A re-determination of the equal-loudness relations for pure tones

1956 Br. J. Appl. Phys. 7 166

(http://iopscience.iop.org/0508-3443/7/5/302)

(7) O'Neill, H. *The Hardness of Metals and its Measurement* (London: Chapman and Hall Ltd., 1934).

(8) Mohs, F. *Grundriss der Mineralogie* (1824). English translation by W. Haidinger, *Treatise on Mineralogy* (Edinburgh: Constable and Co., Ltd., 1825).

(9) Bridgman, P. W. *Studies in Large Plastic Flow and Fracture* (London: McGraw-Hill Publishing Co., Ltd., 1952).

(10) King, R. F., and Tabor, D. *Proc. Roy. Soc. A*, **223**, p. 225 (1954).

(11) Tabor, D. *Proc. Phys. Soc.* [*London*] *B*, **67**, p. 249 (1954).

(12) Taylor, E. W. *Miner. Mag.*, **28**, p. 718 (1949).

(13) Khrushchov, M. M. *Zavod. Labor.*, **15**, p. 243 (1949).

(14) Winchell, H. *Amer. Min.*, **30**, p. 583 (1945).

(15) Knoop, F., Peters, C. G., and Emerson, W. B. *J. Res. Nat. Bur. Stand.*, **23**, p. 39 (1939).

ORIGINAL CONTRIBUTIONS

A re-determination of the equal-loudness relations for pure tones

By D. W. Robinson, B.Sc., A.C.G.I., A.M.I.E.E., and R. S. Dadson, M.A., National Physical Laboratory, Teddington, Middlesex

[*Paper received* 17 *January*, 1956]

The paper describes a new determination of the equal-loudness relations for pure tones in free-field conditions which has been carried out at the National Physical Laboratory as a result of requests from organizations interested in various aspects of the acoustics of hearing. The equal-loudness contours are of considerable importance in this field, being fundamental to a proper understanding of aural judgments of the loudness of sounds of all kinds. They are also concerned in numerous practical applications in the study of noise.

The first set of contours for free-field conditions was given by Fletcher and Munson in 1933, and a second determination was carried out by Churcher and King in 1937, but these two investigations showed considerable discrepancies over parts of the auditory diagram. The present work has been carried out on a more extensive scale, using a large team of otologically normal persons, and new techniques have been introduced enabling reliable measurements to be made over a wider range of intensity than has hitherto been possible.

The new results cover a range of frequency of from 25 to 15 000 c/s and of sound pressure level up to about 130 dB relative to 0·0002 dyn/cm^2. The data show a greater degree of regularity than the former results, and allow the equivalent loudness of a pure tone of any frequency to be expressed by formulae quadratic in the sound pressure level, the coefficients varying smoothly with frequency. The results include a new determination of the normal threshold of hearing in free field, which is highly consistent with the equal-loudness contours. At frequencies above 1000 c/s account needs to be taken of variations due to the age of the observers, which become of particular importance at the upper end of the frequency range.

The new results have indicated the causes of some of the discordant features in the earlier determinations and it is hoped that the work will facilitate agreement on a standard set of equal-loudness contours. On account of its relevance to noise measurement, some extension to equal-loudness relations for bands of noise is being undertaken.

In the study of many aspects of subjective acoustics, and especially in the study of noise, the concept of the "loudness" of a sound is of fundamental importance. It is, of course, clear that loudness is not the only subjective attribute of a complex sound which may be of concern, though it is frequently the primary factor. Experience has shown that equal-loudness judgments are considerably less subject to extraneous psychological factors than are the less tangible qualities such as, for example, "annoyance." Moreover there is a fairly direct correspondence between loudness and the physical properties of the stimulus, which provides a firm basis for a consistent scheme of measurement.

Sounds may be analysed in various ways, but for most purposes analysis into a frequency spectrum, either of individual components or a distribution in terms of frequency bands, is the most convenient. It thus comes about that the loudness relations for pure tones of different frequencies, and as an extension of this the corresponding relations for distributed sound in particular frequency bands, are not only of great interest in themselves but are of basic importance in many applications. Such diverse problems as the study of the behaviour of the ear to complex sounds, and the procedure by which it arrives at a characteristic loudness sensation, the selection of frequency weighting networks for objective noise meters, and the development of methods for computing the loudness levels of complex sounds from objectively measured spectrum analyses, all rest at some stage or other on a knowledge of the equal-loudness relations for pure tones or other simple sounds. The relations are also involved in many aspects of the communication and reproduction of speech and other sounds.

The earliest measurements of the equal-loudness relations for pure tones were those of Kingsbury,[1] but as these were restricted to telephone listening conditions they will not be considered further here. The present paper describes a new determination of the equal-loudness relations for pure tones which has recently been carried out at the National Physical Laboratory. The investigation relates to free-field conditions of listening, under which conditions there have been two previous determinations, namely that of Fletcher and Munson in 1933[2] and Churcher and King in 1937.[3] It has long been apparent that it would be most valuable to have a single set of equal-loudness contours acceptable to all workers in the field. The two previous determinations, however, showed considerable discrepancies over parts of the auditory diagram of such a nature and magnitude that it would have been difficult to derive from them a standard set of contours. It was largely as a result of discussions on this situation at the Committee on Noise Measurement of the British Standards Institution that it was decided that a new investigation should be carried out at the National Physical Laboratory. This had the principal aims of resolving the discrepancies in the previous work and of providing a new and more comprehensive set of data, as a contribution to future standardization. The number of subjects tested is considerably greater than hitherto, and several new techniques have been introduced in order to cover the widest possible range of intensity and frequency.

The equal-loudness relations must for completeness be associated with a threshold curve appropriate to the same conditions of listening, and this forms a natural "zero" for the whole set of curves. The investigation has therefore included a re-determination of the threshold of hearing in free field which shows a high degree of consistency with the equal-loudness relations at supraliminal levels.

Some extensions of this work, to include equal-loudness relations for sound distributed in particular frequency bands, is being undertaken and will be published separately in due course. The main object of this work is to clarify and further develop methods of computing the loudness levels of complex sounds from their spectrum analyses, which seems to be a particularly promising method for use in noise measurement problems.

DESCRIPTION OF EXPERIMENTAL PROCEDURE

(*a*) *General.* In view of the main object of this investigation, namely, to provide data which might assist in arriving at standardization of the equal-loudness contours, it was decided to maintain continuity with the practice in earlier determinations and with the definition of equivalent loudness[4] by carrying out the measurements under the same reference conditions. The sounds were presented to the listeners in the form of free progressive plane waves of sinusoidal waveform, with the observers listening binaurally and directly facing the oncoming wave. The advantages of these somewhat artificial conditions are the simplicity and freedom from ambiguity with which they can be specified. On the other hand, the accurate realization of the reference conditions over the wide frequency and intensity range of normal hearing involves some instrumental difficulties and limitations. The measurements have been carried beyond the range covered by the previous investigations[2,3] by the use of special reproducing arrangements. Except for the highest frequencies, the tests range from threshold to sound pressure levels of about 130 dB, and in frequency from 25 to 15 000 c/s. Altogether the investigation was in progress for nearly four years.

(*b*) *Selection of observers.* A random sample of observers from the entire population would include persons of various degrees of otological abnormality in addition to the great majority having normal hearing. The study and classification of abnormal hearing is, however, the province of the otologist and would form a separate investigation. The present experiments were deliberately restricted to persons of normal, in the sense of unimpaired, hearing. To this end the prospective subjects were submitted to an otological examination to ensure that no one was included having symptoms or signs of ear disease, or wax in the external auditory canals. Only subjects judged as normal at this examination were accepted for the tests. This basis of selection conforms to the usage adopted in a recent British Standard[5] for the threshold of hearing by earphone listening. The observers were selected from the available personnel of the Laboratory to be representative of both sexes and of a wide distribution of adult ages. So far as was known, the observers had not been subjected to prolonged periods of excessive noise such as might have caused permanent loss of hearing.

Experience at the National Physical Laboratory in various kinds of loudness estimation has indicated that stable results are generally obtainable with a homogeneous group of about thirty observers. To obtain truly representative results, however, a larger test group is necessary, particularly in those tests where the age is an important factor. On these grounds it was considered desirable to employ a substantially larger number of subjects than those who took part in the earlier determinations.

The present results are based primarily on the judgments of a team of ninety persons, forty-five of each sex, with ages ranging from sixteen to sixty-three years. This group participated in about thirty tests. In addition, a second team, consisting of seventeen men and thirteen women of average age about thirty years and all distinct from the members of the larger group, took part in over a hundred tests. Their results were barely distinguishable from those of larger group.

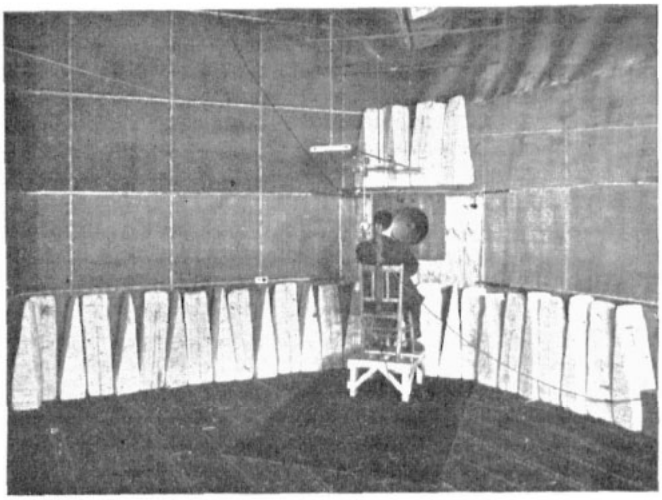
(*c*) *Description of apparatus.* The extensive frequency and intensity range of normal hearing necessitated the use of several distinct experimental arrangements. Except for the lowest frequencies the subjects were seated on the axis of and directly facing various sound sources in a lagged room, a headrest or sighting wires being provided to fix the head position. Measurements of the ambient noise level in the test room showed that it was lower than the normal threshold of hearing.

A large part of the investigation was carried out with conventional cone-type loudspeakers. In order to reach high intensity levels these arrangements were supplemented by the use of a powerful reflex horn loudspeaker.

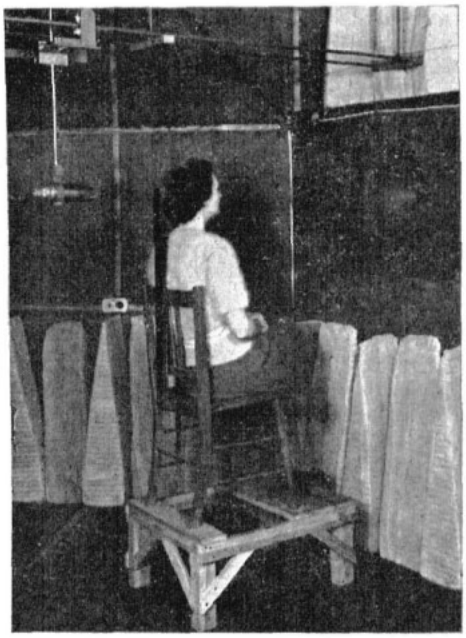
As the investigation was intended to relate to strictly pure tones, various devices were employed to reduce the harmonic distortion and circuit noise. The harmonic distortion of the tones used in tests at moderate levels was generally of the order -50 dB relative to the fundamental. In order to allow for extreme conditions, experiments were made to determine the degree of harmonic distortion which was acceptable without influencing the equal-loudness judgments. These are described in Appendix 1. The results of these tests were taken into account in setting an upper limit to the sound levels at which the different loudspeakers were operated.

For the higher frequencies, a plane array of twenty-eight $2\frac{1}{2}$ in. cone speakers was designed, and this was used in conjunction with the reflex horn, the latter providing the 1000 c/s tone. The twenty-eight loudspeakers were electrically equalized in sensitivity and phase and so connected that a better approximation to a plane wave front was realized than could have been obtained from conventional sources of comparable power. For some work at 15 000 c/s at high intensities a resonant barium titanate transducer in the form of a cylinder of diameter $4 \cdot 7$ cm was employed.

At the lowest frequencies it would be very difficult to obtain the high intensities required with direct radiators in a freespace room. Since one of the objects of the investigation was to extend the range of the equal-loudness contours in the low-frequency region while maintaining progressive plane-wave conditions the difficulty was met by the use of a duct. This was of the same general form as that described by Dadson and Butcher[6] which is in use at the National Physical



(*a*)



(*b*)

Fig. 1. Views of anechoic room showing conventional cone loudspeaker

Laboratory for low-frequency microphone calibration. The duct was of a hard-walled construction with a square cross-section 60 × 60 cm, and overall length 6·8 m of which 5·2 m were occupied by an absorbent termination consisting of a double fibre glass wedge. The driving unit employed was an 18 in. cone-type loudspeaker with a modified coil assembly. The subject's head and shoulders were accommodated in the free section of the duct, with arrangements for easy access. Reflexion from the absorbent wedge and the obstruction to the sound wave caused by the body of the subject were found to be negligible for frequencies from 25 to 200 c/s, and sound pressure levels up to 130 dB could be obtained without appreciable harmonic distortion. The duct was not suitable for the transmission of plane waves at 1000 c/s; accordingly intermediate reference tones of 200 c/s or 50 c/s were employed. Some of the apparatus used is shown in the photographs, Figs. 1, 2 and 3.

(*d*) *Methods for equal-loudness comparisons.* The great majority of the equal-loudness tests have been carried out by the constant-stimulus (CS) method, in which the subject has the sole task of making judgments of loudness inequality but is otherwise passive. Reasons for preferring this method have been described by Robinson in a previous paper.[7] For the most part comparisons were made between a pure tone

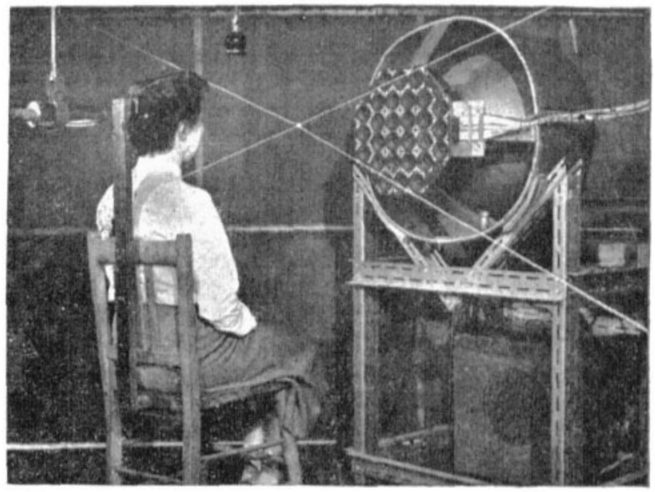
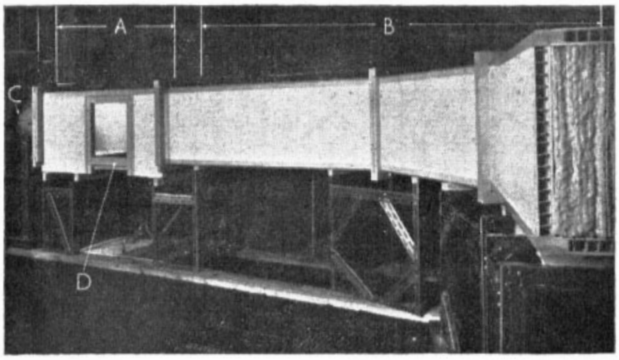
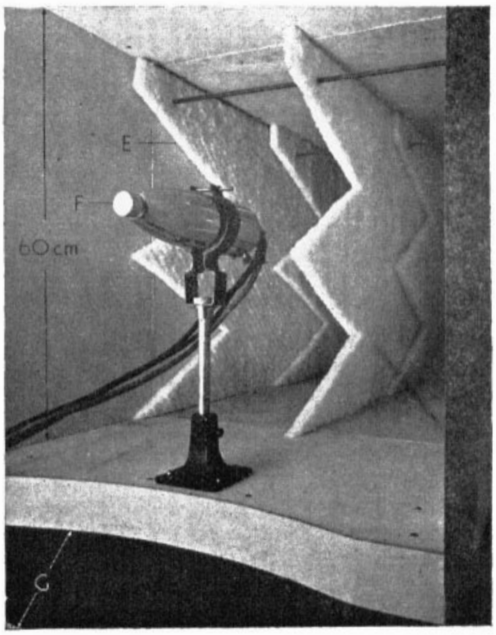


Fig. 2. Compound loudspeaker unit, The high-frequency unit is mounted on a high-efficiency reflex horn type, the latter providing the 1000 c/s tone

of constant intensity and frequency and a pure tone of 1000 c/s at intensity levels which were varied at random. The fixed and variable tones were occasionally interchanged and the distinction which should be drawn between tests carried out in these two ways will be discussed later. A number of comparisons between tones other than 1000 c/s were also made, in part to demonstrate the invariance of the equal-loudness relations but also in cases where instrumental limitations precluded reproduction of the 1000 c/s tone under the reference conditions. In the constant-stimulus tests, the time-sequence patterns *AOBOAOB* and *BOAOBOA* were employed, where *A* represents the fixed stimulus, *B* the randomized variable stimulus and *O* a period of silence. The order of presentation of these sequences was randomized as well as the level of *B*. The duration of *A*, *B* and *O* was based on the available evidence as to the time taken to reach full sensation,[8] periods of from 1 to 3 s being used depending on the comparison frequencies. The growth and decay of the tones were retarded by an electronic switch in order to eliminate the disturbing click which is heard with rapid transients. The switching was complete in 0·15 s. The *B* tone was varied over a wide range of sound pressure level so as to include, with an adequate margin, the transition levels of all observers. There are considerable instrumental difficulties in providing signals variable over a range consistent with these requirements, while allowing adequate signal-to-noise ratio, particularly in tests at high intensities. For very quiet *A*-tones a complication of another kind arises. If the *A*-tone is fixed at a sensation level within the dispersion of the threshold levels, the equal-loudness test must fail for some fraction of the observers. Fortunately, the dispersion of threshold values for otologically normal observers is not very great. At levels down to 20 phons no data were lost in this way. At levels in the region of 10 phons, however, which have been included by previous experimenters, appreciable errors would have arisen from the dispersion of the threshold and the resulting loss of data. It was considered that such tests would have been misleading and in the present work the equal-loudness comparisons were not usually taken below 20 phons.



(*a*)



(*b*)

Fig. 3. Low-frequency duct

(*a*) General exterior view. *A*, working section; *B*, absorbent section; *C*, sound source; *D*, double door for access to listening position. (*b*) Detail of absorbent wedge structure looking towards remote end of duct. *E*, commencement of absorbent termination; *F*, standard microphone at the listening position; *G*, adjustable aperture to suit body dimension of listener.

In order to extend the upper limit of intensity at which tests by the normal CS procedure are possible, use was made of a variant of the constant-stimulus procedure which has been described by Robinson.[7] In this method the fractions of observers judging tone *A* louder than tone *B* for a fixed level of *A* and a very few levels (usually 2 or 3) of tone *B* not far removed from the expected average transition level of the group are recorded, and from these values it is possible to deduce the correct transition point of the group. Generally only one level of *B* was presented at each testing session. It has been found that a modest change of perhaps 5 dB in the level of the *B*-tone produces a critical change in the fraction of judgments. The modified CS method has the double advantage over the usual procedure of extending the practical limit of intensity set by the apparatus and of producing results of comparable reliability in less time. The method was accordingly also used to a limited extent for some comparisons at moderate levels and the opportunity was taken in these cases to test the results against the more usual procedure. Success depends on a fairly accurate prediction of the result by the experimenter but this, of course, in no way influences the observers.

(*e*) *Interpretation of the equal-loudness test records.* The result of each equal-loudness test was obtained by the operator in the form of a chart showing the judgment recorded at each intensity level of the variable tone. These levels were presented at random, as already mentioned, and were recorded in the form shown in Fig. 4, which gives extracts from two typical test records of a constant-stimulus equal-loudness comparison. In the first example the pattern is symmetrical and the transition value corresponds evidently to the attenuation setting 16·5. A substantial proportion of the results conformed to simple patterns such as this, but some of the tests yielded a blurred pattern of which the second example is typical. In such cases the transition value is not obvious, but various formal processes may be adopted as a means of consistent interpretation. For instance, assigning α to the probability that tone *A* is louder and hence $(1 - \alpha)$ to the probability that *B* is louder, the judgments recorded may be regarded as quantized estimate of α or $(1 - \alpha)$ restricted to the values 0 and 1. By computing the attenuation value corresponding to $\alpha = \frac{1}{2}$, one arrives at a consistent interpretation of the data. As a result of applying these methods to a large number of transition patterns occurring in the tests, it was found, as regards averages over thirty or so test results, that they could be replaced by a simpler process. This consisted of averaging the mid-points of the ranges of uncertainty in the upper and lower rows of judgments. For example, in the second illustration the upper row is uncertain from 12 to 20 and the lower row from 13 to 23. The transition value $(16 + 18)/2 = 17$ was therefore assigned.

Attenuation (dB)	6		8		10		12		14		16		18		20		22		24		26		28		30	
AOBOAOB					1		1		1		2		2		2		2		2							
BOAOBOA				2		2		2		2		2		1		1		1		1						

(a)

Attenuation (dB)	8		10		12		14		16		18		20		22		24		26		28		30		32	
AOBOAOB	1		1		1		2		1		1		2		2		2		2		2					
BOAOBOA		2		2		2		1		2		1		2		1		1		1		1		1		

(b)

Fig. 4. Typical records from equal-loudness test by constant-stimulus method

1 signifies first tone heard was louder; i.e. tone A in upper rows, tone B in lower.
2 signifies second tone heard was louder.

Having arrived at these values, the next step is to deduce from them the normal equal-loudness relations for the group. The actual method adopted to represent the group data by a single figure is to some extent arbitrary and a number of different methods were explored to determine which could best be regarded as typical.

One of the ways of treating the data, which has the advantage of being independent of the form of the distributions and of the units employed, is to determine the median value for the group. In order to avoid the quantization error resulting from observations at whole decibel intervals, the position of the central value is better determined by drawing a freehand curve through the cumulative distribution plotted on probability paper and reading off the 50% point. Both the individual transition levels, and the totals of votes cast for, say, "tone *A* louder" at each attenuation step, could be treated in this way and the results were usually indistinguishable.

In the majority of the tests, the results conformed fairly closely to the normal error law provided the levels of attenuation were expressed in logarithmic units. In such cases the process for extracting the median value could be formalized by the method of probit analysis. The median values shown in the tables in Appendix 2 were mainly obtained in this way, including all cases of tests carried out by the modified CS procedure.

With the exception of the modified CS tests, values for the mode and mean of the distributions were also obtained, the attenuation steps being expressed in decibels. These values are also included in the table of test results. It is, of course, always possible to evaluate the mean but the result was occasionally influenced to an appreciable extent by lack of symmetry in the tails of the distribution, especially if the dispersion was large. The determination of modal values often required some care and numerical processes were devised for smoothing the histograms as an aid to consistent interpretation. In a few cases the mode was rather indeterminate or even, apparently, double-valued, and could not be estimated with accuracy. As the tables of test results show, the various processes rarely led to values differing by more than 1 phon. Moreover, any differences occurring were unsystematic in trend.

It should be noted that the ways just described for obtaining a value representative of the group suggest themselves as a natural consequence of a particular experimental procedure, namely the comparison of a constant test tone with a reference tone (usually 1000 c/s) of variable intensity. From a theoretical standpoint, there is no reason to accord the 1000 c/s tone this special place, although this is undoubtedly convenient for the purposes of definition and in practice. The equal-loudness relations ought to be valid for comparisons conducted between any pair of tones, either of which might be the one varied during the experiment. It is shown in Appendix 3 that a symmetrical treatment of the data can be defined which is independent of the use of a conventional reference tone, but it has been found experimentally that the more usual statistical procedures already discussed give results which are not materially different.

To summarize, the results adopted for incorporation in the tables and curves of equal-loudness are, in general, the modal values for the group of observers, corresponding to the usage in the definition of the phon.[4] In a few cases where the mode was ill-defined, for example, the tests at the higher frequencies where the necessary separation into age-groups resulted in smaller numbers of observers per group, the mean value is adopted. The results of tests by the modified CS method are expressed as the median value, obtained by probit analysis.

(*f*) *Method of determining the threshold of hearing.* The threshold curve may be regarded as a special member of the equal-loudness contours corresponding to just-zero loudness. It is appropriate at this point to anticipate the results and to note that the threshold values, though determined in a substantially different way, are entirely consistent with the equal-loudness data and, moreover, could have been predicted from them with considerable accuracy.

The method adopted for the determination of the minimum audible field was to switch the tone on and off at irregular intervals of a few seconds while decreasing by 2 dB steps from a clearly audible level until the observer no longer signalled in synchronism. In the neighbourhood of the minimum audible level the tone was presented several times at each step, and was recorded as audible if signalled for more than half the appropriate periods and as inaudible otherwise. The level was then raised in 2 dB steps from a level below threshold until the observer again responded. This frequently occurred on the same step recorded as minimum audible on the descent. The procedure was next repeated at levels staggered 1 dB with respect to the first set.

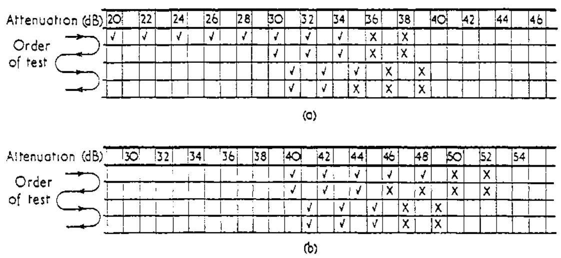


Fig. 5. Typical test records from a threshold determination

Two specimen test records are shown in Fig. 5, and serve to illustrate methods of assessing the threshold level. The ticks and crosses represent "audible" and "inaudible" as just defined. In practice only the steps in the immediate vicinity of the limen would have been presented more than once, for instance, steps 34 to 36 in Fig. 5(*a*). At least two criteria may be adopted to define the threshold, depending on whether an optimistic or pessimistic view is taken of contradictory judgments. Referring to Fig 5(*a*), the top two lines indicate a result between 34 and 36, while the bottom two lines point to a value very close to 35. On the other hand, step 35 has not always been heard. An optimistic estimate of 35 would be assigned in this case. In Fig. 5(*b*), 48 has been heard in the first descent but the subject has then "lost the thread" and only responded reliably at 44 on the ascent. In the second part of the test, the indications are in favour of a value close to 46 which would be the value assigned. On a more conservative view the threshold could be defined as the mean of the four least levels reliably heard, yielding 34 and 45·5 for the two illustrations respectively. The difference between these bases of assessments is commonly 0, 0·5 or 1·0 dB, and the difference averaged over all the threshold test records was 0·7 dB of which 0·5 is due to the use of 1 dB steps. For the purposes of this paper the lower levels have been adopted, but since each frequency is affected to a similar extent, it makes practically no difference to the equal-loudness relations which criterion is used.

RESULTS

(*a*) *General.* The experimental results are shown in Figs. 6 and 7, in which the co-ordinates are sound pressure level in decibels relative to 0·0002 dyn/cm^2 and equivalent loudness in phons, and in Fig. 8 which shows them in the form of equal-loudness contours. Comparisons between tones other than 1000 c/s have been converted to phons by means of the smooth curve at the appropriate frequency; this applies to the data for 25 and 33·3 c/s, the upper levels at 50 c/s, and to miscellaneous points on the curves for 100, 200 and 10 000 c/s. For completeness, the full experimental results have been collected and are tabulated in Appendix 2.

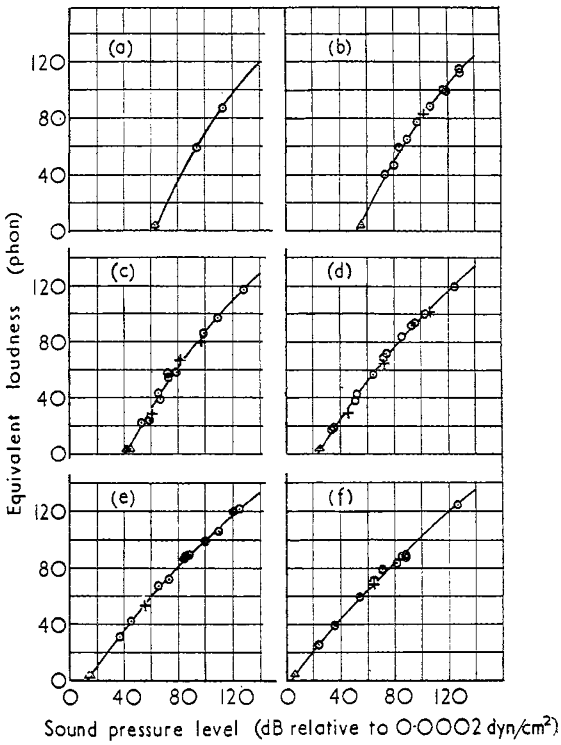


Fig. 6. Equal-loudness relations for pure tones below 1000 c/s

+, team A (ninety subjects); ⊙, team B (thirty subjects); △, combined team (threshold determination).
(*a*) 25 c/s, (*b*) 33⅓ c/s, (*c*) 50 c/s, (*d*) 100 c/s, (*e*) 200 c/s, (*f*) 500 c/s.

It would have been expected from previous work that the equal-loudness relations at the higher frequencies would show a fairly complicated dependence on the age of the observers. In the present work the effects of age were found to be negligible below and including 1000 c/s. Above 1000 c/s it is necessary to divide the observers into age groups in order to obtain a consistent picture of the equal-loudness relations. The curves show an increasing separation as the frequency is raised; the separation between the age groups is greatest at the threshold level and progressively diminishes as the intensity is raised. These effects are illustrated in various ways in Figs. 7, 8 and 9. Figs. 6 and 7 show the equal-loudness relations for each of the principal test frequencies, the results above 1000 c/s having been divided into successive decades of age. These curves show clearly that the maximum spread due to age occurs at the lowest levels of intensity and illustrate the converging tendency of the curves as the level is raised.

Analysis of the results has shown that the variations with age may be represented with considerable accuracy by

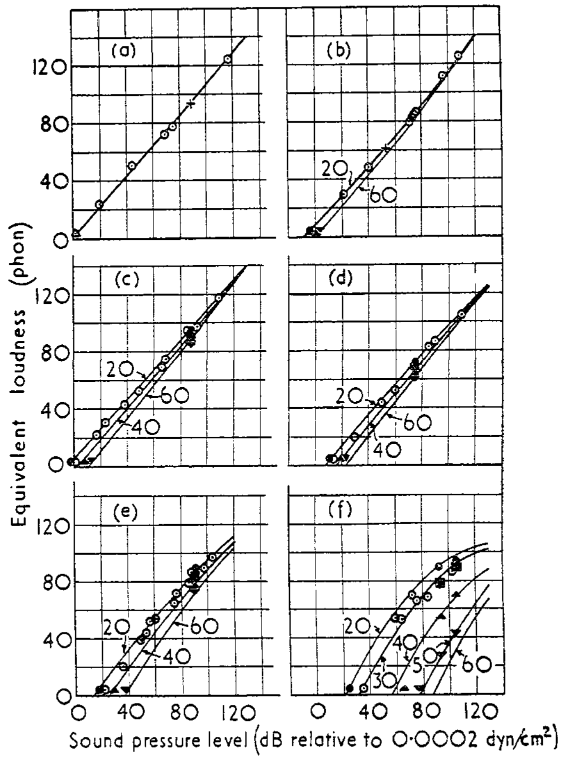


Fig. 7. Equal-loudness relations for pure tones above 1000 c/s

○, team B (average age about thirty years).
+ { ●, team A (average age about twenty years). ■, team A (average age about thirty years). ▲, team A (average age about forty years). ▼, team A (average age about fifty years). }
● ⊙ ▲ ▼, combined team (threshold determinations).
(*a*) 2000 c/s, (*b*) 3000 c/s, (*c*) 5000 c/s, (*d*) 7000 c/s, (*e*) 10 000 c/s, (*f*) 15 000 c/s.

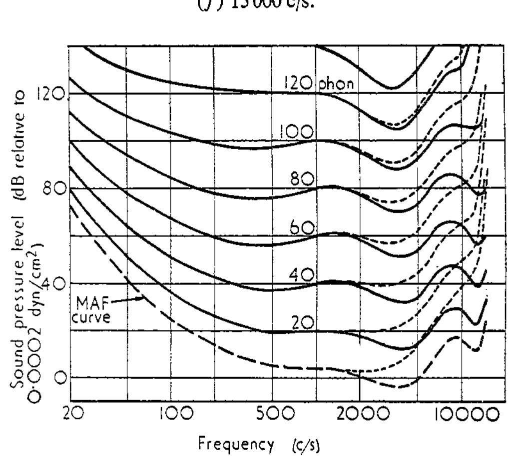


Fig. 8. Equal-loudness contours

——— age twenty years; – – – – age sixty years.

formulae involving two parameters, one of which is a function only of the age and the other a function only of the frequency. The formulae and tables of these parameters are given in Appendix 4, and were used in deducing the final forms of the equal-loudness relations. The equal-loudness contours are shown in Fig. 8 in which the results for age groups centred round twenty to sixty years have been separated in order to illustrate the trend of the results. The threshold level is shown in detail in Fig. 9 for successive age groups at ten year intervals. The dependence of the results on age shows many interesting points of detail which will be further considered in another publication.

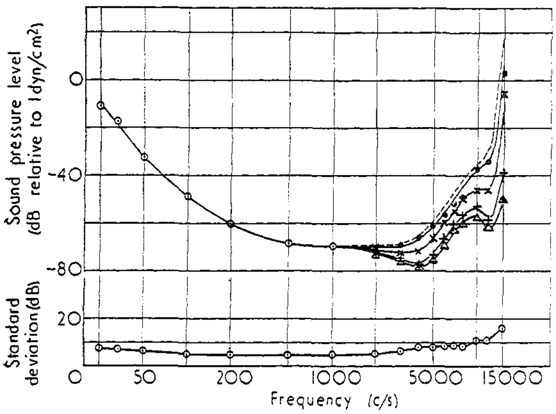


Fig. 9. Minimum audible field. (Binaural, listener facing source)

Ages: ---- 60 years; —•— 50 years; —×— 40 years; —+— 30 years; —△— 20 years.

It will be apparent from Fig. 7(*f*) that the age effects at 15 000 c/s are so large as to be a dominant factor and this gave rise to considerable difficulty in the treatment of the data. At frequencies higher than 15 000 c/s the dispersion would probably be greatly increased, and it is doubtful whether group data could be defined without using a very large number of subjects and a disproportionate expenditure of effort.

The data have also been analysed to see whether there was any substantial difference between the results for men and women. In most of the comparisons little difference was found in the loudness judgments of men and women but this was not always the case. Over the lower half of the frequency range there were occasional differences which, however, showed no signs of any systematic effect. Above 4000 c/s some larger differences were observed of the order ± 2 dB from the mean which appeared to exhibit a systematic trend with frequency, and which may be connected with differences in the diffraction effects of male and female heads. The general trend with frequency is consistent with a change in the diffraction pattern corresponding to a reduction of linear dimension, of the order 10%, of female as compared with male heads. In view of the comparatively small magnitude of the effect, however, it has been ignored in drawing the equal-loudness contours which relate to the average of the male and female data. It may be noted that Dadson and King,[9] commenting on the minimum audible pressure for telephone listening, stated that there was no significant difference between the sexes in that case. Thus, with the possible exception of effects due to external geometry, it is fairly safe to conclude that males and females respond equally to a given sound pressure.

(*b*) *Equal-loudness relations.* Referring to Figs. 6 and 7 it will be seen that in most cases the experimental results fall almost on straight lines and in no case require more than quadratic dependence on the sound pressure level to represent them. This has enabled the full results of the equal-loudness investigation to be expressed in relatively compact tables, giving, for normal observers, the equivalent loudness of a pure tone of given frequency and sound pressure level corresponding to ages from eighteen to sixty years. These tables are given in Appendix 4. The equal-loudness relations found in the present investigation, particularly for the tones of low frequencies, differ considerably from the corresponding results in earlier work and are a good deal more regular in form. This point is illustrated in Figs. 6 and 12 and is discussed in more detail in a later section.

As mentioned above, the equal-loudness relations for tones above 1000 c/s show an increasing tendency to divide into age groups. At any one frequency the curves for different age groups tend to converge as the intensity level is raised, indicating that the rate of growth of loudness with intensity is augmented for those subjects whose threshold level is raised, a phenomenon which is analogous to that known in the clinical literature[10,11] as "recruitment." Bearing in mind that the subjects were classed as otologically normal this effect seems to be present to some extent, apparently in binaural form, among older subjects generally.

(*c*) *The equal-loudness contours.* Fig. 8 shows the results in the form of equal-loudness contours and is simply the data of Figs. 6 and 7 presented in another way. For clarity only the curves for twenty and sixty years of age are drawn. The ordinate scale has been referred to a datum level of 0·0002 dyn/cm². The lowest contour is the minimum audible field, the value being 4·2 dB relative to 0·0002 dyn/cm² at 1000 c/s, or in round numbers, −70 dB relative to 1 dyn/cm². A zero-phon contour, passing through 0·0002 dyn/cm² at 1000 c/s has not been shown as it would seem misleading to indicate a contour audible to only 12% of the population, or even fewer for the higher frequencies.

The general form of the contours resembles those of the preceding investigations, although there are several differences which will be discussed in a later section. A new feature of the results is the dip in the contours in the region 400–500 c/s. In order to confirm this feature a supplementary series of tests was carried out, using frequencies of 333·3, 500 and 750 c/s, those at 500 c/s being repeated with several different loudspeakers, in different test rooms, and at two different listening distances, namely 1 and 2 m. All these confirmed the existence of the dip and its magnitude which is greatest for moderate levels of the intensity. The general appearance of the contours in the upper part of the frequency range where they are influenced by diffraction effects resembles the earlier results, though there are differences in detail.

(*d*) *Dispersion of the results.* The variance of an equal-loudness judgment is composed of three clearly defined components in addition to the residual errors of objective origin including instrumental error. A number of the tests were repeated under identical conditions, both with the object of assessing the relative importance of the various errors and of determining the extent of any interaction between them. These tests showed that the largest component is the variance between persons and the next largest the variance of repeated results by one observer representing on the average 89% and 9% respectively of the total variance. There is also a small component of variance due to the order of presentation, which is discussed in the next paragraph. In high frequency comparisons the variance between the results for different persons can be further resolved by separating the subjects into age groups. The "crude" (i.e. total) standard deviations of the various tests are included in the tables of test results (Appendix 2) and their dependence on frequency separation is illustrated in Fig. 10. It may be noted that in this figure the standard deviations have been expressed in phons which gives a more realistic comparison between the different tests.

In the threshold determinations, the tests at 50, 200, 3000 and 7000 c/s were repeated at intervals of several months, and a similar analysis shows again the largest part (80%) of the total variance is due to the differences between individuals.

(*e*) *Assessment of accuracy.* The overall accuracy of the equal-loudness relations can be assessed from the data contained in Fig. 10. For example, for moderate frequency

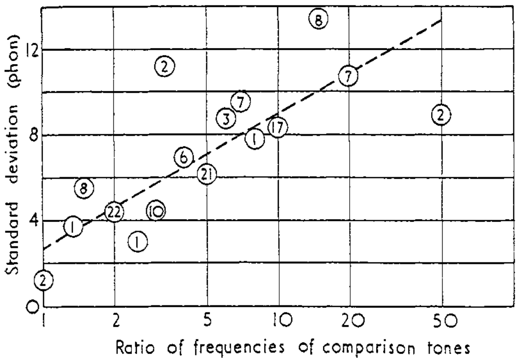


Fig. 10. Standard deviation in equal-loudness tests as a function of the frequency separation of the comparison tones. Number of tests shown in circles

separations, standard deviations of the order of 5 phons were obtained. Bearing in mind the numbers of observers involved, this is equivalent to a standard error of rather less than 1 phon. This figure is subject to appreciable improvement for moderate levels of the intensity due to the fact that the equal-loudness relations Figs. 6 and 7—which are no doubt inherently smooth curves—rely on data distributed over a considerable range of the variables. Using the method developed by Hayes and Vickers,[12] the accuracy of the curves over 85% of the intensity range covered by the measurements was found to be improved approximately two-fold compared with that of the individual values plotted. In addition, since the threshold determinations are based on the maximum number of observers employed and were, moreover, repeated at a number of frequencies, the error of the threshold measurements is already comparatively small. To summarize, over most of the intensity range, and all but the highest frequencies, the standard error of the equal-loudness relations is about 0·4 to 0·9 phon, depending on the frequency separation, the values rising to 1·0 to 1·5 phon for extreme upper levels of the intensity.

CONSISTENCY TESTS FOR THE EQUAL-LOUDNESS RELATIONS

Three types of consistency test were carried out in order to provide evidence as to the degree of coherence of the data as a whole. These were as follows:

(*a*) changing the order of presentation of the two comparison tones;
(*b*) interchanging the fixed and variable tones;
(*c*) triangular comparisons of the type $A - B$, $B - C$, $C - A$.

(*a*) The extent to which a change of order affected the judgments was readily obtained from the test records, illustrated in Fig. 4. Briefly, the results showed a small but persistent tendency to a surplus of "2" judgments in the vicinity of the equal-loudness point, that is to say the observers nearly all show a bias towards overestimation of loudness of the second tone heard, without, of course, being aware of this tendency. The extent of the "order" effect does not seem to depend on the nature of the comparison sounds but is somewhat dependent on the intensity level. It is interesting to note that the "order" effect was also observed in an experiment using two tones of the same frequency (1000 c/s). Between 30 and 100 phons, the "order" effect increases on the average from 0·6 to 1·3 dB. The effect does not appear in the results as presented, however, an average having been taken in all cases.

(*b*) For complete coherence of the data the average judgments recorded in any particular comparison should not depend on which of the two tones is fixed and which is variable. This point has already been mentioned in connexion with the statistical handling of the data. A number of tests, for the pairs of frequencies 50–1000, 50–200 and 33·3–50 c/s, were carried out both ways in order to investigate this effect. As regards individual observers, it was demonstrated that an interchange of the fixed and variable tones was immaterial; for example, in a two-stage comparison 50–1000–50 c/s, thirty observers returned to the initial value to within 1 dB on the average, in spite of a large dispersion at the intermediate stage. Similar tests carried out using the mean values for the group indicated a similar consistency of the order 1 dB, and in no case was it found necessary to take account of the order of the variables in treating the results.

(*c*) The third, and most crucial, test of consistency amounts to a verification of the proposition that sounds which are separately judged equal in loudness to the same reference sound are themselves judged equally loud. That this is substantially true was shown in a number of triangular comparisons, using tones of 50–100–1000 c/s, 33·3–50–200 c/s, 200–500–1000 c/s, 200–1000–10 000 c/s and 1000–3000–10 000 c/s. These triangular comparisons showed on the average a consistency of 1·5 phons, which is comparable with the standard error of the observations involved. The relevant determinations are listed in the tables in Appendix 2, especially Table 6.

The consistency revealed by these tests justifies their amalgamation with the remainder of the data in order to arrive at the final equal-loudness relations.

DISCUSSION

It is useful at this stage to compare the results of the present investigation with those previously available, due to Fletcher and Munson and Churcher and King. An attempt is made below to discuss the possible origin of the various discrepancies.

Considering in the first place the data on minimum audible field (MAF) (see Fig. 11) it is necessary to take note not only of the threshold curves given by the former authors but also of their actual experimental points. Fletcher and Munson preferred to show a conventional minimum contour passing through a round value, namely 10^{-16} W/cm^2 or roughly 0·0002 dyn/cm^2 at 1000 c/s, while recognizing that this represented hearing some 3 to 4 dB more sensitive than the

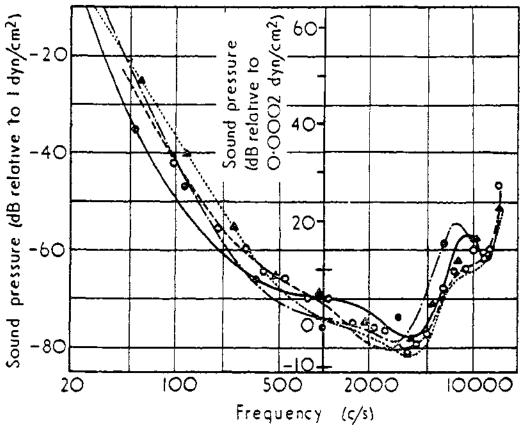


Fig. 11. Comparison of minimum audible field data

——, minimum audible field at age twenty years (this paper).
- - - - -, Fletcher and Munson zero phon contour; ▲, threshold values.
— · —, Churcher and King zero phon contour; ●, threshold values.
- - - -, Sivian and White smoothed MAF curve; ○, experimental values (average: twenty-four years).

average for their observers. It appears, however, that the contour so chosen does not lie uniformly below Fletcher and Munson's experimental threshold values but ascends to meet them at the low frequencies. Churcher and King also do not appear to have drawn their zero-phon contour through their threshold results. For example, their threshold level for 54 c/s (38·5 dB) is in very close agreement with the

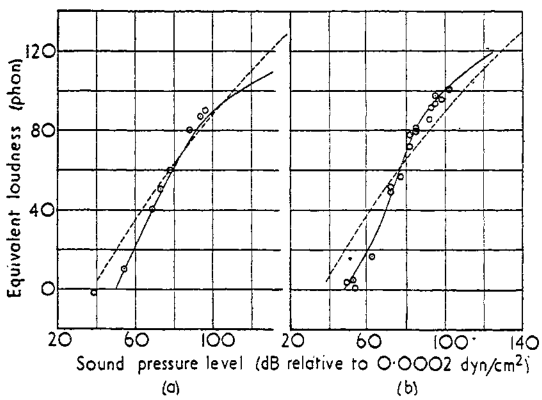


Fig. 12. Comparison of equal-loudness relations

(*a*) 54 c/s Churcher and King

○, experimental data; ——, smooth curve adopted by Churcher and King; - - - -, this paper.

(*b*) 60 and 62·5 c/s. Fletcher and Munson

○, experimental data; ——, smooth curve adopted by Fletcher and Munson; - - -, this paper.

present work but the Churcher–King contour for zero phon passes 11·5 dB above this value. It has already been shown that in the present results the threshold values fall naturally on an extension of the equal-loudness relations which is not the case in Churcher and King's results, as may be seen in Fig. 12(*a*). It would appear that these authors preferred to rely on the equal-loudness results, extrapolating these to a nominal threshold value. This procedure involves placing reliance on equal-loudness judgments conducted at a level of only 10 dB above 0·0002 dyn/cm^2, the objections to which have already been mentioned.

As regards the high frequency end of the scale there appears to be general agreement as to the incidence of the fluctuations due to diffraction effects but some discrepancy in the general levels. Direct comparison with the earlier results is not, however, possible in the absence of information regarding the age of the subjects. The only direct comparison which is available at high frequencies is between the present work and that of Sivian and White,[13] the age of whose observers is known. In this case, once again excepting the low-frequency region, there is fairly good agreement.

To summarize, the principal difference between the present work and the previous investigations as regards the threshold levels, consists in a lowering of the threshold of the order 10 dB at the low frequency end combined with a rise of 4 dB at 1000 c/s. There remain ambiguities at the high frequencies, the resolution of which would require information about the age of the observers in previous investigations.

Recent results due to Dadson and King[9] make it possible to relate the present work on MAF with their results for monaural minimum audible pressure (MAP) by telephone listening (see Fig. 13). It is interesting to note that the

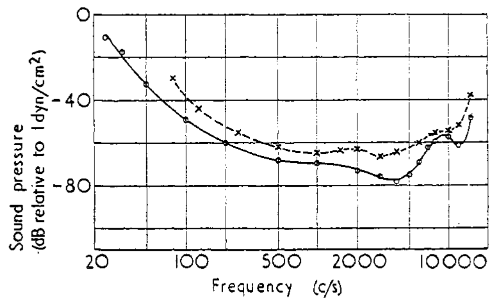


Fig. 13. Threshold of normal hearing for the age group eighteen to twenty-five years

——○——, N.P.L. minimum audible field (binaural).
- - - × - - -, minimum audible pressure (monaural) (B.S. 2497: 1954).

difference between the MAF and MAP values, often referred to as the "missing 6 dB," is very consistent* with the corresponding *difference* in the work of Sivian and White, although the absolute values obtained by the American workers are not in agreement with the British results in the low-frequency region. The cause of the latter is not entirely clear but is probably of objective origin. It may be remarked that the present work, as well as that of Dadson and King, was carried out in exceptionally silent conditions. It is possible, therefore, that the elevation of the threshold level and the

* The average elevation of the MAP relative to the MAF over the frequency range 80–6000 c/s is 10·2 dB for Sivian and White's data and 9·2 dB in the case of the N.P.L. results.

corresponding crowding of the lower level contours at the low frequencies in the previous investigations may be explainable on the basis of masking by ambient noise.

Considering next the data for equal loudness over the remainder of the auditory diagram, the main differences between the present and previous work may first be illustrated by reference to the equal-loudness relations. The differences are most apparent at the low frequencies where the present investigation leads to curves which are considerably more regular in form (see Fig. 12). In particular, the inflected form of equal-loudness relations shown by the work of Fletcher and Munson receives no support. This increased regularity of the equal-loudness relations is reflected in the more uniform distribution of the equal-loudness contours in the low-frequency area, which may be seen by comparing Figs. 8 and 14. There is an interesting difference between

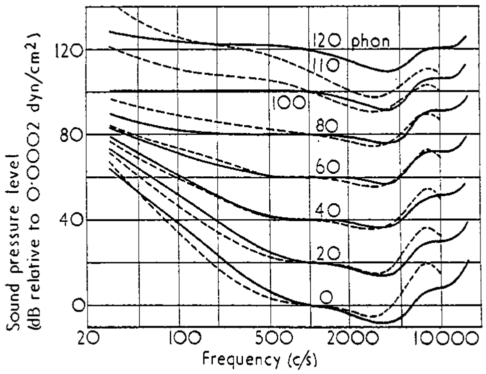


Fig. 14. Equal-loudness contours (previous determinations)

- - - -, Churcher and King; ———, Fletcher and Munson.

the results of Fletcher and Munson and the present work for frequencies above 6000 c/s. Whereas Fletcher and Munson show the loudness level increasing with the intensity level at an increasing rate, the present results show the opposite tendency. This tendency is present in all age groups and suggests some kind of saturation in the processes determining the sensation of loudness at high frequencies.

CONCLUSIONS

A new determination of the equal-loudness relations for pure tones has been carried out at the National Physical Laboratory using a large team of otologically normal observers. The measurements have extended over the frequency range 25 to 15 000 c/s and up to some 130 dB above $0\cdot0002$ dyn/cm^2. The objects of the work were to attempt to reconcile the discrepancies between previous determinations, to extend the range of intensity as far as possible and to provide data which may assist in arriving at a standard set of contours. The work includes a re-determination of the threshold of hearing for free-field listening.

The great majority of the equal-loudness tests were conducted by the constant-stimulus method, in which the observers' task is confined solely to judgments of inequality of loudness in relation to pairs of pure tones, one of fixed intensity and the other variable in random steps. New techniques were used in order to realize the reference free-field conditions at the extremities of the frequency and intensity range.

The relations connecting sound pressure level expressed in decibels with equivalent loudness in phons are accurately expressible by formulae quadratic in the sound pressure level. These formulae are expressed by means of parameters which are functions of the frequency, the sensation level of the tone, and the age of the observers. This has enabled compact tables to be prepared for the equivalent loudness of a pure tone over the range of measurement.

At each of the frequencies concerned, the threshold measurements proved to be highly consistent with the equal-loudness relations.

At frequencies above 1000 c/s, the age of the observers is an important factor, becoming dominant at 15 000 c/s. At the latter frequency the equal-loudness relations suggest a kind of saturation effect in the process of audition. An effect analogous to bilateral recruitment also seems to occur to some degree in normal observers of the higher age groups at high frequencies.

Comparing the results with the previous data, the equal-loudness relations show a considerably greater degree of regularity. At the low frequencies a lower threshold level has been found, and in the region around 500 c/s there is a depression in the contours, which has been carefully investigated. This was not shown in the earlier determinations.

The dispersion of the measurements corresponds generally with a standard deviation of the order of 5 phons, depending somewhat on the separation of the comparison frequencies. The major component of the variance, which arises from the differences between individuals, considerably exceeds the uncertainties of repetition by any given observer. Bearing in mind the number of persons tested, the accuracy of the smoothed equal-loudness relations is estimated to be of the order ± 1 phon over the greater part of the auditory diagram. At the extreme upper frequencies only, a more conservative figure should be assigned.

The minimum audible field values are found to bear a close resemblance to the corresponding minimum audible pressures by earphone listening recently determined at the N.P.L., but lie uniformly below them. This generally confirms the well-known phenomenon loosely called the "missing 6 dB." It is interesting to note that the minimum audible field at 1000 c/s was found to be -70 dB relative to 1 dyn/cm^2 in agreement with the experimental data of Fletcher and Munson and of Sivian and White. Fletcher and Munson, however, adopted a conventional zero of 10^{-16} W/cm^2 (roughly $0\cdot0002$ dyn/cm^2) at 1000 c/s as the reference datum for their equal-loudness contours, this convention having subsequently been sanctioned in the definition of the phon. The present results show that a contour through the latter value would lie outside the range of audibility of the great majority of normal observers.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the patient co-operation of the staff of the N.P.L. who took part in the subjective tests. Acknowledgement is due to Mrs. G. W. Tucker who carried out many of the tests, to Mr. C. W. Clenshaw of the Mathematics Division for assistance in the preparation of the tables, and to Mr. G. Bradfield of the Physics Division who kindly supplied the transducer used for some work at 15 000 c/s. Thanks are also due to Mr. T. Cawthorne and Mr. R. M. Harvey of King's College Hospital, London, who

kindly arranged the otological examination of the subjects. The work described was carried out as part of the research programme of the N.P.L., and the paper is published by permission of the Director of the Laboratory.

REFERENCES

(1) Kingsbury, B. A. *Phys. Rev.*, **29**, p. 588 (1927).

(2) Fletcher, H., and Munson, W. A. *J. Acoust. Soc. Amer.*, **5**, p. 82 (1933).

(3) Churcher, B. G., and King, A. J. *J. Instn Elect. Engrs*, **81**, p. 57 (1937).

(4) *Brit. Stand.* 661: 1955. Clause 3013.

(5) *Brit. Stand.* 2497: 1954.

(6) Dadson, R. S., and Butcher, E. G. *Acustica*, **4**, p. 103 (1954).

(7) Robinson, D. W. *Acustica*, **3**, p. 317 (1953).

(8) Munson, W. A. *J. Acoust. Soc. Amer.*, **19**, p. 584 (1947).

(9) Dadson, R. S., and King, J. H. *J. Laryng. and Otology*, **46**, p. 366 (1952).

(10) Dix, M. R., Hallpike, C. S., and Hood, J. D. *Proc. Roy. Soc. Med.*, **41**, p. 516 (1948).

(11) Denes, P., and Naunton, R. F. *J. Laryng. and Otology*, **64**, p. 375 (1950).

(12) Hayes, J. G., and Vickers, T. *Phil. Mag.*, **42**, p. 1387 (1951).

(13) Sivian, L. J., and White, S. D. *J. Acoust. Soc. Amer.*, **4**, p. 288 (1933).

(14) Gates, B. G. Discussion to Ref. 3.

APPENDIX 1

A note on the permissible harmonic distortion in pure-tone equal-loudness measurements

Harmonic distortion is most likely to be a source of error in the low-frequency region, partly on account of the difficulty in producing a sinusoidal sound waveform and partly because the acuity of hearing is considerably reinforced for the harmonics relative to the fundamental.

Several experiments were carried out with tones of fundamental frequency 25, 33·3, 40, 50, 200 and 1000 c/s to determine the permissible level of harmonic distortion at various intensity levels. Two methods were used:

(*a*) Constant-stimulus loudness comparisons between a reference tone and a tone of the same fundamental frequency containing a controlled admixture of harmonic components, giving a quantitative estimate of the loudness increment due to the added components.

(*b*) A qualitative form of test in which the observer indicated his awareness or otherwise of any effect on the total loudness due to the presence of overtones. The method used was to inject into the sound source an input at a frequency very slightly removed from that of the harmonic already existing and so adjusted in amplitude that it alternately reinforced and cancelled it.

The results showed that for tones of moderate intensity level the loudness was not significantly influenced by harmonic components at a level of −40 dB or lower relative to the fundamental. Applying the method of loudness summation due to Gates,[14] combined with the equal-loudness curves in this paper, to the results obtained by method (*a*), it was found that the increment of loudness due to the presence of harmonics could be calculated with considerable accuracy from the sound pressure levels of the components.

At high intensities, quite high levels of harmonic distortion may be tolerated without effect on the total loudness, which is presumably due to the already considerable effect of aural harmonics. For example, in the case of a 33·3 c/s tone operated at a sound pressure level of 130 dB, and using method (*b*), observers were unable to detect the presence of a third harmonic component only 20 dB below the fundamental.

Method (*b*) was particularly useful in connexion with the low-frequency duct in order to determine the maximum sound levels at which the experiments could be carried on. Provided observers could not detect the fluctuations as the harmonic component cancelled and reinforced the injected signal it was concluded that no error in the loudness comparison was occurring, whatever might be the indicated level of the harmonic.

In the threshold measurements it was only necessary to ensure that each harmonic was well below the normal threshold for the harmonic concerned. This condition was met throughout, the margin being still adequate at 25 c/s.

APPENDIX 2

Experimental results

The results of the experiments are given in Tables 1–7. All values are rounded to the nearest 0·5 dB. In general, the mean, modal and median values are quoted. In some cases the modal value could not be estimated with accuracy, and has been omitted. In experiments conducted by the modified CS procedure, only the median value can be determined. In tests at frequencies above 1000 c/s the effect of age is shown by the results for groups of average age twenty, thirty, forty and fifty years. In these cases the standard deviations are net values, the variation due to age having been allowed for.

Table 1. *Minimum audible field; frequencies up to 1000 c/s*

Frequency (c/s)	*MAF (dB relative to 1 dyn/cm²)* Mean	Mode	Median	*Standard deviation of observations (dB)*
25	−10·5	−11	−10·5	8
33·3	−17·5	−17·5	−17·5	7
50	−31	−32	−31	6·5
50	−32	−32·5	−32·5	6
100	−49	−49	−49	5
200	−59	−59·5	−59	4·5
200	−59·5	−59·5	−59·5	4·5
500	−68·5	−68·5	−69	4·5
1000	−69·5	−70	−70	4·5

Table 2. *Minimum audible field; frequencies above 1000 c/s*

Frequency (c/s)	MAF (dB relative to 1 dyn/cm²) (mean value) Average age of observers (years) 20	30	40	50	Standard deviation of observations within groups (dB)
2 000	−73·5	−72·5	−72	−69·5	5
3 000	−75·5	−74·5	−73·5	−70·5	6
3 000	−75	−74	−72·5	−68·5	6·5
4 000	−79	−78	−73	−65·5	8
5 000	−75	−74	−66	−61·5	8
6 000	−69·5	−67	−59·5	−57·5	8·5
7 000	−63·5	−60	−55	−50	10·5
7 000	−62	−61·5	−53	−54·5	9·5
8 000	−60·5	−56	−50	−50·5	8·5
10 000	−57·5	−53	−46	−38·5	11·5
12 000	−61	−58·5	−46	−37·5	11·5
15 000	−49·5	−39	− 5	+ 2·5	17
Number of observers per group	51	38	17	14	

Table 3. *Direct phon determinations; tones below* 1000 *c/s*

Test tone: Frequency (c/s)	Test tone: Sound pressure level (dB relative to 0·0002 dyn/cm²)	Equivalent loudness (phons): Mean	Mode	Median	Standard deviation of observations (phons)
50	52	21·5	22	21·5	5·5
	60	32	28	30·5	10·5
	65	44	—	44·5	13
	65	44·5	—	44	15
	72·5	55·5	57·5	57	13·5
	80	65·5	67·5	66·5	9
100	32·5	17	17·5	17·5	5
	45·5	29·5	28·5	28·5	6·5
	50·5	39	38	38	7·5
	70·5	65·5	65·5	65·5	7
	73	—	—	72 *	6·5
	85·5	81·5	84	83	7·5
	95·5	92·5	94·5	93·5	6·5
	103	—	—	100 *	7
200	36	30	30·5	30	4·5
	45	42·5	42	42	5·5
	55·5	52	52·5	52·5	6
	65	66·5	67·5	67	5·5
	72	—	—	70·5*	5·5
	86	88	88·5	88·5	5·5
	86	88	88	88	5
	99·5	97·5	98·5	99	4·5
	109	—	—	106·5*	3
	120	—	—	120·5*	5·5
	124	—	—	122·5*	12·5
333	61	62·5	62	62	5
500	24	25	25	24·5	4
	34·5	38	38·5	38	3
	54	59	58·5	59	4
	64·5	70	69·5	69·5	4
	64·5	68·5	68·5	68·5	4
	70·5	—	—	79 *	3·5
	81	83·5	—	83·5	3
	85	88	88·5	88	2·5
	86	89	88·5	88·5	2·5
	86	89	88·5	88·5	4
	127	—	—	124·5*	3
750	53·5	59·5	59·5	59·5	3·5

* Determinations by modified CS procedure. Standard deviations in these cases are estimated values.

Table 4. *Direct phon determinations; tones above* 1000 *c/s*

Test tone		*Equivalent loudness (phons)*						
Frequency (c/s)	*Sound pressure level (dB relative to 0·0002 dyn/cm²)*	*Mean value* *Age* 20	30	40	50	*Mode* *Age* 30	*Median* *Age* 30	*Standard deviation of observations (phons)*
1 500	74·5	—	75·5	—	—	75·5	75·5	3·5
2 000	20·5	—	24·5	—	—	24·5	24·5	4·5
	45·5	—	49	—	—	49·5	49·5	5
	70·5	—	72·5	—	—	72·5	72·5	4·5
	75·5	—	—	—	—	—	76	6
	90	92·5	93	92	93	—	—	4
	118	—	—	—	—	—	124	7
3 000	22	—	29	—	—	29	29	4
	42	—	48	—	—	47·5	48	5·5
	55	—	61	—	—	60·5	61	5
	55	63	62	59·5	62	—	—	4·5
	73	—	—	—	—	—	79·5*	4
	74·5	—	—	—	—	—	82 *	4
	75	—	83·5	—	—	84·5	84	5
	75	—	84	—	—	84	84	4
	97·5	—	—	—	—	—	111 *	4
	110	—	—	—	—	—	124·5*	—
5 000	16·5	—	—	—	—	—	22·5*	5
	23	—	31	—	—	30·5	31	6
	37·5	—	—	—	—	—	42·5*	4·5
	48	—	51	—	—	52·5	51·5	8·5
	66·5	—	—	—	—	—	69·5*	7·5
	68	—	72·5	—	—	75·5	73·5	8
	85·5	—	—	—	—	—	93 *	4·5
	87	91·5	90	85	86	—	—	6
	92	—	96·5	—	—	96·5	96·5	6·5
	109	—	—	—	—	—	117 *	8
7 000	30	—	—	—	—	—	18·5*	9
	50	—	42·5	—	—	42	42	8·5
	60	—	—	—	—	—	51·5*	8
	75	70·5	69	60·5	59·5	—	—	10·5
	84·5	—	79·5	—	—	82·5	80	6·5
	89·5	—	—	—	—	—	86·5*	9·5
	109	—	—	—	—	—	104 *	—
10 000	34·5	—	—	—	—	—	19·5*	9·5
	49·5	—	42·5	—	—	38·5	41	9
	53·5	—	—	—	—	—	43·5*	10
	58·5	—	51·5	—	—	53	52	13
	74	—	—	—	—	—	65 *	9
	89	—	85	—	—	87	86·5	11
	89	82·5	78·5	70	68·5	—	—	12·5
	97·5	—	90·5	—	—	89	90·5	9·5
	103	—	—	—	—	—	96·5*	6·5
15 000	59·5	—	50·5	—	—	54	53	13
	64	—	51	—	—	53·5	52	15
	72·5	—	64	—	—	69	67	10
	76·5	—	68	—	—	65·5	67	13
	83·5	—	—	—	—	—	67·5*	11·5
	94	89	77	53·5	26·5	—	—	20
	103·5	—	—	—	—	—	86 *	11·5
	105	93	89·5	70	43·5	—	—	13

* See footnote to Table 3.

Table 5. *Phon determinations by "average-error" method*

Test tone			
Frequency (c/s)	Sound pressure level (dB relative to 0·0002 dyn/cm²)	Equivalent loudness (phons)	Standard deviation of observations (phons)
100	32·5	18	5·5
	52·5	44	10
	72·5	70	7·5
	92·5	92	5

Table 6. *Indirect determinations*

Test tone		Reference tone equal in loudness to test tone					
Frequency (c/s)	Sound pressure level (dB relative to 0·0002 dyn/cm²)	Frequency (c/s)	Sound pressure level (dB relative to 0·0002 dyn/cm²) Mean	Mode	Median	Standard deviation of observations (dB)	Equivalent loudness (phons)†
25	94·5	50	77	76·5	77	5	57·5
	113·5	200	85	85·5	85	7·5	87
33·3	72·5	50	64·5	64·5	64·5	6·5	37·5
	84	50	77·5	77·5	77·5	5	59
	99	50	92	92	92	4	78
	116·5	50	111·5	111·5	111·5	3·5	101
	129	50	—	—	125·5*	2	116
	80·5	200	49·5	—	49·5	10·5	48
	118	200	99	—	98·5	7	100
	129	200	—	—	113 *	7	113
50	72·5	100	62·5	—	64	7·5	53·5
	58	200	33	30·5	32·5	7·5	25
	78	200	57·5	58	56·5	8	58
	96·5	200	79	79·5	79	6·5	81
	97·5	200	83	84	83·5	5·5	86
	107·5	200	94·5	95	95	6·5	96·5
	127·5	200	—	—	118·5*	4·5	117·5
100	65	200	57·5	58	57·5	4·5	57·5
	106	200	100·5	101	100·5	3·5	102
	124·5	200	—	—	121	2·5	119·5
200	86	500	85	85	85·5	3	88·5
10 000	55·5	3000	45	—	45	11·5	50
	75	3000	64·5	—	66	9	71·5

* See footnote to Table 3.

† The loudness in phons is calculated from the tables in Appendix 4, using the modal value of the equally-loud reference tone where available.

Table 7. *Experiments with interchange of fixed and variable tones*

The entries in this table should be compared with those for 33·3 c/s in Table 6 and for 50 c/s in Table 3

Fixed tone		Variable tone, equal in loudness to fixed tone					
Frequency (c/s)	Sound pressure level (dB relative to 0·0002 dyn/cm²)	Frequency (c/s)	Sound pressure level (dB relative to 0·0002 dyn/cm²) Mean	Mode	Median	Standard deviation of observations (dB)	Equivalent loudness (phons)*
50	81·5	33·3	89·5	90	89·5	4	64·5
	96·5	33·3	102·5	102·5	103	4·5	83·5
	100·5	33·3	107	107	107	4	88·5
1000	39·5	50	67	65·5	66	6	39·5

* The loudness in phons is calculated from the tables in Appendix 4 for the fixed tone in each case.

APPENDIX 3

The effect of interchanging the fixed and variable stimuli in comparisons between two tones and a method for treating the results in a symmetrical manner

It has been mentioned in connexion with the tests of consistency that, so far as individual observers are concerned, it was found to be immaterial which tone was varied in equal-loudness tests: there is a unique correspondence between the two sound pressure levels for the condition of equal-loudness. Thus, expressing the sound pressure level of one tone as a function of that of the other, an equal-loudness relation can be defined for each individual. Many of these relations were drawn out and showed that, though they are fairly diverse in position, slope, and curvature as between one individual and another, the curve for each observer is well-defined and smooth.

To illustrate the problem of determining the symmetrical form of the equal-loudness relation for a group in terms of the individual data already mentioned in connexion with the statistical treatment of the results, it will be sufficient to represent the individual loudness relations by linear functions of the sound pressure levels. In practice, the form of the functions is more nearly quadratic, but this fact, while introducing a complication, does not change the argument in principle. It should be emphasized that the following discussion is based on an algebraic rather than a statistical consideration. We may write, for each individual,

$$y = ax + b \tag{1}$$

where x is the sound pressure level of one tone and y the sound pressure level of a tone at another frequency judged by that individual to be equally loud to the first and a and b are constants for the individual but have statistical distributions over the group, the nature of which is immaterial to the discussion. In determinations of phon values, we identify y with the 1000 c/s tone and x with the fixed tone, and the equal-loudness relation for the group as customarily defined is

$$y = \bar{a}x + \bar{b} \tag{2}$$

in which the barred symbols represent some form of average.

However, let us suppose that the test procedure had been reversed, so that fixed 1000 c/s tones (y) were equated in loudness by adjustment of the other tone. Since the loudness relations (1) are fixed for each observer, they may also be written $x = (1/a)y - (b/a)$, and the results of the reversed tests for the group would be $x = \overline{(1/a)}y - \overline{(b/a)}$ or

$$y = [1/\overline{(1/a)}]x + [\overline{(b/a)}/\overline{(1/a)}] \tag{3}$$

From the theoretical standpoint, equations (2) and (3) ought to represent the same function. It is easy to prove, however, that this condition is not satisfied if the averaging process indicated by the bar symbol corresponds to any of the usual statistical operations for measurement of central tendency such as the mean, mode or median processes.

Viewed as a geometrical problem in the x, y plane, it may be seen that the desired operation represented by the bar symbol must depend only on the ensemble of lines in equation (1) and not on the co-ordinate frame to which they are referred. In particular, the operation must be invariant with respect to translation of the co-ordinate frame since this is equivalent merely to changing the zero level to which the sound pressures are referred. This condition of invariance is satisfied if the bar symbol is identified as the operation defining a line having the orientation

$$(1/n)\sum_{r=1}^{n}\theta_r$$

where $\theta_r = \arctan a_r$, and which passes through the point (x_0, y_0) such that the sum of squares of normals from (x_0, y_0) on to the ensemble of lines in equation (1) is minimal; r is a running index for the different members of the group.

This method was applied to groups of thirty individual equal-loudness relations for the frequency pairs 33·3/1000, 100/1000, and 200/1000 c/s and the resulting functions were compared with the pairs of relations corresponding to equations (2) and (3) with the bar symbol taken, by way of example, to signify the arithmetic mean. The differences were found never to exceed 1 dB and, except at the extremities of the intensity range, the difference was only a small fraction of 1 dB. These examples, especially the first, apply to comparisons between fairly remote frequencies for which the distributions of the parameters of the individual curves had a considerable dispersion. It was therefore concluded that no appreciable error is caused by treating the loudness judgments in this investigation by the more usual statistical methods, and no distinction needs to be made in practice between the fixed and variable stimuli.

APPENDIX 4

Tables of normal threshold of hearing (MAF) and equivalent loudness

(*a*) *Equivalent loudness of pure tones.* The equal loudness relations for pure tones for normal subjects of age twenty years are expressed by the formulae

$$y_{20} = a + bx + cx^2$$

where y_{20} = equivalent loudness in phons

x = sound pressure level of pure tone of frequency f(c/s), expressed in dB relative to $0\cdot0002$ dyn/cm^2

and a, b and c are functions of the frequency f.

Values of a, b and c are given in Table 8. For most accurate use, second differences should be employed when interpolating in the table. The spacing of frequencies has, however, been so chosen that equivalent loudness values may be derived to the nearest phon by linear interpolation. Owing to changes of sign in the second differences, the values of y in such cases should, in general, be rounded up except for frequencies in the range 1000–4000 c/s where they should be rounded down. The experimental results for frequencies from 33·3 to 1000 c/s cover the intensity range from threshold to about 130 dB, and for the higher frequencies to about 110 dB. The accuracy of y calculated for values of x within these ranges is of the order 1 phon, except possibly above 10 000 c/s. The estimated accuracy for $x = 140$ (from 33·3 to 1000 c/s) or for $x = 120$ (from 3000 to 10 000 c/s) is about 2 phons, these values representing an extrapolation of some 10 dB beyond the experimental data. Values in the table for $f = 20$ c/s are extrapolated.

(*b*) *Effect of age on the threshold of hearing of normal subjects.* The threshold level t_n, expressed in decibels, for normal subjects of age n is given by $t_n = t_{20} + \beta\gamma$ where t_{20} is the threshold of hearing (MAF) for normal subjects of age twenty years, β is a function only of the frequency, and γ a function only of the age.

Values of t_{20} are given in Table 8 and values of β and γ in Tables 9 and 10 respectively. The value of γ is arbitrarily chosen to be zero for age twenty.

The tabulated values of t_{20} correspond to the twenty-year MAF curve shown in Fig. 9. This figure also illustrates the accuracy with which the $\beta\gamma$ correction fits the experimental data.

Table 8. *Equivalent loudness parameters*

Frequency (c/s)	a	b	c	t	b′	Frequency (c/s)
20	−217·2	3·669	−0·009 27	+74·3	2·291	20
25	−167·2	3·145	7 82	65·0	2·128	25
30	−135·7	2·801	6 77	58·1	2·014	30
35	−113·2	2·538	5 89	52·7	1·917	35
40	− 96·3	2·326	5 14	48·4	1·828	40
45	− 83·2	2·155	4 54	44·8	1·748	45
50	− 73·0	2·021	4 08	41·7	1·681	50
55	− 65·1	1·917	3 71	39·1	1·627	55
60	− 58·7	1·835	−0·003 43	+36·8	1·583	60
60	− 58·7	1·835	−0·003 43	+36·8	1·583	60
70	− 49·1	1·722	3 05	32·9	1·521	70
80	− 42·5	1·652	2 85	29·8	1·482	80
90	− 37·4	1·603	2 75	27·2	1·453	90
100	− 33·5	1·570	−0·002 68	+25·1	1·435	100
100	− 33·5	1·570	−0·002 68	+25·1	1·435	100
120	− 27·0	1·512	2 60	21·4	1·401	120
140	− 22·7	1·473	2 55	18·9	1·377	140
160	− 19·4	1·444	2 49	16·8	1·360	160
180	− 16·9	1·422	2 45	15·2	1·348	180
200	− 14·7	1·404	−0·002 41	+13·8	1·337	200
200	− 14·7	1·404	−0·002 41	+13·8	1·337	200
250	− 10·8	1·362	2 30	11·2	1·310	250
300	− 8·1	1·325	2 18	9·4	1·284	300
350	− 6·1	1·290	2 04	8·1	1·257	350
400	− 4·7	1·259	−0·001 88	+ 7·2	1·232	400
400	− 4·7	1·259	−0·001 88	+ 7·2	1·232	400
600	− 1·8	1·155	1 16	5·2	1·143	600
800	− 0·5	1·064	−0·000 48	4·4	1·060	800
1 000	0	1	0	+ 4·2	1	1 000
1 000	0	1	0	+ 4·2	1	1 000
1 500	+ 1·4	0·944	+0·000 60	+ 3·0	0·948	1 500
2 000	+ 3·3	0·924	1 00	+ 1·0	0·926	2 000
2 500	+ 5·3	0·928	1 18	− 1·2	0·922	2 500
3 000	+ 6·9	0·937	1 20	− 2·9	0·930	3 000
3 500	+ 7·9	0·946	1 13	− 3·9	0·937	3 500
4 000	+ 7·9	0·954	0 98	− 3·9	0·946	4 000
4 500	+ 7·1	0·963	0 80	− 3·0	0·958	4 500
5 000	+ 5·3	0·973	+0·000 59	− 1·1	0·972	5 000
5 000	+ 5·3	0·973	+0·000 59	− 1·1	0·972	5 000
6 000	− 0·5	1·011	+0·000 14	+ 4·6	1·012	6 000
7 000	− 7·5	1·075	−0·000 35	10·9	1·067	7 000
8 000	− 13·3	1·159	89	15·3	1·132	8 000
9 000	− 16·5	1·242	1 45	17·0	1·193	9 000
10 000	− 16·8	1·314	2 03	16·4	1·247	10 000
11 000	− 14·8	1·377	2 69	14·2	1·301	11 000
12 000	− 12·7	1·450	3 50	12·0	1·366	12 000
13 000	− 13·9	1·566	4 54	12·0	1·457	13 000
14 000	− 22·7	1·777	5 91	16·0	1·588	14 000
15 000	− 43·0	2·146	−0·007 72	+24·1	1·774	15 000

Table 9. *Threshold parameter (age)*

Age	γ	Age	γ	Age	γ	Age	γ
18	−0·9	30	5·9	40	19·9	50	32·7
20	0·0	32	7·6	42	23·8	52	34·0
22	+1·0	34	9·7	44	26·9	54	35·3
24	2·0	36	12·2	46	29·2	56	36·4
26	3·2	38	15·7	48	31·1	58	37·4
28	4·4					60	38·3

Table 10. *Threshold parameter (frequency)*

Frequency (c/s)	β	Frequency (c/s)	β	Frequency (c/s)	β
0	0	5500	0·41	11 000	0·71
500	0	6 000	0·41	11 500	0·78
1000	0	6 500	0·41	12 000	0·85
1500	0·02	7 000	0·40	12 500	0·94
2000	0·06	7 500	0·41	13 000	1·05
2500	0·13	8 000	0·43	13 500	1·17
3000	0·20	8 500	0·45	14 000	1·32
3500	0·26	9 000	0·49	14 500	1·49
4000	0·32	9 500	0·54	15 000	1·72
4500	0·37	10 000	0·59		
5000	0·40	10 500	0·65		

(*c*) *Effect of age on the equivalent loudness at levels above threshold.* The threshold shift has already been given in paragraph (*b*) as the product ($\beta\gamma$) of two independent parameters. The dependence on age of the equal-loudness relations at levels above the threshold naturally requires a more complicated expression. In the first place it must take account of the form of the equal-loudness relations themselves, and, as already stated, these are quadratic functions of the sound pressure level. Secondly, allowance must be made for the converging tendency of the equal-loudness relations for different age groups at the higher levels of intensity. It has been found that this part of the variation due to age is adequately represented by a quantity diminishing in a linear manner with the sound pressure level. From these considerations we arrive at the following approximate formula for the variation of the equivalent loudness with age of pure tones in the frequency range above 1000 c/s. The reduction of equivalent loudness in phons Δy corresponding to age n relative to the value for age twenty is given by

$$\Delta y = y_{20} - y_n = \beta\gamma(b' + c\beta\gamma)\,(140 + t_{20} - x)/(140 - \beta\gamma)$$

In this formula, β, γ, c, t_{20} and x have the same significance as above and $b' = b + 2ct_{20}$. For convenience in computing the age corrections, values of b' are included in Table 8.