof)	(k) g				ng
Glottal (glottis in back)		(h)			

30

1. Speech recognition and speech synthesis are very important applications of acoustic signal processing. Speech recognition involves programming a computer to determine what words are being spoken based on digitized samples of an acoustic signal. Speech synthesis involves programming a computer to generate spoken words on a loudspeaker or earphone based on the text of the message to be spoken, which is especially helpful for people who are sight-impaired.
2. In order to discuss these applications, it is first necessary to understand that sentences consist of words, which consist of syllables, which consist of elemental sounds called phonemes. The 3 tables on this slide show the phonemes used to speak English. Other languages may use different phonemes.
3. All vowel phonemes and some consonant phonemes are "voiced"; that is, they are produced while the vocal cords vibrate to emit pulses of air which resonate through the upper vocal tract. Some consonant phonemes are "unvoiced"; that is, they are produced by blowing air past relaxed vocal cords that do not vibrate significantly. These tables ignore the fact that all whispered phonemes are unvoiced.
4. Vowels are classified as "pure" if their spectral content (that is, power at various frequencies) remains essentially constant as they are spoken, or else vowels are classified as "diphthongs" if their spectral content changes while being spoken.
5. The bottom table shows how consonant phonemes can be classified by whether they are unvoiced or voiced, by the way they are produced (such as "plosives" that suddenly release air, "fricatives" that force air through a constricted gap, or "nasal" phonemes that expel air through the nose while the mouth is closed), and by where in the vocal tract they are produced.

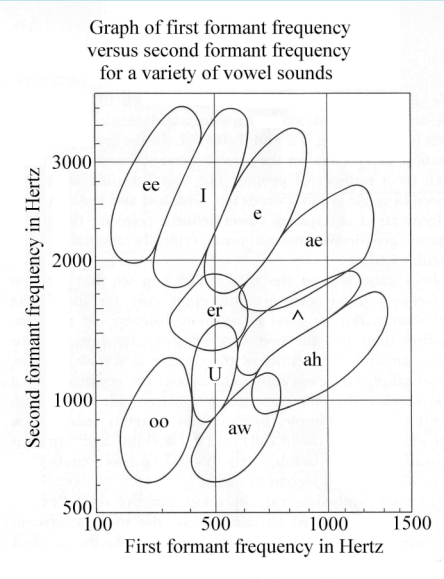
Phonetic speech mouth, formants, wave, and spectrum

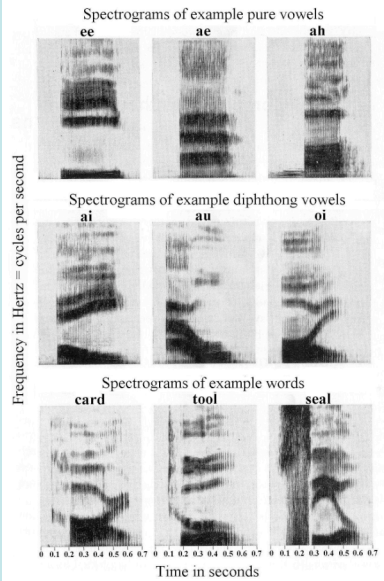


31

1. The left and right figures on this slide are both based on the book *Speech Synthesis: An Experiment in Electronic Speech Production* by Cecil H. Coker, Peter B. Denes, and Elliot N. Pinson, Bell Telephone Laboratories, © 1963.
2. The mouth and resonance spectra on the left side of this slide are from page 66 of that book.
3. The top time-domain waveform and spectrum on the right side of this slide are slightly edited from Fig. 4.18 (b) on page 68 of that book.
4. The lower two time-domain waveforms and spectra on the right side of this slide are slightly edited from Fig. 4.17 (a) and (b) on page 67 of that book.
5. The shape of the vocal tract for the phoneme "I" (as in "hit") shown at the upper left has one formant (or resonance) with high magnitude at low frequency, plus two formants with lower magnitudes at higher frequencies.
6. The phoneme "aw" (as in "call") shown at the middle left has two formants with high magnitude and low frequencies, plus one formant with lower magnitude at higher frequency.
7. The phoneme "oo" (as in "cool") shown at the bottom left has three formants whose magnitudes decrease as their frequencies increase.
8. Each of the 3 plots in the 3rd column shows sound pressure versus time for about 24 milliseconds. Therefore, the top 2 plots each show the response of the vocal tract to about (0.024 s) (90 pulses/s) ≈ 2 pulses of the vocal cords, and the bottom plot shows the response to about (0.024 s) (150 pulses/s) ≈ 4 pulses of the vocal cords.
9. In the spectral plots at the far right column, the vertical lines under the formant curves of the top and middle plots are spaced 90 Hz apart, while the vertical lines of the bottom plot are spaced 150 Hz apart. The 3 formants for "uh" in the top spectral plot (for the phoneme "uh" as in "the") have greater magnitude at higher frequencies compared with the 3 formants in the middle and bottom spectral plots (for the phoneme "ah" as in "father").

Phonetic speech recognition and synthesis





32

1. The left and right figures on this slide are both based on the book *Speech Synthesis: An Experiment in Electronic Speech Production* by Cecil H. Coker, Peter B. Denes, and Elliot N. Pinson, Bell Telephone Laboratories, © 1963.
2. The labeling of the graph on the left side of this slide showing clustering of the first and second formant frequencies for a variety of vowels is slightly edited from page 92 of that book.
3. The labeling of the spectrograms on the right side of this slide is slightly edited from page 98 of that book. The 3 spectrograms in the top row show that the magnitude (shown by the darkness of the spectrogram signal) stays essentially constant with time for each of the 3 pure vowels “ee” as in “heat”, “ae” as in “had”, and “ah” as in “father”. In contrast, the 3 spectrograms in the middle row show that the magnitudes change considerably (increasing or decreasing the magnitudes at various frequencies) with time as each of the 3 diphthong vowels “ai” as in “might”, “au” as in “shout”, and “oi” as in “toil” is pronounced). The 3 spectrograms in the bottom row show how phonemes are combined into words, including the essentially uniform excitation of a broad range of frequencies (that is, “broadband noise”) for the unvoiced fricative consonant “s” in “seal”.