A Least Squares Model for Human Hearing

By Cailtin Roos

The human ear is made up of three distinct parts the outer, middle, and inner ear. Soundwaves enter through the outer ear and move through the middle ear where they are channeled to the eardrum, and which then causes the eardrum to vibrate. Tiny bones link the middle ear to the cochlea in the inner ear, these bones pass those individual vibrations to the liquid containing cochlea, which moves minute hairs in the inner ear that in turn send information about high and low frequencies through the neural pathways into the brain to be interpreted as sound. The higher frequencies produce vibrations and are heard near the middle ear, and low frequencies produce vibrations only being heard deep into the ear, this is why deep bass music can make you feel like your jaw is vibrating, as the sound waves really do go deeper into your head than high pitches do. These soundwaves that our ear picks up are essentially variations in air pressure in time.

A sound wave can be described in a sinusoidal way, as each sound wave has a consistent repetitive oscillation, we can model them as a function of time such that we have a normal atmospheric pressure condition, A_0, the maximum deviation of the air pressure from the normal atmospheric pressure, A, the frequency of the wave in cycles per second, $\frac{w}{2\pi}$, and the phase angle of the wave, δ, to create the function model $q(t)$, of the atmospheric pressure at the eardrum over time. Such that we get the function $q(t) = A_0 + A\sin w(t - \delta)$. The human ear can only perceive waves with frequencies between 20 cycles per second and 20,000 cps, any sinusoidal waves with a frequency outside of this range will not stimulate the minute hairs within the cochlea enough to produce nerve signals to the brain.

We can approximate, to a reasonable degree, the ear as a linear system. As prior research has shown that when the ear is presented with a complex sound wave, it breaks it apart into multiple waves that it can understand, a combination of our function q(t). This is remarkably similar to what we were doing in class, when we were breaking down a vector into different orthogonal components, which would allow us to have an easier time interacting with that vector in equations and problems. So, the ear breaking down this complex sound wave allows for the ear to understand the frequencies and produce the same sound from the individual q_i's as it would from the whole complex wave. Now, we can not necessarily say every soundwave will be a finite sum of functions like q(t), so let's consider a periodic sound wave p(t), with period T, so $p(t) = p(t+T)$ as periodic functions are consistent repeating functions. Through prior research and examination, it has been found that the response of the ear to such a wave will produce the same response as the sum of some sinusoidal waves. So, there must exist some $q(t) = [A_0 + \sum_{i=1}^{n} A_i \sin w_i(t - \delta_i)]$ such that q(t) produces that same sound as p(t), even though these are different functions of time.

To properly approximate our periodic function p(t) by a sum of q_i's we must determine the amplitudes, frequencies, and phase angles of the sinusoidal components of q(t). Now since q(t) produces the same sound as p(t), it is reasonable to assume that the period of the functions is the same, T. Which in turn must mean that the components of q(t) also have period, T, so the frequencies of the sinusoidal components will be multiples of the frequency $\frac{1}{T}$, of p(t), these frequencies need not be identical to produce the desired results (Wen-juan). So, we now have our w defined as $w_k = \frac{2k\pi}{T}, k = 1,2, ...$ Now as mentioned before the ear cannot perceive waves with frequencies greater than 20,000 cps, so we can omit k-values that would cause such frequencies. So, if we define a variable, n, to be the largest integer such that $\frac{n}{T} < 20{,}000$, we can write $q(t) = A_0 + A_1 sin\frac{2\pi}{T}(t - \delta_1) + \cdots + A_n sin\frac{2n\pi}{T}(t - \delta_n)$. Now we must determine the appropriate amplitudes, A_n, and phase angles, δ_n , so that our q(t) and p(t) will produce the same responses.

There exists criteria by which the human auditory just ‘picks’ these values to ensure the correct response, so the most useful step to take now is rather that figuring out exactly our amplitudes and phase angles, we will take an approach of minimizing the total error, difference in out put between our functions, of q(t) and p(t). So, lets define $e(t) = p(t) - q(t)$, and this will be our error function. Note, as we are considering q(t) an approximation of p(t) that the ear has the exact same response too, we must have e(t) arbitrarily small over our period, T, a value less than 20 cps that the ear cannot perceive.

Now when one goes about minimizing the error if we just use our e(t) all parts of the error are equally represented. Which may cause unnecessarily close approximations in some places and large differences in other parts. However, if we want to minimize the larger differences between the functions to a greater degree than the small differences, then it is more helpful to minimize the square of such error, [e(t)]^2. Which gives greater weight to greater differences. To go about minimizing our error as accurately as possible we will integrate over our functions period, T, and use a Fourier series expansion of p(t), and our q(t). So, the integral $e(t) = \int_0^T [p(t) - q(t)]^2 \, dt$, will be the least squares error of our approximation. As the criteria of this approximation is the smallness of the above integral, we can say that q(t) is the least squares approximation to p(t) over the interval [0,T]. To find our needed amplitudes and phase angles we will go through the following process, as q(t) is continuous on our period, T, the trigonometric function $q(t) = a_0 + a_1 \sin\left(\frac{2\pi t}{T}\right) \ldots + a_n \sin\left(\frac{2\pi n t}{T}\right)$ that will minimize the mean square error of e(t), has coefficients $a_k = \frac{2}{T}\int_0^T p(t) \sin\left(\frac{2\pi n t}{T}\right) \, dt$ with $k = 1,2, \ldots, n.$

As an example, consider, we have the ear picking up a non-sinusoidal wave $p(t) = (3t - 4\pi)^2$ on the interval $[0,2\pi]$ and we want to find the q_i’s that the ear will actually be picking up the frequencies of, then our e(t) would look like $e(t) = \int_0^T \left[\{(3t - 4\pi)^2\} - \{a_0 + a_1 \sin\left(\frac{2\pi t}{T}\right) ... + a_n \sin\left(\frac{2\pi n t}{T}\right)\}\right]^2 \ dt$. Now to find our amplitudes and phase angles we will compute $a_k = \frac{2}{2\pi}\int_0^{2\pi}(3t - 4\pi)^2 \sin\left(\frac{2\pi k t}{2\pi}\right) dt$ which yields $a_k = 6k^{-2}\ \forall k$, let's approximate our polynomial to the 4th degree, $a_0 + a_1 \sin(t) + a_2 \sin(2t) + a_3 \sin(3t) + a_4 \sin(4t)$, with the appropriate a_k's yields $q(t) = 0 + 6\sin(t) + \frac{3}{2}\sin(2t) + \frac{2}{3}\sin(3t) + \frac{3}{8}\sin(4t)$. So, we have successfully found a combination of sinusoidal soundwaves that will be perceived the same way our non-sinusoidal p(t) wave will be by the ear.

In summary the least squares approximation of p(t) by a finite sum of q(t)'s is essentially a Fourier transformation, as scientifically speaking: The membranes in the inner ear spread out different frequencies: high frequencies produce a large vibration at the end near the middle ear, and low frequencies a large vibration at the distant end. Thus, the ear performs a sort of frequency analysis, roughly similar to a Fourier transformation. And we can easily break down complex non-sinusoidal waves, into finite combinations of functions of air pressure change in the ear to closely approximate all human heard noises utilizing sine curves.

Bibliography

Anton, H. (2010). Chapter 10, Section 19. In *Elementary Linear Algebra Applications Version* (10th ed.). textbook, Wiley & Sons.

Scheick, J. T. (1997). Section 3.4 Projections in F^m. In *Linear Algebra with Applications* (Ser. International Pure and Applied Mathematics, pp. 133–151). textbook, McGraw-Hill.

Wen-juan, Y., Jian-wei, M., & Bao-lin, H. (2011, October 10). *Numerical Model on Sound-Solid Coupling in Human Ear and Study on Sound Pressure of Tympanic Membrane*. Wiley Online Library. https://onlinelibrary.wiley.com/doi/10.1155/2011/282696

A Least Squares Model for Human Hearing:

The Ear as a Fourier Transformer

By Caitlin Roos

Fall 2024

SUNY Geneseo Mathematics

Anatomy of the Ear

The Outer Ear

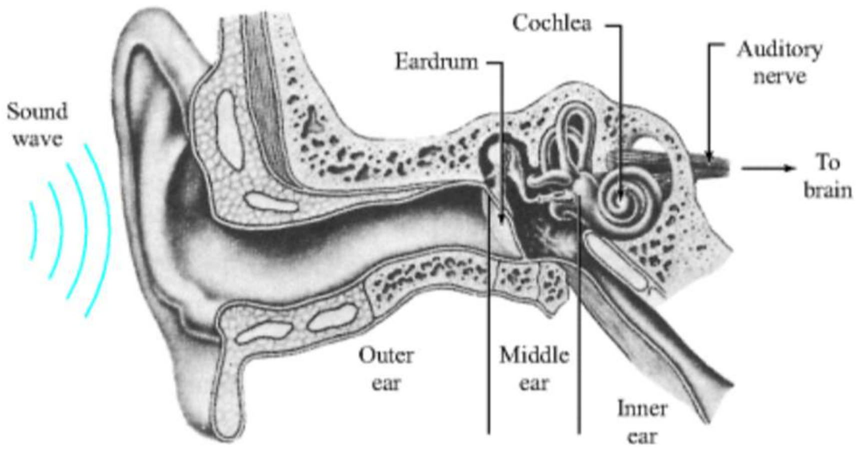
- Where soundwaves enter
- Contains the ear canal

The Middle Ear

- Here soundwaves from the ear canal are channeled to the eardrum causing it to vibrate
- Separated from the inner ear by the eardrum

The Inner Ear

- Contains the cochlea, auditory nerve and minute hairs that are stimulated by soundwave frequencies



How We Hear

Eardrum

- Once soundwaves pass through the outer and middle ear to the eardrum they continue thru the inner ear to the cochlea

Cochlea

- contains a liquid that is altered by soundwaves
- these alterations are picked up by a plethora of tiny hair-like nerves

Nerves

- These nerves can only distinguish frequencies (sounds) if they vary enough (i.e. pass a threshold in variation)
- An ear contains many small hair cells. The hair cells differ in length, tension, and thickness, and therefore respond to different frequencies

Location

- high frequencies produce a large vibration at the front of the inner ear, and low frequencies a large vibration at the distant end

Soundwaves

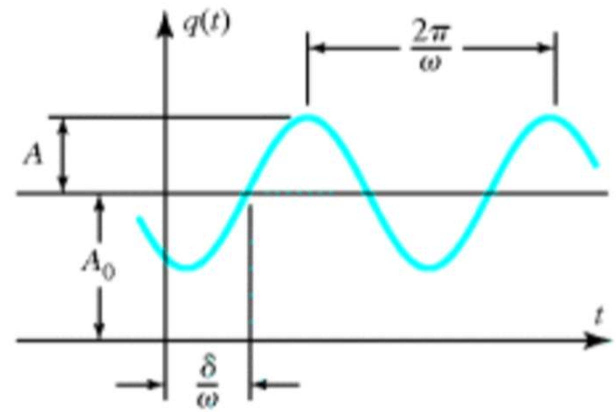
For the auditory system, the most elementary type of sound wave is a **<u>sinusoidal variation</u>** in the air pressure.

$$q(t)=Ao + A\sin(wt - \delta))$$

w = frequency of the wave and
δ = phase angle of the wave,
Ao=Normal atmospheric pressure,
A= maximum deviation of the pressure from the normal atmospheric pressure

Our q(t) represents the atmospheric pressure in the eardrum

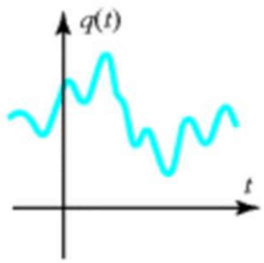
- To be perceived by the ear the frequencies must have a range of 20 cycles per second to 20,000 cycles per second

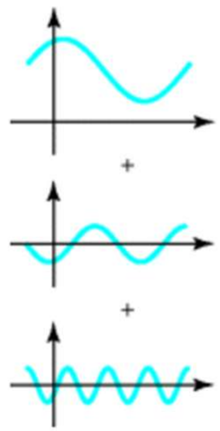


Breaking Down Soundwaves

Complex Soundwaves

- when the ear is presented with a complex sound wave, it breaks it apart into multiple waves that it can understand, a combination of our function q(t)



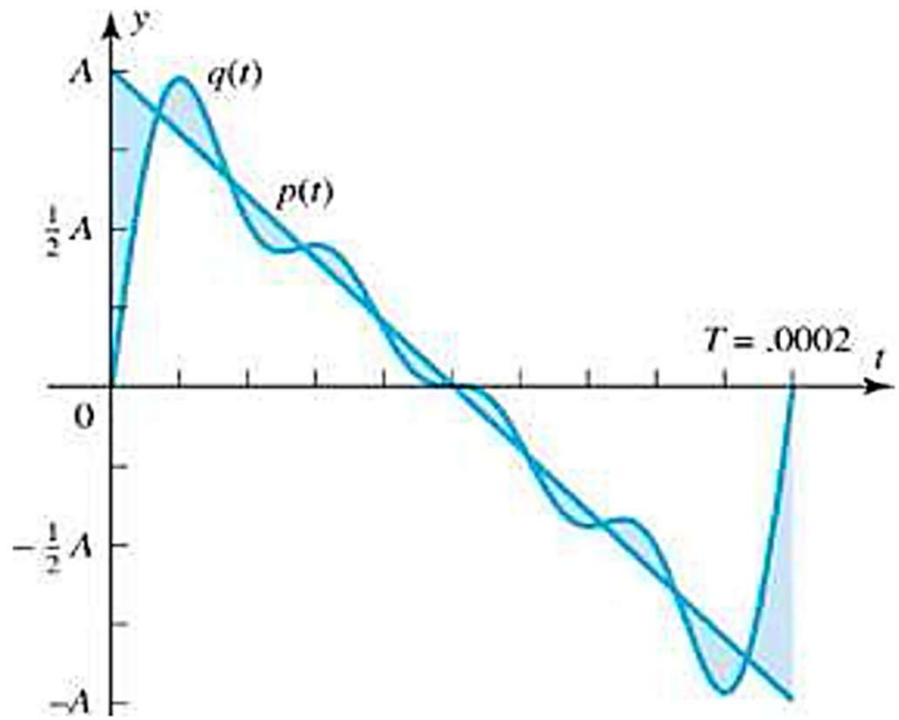


Any Periodic Soundwave

- research has shown that the response of the ear to a wave, p(t) any continuous function with period, T, will produce the same response as the sum of some sinusoidal waves q_i's
- there must exist a function q(t) with such a small error from p(t) that the response of the ear to each wave function is identical

How do we find this q(t) ?

- we will minimize the error, e(t), of the function approximation, q(t), of our periodic function
- When minimizing all parts of the equation are equally represented which may lead to inaccuracies is the approximation is more accurate in some places than others
- To circumvent this we will be minimizing the error function <u>squared</u>



The Error Function

- Because we are dealing with periodic functions, we will sum up all the errors over each interval dt, forming the integral ->

- Now we must find the specific amplitudes, w, and phase angles, δ, of our components of q(t) which will yield the correct LSRL

$$\int_0^T [e(t)]^2 dt = \int_0^T [p(t) - q(t)]^2 dt$$

Minimizing Mean Square Error on [0,T]

Thm 10.19.2, Anton

If $f(t)$ is continuous on $[0, T]$, then the trigonometric function $g(t)$ of the form

$$g(t) = \frac{1}{2}a_0 + a_1 \cos\frac{2\pi}{T}t + \ldots + a_n \cos\frac{2n\pi}{T}t + b_1 \sin\frac{2\pi}{T}t + \ldots + b_n \sin\frac{2n\pi}{T}t$$

that minimizes the mean square error

$$\int_0^T [f(t) - g(t)]^2 \, dt$$

has coefficients

$$\begin{aligned} a_k &= \frac{2}{T}\int_0^T f(t)\cos\frac{2k\pi t}{T}\, dt, \quad k = 0, 1, 2, \ldots, n \\ b_k &= \frac{2}{T}\int_0^T f(t)\sin\frac{2k\pi t}{T}\, dt, \quad k = 1, 2, \ldots, n \end{aligned}$$

Example: least squares approximation to a soundwave

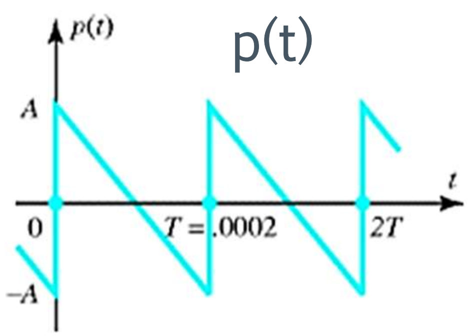
- Let our soundwave p(t) have a saw tooth pattern with basic frequency of 5,000 cps. The period of the wave, $T = \frac{1}{5000}$, so we have the equation $p(t) = \frac{2A}{T}(\frac{T}{2} - t)$
- By Thm 10.19.2
  - $a_0 = \frac{2}{T}\int_0^T \frac{2A}{T}\left(\frac{T}{2} - t\right) dt \;= 0$
  - $a_k = \frac{2}{T}\int_0^T p(t) cos\left(\frac{2k\pi k}{T}\right) dt \;= \frac{2}{T}\int_0^T \frac{2A}{T}\left(\frac{T}{2} - t\right) cos\left(\frac{2k\pi k}{T}\right) dt = 0 \;\forall k \in N$
  - $b_k = \frac{2}{T}\int_0^T p(t) sin\left(\frac{2k\pi k}{T}\right) dt \;= \frac{2}{T}\int_0^T \frac{2A}{T}\left(\frac{T}{2} - t\right) sin\left(\frac{2k\pi k}{T}\right) dt = \frac{2A}{k\pi}$ k = 1,2,…

Ex cont.

- Note $\frac{4}{T} = 20{,}000$cps so our approximation should only be up to 4th degree
- So, our least squares approximation for our p(t) is:

$$q(t) = \frac{2A}{\pi}[sin\frac{2\pi}{T}t + \frac{1}{2}sin\frac{4\pi}{T}t + \frac{1}{3}sin\frac{6\pi}{T}t + \frac{1}{4}sin\frac{8\pi}{T}t]$$

  - Note our four sinusoidal terms have frequencies 5000, 10000, 15000, and 20000 cps, respectively



π

- Now as we know a least squares approximation becomes better as the number of terms becomes larger
  - Our e(t) tends to zero
- We can denote this as $f(t) \sim \frac{1}{2}a_0 + \sum_{k=1}^{\infty}(a_k \cos kt + b_k \sin kt)$
- Where the right-side of this equation is the Fourier expansion of q(t)
- Which shows that interestingly our ear is a Fourier Transformer

Thank You

Sources:

Anton, H. (2010). Chapter 10, Section 19. In *Elementary Linear Algebra Applications Version* (10th ed.). textbook, Wiley & Sons.