The Frequency Discrimination of The Ear

Yoshimutsu. Hirata

November 6, 2014

Summary

What mechanism the ear uses for the remarkable frequency discrimination remains to be solved. The ear detects the frequency change of about 2 Hz for a frequency of 1 kHz. More than 50 years ago, Békésy investigated the cochlea and he concluded that sound of a specific frequency produces waves that flex the basilar membrane at a specific spot, causing the hair cells there to react and send signals to the brain. The location of hair cells would correspond to the frequency. But Békésy explanation is too simple to unravel the mystery of the frequency discrimination of the ear. This paper presents the theory of hearing that accounts for the remarkable frequency discrimination of the ear. For verification, the time-frequency analysis of sound obtained by the method based on the theory is shown.

PACS no. 43.66.Ba, 43.66.Cb, 43.66.Dc, 43.66.Fe

1. Introduction

The ear is made up of the three sections: outer, middle, and inner ears, as they are called. In the middle ear, the smallest bones in the human body, commonly called the hammer, anvil, stirrup, form a bridge linking the eardrum with the oval window, the portal to the inner ear, The job of the middle ear is to transform the acoustical vibration of the sound waves into mechanical vibration and pass it on to the inner ear, see Fig.1.

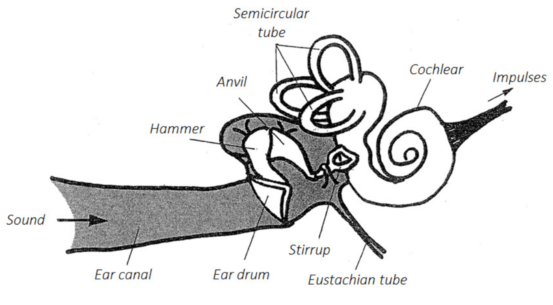

Figure 1. Sketch of the ear.

The linkage of the three little bones of the middle ear functions as the automatic gain control of incoming forces. Thus, the force exerted on the eardrum is amplified just enough to set the fluid in the cochlea in motion. The cochlea is a bundle of three fluid-filled ducts coiled up in a spiral like the shell of a snail. Two of the ducts are connected at the apex of the spiral. When the oval window is set in motion by the stirrup, it moves in and out, setting up hydraulic pressure waves in the fluid. These waves undulate the walls separating the ducts.

Along one of these walls, known as the basilar membrane, is the organ of Corti, the true center of healing. It consists of rows of sensory hair cells, some 7,500 or more. Each hair cell has conections with other hair cells. From these hair cells, thirty thousands of nerve fibers carry information about the frequency and intensity of the sound to the brain, where the sensation of hearing occurs.

Georg von Békésy discovered that as the hydraulic pressure waves travel along the ducts in the cochlea, they reach a peak somewhere along the way and push on the basilar membrane. Waves generated by high-frequency sounds push on the membrane near the base of the cochlea, and waves from low-frequency sounds push on the membrane near the apex. Thus, Békésy concluded that sound of a specific frequency produces waves that flex the basilar membrane at a specific spot, causing the hair cells there to react and send signals to the brain. The location of hair cells would correspond to the frequency and the number of hair cells triggered would correspond to the intensity.

The organ of Corti converts the movement of the basilar membrane into electronical impulses and sends these to the brain. The frequency of nerve impulses is about 1 kHz to 3 kHz and the signals represented by impulses are about the same in duration and strength[1].

Even now, however, scientists are not sure what mechanism the ear uses to discriminate the frequency of sounds. The ear detects the frequency change of about 2 Hz for a frequency of 1 kHz. This paper presents theories which explain the model of the band-pass filter composed of hair cells, the method of converting a filtered signal into bits, and the main topic of the frequency discrimination.

2. Theory of hearing

2.1. Band-pass filter

The model of the band-pass filter in the cochlea is composed of three hair cells. The peak frequency of the response of hair cells excited by the movement of the basilar membrane corresponds to the location of hair cells. The curve $\mathrm{E_A}$ plotted in Fig. 2 shows the response of the hair cell whose peak frequency is 1,000 Hz, which is approximated by the frequency response of a single freedom of vibration of the quality $\mathrm{Q} = 7$.

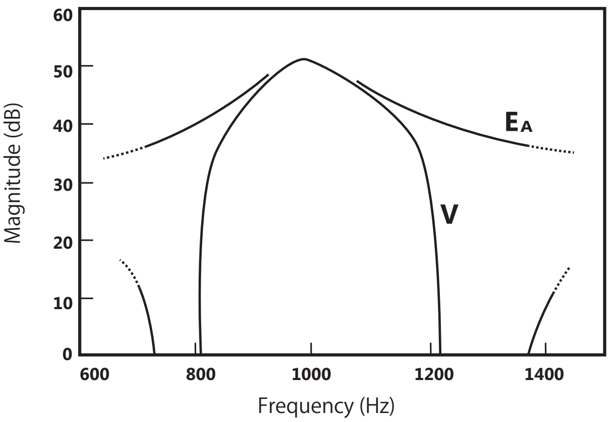

Figure 2. The frequency response of the hair cell approximated by the response of a single freedom of vibration, E_A, and the response of the band-pass filter, V.

Let $\mathrm{E_B}$ and $\mathrm{E_C}$ be the responses of hair cells of which peak frequencies are 800 Hz and 1,250 Hz, respectively, and

$$V = \mathrm{E_A} - \mathrm{k}\left(\mathrm{E_B} + \mathrm{E_C}\right), \tag{1}$$

then the response of band-pass filter, V, gives $V = 0$ at the frequencies 820 Hz and 1,220 Hz, where $k = 0.125$. On the assumption that the band-pass filter is composed of three hair cells, the number of the filter is about 2,500 which is approximately equals to the estimated number of spectra which is given by the smallest detectable changes in frequency.

2.2. Coding

It is known that the brain does not respond to mechanical vibrations but only to electro-chemical changes. Thousands of nerve fibers send information about the intensity of sound to the brain. Thus, the organ of Corti must convert the magnitude of the filtered signal into a code of binary bits, namely, impulses.

The smallest detectable change of sound level is 1 dB. The ear can cope with a range of about 130dB in loudness. Thus, the maximum number of impulses of a code is about 130. The frequency of nerve pulse increases with the increase of the strength of excitation. Assuming that the frequency of nerve pulse is 3 kHz, the time length of a code representing the sound level of 130 dB is 43 ms. Also assuming the frequency of nerve pulse is 1 kHz, the time length of a code representing the sound level of 40 dB is 40 ms. Namely, from the soft whisper to the roar of a jet plane taking off, the signals sent by the organ of Corti are about the same in duration.

The band-pass filter sends the signal of time varying intensity at intervals of about 40 ms, i.e., the sampling frequency is about 25 Hz. The level difference near the apex of the frequency response V from 975 Hz to 1,025 Hz is , as shown in Fig. 2, less than 0.5 dB. Rather than increasing the sampling frequency, the ear sends plural signals from the neighbouring band-pass filters. For example, the time varying intensity of 1,000 Hz is represented by the sum of the sampled signals from band-pass filters specified by the center frequency of 1,000 Hz, 1,002 Hz, 1,004Hz and 1,006 Hz, where the sampling lags are 0 ms, 10 ms, 20 ms, and 30 ms, respectively. Then, the signal is sent to the brain at intervals of 10 ms.

2.3. Frequency discrimination

A basic problem for the theory of hearing is to account for the remarkable frequency discrimination of the ear[2]. Fig.3 shows the schematic frequency responses of band-pass filters F_P, F_A, and F_Q. Assuming the patterns of these frequency responses (log scale) are same and f_P, f_A, and f_Q the center frequencies of the filter F_P, F_A, and F_Q, respectively, where $f_P f_Q = {f_A}^2$, the output level, L_2, of the filters F_P coincides with that of the filter F_Q for the input signal of the frequency f_A, and then the magnitude of the input signal is given by the output level, L_1, of the filter F_A. A small frequency change Δf of the input signal, as shown in Fig.3, causes the difference D_2 which is large compared with the difference D_1, i, e, the change of the output of the filter F_A. Thus, if D_2 is larger than 1 dB, the ear can discriminate the frequency change Δf.

Now, consider the values described above using the frequency response V shown in Fig. 2. The pair of frequencies, f_P and f_Q, are given by $f_P = (1.002)^{-n} f_A$ and $f_Q = (1.002)^n f_A$ in the frequency range from 500 Hz to 2 kHz, where n is an integer and f_A is 1 kHz. Putting $D_2 = 1$ (dB) and $\Delta f = 2$ (Hz), then $n = 103$ which gives the pair of frequencies 832 Hz and 1,202 Hz. Similarly, putting $D2 = 1$ (dB) and $\Delta f = 10$

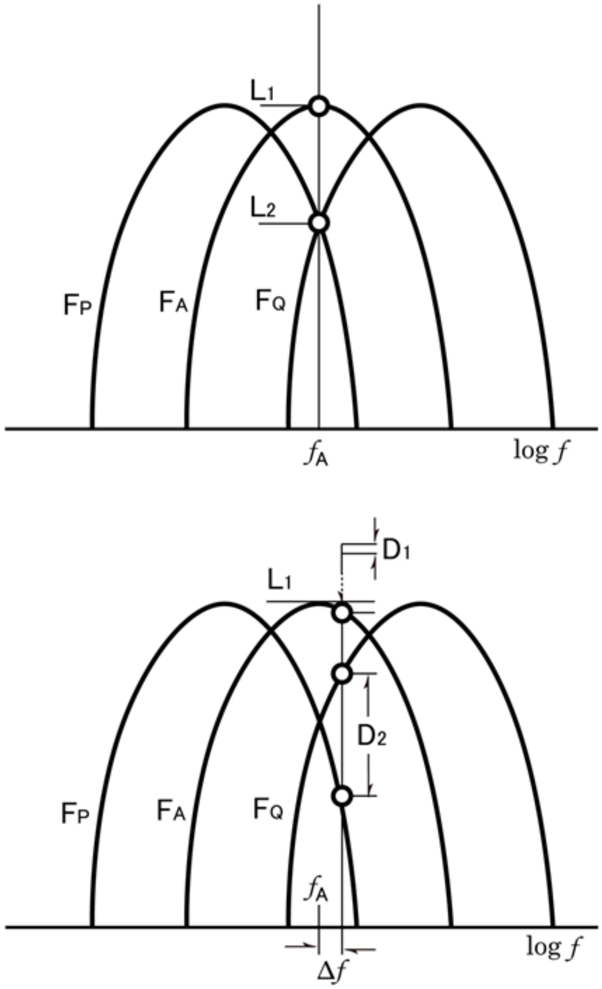

Figure 3. Illustration of the frequency responses of band-pass filters F_P, F_A and F_Q pertaining to the discrimination of the frequency f_A (upper) and $f_A + \Delta f$(lower).

(Hz), then $n = 53$ which gives the pair of frequencies 900 Hz and 1,110 Hz. The detectable frequency change Δf increases as n decreases.

Our ears focus on a tone played by one instrument in an orchestra of a hundred. The empirical fact suggests that the precision of frequency discrimination is selective and, thus, the signals from plural pairs of band-pass filters, F_P and F_Q, pertaining to the band-pass filter F_A are sent to the brain where the sensation of hearing occurs. Signals from about 2,500 band-pass filters are carried by 30,000 nerve fibers and sent to the brain. Thus, the brain uses the signals from 12 band-pass filters to identify one spectrum, *e.g.*, 4 signals to get the time varying intensity and 4 pairs of signals to discriminate the frequency.

3. Experiments

Outputs of the thousands of band-pass filters are represented by the set of the non-harmonic Fourier spectrum $y(f_n)$ where f_n is a frequency such that $f_n = nf_1$ where f_1 is smaller than the reciprocal of the time length of sample data, T, for analysis. Since the smallest detectable change of sound level is 1 dB, the spectrum $y(nf_1)$ is transformed into the discrete magnitude $Y(nf_1)$, i.e., the magnitude in decibel is expressed by an integer. The pair of frequencies are expressed by$f_P = f_A n f_1$ and $f_Q = f_A + nf_1$. Thus, for $1/2T < nf_1 < 1/T$, if

$$Y(f_A n f_1) = Y(f_A + n f_1) \tag{2}$$

and

$$Y(f_A) > Y(fAnf_1), \tag{3}$$

the spectrum of sample data is given by the magnitude $Y(f_A)$ and the frequency f_A. The center frequencies of band-pass filters in the organ of Corti are arrayed with constant spaces in log scale from about 500 Hz to 2 kHz, and the pair of frequencies are related by $f_P f_Q = {f_A}^2$. The center frequencies of band-pass filters represented by the Fourier spectrum are arrayed with constant spaces in linear scale, and $f_P + f_Q = 2f_A$. These differences are not substantial problems for appreciating the theory of hearing by experiments.

According to the method of hearing model described above, the spectrum of sample data was estimated at intervals of 10 ms to get the time-varying spectrum of sound. In the experiment, the sample of MIS, the university of Iowa Musical Instrument Samples, was used to check the spectral estimation[4]. Fig. 4 shows the time-varying spectrum of a tone (C5) played by a piano in the frequency range less than 2 kHz, where $n = 70$ and $f_1 = 1$ (Hz). The method removes the obscurity of the spectrum that occurs when analyzing short data by traditional spectral estimation techniques. The accuracy of the frequency discrimination is about 2 Hz for $nf_1 = 70$. Similarly, the same set of sample data was analyzed by the non-harmonic Fourier analysis proposed by the author[3]. The result of the spectral estimation is shown in Fig. 5. The accuracy of the frequency discrimination is 1 Hz. Compared with the procedure of the non-harmonic Fourier analysis, the computational requirement of the hearing model is significantly low, *i.e.*, the method is suitable for real-time implementation.

The theory of hearing accounts for the remarkable frequency discrimination of the ear. The theory mentions several numbers such as 7,500 hair cells, 2,500 band-pass filters, 30,000 nerve fibers and so on. These are key numbers to unravel the mystery of the ear. The theory suggests that the ear uses principles of acoustics, mechanics, electronics, and higher mathematics to accomplish the spectral estimation. How did the ear get all of them?

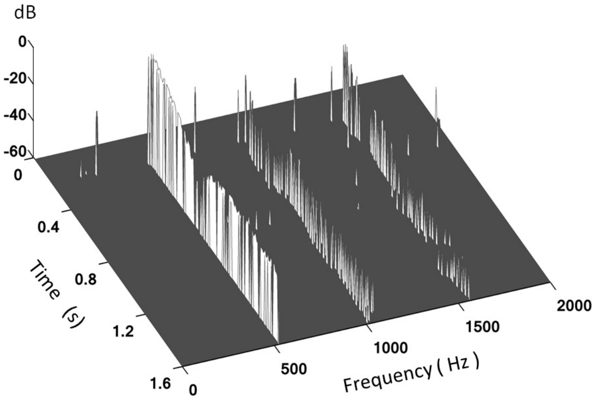

Figure 4. The time-varying spectrum of a tone (C5) played by a piano given by the hearing model..

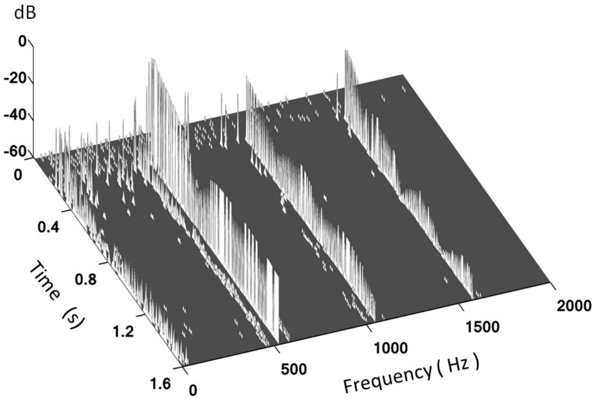

Figure 5. The time-varying spectrum of a tone (C5) played by a piano given by the non-harmonic Fourier analysis.

4. Acknowledgement

The author thanks Mikio Tohyama for his useful discussion. The author also thanks Satoru Goto for his help in the experiment.

References

[1] S. S. Stevens and F. Warshofsky, Sound and Hearing. TIME Inc. 1965

[2] B. C. T. Moore, An Introduction to The Psychology of Hearing. Academic Press. 1997.

[3] Y. Hirata, Non-harmonic Fourier analysis available for detecting very low-frequency components. J. Sound and Vibration 287, (2005) 611-613.

[4] Y. Hirata and S. Goto, Pictorial of Sounds. http://www7b.biglobe.ne.jp/ wavesciencestudy/