LING 450/550
3 – Acoustics

Limitations of Transcription

Transcription uses discrete units (segments) to describe a dynamic, continuous phenomenon (speech).

- As such, even the narrowest transcription is an imperfect record of the speech signal it represents. In reality, the boundaries between segments are not always clear, and the relative durations of segments can vary quite a bit.
- Nonetheless, transcription is a powerful tool for representing certain aspects of the speech signal (e.g., the sequence of sounds, dialectal variation, which sounds are phonemic in a given language, etc.)

http://www.youtube.com/watch?v=0G01HNAkIbg

https://www.youtube.com/watch?v=2Q9MTJokc4A

SocioPhonetics at UW
Pacific Northwest English

https://zeos.ling.washington.edu/~PNWEnglish/index.php

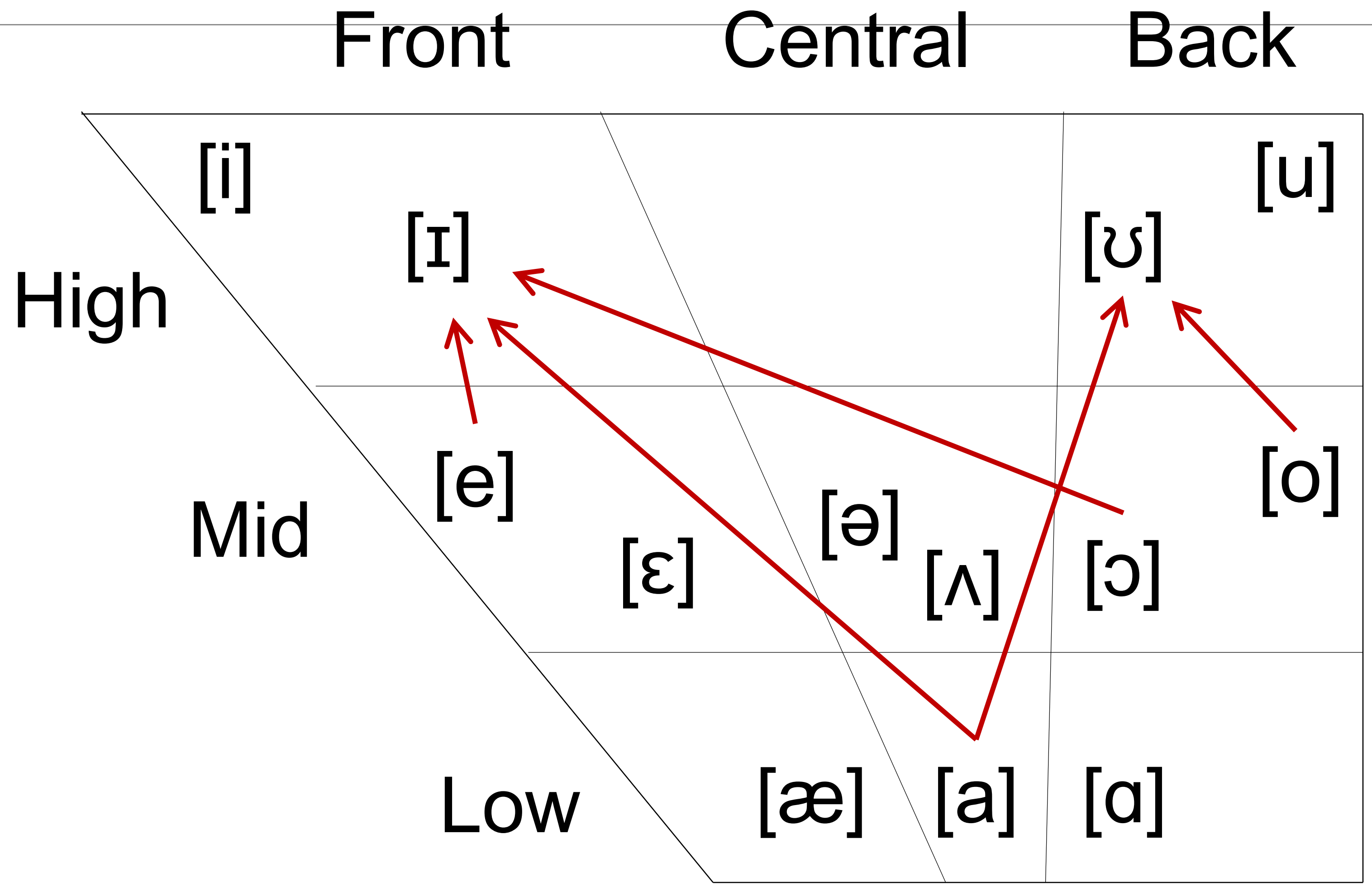
American English

WARMING UP

English consonants

	Bilabial	Labio-dental	Dental	Alveolar	Palato-alveolar	Retroflex	Palatal	Velar	Glottal
Stop	p b			t d				k g	ʔ
Nasal	m			n				ŋ	
Tap				ɾ					
Fricative		f v	θ ð	s z	ʃ ʒ				h
Affricate					ʧ ʤ				
Approximant	w					ɹ	j		
Lateral				l					

American English Vowels



English transcription practice: Consonants & Vowels

1.
2.
3.
4.
5.
6.
7.
8.
9.
10.

English transcription practice: Consonants & Vowels

1. /ˈbæd/
2. /ˈsɪti/
3. /dʒɪn/
4. /hæŋ/
5. /ˈkaɪt/
6. /ˈłɛŋϑ/
7. /ˈɹaɪ/
8. /ˈʃaɪ/
9. /ˈϑaɪ/
10. /ˈðaɪ/

Project

Acoustic Theory

PHYSICAL PROPERTIES OF SOUND

READ LADEFOGED & JOHNSON, CHAPTER 8 (PP. 197-208)

adapted from slides by Richard Wright and Dan McCloy

Pressure

Air is a *fluid*, or something that flows.

- Note: a *fluid* is not the same as a *liquid*! Liquids can be fluids, but gases (like air) and plasmas are also fluids.

When objects move in a fluid medium such as air, they create localized changes in the *pressure* of the fluid.

- This might be easier to imagine with a liquid, which is generally denser than air.

In a *homogeneous fluid* (one that has the same properties throughout), pressure changes begin at the source of movement and propagate outward spherically.

- In other words, there is a chain reaction of pressure changes through the homogeneous fluid in all directions.

What is Sound?

Sound is the perceptual response to pressure fluctuations in the atmosphere.

- Technically, *sound* does not exist in the atmosphere; *pressure waves* do. Sound exists in the awareness of the hearer. In practice, however, it is common to talk about *sound waves* and *sound pressure*.

http://www.youtube.com/watch?v=Si-OYX20FRs

Waves

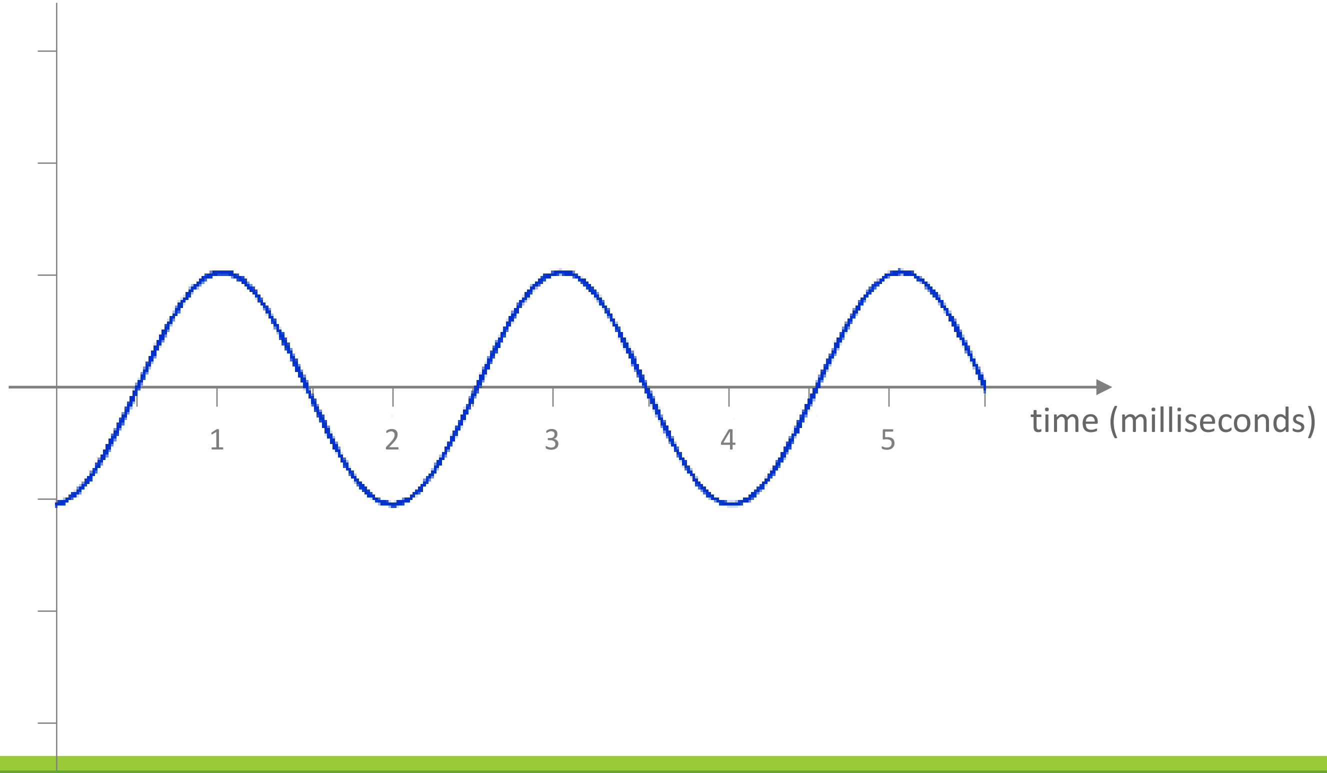
Many sounds, including most sounds of speech, are caused by *vibration*.

- For instance, musical instruments generally involve something that vibrates (a string, a reed, a membrane, …).
- Your vocal folds also vibrate.

When objects are vibrating in a fluid medium, the resulting pressure changes are *periodic*, meaning the changes fluctuate in a pattern that repeats over time. Such repeating patterns are called *waves*.

http://www.acs.psu.edu/drussell/Demos/waves-intro/peoplewave.gif

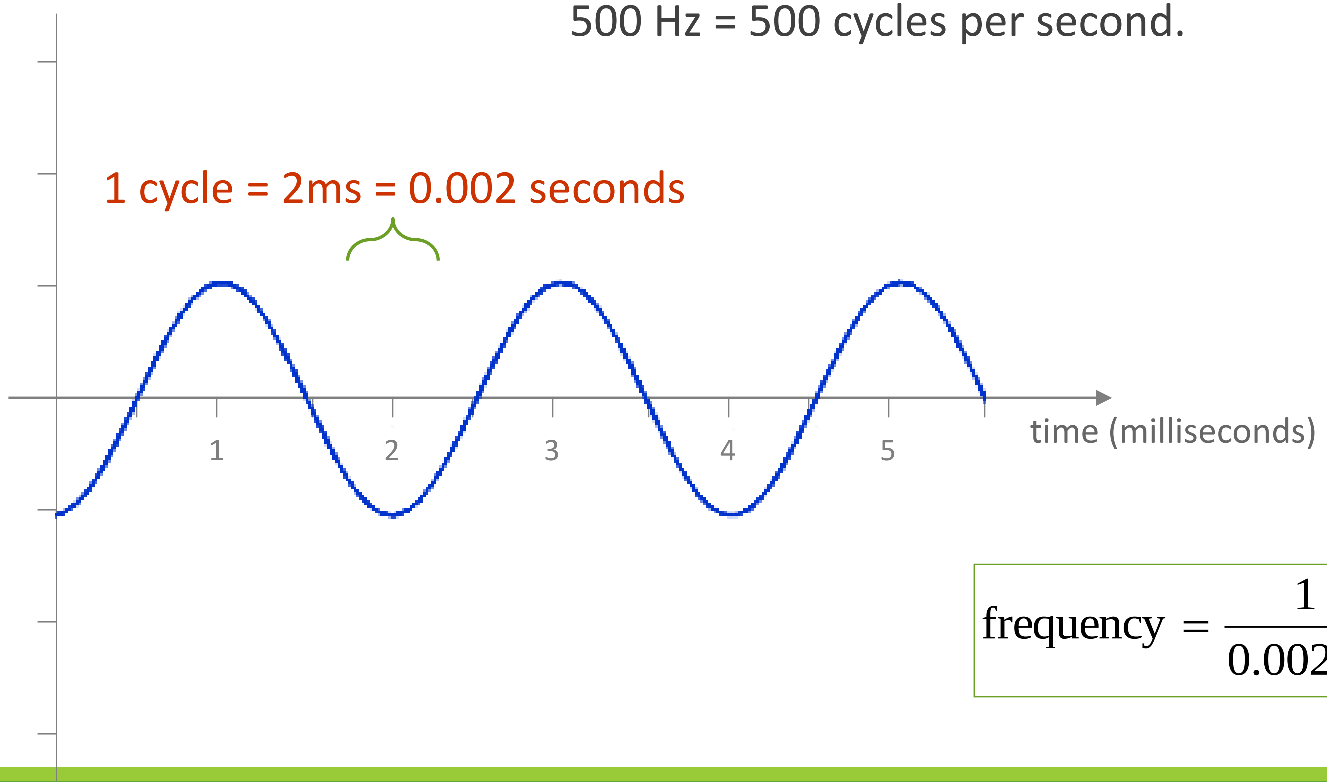
Plotting a Waveform



Frequency

The *frequency* of a wave is a measure of how many times it repeats in a given amount of time.

- Wave repetitions are usually called *cycles*.
- Frequency is measured in *Hertz* (Hz): 500 Hz = 500 cycles per second.



$$\text{frequency} = \frac{1\ \text{cycle}}{0.002\ \text{seconds}} = 500\,\text{Hz}$$

(Bad) Acoustics Joke

If a tree falls in the forest and no one is around to hear it, does it make a sound?

- No, it just causes pressure fluctuations in the air. Pressure fluctuations aren't "sound" until they reach someone's ear.

http://www.youtube.com/watch?v=MFLeGJclQiI

Frequency Sensitivity

Different species respond to different frequencies of pressure fluctuations. Young healthy humans typically are sensitive to frequencies between 20 Hz and 20,000 Hz.

- Pressure fluctuations within this range will register with humans as sound.
- Speech sounds only go up to 10,000 Hz.
- Pressure fluctuations outside the 20-20,000 Hz range will not be heard by humans (though vibrations in the air at frequencies lower than 20 Hz may sometimes be *felt* with the skin).

http://www.youtube.com/watch?v=VxcbppCX6Rk

http://www.youtube.com/watch?v=qNf9nzvnd1k

Frequency Sensitivity and Pitch Perception

Human sensitivity to pressure wave frequency is *nonlinear* — sounds in the middle ranges are heard more easily than very high or very low frequency sounds.

The human percept of frequency is called *pitch*. In general, higher frequencies are perceived as higher pitches. However, because of the nonlinear nature of frequency sensitivity, frequencies that are measured to be equally spaced may not sound as though they are equally spaced pitches.

- Measures of frequency that attempt to account for the nonlinearities of human perception include the *Bark* and the *mel*.

Amplitude

Speech causes pressure fluctuations in the air. Air molecules get pushed together (*compression*), and then they spring apart (*rarefaction*), attempting to return to a state of rest.

http://www.acs.psu.edu/drussell/Demos/waves-intro/wavepulse.gif

Amplitude is a measure of the magnitude of the pressure fluctuation (i.e., the difference between the wave pressure and the pressure of the medium at rest). For waves in air this is also called *sound pressure*.

- The standard units for sound pressure are *Pascals* (Pa).
- When sound waves are converted to electrical signals by a microphone, the amplitude is measured as difference in electrical potential (i.e., *voltage*; the unit is *Volts* (V)).

Measuring Amplitude

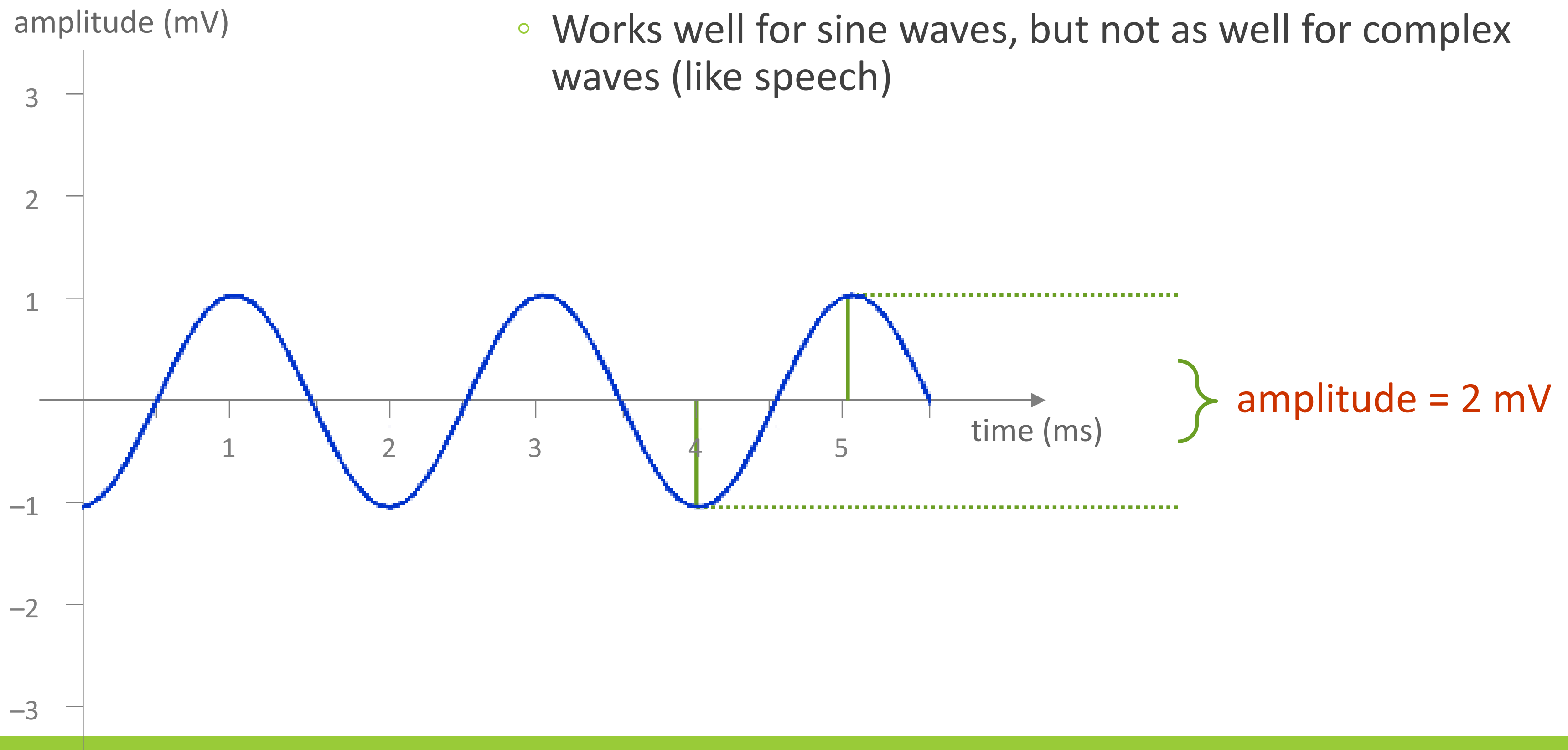
There are two ways that the amplitude of a wave is commonly calculated:

- *Peak-to-peak*: the vertical difference between the highest point and lowest point of the wave.
- *Root Mean Square (RMS)*: the average absolute distance from zero of the wave over the course of one cycle.

Peak-to-Peak Amplitude

The peak-to-peak amplitude is calculated by finding the difference between the highest and lowest values of the wave.

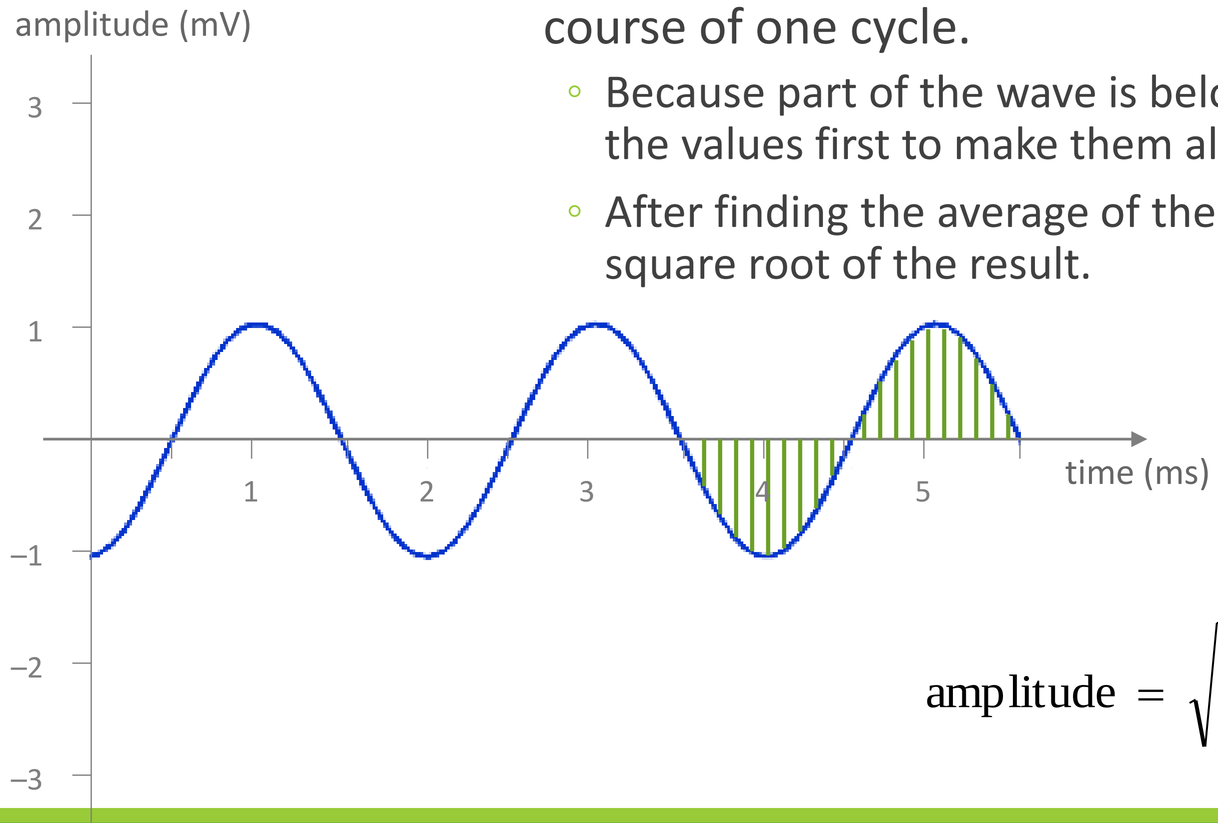
- Works well for sine waves, but not as well for complex waves (like speech)



Root Mean Square (RMS) Amplitude

The RMS is the average value of the wave over the course of one cycle.

- Because part of the wave is below zero, we square each of the values first to make them all positive.
- After finding the average of the squared values, we take the square root of the result.

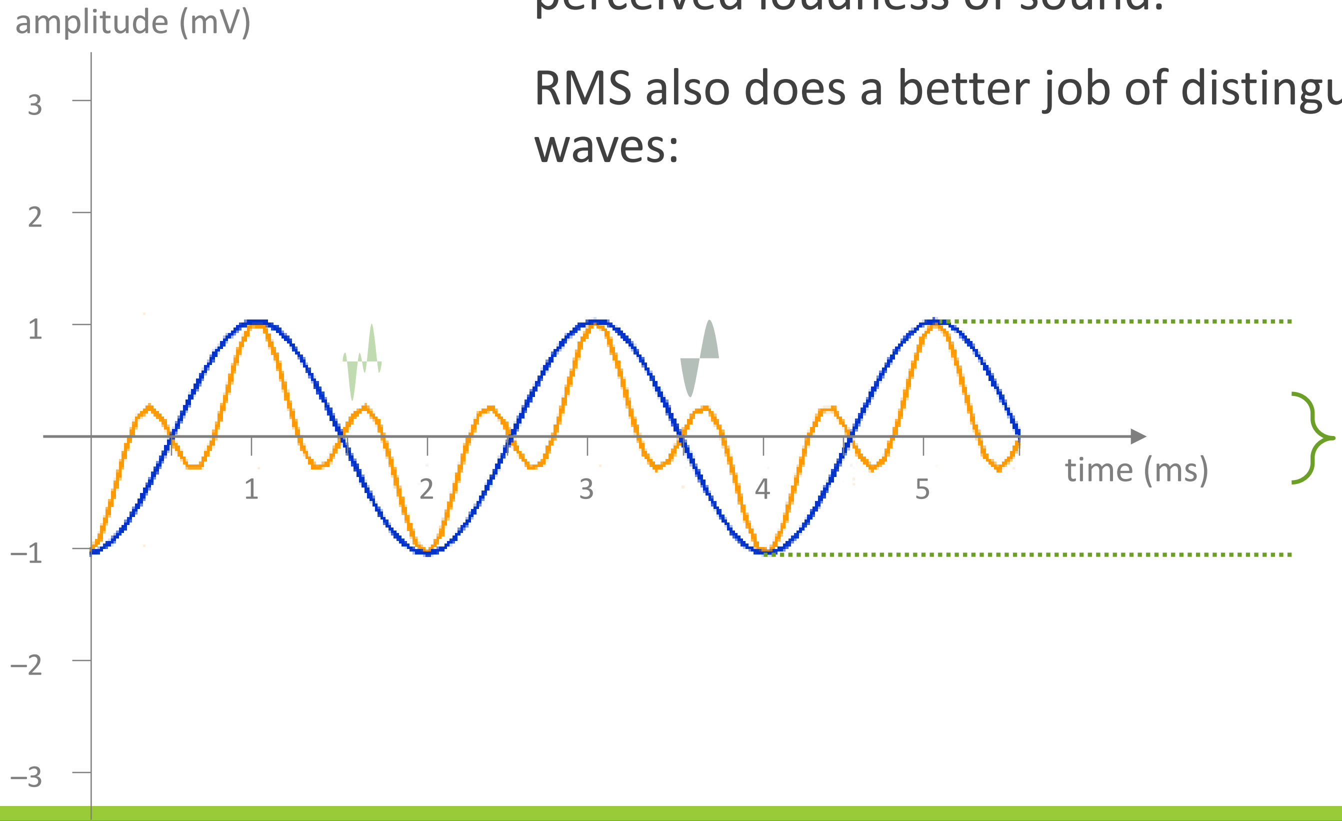


$$\text{amplitude} = \sqrt{\frac{a_1^2 + a_2^2 + a_3^2 \ldots + a_n^2}{n}}$$

Calculating Amplitude: Which Method is Better?

Phonetic researchers typically use RMS as a measure of amplitude because it more closely correlates to humans' perceived loudness of sound.

RMS also does a better job of distinguishing complex waves:



same peak-to-peak amplitude

different RMS amplitudes

Intensity and Loudness

The *intensity* of a wave is a measure of *average energy flux* (energy per unit area per unit time). It is proportional to the square of the wave's amplitude.

- The official unit of intensity is W/m^2 (Watts per square meter), but in practice, intensity is usually expressed in dB (decibels), which is a dimensionless unit representing the logarithm of the ratio between the intensity of the measured sound and the intensity of a reference sound.
- For our purposes, all you need remember is that a 1dB increase in intensity corresponds to a tenfold increase in RMS amplitude.

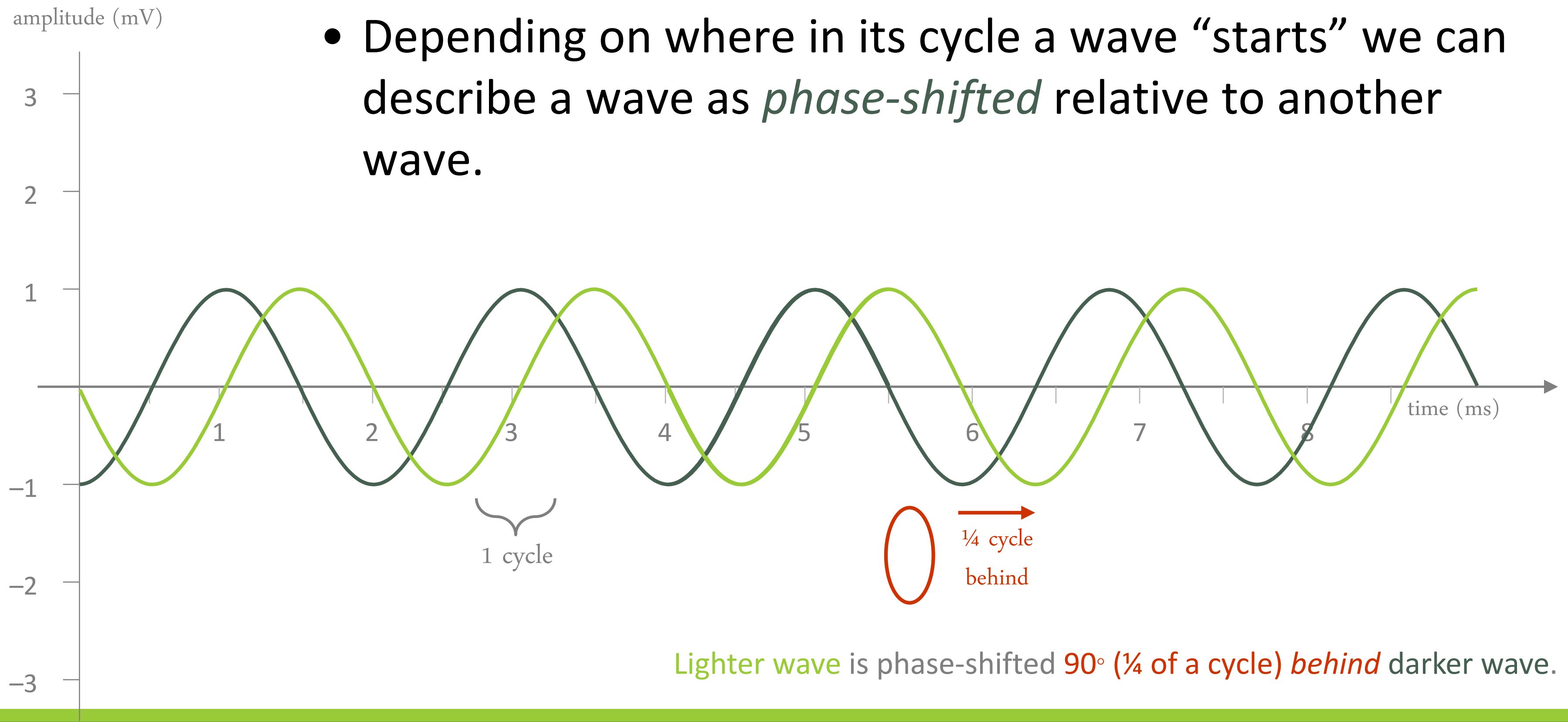
Loudness is the human percept of intensity. Like pitch, loudness perception is nonlinear.

- Measures of loudness that attempt to account for the nonlinearities of human perception include the *sone* and the *phon*.

http://www.youtube.com/watch?v=_vYYqRVi8vY

Phase

- Progress through a wave's cycle is measured in degrees. Like circles, waves have 360◦ per cycle.
- Depending on where in its cycle a wave "starts" we can describe a wave as *phase-shifted* relative to another wave.



Lighter wave is phase-shifted 90◦ (¼ of a cycle) *behind* darker wave.

Phase Perception

- Humans cannot hear the difference between two waves that differ only in their phase.
  - However, when wave travels different distances to reach each of our ears, the waves striking our two eardrums will be *out of phase* (i.e., phase-shifted relative to each other).
  - Comparing the phase of the two waves helps us perceive where a sound originated in space.

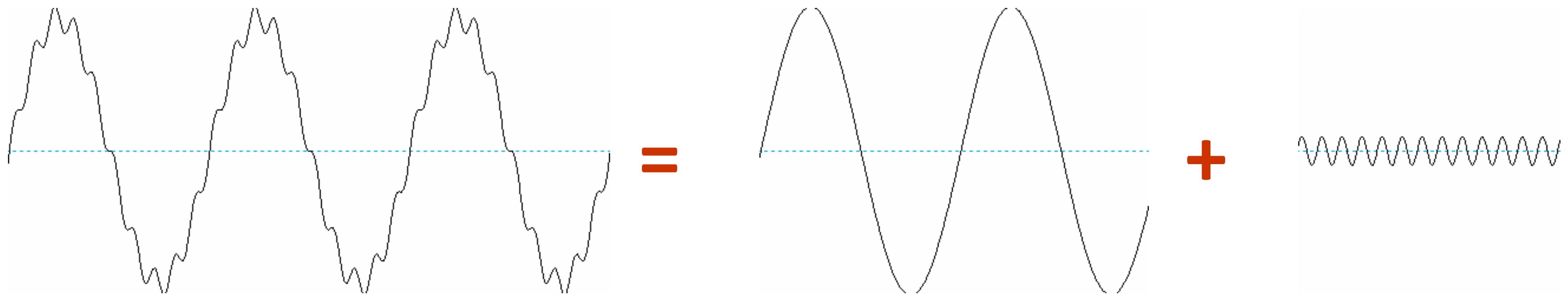
Summary

- *Frequency* is measured in Hz (cycles per second) and is a property of waves.
  - *Pitch* is the percept of frequency, measured in *Bark* or *mels*.
- *Amplitude* is a measure of the magnitude of pressure fluctuations (measured in *Pa* or *mV*).
  - *Intensity* (measured in dB) is usually calculated by the root mean square method of measuring amplitude over a span of time.
  - *Loudness* is the percept of intensity, measured in *sones* or *phon*.
- *Phase* does not affect our perception of sound *per se*, but it does help us identify a sound's directional source.

Complex Waves

In reality, nearly every material vibrates irregularly, and thus the sound waves created by vibration are almost always *complex waves*.

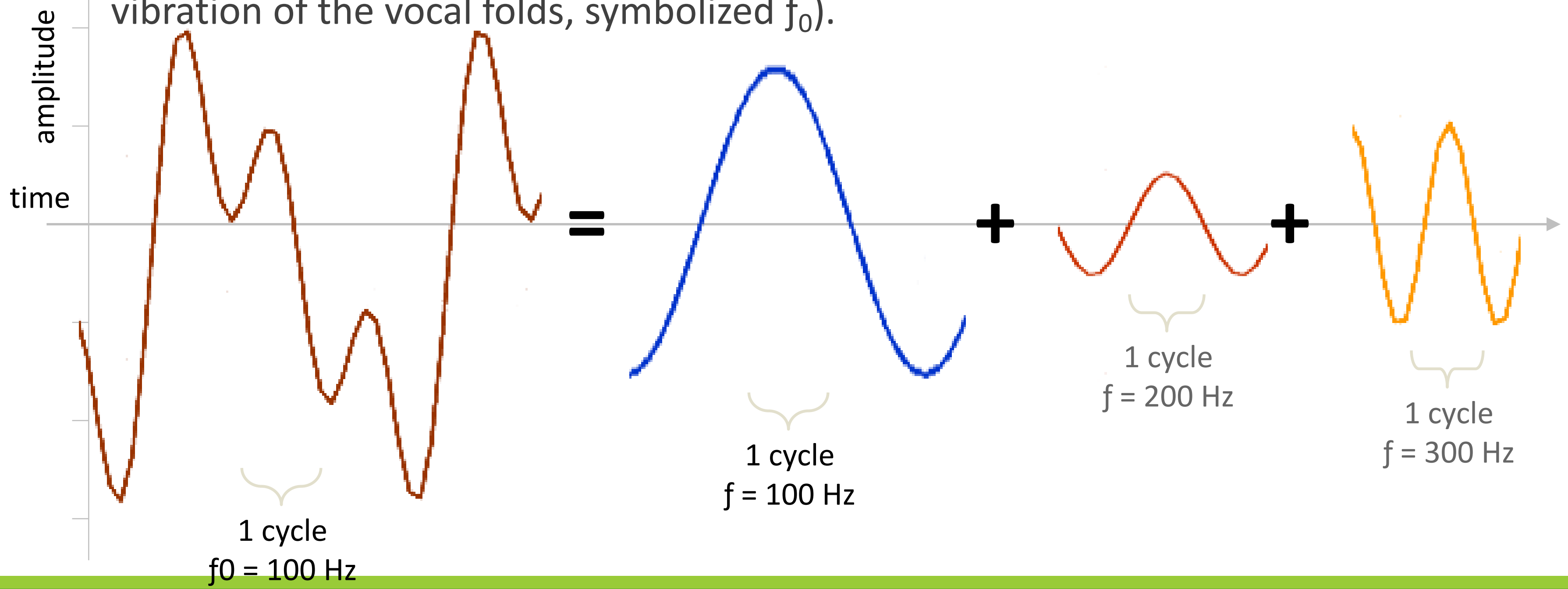
Pressure waves are *additive*, so that a complex wave can be said to be made up of several *component waves*.



Fourier's Theorem

Fourier's theorem states that all complex periodic waves can be broken down into a sum of simple *sine* waves of varying amplitudes, frequencies, and phases.

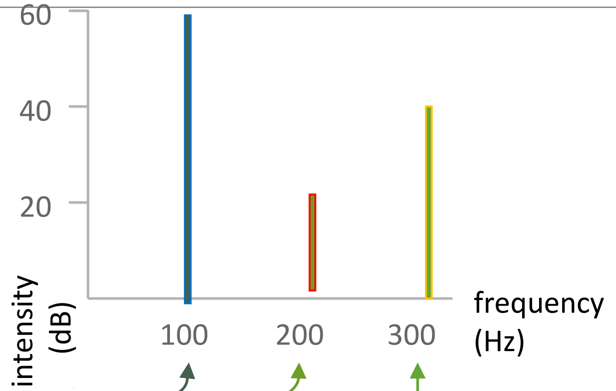
- The frequencies of the component waves will always be integer multiples of the frequency of the original complex wave (the *fundamental frequency*, the rate of vibration of the vocal folds, symbolized f_0).

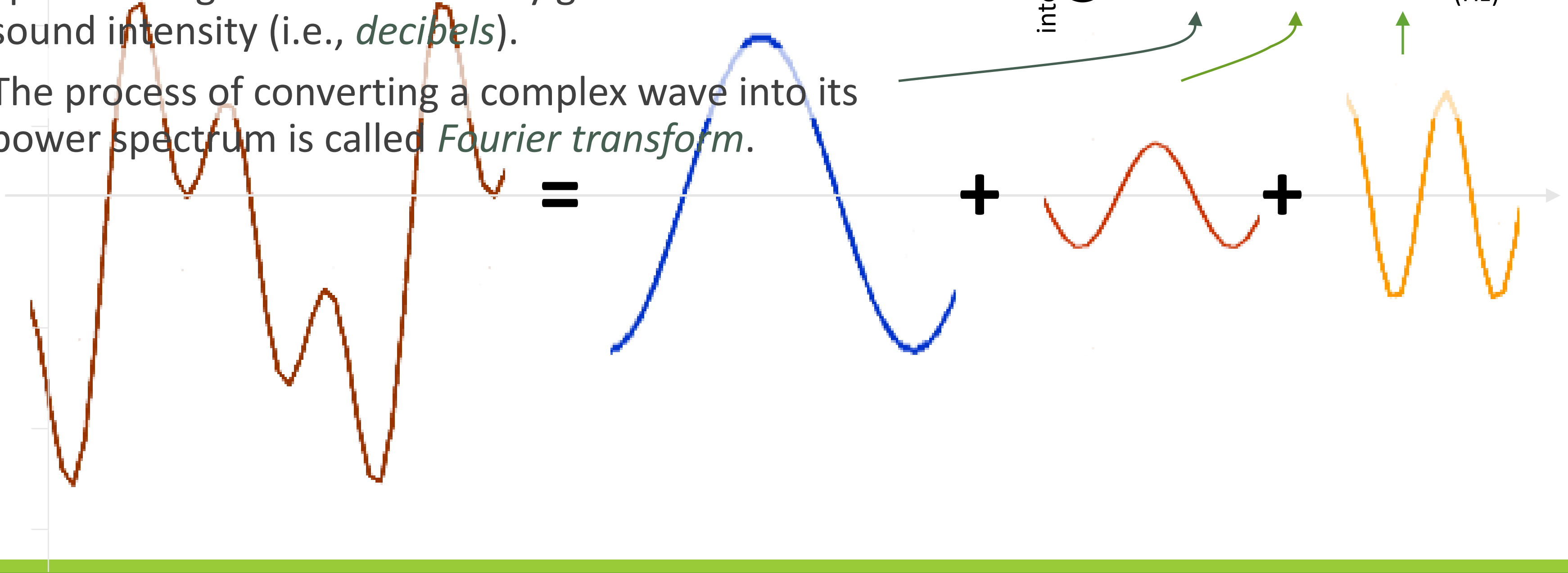


Power Spectra

A *power spectrum* (plural *spectra*) is a graph showing the different frequencies of sine wave components that make up a complex wave, and the different magnitudes of each contributing component wave.

- Spectral magnitudes are usually given in units of sound intensity (i.e., *decibels*).
- The process of converting a complex wave into its power spectrum is called *Fourier transform*.





Quality

When two complex waves have the same fundamental frequency but vary in the relative amplitudes of their components, we will perceive them to have the same pitch, but differ in *quality*.

- Quality (e.g., *vowel quality*) is the human percept of a power spectrum (i.e., the perception of varied intensities at several frequencies simultaneously).
- In other words, changing the shape of the vocal tract (e.g., to make a different vowel quality) changes the amplitudes and frequencies of the component waves, so power spectra will differ between vowel qualities.
- In the study of musical sound, quality is often called *timbre*.

Reminders

Do HW 2

Read Ladefoged & Johnson chapter 8

Get started on Project Transcriptions