A DIRECT COMPARISON OF THE LOUDNESS OF PURE TONES

By B. A. Kingsbury

Abstract

The loudness of pure tones of frequencies 60 to 4000 cycles and for intensities from the threshold to 90 T.U. above the 700 cycle threshold.—The loudness of eleven pure tones was studied by adjusting the voltage applied to a telephone receiver to make these tones as loud as certain fixed levels of a 700 cycle tone. The average results of 22 observers, 11 men and 11 women, were arranged as contour lines of equal loudness through the normal auditory sensation area in terms of r.m.s. pressure in ear canal as a function of frequency. Frequencies from 60 to 4000 cycles were used and intensities from threshold of audibility to 90 T.U. above the 700 cycle threshold. It was found that if the amplitudes of pure tones are increased in equal ratios the loudness of low frequency tones increases much more rapidly than that of high frequency tones. For frequencies above 700 cycles the rate is nearly uniform.

A loudness unit.—As a loudness unit the least perceptible increment of loudness of a 1000 cycle tone was employed. In absolute magnitude this varies from level to level, but in the ordinary range of loudness it becomes constant. This unit takes into account the subjective character of loudness.

Sources of variation in data on loudness.—The variability of the data from which the averages were computed was separated into a factor expressing dissimilarity of ears and another expressing errors of observers' judgment. There was no level at which the variances were a minimum. Dissimilarity of ears causes more variation than errors of observers' judgment. The variances showed no significant sex difference.

THE purpose of this paper is to give data on the relative loudness of pure tones as judged by a group of eleven men and eleven women whose average threshold of audibility approximated former measurements. Since most ears probably can perceive nearly the same minimum amounts of sound energy and recognize about equal changes in pitch and intensity, measurements made by the above group should apply to the hypothetical average normal ear.

In Fig. 1 are shown the usual curves for the threshold of feeling and audition expressed in 20 times the common logarithm of the r.m.s. pressure on the ear drum (dynes per square centimeter) as a function of frequency. The pressure is that which would be produced in the ear canal if the walls and drums were rigid. The difference between any two ordinates is then the ratio of the pressures in transmission units (T.U.). This is the common unit for expressing amplitude ratios in telephony and at the same time takes into account the logarithmic relation between sensation and stimulus. The lower of the curves is the average threshold of audibility for the above group.

Let us imagine that the r.m.s. pressure of a test tone at the threshold of audibility is increased until it comes up to a level P. If one listens first to this tone and then to a tone of frequency f_1, he can tell which appears the louder and by adjusting the level of the tone f_1, he can make the two appear equally loud. By comparing several tones with a test tone in this way a contour line of equal loudness through the region of audition may be determined. For instance, P_1, P_2, P_3, etc. are points which were experimentally determined for one of these contours. In the work to be described 700 cycles was made the test tone, while the frequencies compared with it were distributed at approximately equal logarithmic intervals through the more important range of audition. This seemed the best way to secure results applicable in general to pure tones with measurements at a minimum number of frequencies.

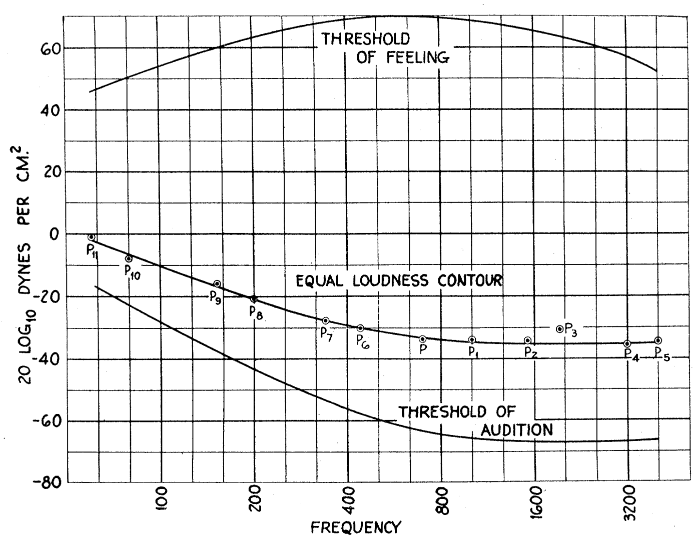

Fig. 1. Normal auditory sensation area.

In order to cover most of the intensities of ordinary sounds contours of equal loudness 10 T. U. apart at 700 cycles were selected up to levels somewhat above that of ordinary conversation. The 10 T. U. interval means that the amplitude of the sound wave for the 700 cycle tone at any contour was 3.16 times its value at the contour below this. There is no necessity for smaller intervals because a few fairly widely spaced contours should show quite definitely how the loudness of pure tones varies with level and frequency.

A schematic diagram of the apparatus used in the tests is shown in Fig. 2. It will be noted that the set-up consists of two symmetrical systems, which we have labelled A and B. The test tone was always furnished by the A oscillator. The filters in each case attenuated the first harmonic of the tone used about 75 transmission units. The thermocouples and galvanometers shown were used to measure the input currents of the attenuators. The resistance network attenuators which are calibrated to read the attenuation in transmission units could be readily connected by a switch to a special iron-clamped high impedance receiver.

The outline of the procedure in making tone balances was as follows: The A attenuator was placed in such a position that the observer could not see the dial; the dial of the B attenuator was covered with a cardboard screen, which prevented the observer from seeing the position of the pointer while making the adjustment. The 700 cycle test tone was set up on the A system and the frequency being compared with this on the B system. First three independent settings of the A threshold were made, then three for the B threshold. Next the experimenter set the A attenuator at one of the selected comparison levels and allowed the observer to adjust the B attenuator until when listening alternately to the two tones they seemed equally loud. The attenuator settings and the deflections of the meters

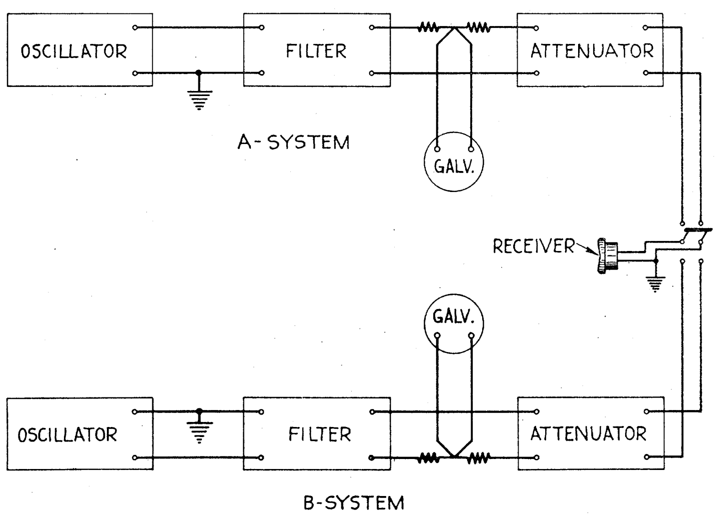

Fig. 2. Schematic diagram of apparatus.

were recorded for both systems. This process was repeated for the other fixed levels of the A tone, until three independent determinations had been made for each level. The order of taking the fixed levels of the A tone was made as random as possible and two successive determinations were seldom made at the same level. When the comparison of the two frequencies was finished, the A and B thresholds were once more secured.

From a calibration and check of the receiver and the input attenuator voltage, attenuator settings were reduced to the r.m.s. pressure in the ear canal in T.U. above or below one dyne/cm^2. It was found that r.m.s. pressures at the average *threshold of audibility* for this group coincided very closely with the values which represent the best estimates[1] for normal ears. Hence this threshold curve is taken as a datum from which all the other points on the diagram are to be reckoned.

[1] H. Fletcher, Useful Numerical Constants of Speech and Hearing, The Bell System Technical Journal, Vol. IV, No. 3, pp. 375, July, 1925.

In Fig. 1 were given the curves for the thresholds of feeling and audition. In Fig. 3, the data from loudness comparisons are plotted in a similar fashion

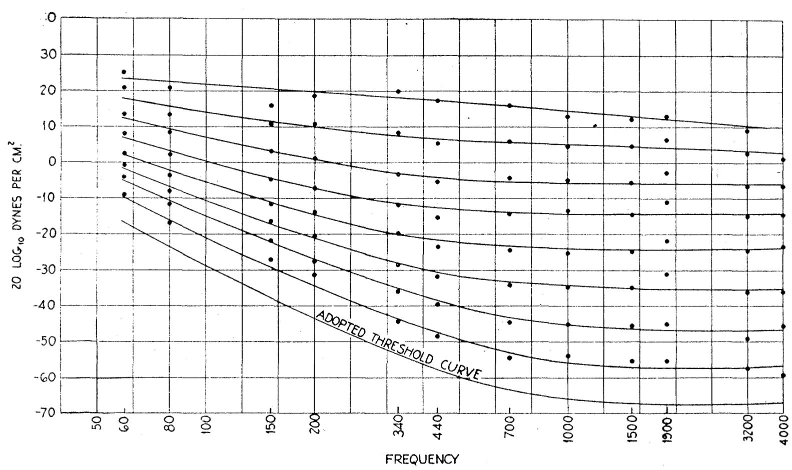

Fig. 3. Contour lines of equal loudness.

to demonstrate contour lines of equal loudness, the lowest curve being the threshold curve of audition.

These data may also be stated very conveniently in terms of *sensation level*, defined by the equation

$$S = 20 \log_{10}(P/P_0)$$

where P is the r.m.s. pressure of the sound wave and P_0 the minimum audible pressure for the average normal ear. That is, to get the sensation

TABLE I

SENSATION LEVEL (T. U.)

Each column represents 20 $\log_{10} P/P_0$ of equally loud tones
Threshold of audibility adopted for these data (P_0 in dynes/cm²).

Frequency:	60	80	150	200	340	440	700	1000	1500	1900	3200	4000
$20 \log_{10} P_0$:	−16.5	−23.0	−37.0	−43.6	−53.5	−57.5	−63.4	−65.5	−67.0	−67.4	−67.6	−67.5

Frequency	Thres	Sensation level ($20 \log_0 P/P_0$)								
60	0	8.4	12.5	15.7	19.2	24.8	29.9	37.3	41.9	
80	0	6.6	11.4	14.9	19.6	25.1	31.7	36.1	43.9	
150	0	9.9	15.2	20.8	25.6	32.5	39.7	47.8	52.9	
200	0	12.7	16.4	23.2	29.9	36.5	44.8	54.4	62.2	
340	0	9.4	18.0	25.1	33.8	41.7	50.7	61.6	73.5	
440	0	9.3	17.7	25.7	33.8	42.2	52.5	62.9	75.3	83.4
700	0	9.3	19.3	29.3	39.3	49.3	59.3	69.3	79.3	89.3
1000	0	11.7	20.7	31.1	40.7	52.1	60.9	69.8	78.5	86.8
1500	0	11.6†	21.4	32.5	42.5	52.5	61.7	71.3	79.2	85.0
1900	0	11.9†	22.4	36.3	45.7	56.7	65.0	73.7	80.2‡	
3200	0	9.9	18.5	31.6	43.1	52.8	61.1	70.0	77.0	
4000	0	8.0	22.1	31.6	44.0	53.2	61.1	68.9		
Loudness*	0	8.5	20.3	39.0	60.0	83.5	107.0	130.0	151.5	172.0

* See 1000 cycle curve, Fig. 7.
† Average of 57 observations.
‡ Average of 63 observations.

level of any point in Fig. 3, it is merely necessary to notice how many T.U. the amplitude of the tone must be reduced to reach the threshold of audibility. In Table I, the data of Fig. 3 are expressed in terms of sensation level.

It has generally been assumed that loudness is a function both of the amplitude and frequency of the pressure wave, but the form of the function is not definitely known. The arbitrary levels of the 700 cycle tone had been chosen 10 T.U. apart, and loudness contours determined on this basis. At the level where the contours are closest together, the change in loudness for the same number of T.U. is greatest. The average interval between the contours was computed at each level, by adding together the intervals for each frequency and dividing by the number of frequencies. The first interval from the threshold was omitted since it was an irregular interval. The average interval between contours for each frequency was also computed. In Fig. 4 these data are plotted first with interval as a function of frequency and

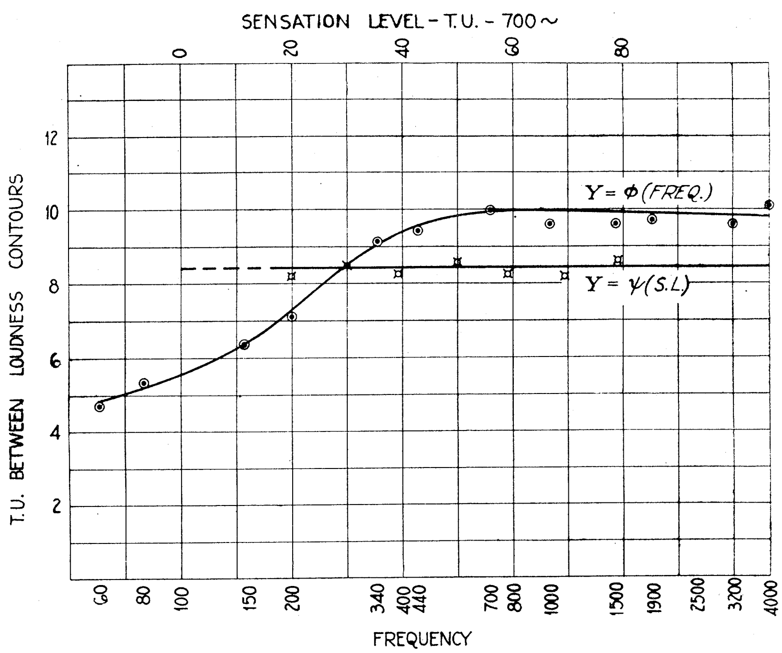

Fig. 4. Effect of sensation level and frequency upon sensation level increment between contours.

then with interval as a function of the sensation level of the 700 cycle tone. From these two curves it would appear that the interval between contours is about the same number of T.U. for all sensation levels, but that the loudness of low frequency tones increases much more rapidly than that of high frequency tones.

It seems desirable now to turn to another aspect of these results which should occupy a place of major interest in all hearing studies. This is the variation in observations which arises from the variability of an observer's judgment and the differences between ears. The usual procedure in reducing data has been to take the arithmetic average of a number of observations and compute the standard deviation. In this case, because of the seemingly large importance of this variation, it appeared desirable to give it more careful treatment.

A particular value in Table I is the average of 66 observations, three of which were made by each of twenty-two observers. The deviation between any determination of an observer and his average can be readily calculated. The sum of the squares of these 66 deviations divided by 44 yields a quantity β which expresses the variance in observers' judgment. If X is an individual observation and $\overline{X}_p$ the average for that person this may be written: (Appendix I)

$$\Sigma(X-\overline{X}_p)^2/44=\beta$$

These variances, one for each value in Table I are included in Table II.

TABLE II

Variances of data averaged for values of Table I in (T. U.)²

20 $\log_{10} P_{700}$: Frequency		−63.4 Thres$_A$	−63.4 Thres$_B$	−54.1	−44.1	−34.1	−24.1	−14.1	−4.1	+5.9	15.9
60	α:	29.9	54.3	35.4	34.9	33.4	34.2	50.7	67.5	59.1	27.0
	β:	3.9	4.8	5.3	6.5	11.0	5.5	12.6	8.6	12.1	9.9
80	α:	23.4	24.8	20.1	26.5	21.4	28.3	37.4	36.8	21.5	17.8
	β:	4.1	8.3	4.8	8.5	8.8	17.6	14.2	13.4	15.6	10.1
150	α:	30.1	56.9	34.2	34.0	34.0	39.7	47.8	71.8	109.9	80.4
	β:	4.3	16.8	8.7	8.7	9.5	10.2	8.3	9.8	12.1	10.1
200	α:	32.6	36.7	11.4	40.7	27.0	33.7	34.8	64.5	67.4	26.2
	β:	3.3	8.0	8.3	7.3	11.7	12.6	13.0	18.3	11.1	14.1
340	α:	30.3	34.5	22.2	13.7	24.0	29.9	42.4	52.1	64.5	24.4
	β:	3.4	6.7	15.8	8.9	9.3	16.4	17.1	25.7	19.4	18.1
440	α:	31.7	54.3	21.8	22.6	31.0	38.9	43.8	48.9	41.5	31.3
	β:	3.3	9.5	11.0	10.5	20.1	7.9	12.6	14.7	14.5	8.5
700	α:*					1.1	0.2	2.0	1.1	.5	.8
	β:					2.2	1.4	1.6	3.	3.6	1.4
1000	α:	30.3	42.9		27.5	21.2	22.2	15.7	13.2	16.4	17.9
	β:	2.9	3.2		18.6	22.1	28.7	13.1	10.2	11.5	7.5
1500	α:	49.1	44.8		59.3	38.9	55.2	48.0	46.8	34.3	31.5
	β:	3.2	2.8		14.4	26.2	17.7	16.6	13.0	4.7	4.2
1900	α:	30.6	70.0		97.4	48.1	56.1	53.8	43.0	36.6	
	β:	2.9	7.0		11.4	27.5	28.0	16.8	11.6	10.8	
3200	α:	29.2	60.5		81.4	83.6	28.7	50.0	46.5	37.5	34.7
	β:	3.7	3.3		36.3	22.3	29.4	33.6	16.5	11.3	12.1
4000	α:	18.7	100.8		96.0	71.9	73.1	44.3	51.6	25.4	
	β:	2.5	5.2		26.8	27.4	31.2	24.4	12.5	8.9	

* From the observations of 20 observers comparing 700 cycles with 700 cycles.

In all this work, measurements were made to the nearest T.U. and arithmetic averages and variances computed to the nearest 0.1 T.U.

The other factor which contributes to variation is the variance between ears. Let us call this variance α. A particular value in Table I is the arithmetic average of 22 averages, each of which is the arithmetic average of an observer's three settings. Let the general average be $\overline{X}$. Then we can make the following equation define α and β: (Appendix I)

$$3\Sigma(\overline{X}_p-\overline{X})^2/21=3\alpha+\beta$$

Since the value of β is known, α can be computed. The values are also given in Table II.

The standard deviation of the observations whose average is an individual entry in Table I is defined by the relation: (Appendix I)

$$S.D.=(0.97\alpha+\beta)^{1/2}$$

If Z equals the ratio of the error of the average to the observed standard deviation, it is a function only of the number of observations. Choosing

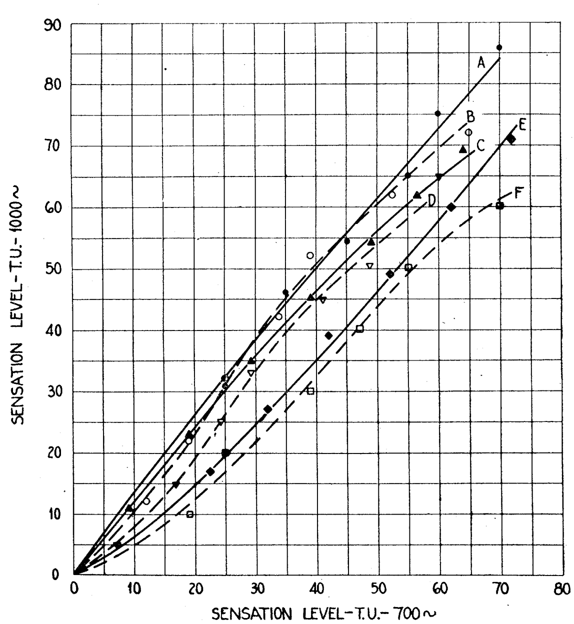

Fig. 5. Comparison of loudness balance methods.

the 95 percent error, a value in Table I will vary more than Z times its standard deviation only five times in a hundred, and another average outside of these limits would very likely be from another population of measurements. Z has in our case a value[2] of 0.25, and an entry in Table I is therefore determined to 1.5 or 2 T.U.

[2] W. A. Shewhart, Correction of data for errors of averages obtained from small samples, Bell System Technical Journal, **V**, No. 2, pp. 314, April, 1926.

All the experimental work up to this point had been carried on by making the levels of the A tone fixed and allowing the observer to adjust the B tone. In Fig. 5 are shown three pairs of representative comparisons, where each pair of curves A and B, C and D, E and F were made by the same individual. Every point shown is the average of 3 observations. A, C and E were made by fixing the level of the A tone as already described. For B, D, and F, the experimenter set the B tone and the observer adjusted the A tone for equal loudness. The results check rather closely; in fact two successive comparisons by the same method would not be more alike. The two comparisons were made about three months apart. It therefore seems reasonable in loudness comparisons to set the test tone at arbitrary levels.

In order to show conclusively that all points on a given contour represent equally loud tones, an additional simple check can be applied. If two B tones are compared directly will the result be the same as when each is compared with the A tone? To test this a direct comparison of 200 cycles and 3200 cycles was made by eleven male observers using 200 cycles as the A tone with fixed levels. Plotted in terms of 20 $\log_{10}$ (volts on receiver) the results of the two methods are shown in Fig. 6. The direct comparison is

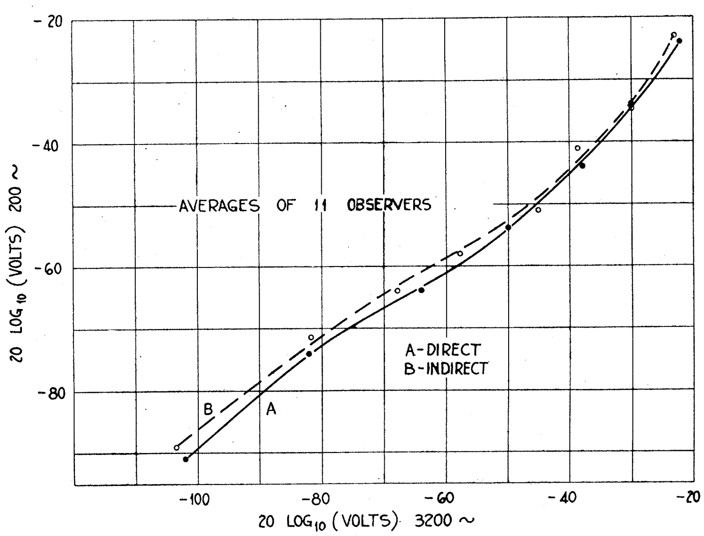

Fig. 6. Direct and indirect loudness comparisons.

the more accurate, since in the indirect comparison, variances are involved twice. In view of the values of α and β already given, the two curves do not appear significantly different.

Discussion of the Nature of Loudness

Since loudness is a subjective matter, it is rather difficult to determine what should be made the unit of loudness.[3] Of course, any arbitrary scale

[3] J. C. Steinberg, The Loudness of a Sound and Its Physical Stimulus, Phys. Rev., Second Series, Vol. 26, No. 4, pp. 507–523, October, 1925.

of loudness may be used provided that sounds which are equally loud have equal numbers on the scale and that loudness numbers vary progressively as we go from fainter to louder tones, but a system which recognizes some of the characteristics of loudness arising from the physical characteristics of audition would be preferable. Fig. 4 seems to show that in the ordinary range of hearing loudness is linear with sensation level if the Weber-Fechner Law holds. In light the unit of brightness has been made the least perceptible increment of brightness. Knudsen[4] has data for several pure tones which show the relation between the Fechner ratio $\Delta E/E$ and the r.m.s. pressure in dynes per cm^2 in the ear canal. This ratio, "intensity sensibility of the ear" expresses the ratio of the smallest perceptible difference in the energy of a tone to its total energy. At 1000 cycles from these data, we were able to plot a curve in which the Fechner ratio was a function of the 1000 cycle sensation level. This curve was then extrapolated to the threshold on the assumption that there the ratio was unity.[5] By making this ratio for the

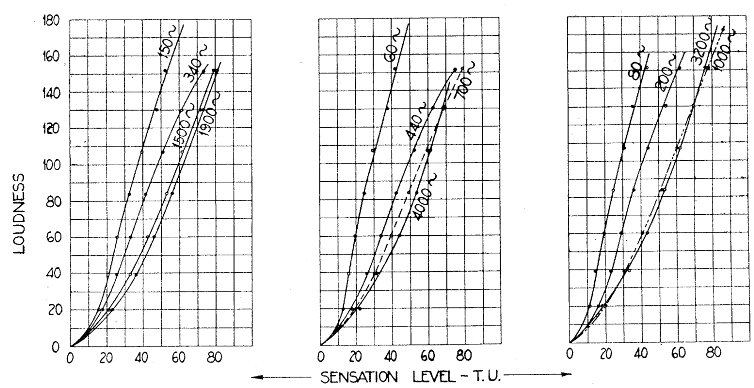

Fig. 7. Loudness of single frequency tones at various sensation levels.

1000 cycle tone our unit of loudness and calling the threshold of audibility zero loudness, it was possible to integrate from the above curve, a curve with the loudness of the 1000 cycle tone as ordinate and its sensation level as abscissa. That is, a 1000 cycle note just perceptibly above the 1000 cycle threshold has loudness unity; if it was just perceptibly louder than this its loudness is two and so on. This curve is marked 1000 cycles in Fig. 7.

The next step was to apply this unit to the data of Table I. Loudness numbers corresponding to the 1000 cycle sensation level given in this table were read from the 1000 cycle loudness curve. Each column in the table represented sounds which were equally loud in the ordinary sense, and the "loudness numbers" read from the 1000 cycle curve applied by the above definition to the other frequencies as well. Loudness for each frequency is thus numerically defined at several sensation levels. The remaining curves

[4] V. O. Knudsen, Sensibility of the Ear to Small Differences of Intensity and Frequency, PHYS. REV., Second Series, Vol. 21, pp. 84–103, January, 1923.

[5] P. G. Nutting, The Complete Form of Fechner's Law, Bulletin of the Bureau of Standards, Vol. 3, No 1, 1907.

of Fig. 7 were drawn up on this basis. A unit of loudness somewhat better for our purpose would have been the least perceptible increment of loudness of a 700 cycle tone compared with its sensation level, but data for such a unit are not available and there is reason to suspect that a curve of this kind between loudness and sensation level is much the same for 700 and 1000 cycles. Moreover, the unit used has been given a very definite quantitative value which makes it possible to compare it with any other unit. (Appendix II)

The Variability of the Data

An inspection of Table II fails to show any level of loudness where the similarity of ears is a maximum for all frequencies. In general the variances are smaller when the tones being compared are more nearly of the same frequency. Since α is usually greater than β, errors of judgment cause less variation than dissimilarity of ears. The differences in the values of α and β for the 700 cycle threshold are not significant. The statistical development on which these quantities rest makes them entirely independent quantities. If β is averaged for level, that is if the values in each column are added and the sum divided by the number in the column, it will be found that β reaches a maximum at about 40 T.U. above the 700 cycle threshold. The variance of judgment also varies with frequency being progressively smaller as we go from higher to lower frequencies. This makes it seem that the sensibility of the ear must be greater at low frequencies, which is equivalent to saying that the loudness sensation level curves will be steeper there.

The variability of ears does not seem to be correlated with frequency although it is somewhat less at the 1000 cycle comparison than elsewhere. The average α of equally loud tones below 700 cycles seems gradually to increase as loudness increases, while the average for equally loud tones above 700 cycles gradually decreases. In other words, at low frequency comparisons, ears are more alike for rather faint tones, while in the comparison of higher frequencies they are more alike at louder levels. This variability of ears is more nearly the same for all frequencies at about 40 T.U. above the 700 cycle threshold, but the error of judgment is here a maximum. It is therefore doubtful whether there is any best level for loudness comparisons. Practice seems to lower the 700 cycle threshold since it becomes lower in the same order as the balances were made. However, it appears probable that this effect is connected with other factors.

At 1900 cycles separate variances were computed from the observations made by the men and the women but no definite trend was found which would suggest sex differences in hearing.

Conclusion

When the amplitudes of single frequency tones are increased by equal ratios, high frequency tones increase in loudness more slowly than do low frequency tones. However, for frequencies above 700 cycles, the idea that tones are equally loud when they are an equal number of T.U. above the threshold is a very good approximation.

The study of the goodness of fit of regression lines which relate the sensation levels of equally loud tones may be made the subject of a later paper. An interpretation of such lines would be very much facilitated by a study of sensibility of the ear to intensity changes of pure tones because it is logical to suppose that in terms of this discrimination loudness matches are made.

Appendix I

In order to make the statistical analysis on which this paper rests easily available to the reader, a brief résumé[6] of the method is included here.

If in a group of measurements two factors cause variation an analysis of variance offers a good method of treatment. Let X be an individual measurement, $\overline{X}_p$ average of each class (determinations of each individual), $\overline{X}$ arithmetic average of all X's, k the number in each class and N' number of classes (observers). It can be shown that the following equation is true:

$$\sum_1^{kN'} (X-\overline{X})^2 = k\sum_1^{N'}(\overline{X}_p-\overline{X})^2 + \sum_1^{kN'}(X-\overline{X}_p)^2 \quad (1)$$

Let β represent the variance within classes. Here there are N' $(k-1)$ degrees of freedom, and we may write:

$$\sum_1^{kN'} (X-\overline{X}_p)^2 = N'(k-1)\beta \quad (2)$$

As twenty-two observers each making three observations were employed in determining the average values of Table I, we have, in this case, to put $k=3 \cdot \cdot \cdot N'=22$.

The mean of the observations in each class is affected by the variance β divided by k as there are k observations in each class, and also by a variance α related to the difference between classes. As there are here $(N'-1)$ degrees of freedom, we write:

$$\sum_1^{N'} (\overline{X}_p-\overline{X})^2 = (N'-1)(\alpha+\beta/k) \quad (3)$$

where again for the special case of this paper $N'=22$ and $k=3$.

From (2) and (3) the value of α and β may be readily secured. It is only necessary to compute two of the three terms in Eq. (1). The usual procedure is to solve Eq. (1) for the summation on the left of Eq. (2) after evaluating the other two summations.

The standard deviation of our measurements from the general mean can be easily calculated from the above equations. The standard deviation is defined as:

$$(SD)^2 = \sum_1^{kN'} (X-\overline{X})^2/(kN'-1) \quad (4)$$

[6] R. A. Fisher, Statistical Methods for Research Workers, 1925.

By the use of (1), (2) and (3) this reduces to:

$$(SD)^2 = [1-(k-1)/(kN'-1)]\alpha+\beta$$
$$= .97\alpha+\beta \qquad (N'=22,\ k=3) \tag{5}$$

We define Z by the relation

$$Z = (\overline{X}-m)/(SD) \tag{6}$$

where m is the true mean of all observations.[2] Z then is a function of the number of observations made and its value for any particular error may be determined from statistical tables. When this has been done, $(\overline{X}-m)$ can be computed.

Appendix II

The primary effort in this paper has been to discuss the loudness of pure tones in a simple experimental fashion without much regard to formulas of loudness. The least perceptible increment in the energy of a 1000 cycle tone was made the unit of loudness in the consideration of the pure tone

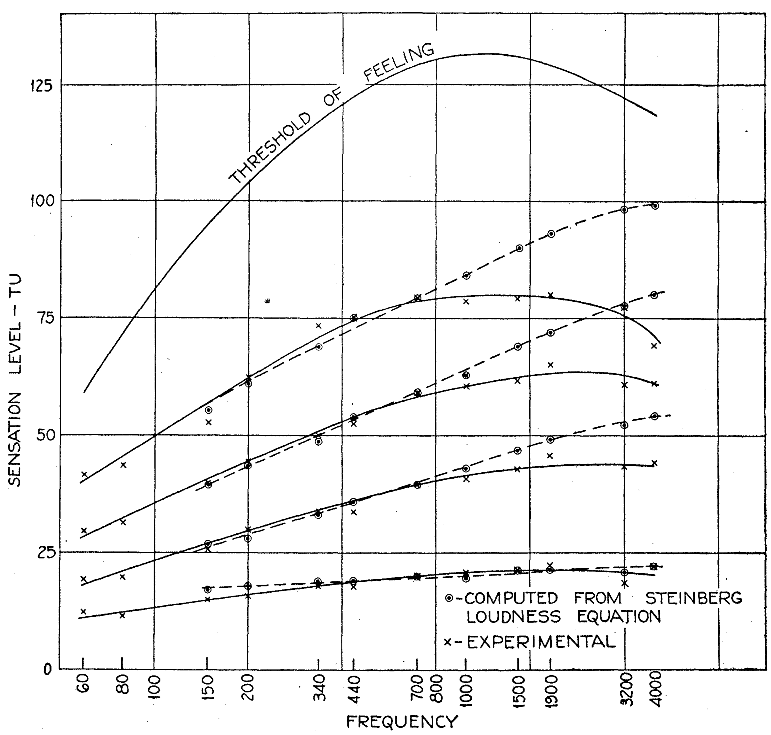

Fig. 8. Sensation levels of equally loud tones.

data because of its simplicity and direct application. However, this procedure did not justify the neglect of other methods of treating the data. Dr. J. C. Steinberg has developed a formula[3] for loudness with the primary purpose of computing the loudness of complex sounds when the frequency spectrum is known. This method involves essentially the summation of the loudness of the various frequency components. It can be applied readily to the loudness of pure tones. The formula employs two factors, first a weight factor which depends on the sensation level of the sound and its frequency, secondly a root factor which depends only on the sensation level of the sound. Using the values of the weight and root factors given by Dr. Steinberg, the sensation levels of tones of the frequencies studied as loud as 700 cycles at 19.3, 39.3, 59.3, and 79.3 T.U. respectively above its threshold value were determined. The results of these computations and the experimental results of this study are shown in Fig. 8 as curves for the sensation levels of equally loud tones. Computations were not made at 60 and 80 cycles because these weight factors were not stated. The procedure employed makes the computed and experimental values for the sensation levels of the 700 cycle tone coincide. For the lower frequency tones the computed values fit the experimental data rather well. If the values of the root factor are considered known and independent of frequency, weight factors may be computed on the basis of the pure tone data in this report.

The justification of this formula as well as any other which attempts to express loudness, must be substantiated both by the study of pure tone loudness and the Fechner ratio. The accuracy with which weight factors calculated from these pure tone data fit measurements on complex sounds has not yet been determined.

Bell Telephone Laboratories, Incorporated
December 1, 1926.