Sound waves and the peripheral auditory system

David Meredith
Department of Computing,
City University, London.
dave@titanmusic.com
www.titanmusic.com

BMus/BSc in Music, Part Two Module: Music Perception and Cognition
Department of Music, City University, London
Tuesday, 6 May 2003.

1. Physical properties of sound waves (Moore, 1989, pp. 1–3; Whelan and Hodgson, 1978, pp. 91–94,98,296–297)

- Sound waves can be either
  - *periodic*, e.g., tuning fork, violin; or
  - *non-periodic*, e.g., sonic boom, hand-clap.
- Sound waves can be either
  - *progressive*, e.g., tuning fork, violin; or
  - *stationary*, e.g., inside sounding organ pipe.
- The *source* of a sound wave is a vibrating object.
- The *disturbance* in a sound wave is a variation in pressure and density of the medium.
- Sound waves are *mechanical waves*—they can only be transmitted through a material medium (not a vacuum).

1. Physical properties of sound waves (Moore, 1989, pp. 1–3; Whelan and Hodgson, 1978, pp. 91–94,98,296–297)

1. I'm David Meredith and I'm going to be giving you 10 lectures on music perception and cognition.
2. You should each have a copy of this 'extra information' sheet and I just want to make sure that you all understand everything on it.
3. [READ THROUGH SHEET AND ASK IF THERE ARE ANY QUESTIONS.]
4. [TAKE ROLL CALL AND TRY TO LEARN NAMES]
5. This week I'm going to review the physics of sound waves, and then I'm going to talk a bit about the anatomy and physiology of the auditory system.
6. As reading for today's lecture, I would recommend Chapters 2, 3 and 7 from Pierce's (1992) *The Science of Musical Sound*, Chapter 2 of Dowling and Harwood's (1986) *Music Cognition*, Chapter 1 of Moore's (1989) *An Introduction to the Psychology of Hearing* and Chapters 1 and 2 of Campbell and Greated's (1987) *The Musician's Guide to Acoustics*.
7. You don't have to read all of these. If you just look at the relevant bits from, say, two of the books, you should be fine.
8. Sound waves can be either *periodic* or *non-periodic*.
9. (a) A periodic sound wave is produced by a vibrating source such as a tuning fork which gives rise to a succession of disturbances called a *wavetrain*.
   (b) A non-periodic sound wave consists of just a single wavefront such as that which occurs in the shock wave emitted by an aircraft when it breaks the sound barrier or in the wave that results from a hand-clap.
10. Sound waves can be either *progressive* or *stationary*.

(a) A progressive wave (also known as a *travelling* wave) is one in which energy is transferred from one location (the source) to another. The wave generated by a tuning fork is an example of a progressive sound wave.

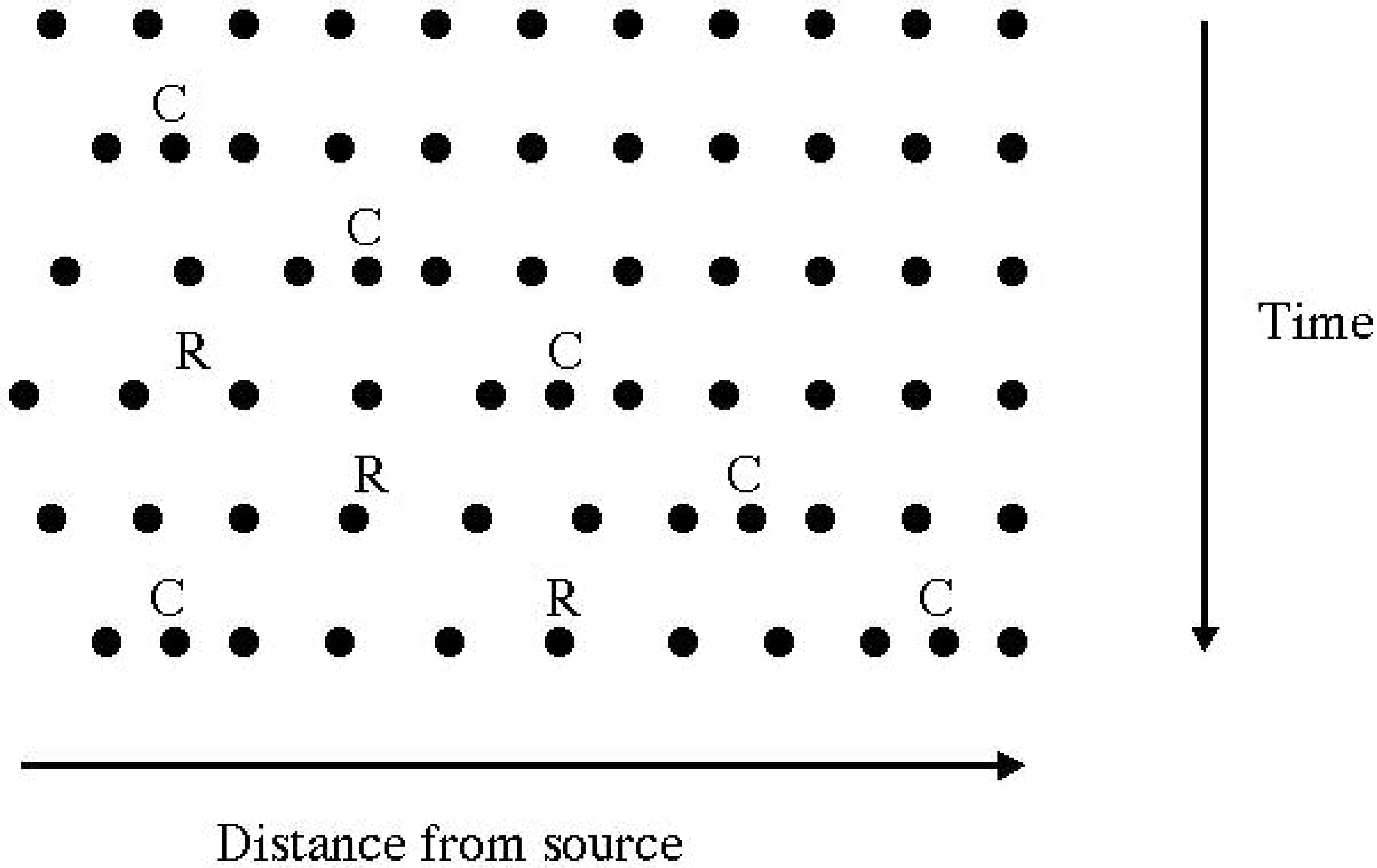
(b) In a stationary or *standing* wave, there is no *net* transfer of energy from one place to another. However, over time, the energy in a standing wave is repeatedly converted from kinetic energy into potential energy and back again. The sound wave *inside* a sounding organ pipe is an example of a standing or stationary sound wave.

11. A sound wave is usually produced by a vibrating object which sends out waves through the medium (usually air) in which it is immersed.

12. The *disturbance* in a sound wave is a variation in the pressure and density of the medium through which the wave passes.

13. A sound wave can therefore only be transmitted through a *material* medium which may be solid, liquid or gas. Sound waves are therefore mechanical waves and, unlike electromagnetic waves, they cannot be transmitted through a vacuum.

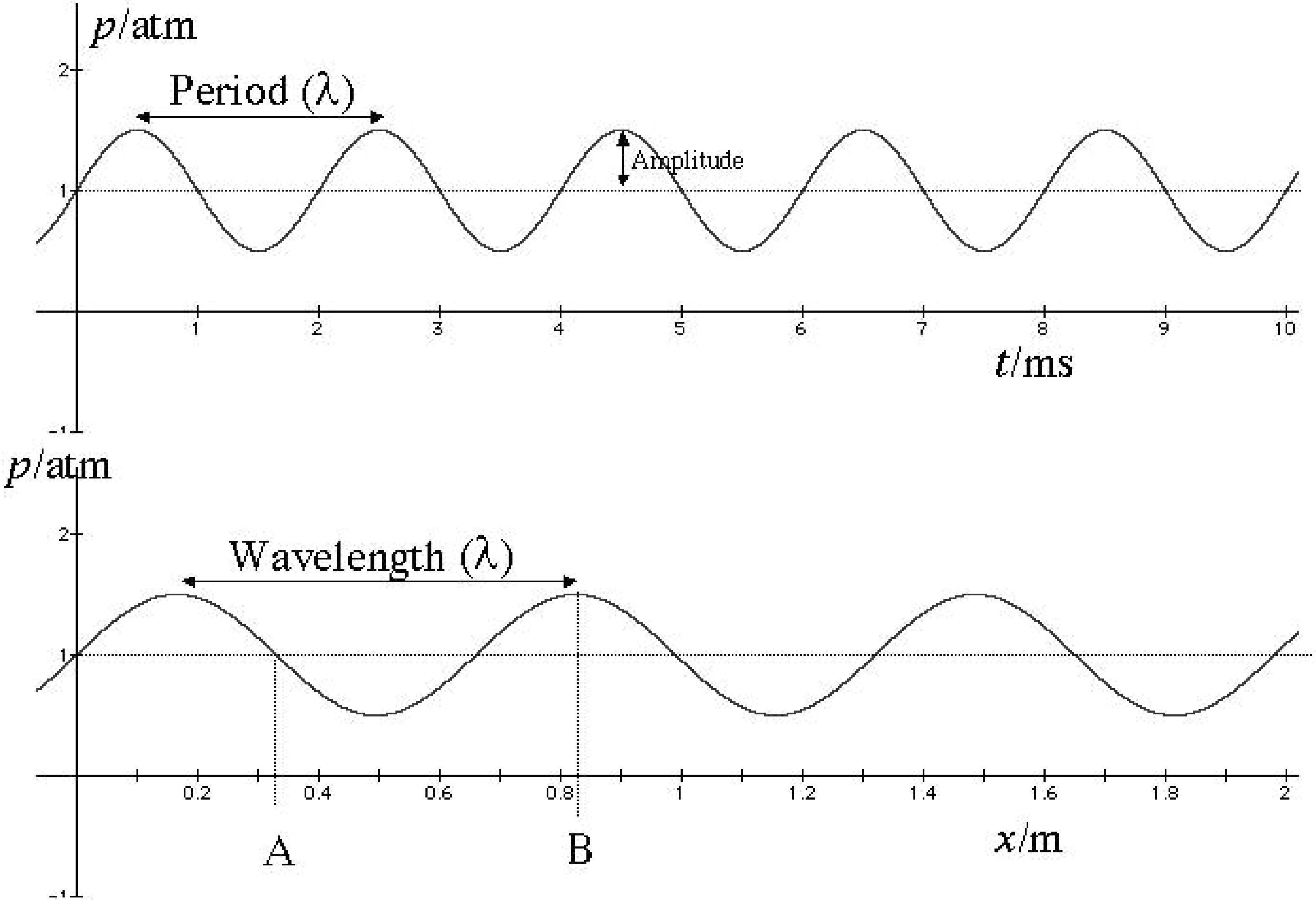
2. Oscillations of molecules in a sound wave



2. Oscillations of molecules in a sound wave

1. When a sound wave travels through a medium, what happens at a particular point in space is that the molecules in the medium are alternately squeezed together (compressed) and spread out (or rarefied).
2. This means that at a particular point in space in a sound wave, the pressure and density of the medium oscillate around their normal, undisturbed values because the molecules at that point are alternately more and less densely packed together than normal.
3. Even in the absence of a sound wave, gas molecules exhibit constant random motion. However, when a sound wave is present, an oscillatory motion is superimposed on this random motion.
4. In a sound wave, the direction of this oscillatory motion is parallel to the direction of propagation of the wave so sound waves are *longitudinal* mechanical waves.
5. You must remember, however, that the molecules in the medium *oscillate* around an average position and do not advance with the wave: it is the *disturbance* that is propagated, not the material of the medium itself.

3. Waveforms and simple tones



3. Waveforms and simple tones

1. The *waveform* of a wave is a graph showing either

   (a) the pressure variation at a particular point in space plotted against time (as shown in the top graph); or

   (b) the pressure variation at a particular instant in time plotted against position in space (as shown in the bottom graph).

2. One of the simplest types of sound wave is one whose waveform is sinusoidal, as shown here. Such a sound is called a simple tone or a pure tone.

3. Simple tones are not only simple mathematically, they also evoke particularly simple responses in the auditory system (Moore, 1989, p. 2).

4. This graph here shows the pressure variation at a point in space plotted against time for a simple tone.

5. The time that the pressure variation at a given point takes to complete one cycle is called the *period* of the wave and the number of periods completed per unit time is called the *frequency* of the wave.

6. The frequency of a periodic sound wave whose waveform is more complex than that of a simple tone is usually called its *periodicity*.

7. The frequency of a wave is usually expressed in hertz (Hz), 1 Hz being equal to 1 cycle per second.

8. The *amplitude* of a sound wave at a point is the difference between the mean pressure at that point and the maximum pressure at that point, as shown here in this graph. In other words, the amplitude of a wave is its *maximum* pressure variation.

9. Beware that in some texts (e.g., Moore, 1989, p. 2), the term ‘amplitude’ is incorrectly used to denote the pressure variation and the term ‘maximum amplitude’ is used to denote the maximum pressure variation (i.e., the amplitude).

10. This graph here shows the pressure variation at a particular instant in time plotted against position in space for a simple tone.

11. The *phase* at a given point in space and a given instant in time is the number of cycles through which the wave has advanced relative to some fixed point in time, expressed in angular measure. The phase is usually given in radians. A full cycle therefore corresponds to 2π radians, half a cycle corresponds to π radians and so on.

12. For example, in this waveform, the point B is lagging 3/4 of a cycle behind point A. Point B therefore exhibits a phase lag of $\frac{3}{4} \times 2\pi = \frac{3\pi}{2}$ relative to point A.

13. The distance between two consecutive points in the wave with the same phase (e.g., two consecutive wave crests) is called the *wavelength* of the wave.

14. The speed of a mechanical wave is the speed at which an observer would have to travel in order to see the wave as being stationary. The speed of a wave is equal to the number of wavelengths per unit time. It is therefore equal to the product of the wave's frequency and wavelength.

4. Power, intensity, pitch and tones

- *Power* of a wave is amount of energy transmitted per unit time.

$$\text{Power} \propto (\text{frequency})^2 \times (\text{amplitude})^2 \times \text{speed}$$

- *Intensity* is energy transmitted per unit time per unit area of the wavefront

$$\text{Intensity} = \frac{\text{Power}}{\text{Area of wavefront}}$$

- *Loudness* depends on intensity and sensitivity of listener to frequency of the sound.
- Periodic sounds with a periodicity between 20Hz and 5000Hz evoke sensation of pitch.
- Pitch is that perceptual attribute of a sound in terms of which it may be ordered on a musical scale.
- A *tone* is a sound that has a pitch.

4. Power, intensity, pitch and tones

1. A progressive wave such as the sound wave produced by a musical instrument transfers energy and momentum from the source to places around it.

2. The amount of energy transmitted by a wave in unit time is called the *power* of the wave.

3. The *power* of a wave is proportional to the square of the frequency, the square of the amplitude and the wave speed.

4. The *intensity* of a wave is the amount of energy transmitted by the wave perpendicularly to its wavefront in unit time per unit area. In other words, it is the power per unit area of the wavefront. The loudness of a sound depends upon the intensity of the sound and the sensitivity of the listener to the frequency of the sound.

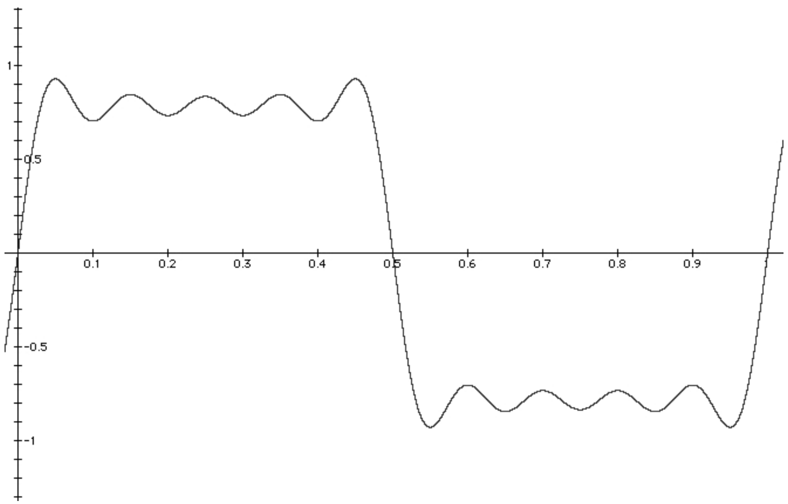
5. Any periodic sound wave with a frequency greater than 20Hz and less than about 5000Hz will evoke a sensation of pitch.

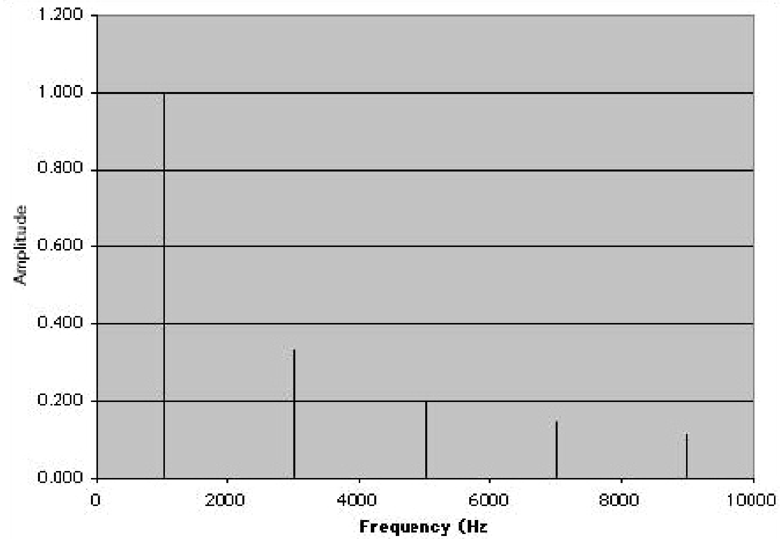
6. Moore (1989, p. 3) defines pitch to be ‘that attribute of auditory sensation in terms of which sounds may be ordered on a musical scale.’

7. Any sound that evokes a sensation of pitch is called a *tone*. Although most tones are periodic, not all of them are: some non-periodic tones are perceived to have pitch.

8. The pitch of a tone is related to its frequency or periodicity. The pitch of a simple tone is usually indicated by specifying its frequency in Hz. The pitch of any other tone S is usually indicated by giving the frequency of a simple tone whose pitch would be perceived to be the same as that of S.

5. Fourier analysis and spectral representations (Moore, 1989, pp. 3–7)

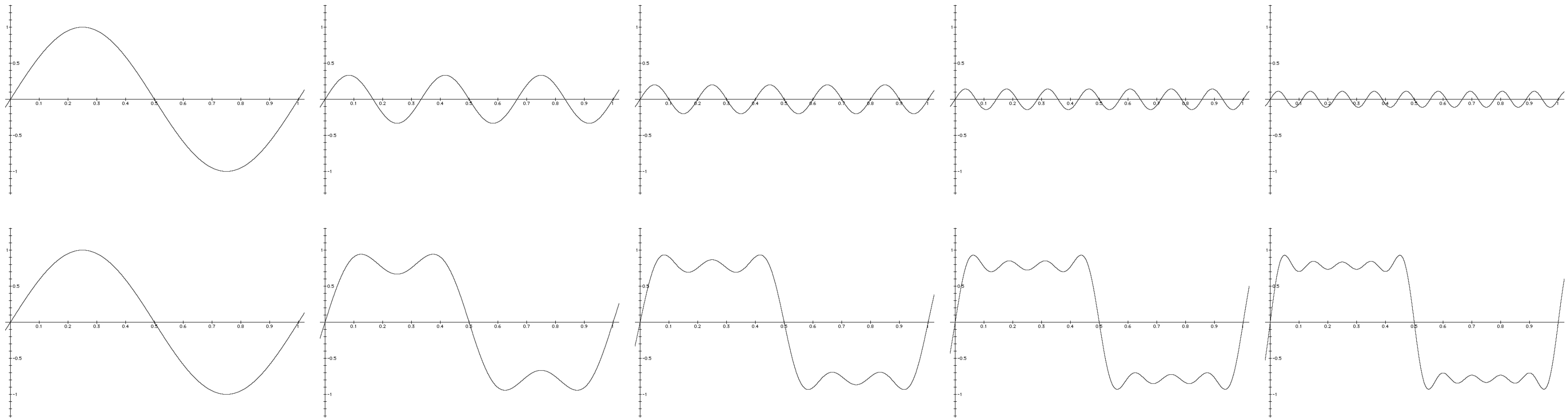




5. Fourier analysis and spectral representations (Moore, 1989, pp. 3–7)

1. Any sound can be represented by a waveform that plots pressure variation against time or spatial position.
2. However, when the sounds are more complicated than sinusoidal simple tones, it is usually more useful to represent the wave by its so-called *Fourier spectrum*.
3. This graph on the right shows the Fourier spectrum that corresponds to this waveform here on the left.
4. The Fourier spectrum is so-called because it is based on a theorem discovered by the French mathematician Jean Baptiste Joseph Fourier (1768–1830) who proved that any complex waveform (with certain restrictions) can be expressed as the sum of a series of sinusoidal simple tones with particular frequencies, amplitudes and phases.
5. Beware that Pierce (1992, p. 41) seems to think that Fourier analysis was invented by the French social reformer, Fran cois Marie Charles Fourier (1772–1837) who was an almost exact contemporary of the mathematician.
6. So this spectrum on the right tells us that we can construct this waveform on the left by adding together five simple tones whose frequencies are 1000Hz, 3000Hz, 5000Hz, 7000Hz and 9000Hz.
7. Moreover, it tells us that the relative amplitude of each of these sinusoidal simple tones must be as shown on the graph. That is, the tone at 3000Hz must have an amplitude that is 1/3 that of the tone at 1000Hz and so on.
8. When a complex waveform is expressed in this way it is called a *Fourier analysis* of the waveform and each sinusoidal simple tone in the summation is called a *Fourier component* or *partial* of the complex waveform.
9. When the frequencies of all the Fourier components of a complex tone are integer multiples of the frequency of the lowest component, the complex tone is called a *harmonic complex tone*, each of the Fourier components of the tone is called a *harmonic* and the lowest harmonic is called the *fundamental*.
10. The fundamental frequency of a harmonic complex tone is equal to the periodicity of the tone.
11. The Fourier component of a harmonic complex tone whose frequency is n times that of the fundamental frequency is called the nth harmonic of the tone. So, in this case, the component at 3000Hz is the third harmonic of the complex tone.

6. Constructing a harmonic complex tone

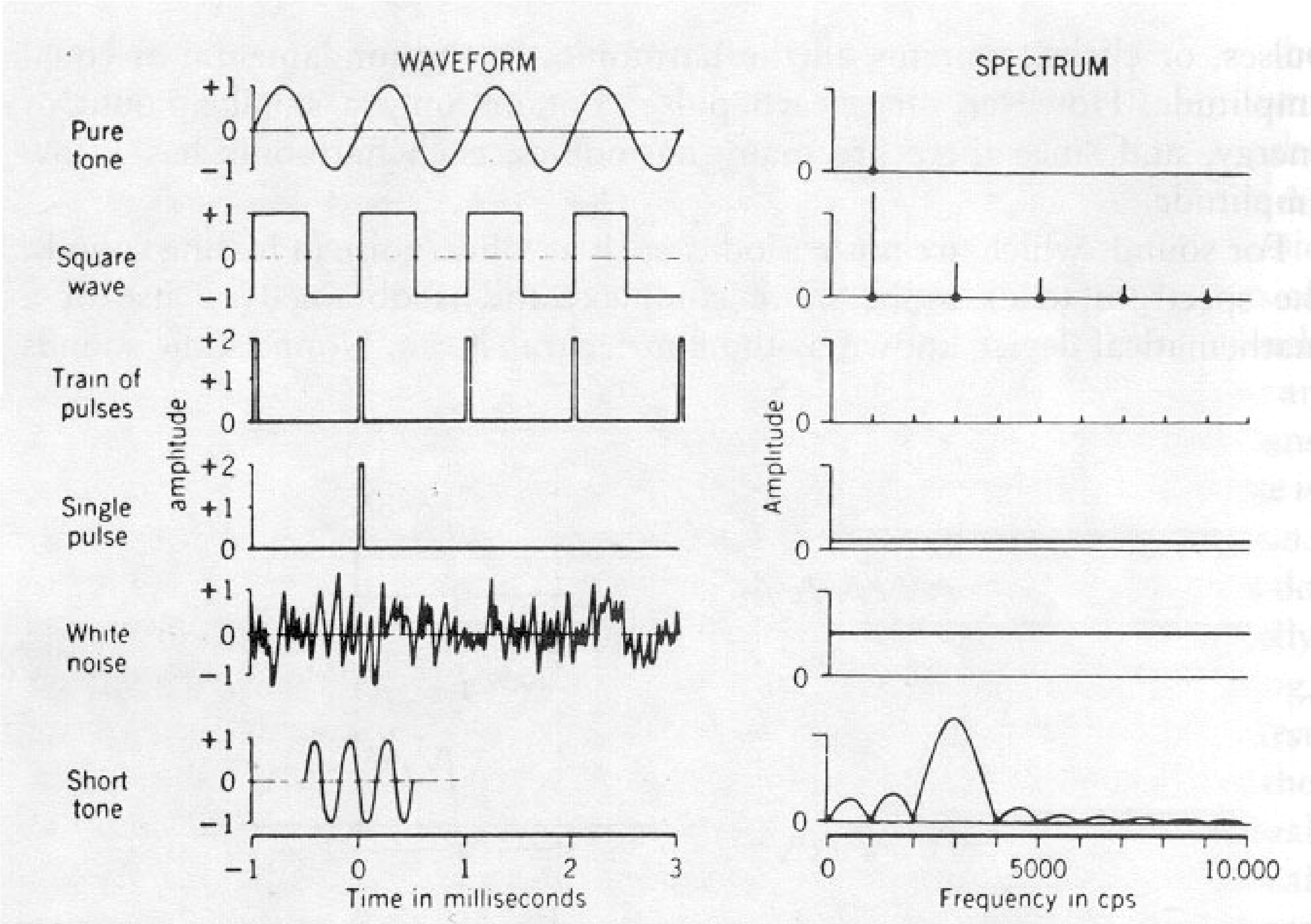


6. Constructing a harmonic complex tone

1. This figure shows how a harmonic complex tone can be built up by adding together sinusoidal components whose frequencies are all integer multiples of a fundamental frequency.
2. This particular example shows how a square wave can be built up by adding together odd-numbered harmonics of a fundamental, each harmonic having an amplitude that is inversely proportional to its harmonic number. So the third harmonic has an amplitude that is 1/3 of the amplitude of the fundamental, the fifth harmonic has an amplitude that is 1/5 of that of the fundamental and so on.
3. Note how all the harmonics in the diagram have the same phase at time zero.
4. Let's hear what each of these tones sound like.
5. [PLAY EXAMPLES OF BUILDING UP A SQUARE WAVE]
6. Fourier's theorem tells us that we can build up any complex tone by adding together the right combination of simple tones. This is very similar to the situation with light where we can produce light of any given colour by adding together different amounts of the primary colours.
7. Our visual system is incapable of performing the reverse process—that is, we cannot attend to just one of the component hues in a beam of light. If lights of two different frequencies (colours) are mixed, we see light of a single colour that corresponds to the mixture of frequencies it contains, we cannot choose to perceive each of the component hues separately.
8. However, when we listen to a complex harmonic tone we can choose either to hear the tone as a unitary percept with a particular timbre, loudness and pitch; *or* we can attend to each of the low harmonics of the tone individually.
9. For example, if I play you the third harmonic of this complex tone on its own and then I play you the complex tone, you should be able to focus your attention on the third harmonic in the complex tone.
10. Our ability to do this was observed by Mersenne in 1636 and formalized in Ohm's Acoustical Law.

11. Our ability to hear out the low harmonics of a complex tone implies that our auditory system performs a kind of Fourier analysis on the incoming sound signal.

7. Examples of Fourier spectra



7. Examples of Fourier spectra

1. We can represent a Fourier analysis of a complex sound by means of a *spectrum*, like the ones shown here.
2. The Fourier spectrum of a complex sound is a graph that shows the amplitude and frequency of each of the sinusoidal Fourier components of the sound. Note that a Fourier spectrum tells us nothing about the phases of the sinusoidal components. This information is usually represented separately in another type of spectrum called a *phase spectrum*.
3. Here we have six waveforms with their corresponding spectra taken from Moore (1989, p. 6).
4. Each of the top three spectra contains a set of isolated vertical lines indicating that the energy falls at highly specific, discrete frequencies. Such a spectrum is called a *line spectrum*. In fact, a sound only has a perfect line spectrum if it lasts forever!
5. So, an everlasting simple tone would have a perfect line spectrum containing just one vertical line at the frequency of the tone, as shown here.
6. A perfect and everlasting square wave would have a perfect line spectrum containing vertical lines at all the odd-numbered multiples of the fundamental frequency, with the amplitude of the nth harmonic being proportional to $1/n$, as shown here.
7. The third example shows a pulse train—that is, an infinite series of equally spaced clicks. This is a harmonic complex tone that contains harmonics at all multiples of the fundamental, all with the same amplitude. Note that because there is energy at every integer multiple of the fundamental, each harmonic only has a very small amplitude.
8. Both the brief click and the white noise (which sounds like a hiss) have continuous flat spectra indicating that they contain energy at all possible frequencies. However, the phase spectrum of the click is different from that of the white noise. For the white noise, the phase at which each component begins is random whereas for the click, all the components have phase $\pi/2$ at time zero. This means that at time zero, the pressure variation is at a maximum for all the frequency components.

9. The sixth example shows the spectrum for a simple tone that does not last forever. Such a tone is called a *tone pulse* or a *tone burst*. Note that whereas the everlasting simple tone has a line spectrum, the sinusoidal tone burst has a continuous spectrum, consisting of a peak at the frequency of the steady-state portion of the burst and surrounded by smaller and smaller peaks spreading out away from this frequency.

10. As the tone burst gets shorter and shorter, the spectrum gets more and more spread out until eventually you get the click here which has a completely flat spectrum.

8. The measurement of sound level (Moore, 1989, pp. 7–10)

- Intensity is energy transmitted per second perpendicularly through 1m^2 of the wavefront. That is,

$$\text{Intensity} = \frac{\text{Power}}{\text{Area of wavefront}}.$$

- Auditory system can deal with huge range of intensity (e.g., gunshot is 10 000 000 000 000 times intensity of quietest detectable sound).
- We generally use a *logarithmic* scale for intensity.
- If we have two sounds with intensities I_1 and I_2 then the *sound level* of I_1 is

$$\log_{10}(I_1/I_2) \text{ Bels}$$

  greater than I_2.
- For example, if $I_1 = 100I_2$ then

$$\log_{10}(I_1/I_2) = \log_{10}(100) = 2 \text{ Bels}.$$

- If we have two sounds with intensities I_1 and I_2 then the *sound level* of I_1 is

$$10\log_{10}(I_1/I_2) \text{ decibels (dB)}$$

  greater than I_2.
- For example, if $I_1 = 100I_2$ then

$$10\log_{10}(I_1/I_2) = 10\log_{10}(100) = 20 \text{ dB}.$$

- Increase in level of 10 dB corresponds to multiplying intensity by 10.
- Increase in level by 3 dB corresponds to doubling intensity.
- If $I_1 = I_2/10$ then

$$10\log_{10}(I_1/I_2) = 10\log_{10}(.1) = -10 \text{ dB}.$$

  That is, the sound level of I_1 is 10 dB less than that of I_2.

8. The measurement of sound level (Moore, 1989, pp. 7–10)

1. We've seen that the intensity of a sound—that is, the amount of energy transmitted by the sound per second through 1 square metre of the wavefront—is proportional to the square of the maximum pressure variation.

2. Our auditory systems can deal with a huge range of sound intensities. For example, the sound of a gunshot at close range has roughly 10 000 000 000 000 times the intensity of the quietest sound that we can hear. The sound of someone shouting at close range has roughly 10 000 000 times the intensity of a whisper.

3. If we have a sound with a low intensity and we increase its intensity linearly over time—say, for example, by 10^{-12}W/m^2 every second—then the loudness of the sound will be perceived to increase rapidly at first but the speed at which the loudness increases will become slower and slower as the sound becomes louder and louder.

4. Because of this, it is more useful to use a logarithmic scale for expressing intensities than a linear one.

5. So, if we have two sounds with intensities I_1 and I_2, then we express the difference between the intensities by the logarithm to the base 10 of the ratio of I_1 to I_2, thus $\log_{10}(I_1/I_2)$.

6. This expression gives the difference between the intensities in Bels: we say that the *level* of I_1 is $\log_{10}(I_1/I_2)$ Bels greater than that of I_2. So, for example, if I_1 is ten times I_2, then the *level* of I_1 is 1 Bel greater than that of I_2. If I_1 is 100 times I_2, then the level of I_1 is 2 Bels greater than that of I_2 and so on. Usually, we actually express level differences in *decibels* (dB), one decibel being one tenth of a Bel.

7. So, if we have two sounds with intensity I_1 and I_2 then we say that the level of I_1 is $10\log_{10}(I_1/I_2)$ decibels greater than that of I_2. An increase in level of 10dB corresponds to an increase in intensity by a factor of 10. A doubling of the intensity corresponds to an increase in level of 3 dB.

8. When I_1 is actually less intense than I_2, then the expression$10\log_{10}(I_1/I_2)$ is negative. For example, if I_2 is ten times as intense as I_1 then the level of I_1 is -10dB greater than I_2, or, equivalently, 10dB less than I_2.

9. Sound pressure level and sensation level

- To express *absolute* sound levels, we need to define a standard reference intensity.
- Most commonly used standard reference intensity is 10^{-12} watts per square metre ($\mathrm{W/m^2}$) which corresponds to pressure variation of $2 \times 10^{-5}\mathrm{N/m}^2$ or 20 μPa (micropascal).
- The sound level of a sound relative to 10^{-12} watts per square metre is called the *sound pressure level* (SPL) of the sound.
- If SPL of a sound with intensity I is 60 dB SPL, then this tells us that

$$10\log_{10}(I/(10^{-12})) = 60 \Rightarrow \log_{10}(I/(10^{-12})) = 6 \Rightarrow \frac{I}{10^{-12}} = 10^6 \Rightarrow I = 10^{-6}\mathrm{W/m}^2.$$

- O db SPL is close to human absolute threshold for 1000Hz tone (actually about 6.5 dB SPL on average).
- *Sensation level* of a sound is the intensity of the sound relative to the absolute threshold for that sound for a given individual, expressed in dB.

9. Sound pressure level and sensation level

1. So far we've only learnt how to express intensity ratios and level differences in decibels. If we want to express an absolute intensity I using decibels then we have to define some standard reference intensity I_0 and then find the number of decibels corresponding to the intensity ratio I/I_0.

2. The standard reference intensity most commonly used is 10^{-12} watts per square metre ($\mathrm{W/m^2}$) which corresponds to a pressure of $2 \times 10^{-5}\mathrm{N/m}^2$ or 20 μPa (micropascal). The sound level of a sound relative to this standard intensity is called the *sound pressure level* or SPL of the sound.

3. For example, the absolute intensity I of a sound whose sound pressure level is 60 dB SPL is given by

$$10 \log_{10}(I/(10^{-12})) = 60$$

   which implies that

$$\frac{I}{10^{-12}} = 10^6$$

   which implies that

$$I = 10^{-6}\mathrm{W/m}^2.$$

4. The standard reference intensity of $10^{-12}\mathrm{W/m}^2$ which corresponds to O dB SPL was chosen because it is very close to the absolute human threshold for a 1000Hz sinusoidal simple tone.

5. The absolute human threshold for a sound is the minimum detectable level of the sound in the absence of any other external sound (Moore, 1989, p. 8). In fact, the average human absolute threshold at 1000Hz is about 6.5 dB SPL.

6. Sometimes, instead of using the standard reference intensity of $10^{-12}\mathrm{W/m}^2$ we use the absolute threshold of the sound being measured for a particular individual. When the sound level is expressed in this way it is called a *sensation level* (SL). For example, if a sound is said to have a level of 60dB SL for a given subject, this means that the level of the sound is 60dB above the absolute threshold of that sound for that subject.

7. The physical intensity that corresponds to a given sensation level therefore varies from sound to sound and from subject to subject.

10. Linearity (Moore, 1989, pp. 10–11)