THE MECHANISM OF HEARING AS REVEALED THROUGH EXPERIMENT ON THE MASKING EFFECT OF THERMAL NOISE

BY HARVEY FLETCHER

BELL TELEPHONE LABORATORIES

Read before the Academy April 26, 1938

Considerable information concerning the dynamical properties of the hearing mechanism can be obtained from physical measurements on audition. In fact, probably more precise information can be obtained by such measurements than from data obtained by animal experimentation, or from data on post mortem sections of human temporal bones.

It is well known that when a person is immersed in a noise his ability to hear other sounds is decreased. The noise is said to produce a masking effect. Before describing the experiments on the masking effect of thermal noise I will first describe the method of obtaining thermal noise and how it is quantitatively defined.

In an electrical conductor there is a statistical variation of the electrical potential difference between its two ends which is due to the thermal agitation of the atoms, including the electrons. This fluctuating voltage may be amplified by means of a vacuum tube amplifier. Such an electrical disturbance is called an electrical thermal noise. If this electrical thermal noise is sent through a telephone receiver, or a loud speaker, it is then converted into acoustical thermal noise. Such an acoustical thermal noise was used in the masking experiments discussed in this paper.

Let I_f be defined as the intensity per cycle of thermal noise using 10^{-16} watts* per square centimeter as a unit of intensity. If ΔI is the intensity due to the thermal noise in the band of frequencies between f and $f + \Delta f$, then

$$I_f = \frac{\Delta I}{\Delta f}. \tag{1}$$

The spectrum level B in decibels is defined by

$$B = 10 \log I_f. \tag{2}$$

In general I_f and B will vary with the frequency. The curve showing the variation of B is called a spectrogram of the noise.

Let us now examine more closely the relationship between the intensity of the noise and its masking effects. When a thermal noise having a given spectrogram is impressed upon the listener's ears it causes the hearing mechanism to vibrate some parts more and other parts less. These vibrations stimulate the nerve endings. The ear has a selective action for sounds of different frequencies so that the amount of agitation or stimulation given any set of nerve endings will depend upon their position.

It is well known that the nerve endings responsible for auditory effects are located along the basilar membrane. According to the measurements of anatomists this membrane varies in width from 0.2 millimeter at the oval window end to 0.5 millimeter at the helicotrema end, and is about 30 millimeters long. It will be seen that it is about 100 times as long as it is wide. Therefore, a single coördinate x will define with sufficient accuracy the position of a patch of nerve endings. Then, if we take one per cent of the length it will contain a patch of nerve endings about as long as it is wide.

To make this concept more concrete, a spiral line having approximately the same shape as the human cochlea along which the basilar membrane runs is shown in figure 1. The numbers along the spiral line give the ap-

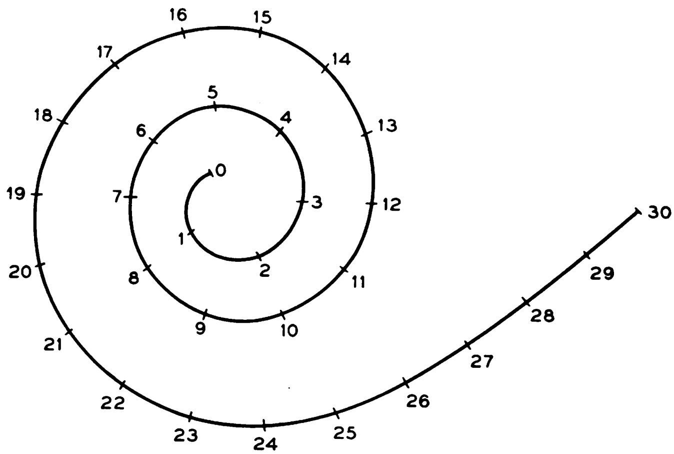

FIGURE 1

proximate distances in millimeters as you go from the helicotrema to the oval window. These dimensions vary in the ears of different persons. Also, the density of the nerve endings, that is, the number of endings per millimeter, is not accurately known. To avoid the necessity of using uncertain anatomical data, the position coördinate x will not be taken in millimeters from the helicotrema but as the per cent of total nerve endings that is passed over in going from the position of maximum stimulation for the lowest audible frequencies (the helicotrema) to a position designated by the coördinate x. The position coördinate thus defined is equal to zero at the helicotrema end and equal to 100 at the oval window end. For example, as will be shown later, the position x of maximum stimulation for a pure tone of frequency 5000 cycles per second is equal approximately to 75. This means that 75 per cent of the nerve endings are in the direction toward the helicotrema and 25 per cent in the direction toward the oval window.

Let us designate the distance in millimeters from the helicotrema as x'. There is then a simple relation between the position coördinate x and the distance in millimeters x' if we know the density σ, that is, the number of nerve endings per millimeter at each distance x'. This relation is given by

$$x = \frac{\int_0^{x'} \sigma \, dx'}{\int_0^{30} \sigma \, dx'}. \tag{3}$$

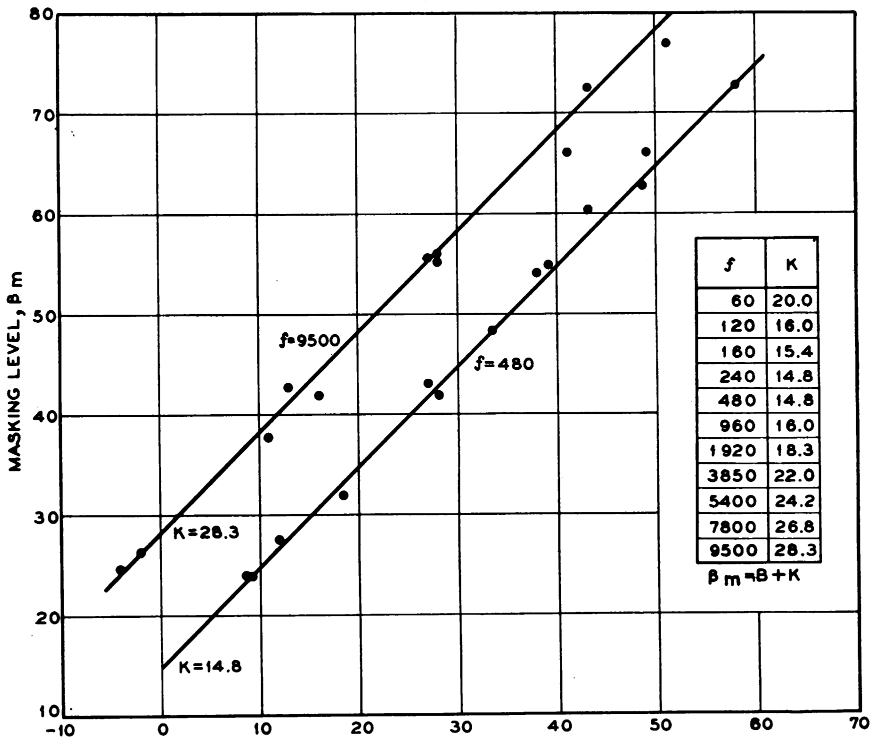

FIGURE 2

In our discussions we will use the coördinate x which can be determined directly from auditory experiments.

When a continuous noise is impressed upon the ear it produces a certain stimulation at the position x. The stimulating agent will be taken as the vibrational power acting upon the nerve endings. Let the amount of this power between positions x and $x + \Delta x$ be designated as ΔJ. Then the power J_x for one per cent of the nerve endings is given by

$$J_x = \frac{\Delta J}{\Delta x}. \tag{4}$$

The stimulation level S at any position x is defined by

$$S_x = 10 \log \frac{J_x}{J_0}, \tag{5}$$

where J_0 is the threshold stimulation. Since here we are dealing with a ratio of vibrational powers, the uncertainty of determining whether to use amplitude velocity or acceleration or any function of them as the stimulating agent does not arise. The value S_x can then be taken as the stimulation of the one per cent patch of nerve endings at the position x. This stimulating power can be spread over a small or a large patch of nerve endings and it will still produce just the threshold value. This is true because our experiments on loudness have shown that the nerve discharges near the threshold are proportional to the intensity of the sound.

Let the following experiments be performed. A broad band of thermal noise whose spectrum level B is known at each frequency is impressed upon the ear of an observer. While this noise is present, a pure tone of known frequency f is raised to the intensity I_m so that the observers just perceive its presence. The intensity level of this tone which is masked will be designated β_m and it is related to I_m by the equation

$$\beta_m = 10 \log I_m. \tag{6}$$

For a fixed frequency a set of values of β_m and B were obtained using different intensity levels of the noise.

The experimental apparatus and methods have been described elsewhere. A typical set of results is shown in figure 2 for two frequencies 9500 cycles per second and 480 cycles per second. The spectrum level B of the noise at each of these frequencies is plotted as abscissae and the intensity level of the pure tone just masked by the noise is plotted as ordinates. It was found that this relation was independent of the frequency width of the band as long as this exceeded a critical value ranging from 100 c. p. s. in the low frequencies to 1000 c. p. s. in the high frequencies. Points above $\beta_m =$ 80 and below $\beta_m = 20$ are usually above the straight line which goes through the points between these levels. There is a reason for both of these departures which cannot be dealt with here. It is the straight portions of the curves which give us the information we want to use in this discussion. Since they are straight lines, equations of the form

$$\beta_m = B + K \tag{7}$$

will represent them, where K is a parameter varying with the frequency regions used. It is this variation that enables us to determine the relation of the position of maximum stimulation x, to the frequency f of the impressed tone. Eleven frequency regions were tested and the results plotted on eleven straight lines like those of figure 2.† The resulting values of K are shown in figure 2 for the various frequencies used. The relation between the acoustic pattern and the corresponding stimulation pattern on the nerves is illustrated in figure 3. A 500-cycle band of noise between 500 and 1000 is spread into a wide stimulation pattern of 16 per cent of the nerve endings, while this same band width between 7500 and 8000 of the same acoustic intensity acts upon only one per cent of the nerve endings at position 88 as indicated. Consequently, one must raise the level of a pure tone in this high frequency region to a higher value to be perceived than if

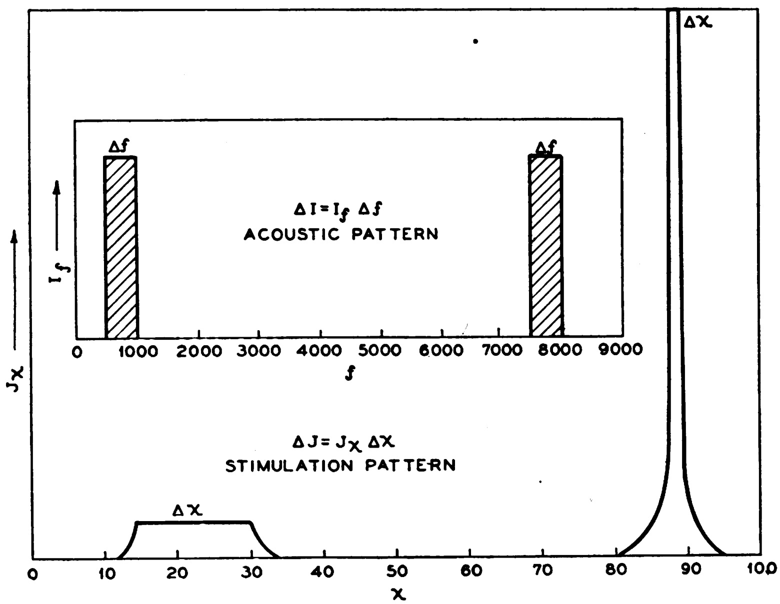

FIGURE 3

one were perceived in the low frequency region where the stimulation per nerve is not nearly so great. Conversely, if the masking level, compared to the spectrum level, is higher for the high frequency region than for the low frequency region, it shows that the frequencies are less spread out in the high and more spread out in the low frequency regions. For example, in the case illustrated the masking tone would be sixteen times greater for one than for the other.

Let us consider this relation in a quantitative way. The intensity ΔI in the thermal noise band between f and $f + \Delta f$ goes to stimulate mostly those nerve endings between x and $x + \Delta x$, where x corresponds to the position of maximum stimulation for a pure tone having a frequency f, and $x + \Delta x$ to the position of maximum stimulation for a pure tone having a frequency of $f + \Delta f$. In other words, the band Δf located at the frequency position f will be chosen so that the power in this particular noise band is spread over the particular patch of nerves Δx located at the position x.

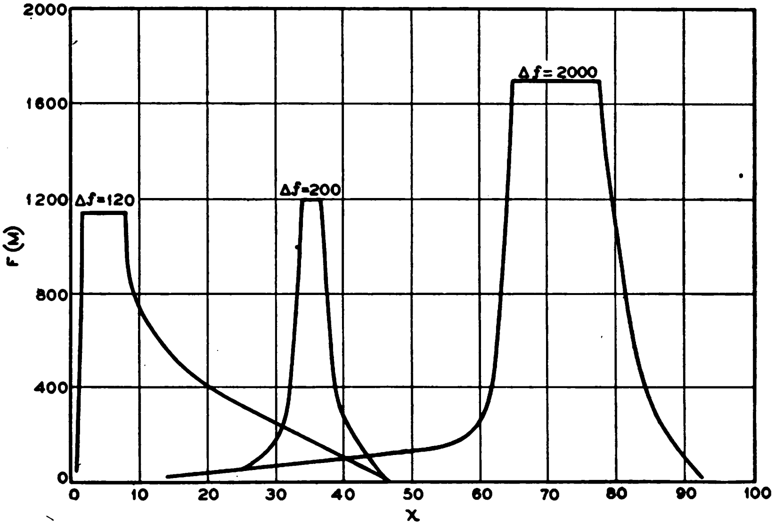

FIGURE 4

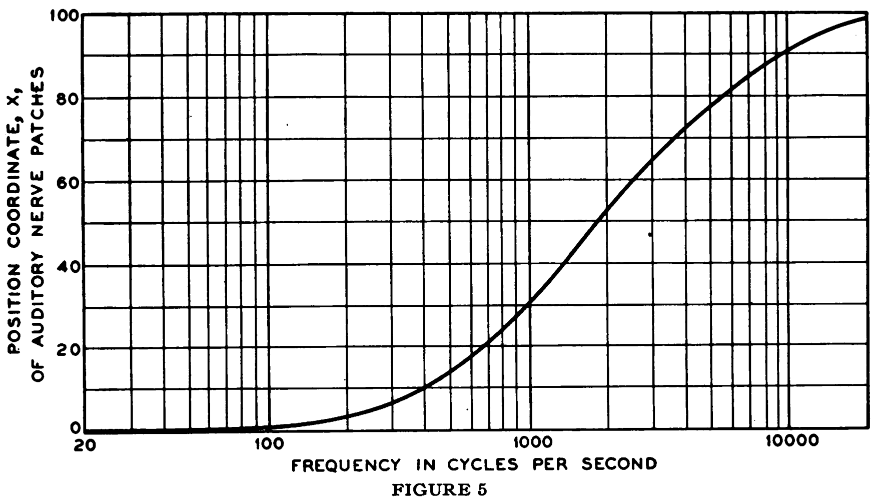

FIGURE 5

As illustrated in figure 3 the patches adjacent to Δx will also be stimulated if a single band Δf is impressed upon the ear. However, for any type of noise whose spectrum is not changing rapidly with frequency these adjacent patches will already be stimulated by the adjacent bands. For this type of noise it is fairly accurate to assume that all the intensity in the band Δf goes to stimulate the nerve patch Δx. The response of the nerves stimulated by small bands of thermal noise has been calculated. The results of this calculation are shown in figure 4 for three different positions. It is seen that the ear is very selective with frequency and that the above assumption is a very good approximation.

Since the thermal noise band and also the masked tone are in the same frequency region, they will have approximately the same frequency discrimination produced by the electro-acoustical system generating the sounds and also by the transmission system in the ear. The acoustical intensity ΔI

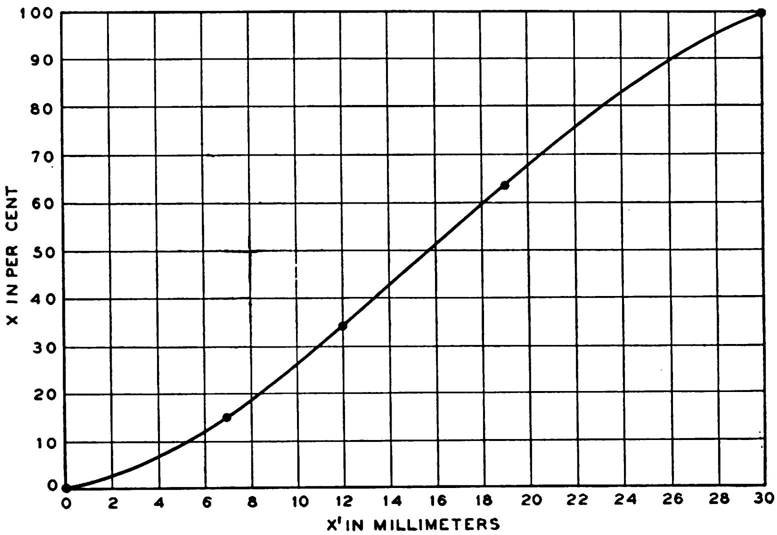

FIGURE 6

produces the stimulation ΔJ. The acoustical intensity I_m produces the stimulation J_m at the same nerve position x as for ΔJ. Therefore in the range where the system is linear

$$\frac{\Delta I}{\Delta J} = \frac{I_m}{J_m} = \frac{I_f \Delta f}{J_x \Delta x}. \tag{8}$$

If we consider Δf and Δx as differentials and integrate we have

$$x = \int_0^f \frac{I_f J_m}{I_m J_x}\, df. \tag{9}$$

It is now assumed that $\frac{J_m}{J_x}$ is a constant C independent of position x and consequently of the frequency f.

$$C = \frac{J_m}{J_x}. \tag{10}$$

Also, equation (9) must fulfil the condition that $x = 100$ when $f = \infty$. If these conditions are applied and the values of B and β_m from equations (2), (6) and (7), then equation (9) reduces to

$$x = C\int_0^f \frac{I_f}{I_m}\,df = C\int_0^f 10^{-K/10}df \tag{11}$$

where

$$C = \frac{100}{\int_0^\infty \frac{I_f}{I_m}\,df} = \frac{100}{\int 10^{-K/10}df}. \tag{12}$$

So we can substitute the values of K from figure 2 and integrate these equations graphically.

In the region where the system is linear

$$\frac{I_m}{I_0} = \frac{J_m}{J_0} \tag{13}$$

where I_0 is the threshold intensity of a band of thermal noise of small width located at the frequency position f corresponding to that of the tone which is masked, and J_0 the stimulation necessary for threshold at the position x corresponding to f. If we multiply the numerator and denominator of the left-hand side of this equation by J_x and then reduce the equation to decibels by taking 10 log of both sides, there results

$$\beta_m - \beta_0 = S + 10\log\frac{J_m}{J_x} = S + 10\log C. \tag{14}$$

The left-hand side of this equation is, by definition, the masking M. This shows that for a linear system the assumption that J_m/J_x is a constant is equivalent to assuming that for a constant masking there is a constant stimulation. In other words, the amount of masking M is a measure of the stimulation S and the difference between these two quantities is simply 10 log of the constant given by equation (10). If we express equation (7) in decibels and then rearrange the terms, there results

$$\beta_m = B + 10\log C + 10\log\frac{df}{dx}. \tag{15}$$

From this equation it is seen that the intensity level β_m of the masked tone can be calculated from the spectrum level B of the noise provided the relation between x and f is known.

Using the values of K shown in figure 3, equations (11) and (12) were integrated graphically. The value for C was found to be 1.1, or within the experimental error equal to unity. This means that when the stimulation J_m due to the tone being masked is equal to the stimulation J_x upon one per cent of the nerves then it will just be perceived by a typical observer. Also, this indicates that the masking M is directly equal to the stimulation S and may be considered a means of measuring it. The resulting values of f and x obtained from equations (11) and (12) are shown in the curve of figure 5. The relationship between f and x indicates that frequencies below about 150 cycles per second all have the position of maximum stimulation on the first one per cent patch of nerve endings. When the frequency of the stimulating tone reaches 1000 cycles per second the position of maxi-

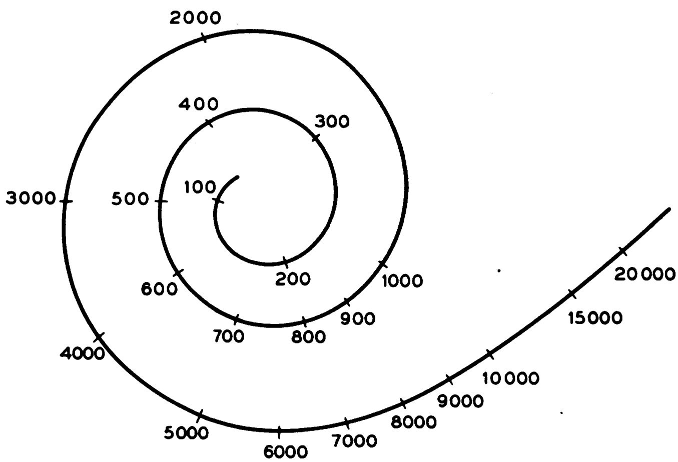

FIGURE 7

mum stimulation has passed over approximately 30 per cent of the nerve endings.

It will be remembered that the values of x refer to nerve patches in such a way that for equal intervals Δx there are an equal number of nerve endings. If the nerve endings are distributed uniformly along the basilar membrane then the values of x used in this discussion will correspond to one per cent of the total length of the basilar membrane. If, however, the nerve endings are not distributed uniformly, then the length dx must be stretched or contracted depending upon the nerve density being less or greater than the average. Let x' be the distance from the helicotrema in millimeters.

Guild‡ has made measurements of the number of ganglion cells along the basilar membrane. He divided the cochlea into four parts, namely (1) upper middle and apical, (2) lower middle, (3) upper basal, (4) lower basal.

He made a count of the total number of ganglion cells in each of these parts. They are given below in column 3 of table 1. Steinberg has estimated the distance in millimeters along the basilar membrane for each of these parts to be that given in the column under x'. The corresponding values of x are then easily calculated as the per cent of the total cells up to the position corresponding to x'. They are given in the last column.

TABLE 1

POSITIONS	x'	COUNT	x
Upper Middle plus apical	0–7	4,200	0–14.5
Lower Middle	7–12	5,600	14.5–33.8
Upper Basal	12–19	8,500	33.8–63.1
Lower Basal	19–30	10,700	63.1–100
Total	0–30	29,000	0–100

It is seen that these data on the number of ganglion cells allow us to calculate only three points besides the two end-points. These values of x and x' are plotted on figure 6. A smooth curve passed through these points. Using this relation, then the values of x and f from figure 5 can be transferred to values of x' and f. In figure 7 this relation is plotted on a spiral having the dimensions shown in table 1. The positions given in this spiral agree with previous determinations within the accuracy of those determinations.

* All intensity values used in this paper are expressed by this same unit.

† I am indebted to Mr. W. A. Munson for doing this experimental work.

‡ Guild, Stacy R., *Acta Oto-Laryng.*, **17**, 207–245 (1932).

THE DIMENSIONLESS CONSTANTS OF PHYSICS

By ARTHUR E. HAAS

UNIVERSITY OF NOTRE DAME

Communicated June 8, 1938

Physics appears to be characterized by several universal constants which are pure numbers. One is the ratio of the mass of the proton to the mass of the electron, i.e., $m_p/m = 1835$, another the ratio $e^2/(h\,c)$, where e is the fundamental charge, h Planck's constant and c the velocity of light. A third is the ratio $e^2/(G\,m_p\,m)$, where G represents the gravitational constant. An important fourth number, under the assumption of a finite uni-