Dana Spiegel
danas@mit.edu

8.03
10/24/97

The Physics and Psychophysics of Music

Introduction

Listening to music and conversing with friends are processes that are made possible by human audition. The primary conduit for such aural perception in humans is the ear. The human ear is a complex system that encodes sounds that travel through it. Due to its physical construction, the ear gives rise to many interesting aural phenomena. In this paper, the physics of the ear are discussed, and some phenomena using simple sounds are explored.

The Ear

The ear is comprised of three parts: the outer, middle, and inner ear (Figure 1). Each part plays a crucial role in the reception and basic processing of sounds. Together, they allow animals to sense sounds (pressure waves) and react to them. While the exact limits of the ear's sensitivity vary with age — especially at the high end of the spectrum — the average human ear is sensitive to sounds between 20 Hz and 20,000 Hz. Human speech falls within the 1,000 Hz to 5,000 Hz band, and almost all important sounds in our environment fall within the bounds of the ear's sensitivity.

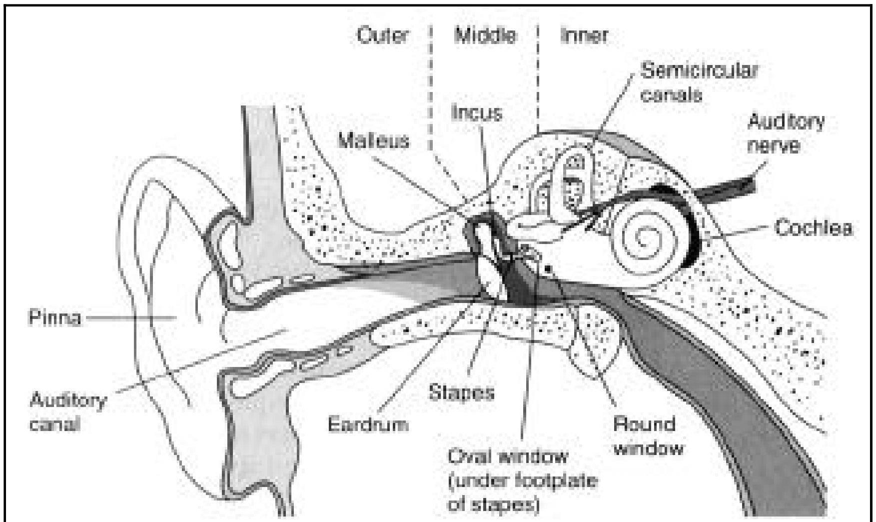

Figure 1 *The human ear*

1 *All images are taken from* Sensation & Perception *(referenced in bibliography)*

The response of the human ear to sound is not linear. This is due in part to the auditory canal, a section of the outer ear, which reinforces certain frequencies. The auditory canal operates much like a tube of air closed on one side: its resonant frequency, which depends on the length of the canal, is around 3,400 Hz, although there is some resonance between the frequencies of 2,000 Hz and 5,000 Hz (Goldstein, 1996). Resonance causes frequencies in this range to be heard at close to their incident intensity, while other frequencies are reduced in intensity.

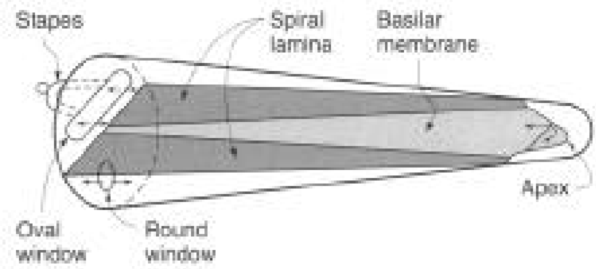

Figure 2 *The Basilar Membrane*

Note that the band of frequencies that cause resonance are those of human speech. Since the ear is tuned to respond most to these frequencies, they are easiest for us to hear: we are most sensitive to sounds that are in the frequency range of human speech. Speech allows us to be warned about harmful events and is key to survival.

The inner ear — more specifically, the basilar membrane of the cochlea — is considered the most important part of the ear involved in the reception and encoding of auditory stimuli. The basilar membrane (Figure 2), a long, stretched membrane wound within the cochlea, responds to sound much like a string bound at both ends: a disturbance (from a pressure wave) causes a transverse traveling wave to propagate down the membrane. (The conversion from a longitudinal wave to a transverse wave is accomplished by changes in pressure above the membrane by the stapes footplate.)

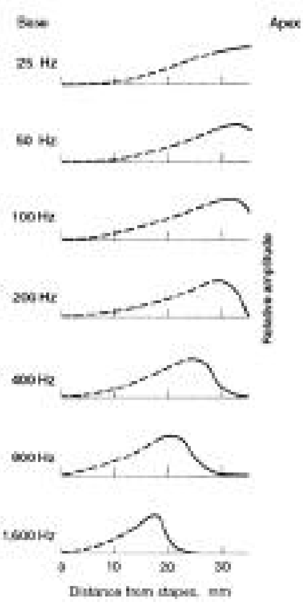

Figure 3 *Wave envelopes on the basilar membrane*

Since the response curve of the membrane is not linear — the base (the end closest to the outside of the spiral) is three to four times narrower than the apex (the innermost part of the spiral) — the envelope of the traveling wave is not laterally symmetric. As a result, each frequency causes a unique point of maximum amplitude along the membrane (Figure 3).

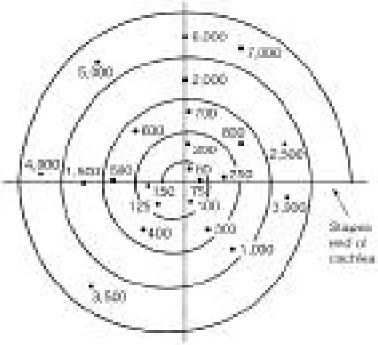

Figure 4 Tonotopic map on the basilar membrane

Nerves below the membrane, which are activated by the transverse displacement of the membrane, code for specific frequencies. This type of coding, in which sounds of increasing frequencies have maximum displacement at increasing distances from the apex of the membrane, is called a tonotopic map. This map isn't linear: the higher the frequency, the closer it is to its neighbors. Frequencies to about 5000 Hz account for more than 80% of the area on the basilar membrane (Figure 4). Since the membrane is also bound at the apex, very low frequencies (those from 1 Hz to 500 Hz), which account for less than 20% of the area on the membrane, cannot cause large transverse displacements. Therefore, neither high nor low frequencies can cause large responses; the system again favors the middle frequencies — those between 1000 Hz and 5000 Hz.

The physiology of the ear predicts sensitivity curves (Figure 5). Aural sensitivity, which is a measure of the threshold above which a sound can be heard, is low for both small and large frequencies. The ear is much more sensitive to sounds near the middle of the spectrum — again, those within the 1,000 Hz to 5,000 Hz range.

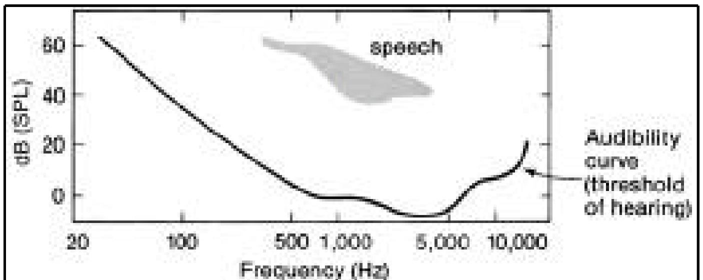

Figure 5 Audibility curve for the human ear

Experimental Setup

The experiments below were executed using a program called *soundmaker*, which generates pure sinusoidal tones or linear superpositions of sinusoidal tones, on Silicon Graphics Octane and O2 workstations. The volume for the threshold experiments was set to -30 decibels in the Sound Manager control panel. The earphones used were made by Sony and were of the in–the–ear style.

All of the experiments were performed using the method of limits: different stimuli are presented and the subject indicates whether each stimulus can be detected. The stimuli are presented in order, either ascending or descending (Fechner, 1860). The data collected are not exhaustively given; rather, the results, along with comments on the experiments, are discussed. The experiments are not meant as thorough investigations into the phenomena studied, but as explorations, which can be used as a guide for deeper study.

The choice of listening device (earphones, in this case) can have a significant impact on the experiments performed, and should therefore be taken into account. The quality of the Sony earphones was good compared to Walkman headphones, but rather low compared to studio–quality headphones. The low–end frequency response was poor; the high–end response was noisy; the frequency of the tones was not very accurate and somewhat variable.Considering their size, however, they are rather good for their designed use: playback from a portable CD or tape player.

Low frequency tones require that the speaker cone move relatively slowly — as slow as 50 Hz. That, and the diminutive size of the earphones (about .4 inches in diameter), means that the energy transferred to pressure waves is low (small amplitude). Turning the volume up on the computer results very quickly in saturation of the speaker cones; the earphones are pushed up against their amplitude limit at about -10 decibels. This results in the vibration of the entire earphone casing, as opposed to just the speaker cone.

Since the "earphones" are in–the–ear style, they are not attached to any rigid body and therefore vibrate freely. (Over–the–ear style "headphones", which are attached to a bracket that goes over the listener's head, provide support for the speaker casings.) As a result, more energy goes to the vibration of the earphone casing in the opposite direction than in the case of a regular floor–standing speaker or over–the–ear headphones. Furthermore, the placement of the earphones at the entrance of the auditory canal causes the canal to act as a column of air closed at both ends, which increases the resonant frequency by half (as opposed to over–the–ear headphones, which maintain the column of air as a system with one end closed).

Experiment 1: Threshold of hearing based on the frequency of pure tones.

Pure tones of decreasing frequency were played. Each frequency was rated as perceptible or not perceptible (with some clarifying adjectives added as needed). Figure 6 show sample results.

The cutoff for low frequencies is about 30 Hz. This result agrees previous discussion on frequency limits of the human ear. The limit of the ear's high frequency response was unable to be tested because soundmaker generates sound at a sampling rate of 22 kHz. Any sound frequency above 11,000 Hz cannot be accurately generated by this program. 11,000 Hz tones, however, were easily heard.

The onset of the sound was noticeable: a distinct clicking sound was heard immediately before the tone was played. This was most likely due to the activation of the workstation's sound circuitry. While the clicking sound was ignored, its presence none–the–less may have tainted the results slightly. One modification that would make this experiment more accurate would be to eliminate this onset clicking noise by having the sound circuitry on the workstation switched on continuously.

Frequency (Hz)	Perceptible?
100	heard
94	heard
88	heard
82	heard
73	heard
70	heard
64	heard
58	heard barely
52	heard barely
46	heard barely
43	heard barely
40	heard barely
34	almost imperceptible
31	not noticeable

Figure 6 *Perception of tone based on frequency*

Experiment 2: Subjective tone generation from the monaural superposition of 2 pure tones of regularly differing frequencies.

Two tones, one of a base frequency and the other a fractional multiple, n/m, of the base frequency (where n and m are small integers), were superimposed and played monaurally. The experiment was repeated for three base tones — 100 Hz, 440 Hz, and 10,000 Hz — representing low, middle, and high frequencies. The perception of a subjective tone was rated.

Subjective tones are perceived tones at the frequency of the greatest common divisor for the two superimposed tones. This is also the cycle repeat frequency for the superimposed tone. The ear, which hears the superimposed tone from summing 440 Hz and 660 Hz tones as separate tones (performing a Fourier analysis of the superimposed tone), also perceives a subjective tone at 220 Hz, even though a tone of that frequency is not actually in the superimposed tone (Roederer, 1975).

At 440 Hz and 660 Hz (3/2 of 440), a tone of 220 Hz was easily heard. At 440 Hz and 550 Hz (5/4 of 440), a tone of 110 Hz was heard. At 10,000 Hz and 15,000 Hz (3/2 of 10,000), a tone at 5,000 Hz was very difficult to hear: the effect was almost imperceptible. At 100 Hz and 150 Hz (3/2 of 100), a tone at 50 Hz was also difficult to hear, but more apparent than for high frequencies.

The perception of subjective tones may be due to the method the ear uses to encode sounds. Neurons along the basilar membrane are not only tuned for specific frequencies, they also encode frequencies in their firing rates. While the subjective tone doesn't appear in the superimposed tone that distorts the basilar membrane, the cycle frequency for the superimposed tone is encoded in the rate of firing of the neurons along the membrane.

The ear's sensitivity to subjective tones is similar to the ear's sensitivity to pure tones. This is likely due to limits on the firing rates of neurons. Below a certain frequency, neurons can fire once per activation. Above this frequency, a neuron can only fire some or most of the time, dependent on the refractory period of that particular cell. A set of similar neurons can encode higher frequencies in their collective firing rate: each fires in turn, resetting itself while other neurons in the group fire. There is also a frequency limit to this synchronicity as well; this accounts for the lack of sensitivity at high frequencies.

Experiment 3: Subjective tone generation from the binaural superposition of 2 pure tones of regularly differing frequencies.

Experiment 2 was repeated, but the two tones that were previously superimposed and presented monaurally were instead presented binaurally, one to each ear.

The subjective tones that were generated when the sounds were presented monaurally were absent when presented binaurally. This result points to the conclusion that subjective tones are dependent on stimulus from the ear, and are not simply a result of higher–level processing.

This result is in line with the conclusions drawn in Experiment 2 regarding the sensing of subjective tones. If the two tones are no longer superimposed and presented together to the ear, there is no way for neurons to encode the frequency of the repetitive cycle of the superimposed tone with their firing rates. Therefore, the phenomenon of subjective tones should be absent when tones are presented binaurally.

Experiment 4: The frequency threshold for detection of beats.

A base tone was superimposed with a tone of slightly different frequency, and the existence of beats was tested. The frequency of the other tone was increased until the beating effect disappeared. The experiment was repeated for 80 Hz, 1,000 Hz, and 10,000 Hz, representing the low, middle, and high frequencies. The frequencies at which the beats disappeared were recorded.

At a base rate of 80 Hz, the accompanying tone had to be 93 Hz before it sounded like two distinct tones. At a base rate of 1,000 Hz, the accompanying tone had to be 1140 Hz before it sounded like two distinct tones. At a base rate of 10,000 Hz, the accompanying tone had to be 10,130 Hz before it sounded like two distinct tones.

Experiment 5: Beats with secondary tones.

Tones similar to those presented in Experiment 2 were presented, this time with a third tone, at a frequency slightly different from the frequency of the expected subjective tone, superimposed. The existence of beats was tested. To ensure that there was no effect from interaction between one of the original tones and the third tone, both the first and the second tones were presented separately superimposed with the third tone, and the existence of beats was again tested.

There is a pronounced beating effect between the subjective tone at 220 Hz produced by superimposing a 440 Hz tone and a 660 Hz tone, and a pure tone of slightly differing frequency at 225 Hz. The beating effect is also apparent when 440 Hz and 550 Hz tones (which creates a subjective tone at 110 Hz) and a pure tone at 112 Hz are superimposed. When the 440 Hz tone and the 225 Hz tone were superimposed and presented, two separate tones were heard corresponding to those frequencies, and no beating was heard (as expected). This was also the case for 660 Hz and 225 Hz superimposed tones.

This is by far the most interesting result of these experiments. Essentially, beating is perceived between a actual tone (225 Hz) and a non-existent tone (220 Hz). This indicates that beating can result from higher–level neural processing as well as directly from the inner ear. Since the 220 Hz and 225 Hz “tones” are encoded through two different channels, there must be higher–level processing that results in our hearing beating between these “tones”.

Conclusion

The physics of the ear provides a wealth of interesting phenomena. By design, the ear lends itself to simple physical modelling that allows many of these phenomena to be understood. Modeling the ear first as a column of air and then as a simple harmonic oscillator, its functionality and response can be predicted. Through these physical models, the process by which animals perceive sounds can more fully understood.

Bibliography

"ear." Britannica Online. <http://www.eb.com:180/cgi-bin/g?DocF=micro/183/25.html> [Accessed 25 October 1997].

Goldstein, Bruce E. *Sensation and Perception*. Brooks/Cole Publishing Co., 1996.

Fechner, Gustav Theodor. *Elements of Psychophysics*. New York: Hold, Rinehart, and Winston, 1966.

Roederer, Juan G. *Introduction to the Physics and Psychophysics of Music*. New York: Springer-Yerlag, 1975.