BRITISH ACOUSTICAL SOCIETY

ADDRESS BY THE 1971 RAYLEIGH GOLD MEDALLIST

"Calculating the Perceived Level of Light and Sound."

Professor S.S.Stevens.

Laboratory of Psychoacoustics, Harvard University, Cambridge, Massachusetts, U.S.A.

In response to the honour of the Rayleigh Medal I should like to comment on a recent attempt to describe the quantitative rules that govern the response of the ear to complex sounds. The outcome of the inquiry has led to a revised procedure, called Mark VII (17) designed to improve our ability to calculate the perceived level of complex noise. I should like also to comment on certain remarkable similarities between vision and hearing, similarities that we encounter in our attempts to calculate perceived level in those two modalities.

Lord Rayleigh was among the first to determine the quantitative similarity in the sensitivities of the two sense organs."It appears," he wrote, " that the streams of energy required to influence the eye and the ear are of the same order of magnitude, a conclusion already drawn by Toepler and Boltzmann" (2,II, p 438). We now know that the two modalities also exhibit many other instructive similarities.

The stimulus energies of light and sound are both basically two-dimensional, varying as they do in frequency and intensity. There is phase too, of course, but phase does not usually play a large role in the subjective magnitude of what we see or hear. In what follows, it is the magnitude of our perceptions, what we call bright-ness and loudness, that will concern us most. In particular, attention will be addressed to the manner in which the eye and the ear integrate across frequency. The problem is this, can we discover the rules by which to predict the visual response, or the auditory response, when the stimulus is composed of a rich array of frequencies? How, in short, do we go from measured spectrum to predicted sensation?

Reference Stimulus

A common basic strategy emerged quite independently in optics and acoustics. The guiding idea called for the choice of a reasonably well-defined and reproducible stimulus to which other stimuli could be compared. In early times the standard light source was a sperm candle. Following Edison's invention it became an incandescent lamp; and it is now defined in terms of a black body at the temperature of the freezing point of platinum. The purpose of having a stan candle was to make possible the specification of other, unknown light sources in terms of the number of candles to which they were equivalent. A direct visual comparison was required, of course, because equivalence was defined in terms of appearance, how bright or how luminous. A visual photometer is a device that provides a convenient arrangement for making a visual match between the output of a calibrated lamp and that of an unknown source.

Acoustics had nothing as convenient as a candle. There was no standard noisemeter. Various investigators tried different types of sounds. Bekesy (3, p.378) tells how Barkhausen in the 1920's used a click which he matched to the loudness of music and other sounds. He built a simple device that he could carry inconspicuously in his pocket. When he closed a switch it produced a click that was transduced by a small receiver inserted in his ear. He could adjust the level of the click to match that of the music.

When the state of the art permitted the use of steady tones, it became customary to use a pure tone as the reference stimulus. Thus Kingsbury (4, 1927) used 700 Hz as his standard "candle", and he matched it tp other pure tones at various levels. In that way he mapped out a set of equal loudness contours. Churcher, King, and Davi (5, 1934) used 800 Hz as their standard, but the frequency 1000 Hz us by Fletcher and Munson (6, 1933) became the reference frequency most used.

I have recently proposed, however, that a better reference "candle" for measurement purposes is a noise 1/3 octave wide and centred at 31 Hz (1). That noise band is located in the frequency region to which t ear is most sensitive. In addition, when its level is increased from threshold, the 3150 band grows its loudness in a more orderly manner than do the lower frequencies.

Given a reference light, and a reference sound, the equivalent level of other lights and sounds can be determined by way of a simple matghing procedure. In principle, the method is straightforward; we merely find at what level the reference stimulus appears to match the unknown. In practice, though, we now use a different procedure entirely.

Before turning to the procedures used in practice, let us eqquire further about the measurement of sensation.

The Psychophysical Law

It is one thing to choose a reference stimulus, and even to map equal loudness and equal brightness contours, but that endeavour by itself does not provide a measure of loudness or brightness. We wa also to know how sensation grows with stimulus intensity.For example, if we double the stimulus intensity, does the magnitude of the sensat double? In other words, what function relates the stimulus level to t subjective effect?

This was the question posed by Fechner more than a centruy ago. He answered it by conjecturing a logarithmic relation, a relation that accords fairly well with the scale we obtain if we count off just noticeable differences. Fechner's logarithmic law dominated the field for many decades, and in some quarters it still appears to reign. It is interesting to note that in 1927 Kingsbury "measured" the loudness contours by counting up just noticeable differences. The results, he pointed out, agreed with Fechner's law.

Shortly thereafter, however, a change began to take place. The stimu measure used by Kingsbury, which was called a transmission unit, or T was rechristened a decibel. It was thought at first that the decibel, being a logarithmic unit, would provide a unit for the loudness scale, and accordingly the number of decibels above threshold came to be called sensation level. It became obvious to anyone else who listened, however, that a sound 100 dB above theshold seemed vastly more than twice as loud as a sound 50 dB above threshold. Listening tests suggested that the Fechner scale was clearly off target.

In the early 1930's the experimental search for a better loudness scale moved along in several laboratories and produced results that showed promising agreement. By 1935 I was led to propose the name sone for the loudness unit (7). That name still serves its original purpose, but the form of the function relating loudness to stimulus intensity has undergone revision and refinement.

In 1953 I had occasion to apply three different procedures, bisection, fractionation, and magnitude estimation, to the measurement of both brightness and loudness (8). The results for the two sense modalities showed a striking similarity. What impressed me most, however, was that the results suggested that both brightness and loudness grow as the cube root of the stimulus intensity. (In terms of sound pressure the exponent for loudness becomes 2/3.) Figure 1 shows two examples of the measurements made in 1953.

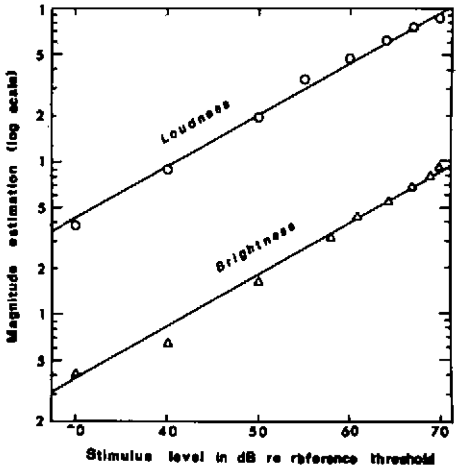


Figure 1.

In the years since 1953, subjective scaling procedures, especially the procedure I called magnitude estimation, have been applied to a score of different sense modalities. Quite uniformly, the meansurements have produced evidence that the psychophysical law is a power function, not a logarithmic function. In some of the textbooks, in fact, Fechner's law is giving way to Stevens' law.

The power law has wide application, to be sure, but when we look closely we find that the loudness of certain sounds does not follow a perfect power function (a stright line in log-log coordinates).

Even at 1000 Hz the downward concavity of the power function is detectable. That concavity constitutes one of the reasons for selecting a reference sound at the higher frequency of 3150 Hz. At 3150 Hz the loudness grows quite closely in proportion to the sound pressure to the two-thirds power. When the level at 3150 Hz is increased, presumably the excitation in the cochlea does not spread out as widely as it does at 1000 Hz.

In vision the power function does seem to encounter fewer perturbations that it does in hearing. In the ear a low-frequency stimulus tends to spill over into other regions of the cochlea when the level of the stimulus is raised. That type of spill-over does not occur in the eye. There are other minor problems in the eye, however, such as light scattering, but presumably the ratio of target intensity to scatter intensity does not change with level. When the light level is reduced to the scotopic region where the Purkinje shift occurs, there is some differential effect on the brightness functions for the different hues. Generally speaking, though, white light and light of various hues grow in brightness according to a power function with the exponent 1/3.

White noise, on the other hand, shows a significant departure from a power function, as do other broad-band signals, so much so that explicit account needs to be taken of the departure when we formulate procedures to calculate subjective values. Thus I have allowed for a departure from the power law in the new procedure, Mark VII, which permits us to calculate the perceived level of broad-band noises.

Cross-Modality Confirmation

If, to a first approximation, the subjective magnitude of both loudness and brightness increase as the cube root of the stimulus energy, we can make a quantitative prediction concerning an interesting bit of human behaviou. If the visual stimulus is set at several levels, and if the subject is instructed to adjust the level of asound to make the apparent loudness match the apparent brightness, we predict that the matching function will have an exponent equal to one. In other words, in decibel coordinates, the matching levels describe a straight-line function with a slope of one. That experiment, properly counterbalanced so that sound was matched to light and light was matched to sound, was carried out by my colleagues J.C.Stevens and L.E.Marks (9). The results confirmed the prediction inherent in Figure 1.

The general principle that governs cross-modality matching is this: the slope (exponent) of the matching function has a value equal to the ratio of the exponents of the two continua that are matched. The principle has been confirmed in numerous experiments (10, 11). As a matter of fact, in one or another laboratory, loudness has been matched to more than a score of different continua. From the consensus of those many experiments it becomes more and more firmly established that loudness grows approximately as the two-thirds power of the sound pressure.

Examples drawn from 10 experiments on the cross-modality matching of loudness to various other criterion stimuli are shown in Figure 2, taken from Stevens (12). The slopes of the lines in the log-log

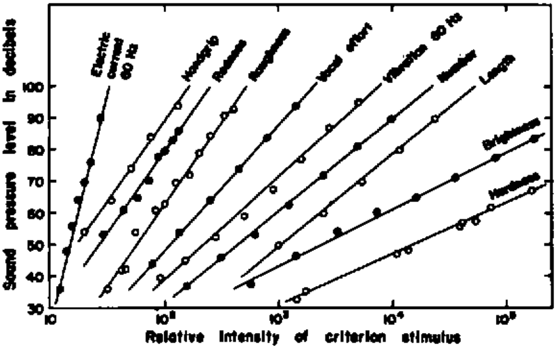


Figure 2.

coordinates of Figure 2 give a direct indication of the wide range of exponents that givern the various sense modalities.

Since each of the 10 continua had also been scaled by magnitude estimation, which requires the subject to match numbers to stimuli, the exponents of the 10 continua were known. Given these exponents and the slope of the lines in Figure 2, we can calculate 10 separate exponents for loudness, one exponent by way of each matching function. The geometric mean of the 10 calculated exponents turned out to be 0.67 with a standard deviation of 0.69 decilog.

One further fact should be noted regarding loudness matching. The procedure has begun to show promise of great practical utility in those circumstances where subjective variables need to be appraised. For example, subjects have matched loudness to their conception of the prestige that attaches to various occupations. Other subjects have adjusted loudness to match the roughness of an automobile ride on a bumpy road. An instrument sold commercially provides a sound whose loudness the hospital patient can adjust to match the degree of his pain or distress. By means of cross-modality matching, the effects of medication may be studied in a quanitative manner.

Summation over Frequency

How the eye integrates energy when the energy is spread over different wavelengths has been studied by many investigators, and the answer is generally taken to be given by Abney's law, so-called. It is assumed that the eye responds to a weighted sum of the radiant energy. The weighting function is the luminosity curve, which is simply an equal brightness contour. Thus the eye is assumed to integrate over the band of energy that reaches the eye through a filter that passes a band of a specified shape extending from about 400 to 700 nanometers. Partly becasue the eye filter is only about one octave wide, Abney's rule serves reasonably well for many purposes. It is conventional to calculate the luminance of a surface as a weighted sum across the energy spectgum.

Does the eye in fact integrate energy in accord with Abney's law? Not quite, it seems. "The deficiencies of Abney's law have been known for some time", said Graham (13, p. 370), "but they have been tolerated or evaded until recently."

An experimental demonstration carried out by Chapanis and Halsey (14) illustrates the problem. With the aid of differently coloured filters, they required observers to make all the colours appear equally bright by adjusting the intensity of the light through each filter. After the colours, some 342 in all, had been matched in apparent brightness, their luminances were calculated by the conventional procedure. The calculated luminances showed systematic departures from equality. For colours of equal brightness, highly saturated colours had lower calculated luminance. The difference between saturated and unsaturated colours reached two-to-one. In other words, two colours that look equally bright may differ in calculated luminance by 3 dB.

It appears, then, that the eye does more than integrate energy over wavelength or frequency.

With the ear also, we face a problem more complex than a wighted energy summation. The weighted sound level meter has its uses, to be sure, and its design can be improved, as I have suggested (1), if an "ear weighting" is incorporated to make it read dB(E), but the ear itself is a more subtle device than a weighted meter.

The approach to the summation problem that has proved most rewarding has usually been the one that transforms the stimulus variables, frequency and sound pressure, into subjective or sensation variables, such as pitch in mels (or Fletcher's position coordinates based on his critical bands) and loudness in sones (or excitation). The trans -formation of the stimulus values to excitation values was essemtially the approach explored in the 1930's at the Bell Telephone Laboratories.

In the late 1940's one of my students achieved a noteworthy simplification of the multitone loudness problem by expressing the specific loudness of each tone in sones per mel (15). He then plotted the loudness or excitation pattern of each tone against the mel scale of pitch, as shown in Figure 3. There it can be seen that, in the subjective coordinates, all the equally loud tones can be represented by similar patterns.

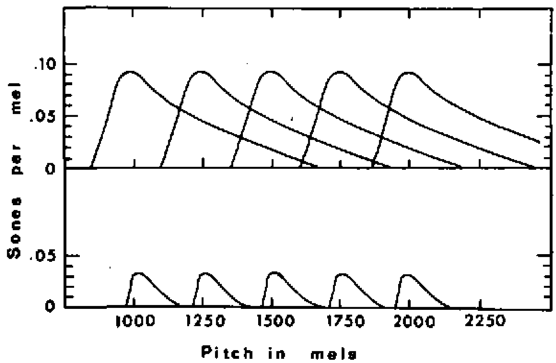


Figure 3.

That becomes possible because the mel scale provides a linear map of the cochlea (16). When the patterns do not overlap, s in the lower part of Figure 3, the loudnesses (areas) sum. When the patterns overlap, the tones exert mutual inhibition on one another, and their total loudness becomes less than the sum of their separate loudnesses.

When the tones lie close together, however, another rule obtains. If the tones fall within a critical band, which is about 100 mels wide (a hectomel), then the energies of the tones summate. In that case, as Fletcher said, "one must first add together all the intensities of the components in a critical band width and treat the combined intensity as a single compnent ...." (17, 1953, p. 195).

For tones separated by more than a hectomel, the simple formula developed by Howes gives a good prediction of the loudness of the multicomponent tones.

Most of the sounds whose perceived level we would like to calculate are not pure tones. Nevertheless, some of the same principles apply. We can think of a continuous spectrum as divided into critical bands or into some useful approximation thereto, such as 1/3 octave bands. If we start with one such band at a given loudness and add an adjacent band of the same loudness, we do not double the loudness. The principles portrayed in Figure 3 tell us that some degree of mutual inhibition must occur, so that the total loudness becomes equal to the loudness of the first band plus some fraction of the loudness of the added band. By pursuing that line of argument in 1955 I managed to develop a useful formula to express the loudness in sones of a complex noise as a function of the loudness of the individual bands (18). That summation formula, backed by extensive experimental work, has become the basis for a standard procedure for the calculation of loudness (19, 20, ISO 532) and also for the calculation of perceived noise level (21, ISO 507). In modified form it remains the basis of the new calculation procedure Mark VII (1).

In Mark VII the summation formula remains formally the same, but in order to describe more accurately the behaviour of the ear, the fractional loudness contributed by each added band is made to depend upon the loudness of the loudest band in the noise. Many experiments have shown that loudness adds more effectively at some levels than at others. Figure 4 shows how the fractional loudness contributed by

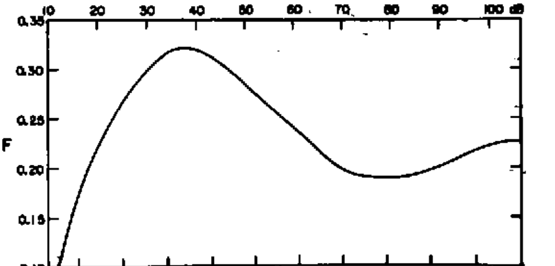


Figure 4.

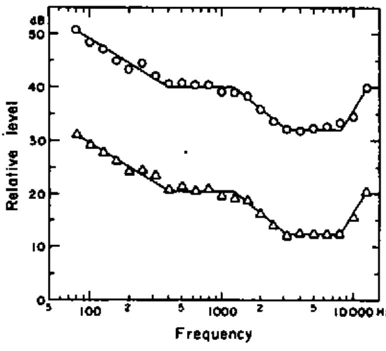
each added 1/3 octave band varies, depending on the general level of the noise. The abscissa represents the perceived magnitude (loud -ness or noisiness) of the bandthat is maximally loud or noisy. The loudness value in the maximum band determines what fraction of all the other bands must be added to the loudness of the loudest band.

Frequency Weighting

Beginning with the equal loudness contours determined by Kingsbury (4) many studies have been made to assess the relative sensitivity of the ear to sounds of different frequency. The most thorough studies of equal loudness relations for pure tones have been carried out in England at the National Physical Laboratory. Studies involving bands of noise have also been carried out in several laboratories, and in 1969 I was able to assemble some 25 sets of frequency-weighting contours measured in 11 separate laboratories during the quarter century preceding (1). Despite an occasional discordant value, the data showed remarkable agreement. The agreement persisted even though the subjects were not always instructed to judge loudness. Whether the subjects were asked to judge loudness, annoyance, or acceptability seemed to make no systematic difference to the contour of equal perceived magnitude.

By a process of iteration and adjustment I determined a contour that appeared to represent the consensus. The average and the median values for the 25 contours were also determined. Those values, along with the consensus contour itself, are shown in Figure 5.

Figure 5.



The contour is shown in segmented form, with discontinuities at regions of maximum curvature. The "true" contour, if we could ever determine it, would probably be smooth, but its representation by a segmented contour provides so many conveniences that I have retained it in Mark VII. The errors entailed by the discontinuities lie well within the bounds of experimental uncertainty.

In order to calculate the perceived level of sonic bangs and explosions it is desirable to extend the contours to the very low frequency. For that purpose we can make use of the informative measurements carried out by D.W.Robinson and his colleagues at NPL. They have extended some of the equal loudness contours down to frequencies as low as 3.15 Hz - a most remarkable achievement. While working with the NPL data, I was struck by a remarkable fact. Below 80 Hz the equally loudness contours can be quite well represented by straight lines. Even more remarkable the contours seem to converge toward a common point located at about 1.0 Hz and at a sound pressure level of 160 dB. The data themselves are presented in the paper on Mark VII (1).

Perhaps I should not have been surprised by the convergence of the equal loudness contours, because the shape of the so-called auditory area, the area bounded by the thresholds of audibility and feeling, have long suggested such a possibility. For example, on page 995 of my Handbook of Experimental Psychology (22) there is a compilation of thresholds that clearly forecasts a necessity for the convergence demonstrated in the data from NPL. I conjecture, however, that the contours may begin to turn downward at frequencies below about 5 Hz, so that a true convergence, with its rather curious implications, may prove to be a snare that nature has sidestepped. Nevertheless, the approximate convergence of the contours toward 1 Hz at 160 dB adds a beautiful simplicity to our attempts to portray the functioning of the auditory system.

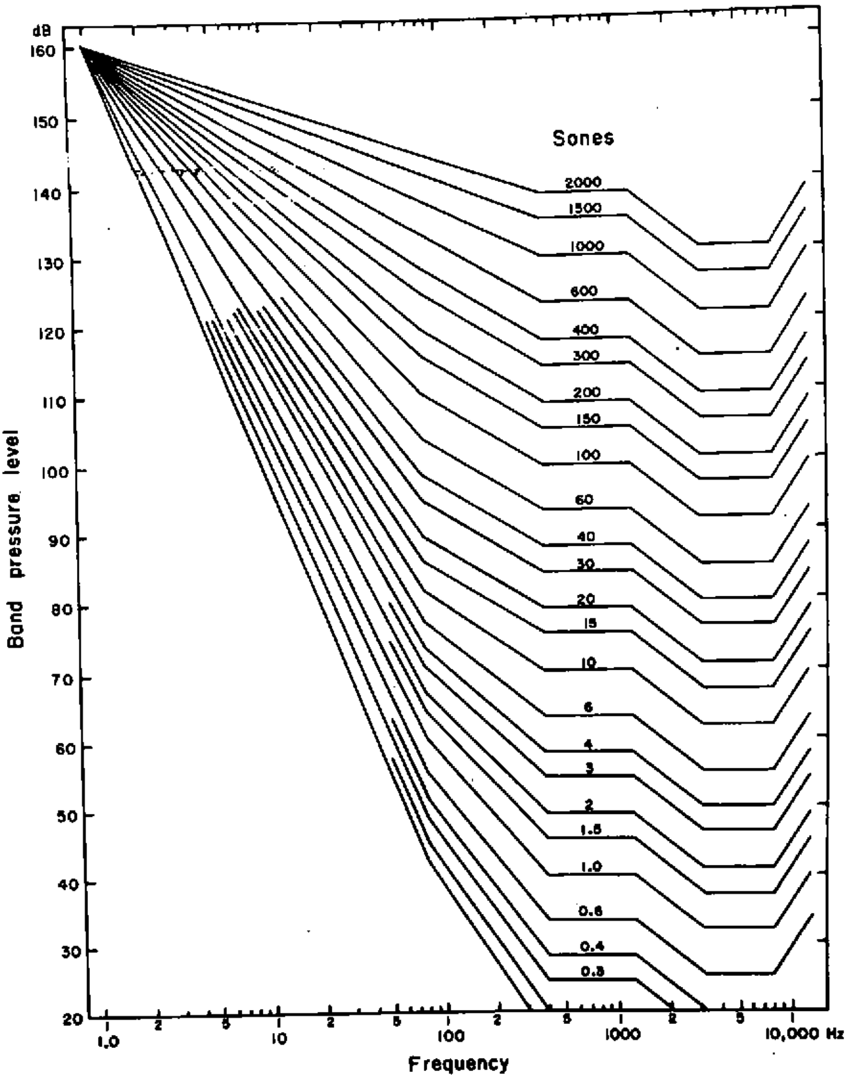
Figure 6 pictures a full set of contours representing equal apparent magnitude. The parameter on each contour is its value in sones. Since there appears to be no substantial difference between loudness and noisiness, it has seemed advisable to use the older unit, the sone, even though the contours may be used for the calculation of so-called noisiness as well as loudness. The sone can be defined as the perceived magnitude of a 1/3 octave band of noise centred at 3150 Hz and having a sound pressure level of 32 dB re 20 $\mu N/m^2$. The form of the equal sensation contour is such that the sone has the same value for the new reference sound that it had when the reference tone was 1000 Hz.

The basic ingredients of Mark VII are contained in Figures 4 and 6. If need be, those two figures can be used to calculate the perceived level in PLdB of any perceptually homogeneous sound, provided its spectrum has been measured in octave or 1/3 octave bands. In practice, however, the instructions and tables published elsewhere (1) would prove useful.

An Ear-Weighted Meter to Read dB(E)

In both optics and acoustics we have made use of meters whose response has been given a frequency weightin to reflect the form of the equal sensation contour. The traditional photometer that makes use of the human eye has been supplemented with photometers that provide a direct meter reading. Attempts have been made to incorporate in direct-reading photometers a filter that matches the pass band of the human eye. If Abney's law were correct, and if such a filter were available, there would be little further use for the kind of visual photometry that involves the matching of visual stimuli.

In sound-measuring instruments we have also used frequency weightings, both to reflect the equal loudness contour and to reflect the effect of level on the form of the contour. Many years ago the effects of level



dB
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
Band pressure level
Sones
2000
1500
1000
600
400
300
200
150
100
60
40
30
20
15
10
6
4
3
2
1.5
1.0
0.6
0.4
0.3
1.0
10
100
1000
10,000 Hz
Frequency

Figure 6.

gave rise to the weightings designated A, B, and C. Perhaps the time has come now to reconsider the weighting problem. Now that a consensus contour has become available, there is an opportunity to give the sound level meter a representative ear weighting. The form that such a meter weighting might have is shown in Figure 7.

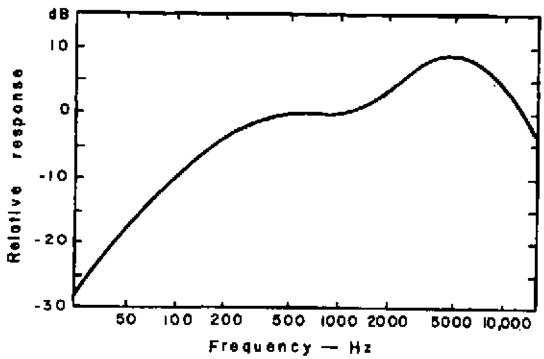


Figure 7.

A major advantage of an ear weighting would be that, for typical noise spectra, the ear-weighted sound level meter, reading in terms of dB(E), would agree within a decibel or two with the perceived level calculated by Mark VII and expressed in PLdB. The good agreement is made possible by the combination of two factors, the form of the consensus weighting contour plus the new reference sound at 3150 Hz. The sensitivity of the ear to the new reference sound serves to lower the calculated perceived level by 8 dB compared to the older loudness level, which referred to the 1000 Hz reference tone.

Meter readings in terms of dB(A) would fall about 5 dB lower than dB(E) but the difference would vary considerably with the shape of the noise spectrum.

The prospect at this point in time is for an ear-weighted sound level meter, reading dB(E), that falls within a tolerably close range of the exact perceived level in PLdB.

Figure Legends

Figure 1. Magnitude estimations of the loudness of a 1000-Hz tone and the brightness of a luminous, achromatic 5-degree disc. In both these early experiments (1953) the subject was first prepared with the strongest stimulus (70 dB) and was told to call it 100. He was then given the various stimulus levels in random order and asked to assign numbers proportional to the apparent magnitude. The lines through the data have a slope equal to 1/3 in the log-log coordinates, i.e., they are cube-root functions. The circles are mean values for 16 observers; the triangles are mean values for 11 observers. The reference levels for the decibel scales are 20 $\mu N/m^2$ and 10^{-10} lambert. Later studies with the method of magnitude estimation showed that it is usually better to omit the naming of a standard.

Figure 2. Equal sensation functions obtained by matches between loudness and various criterion stimuli. The relative positions of the functions are arbitrary, but the slopes are those determined by the matching data. (From 2, Stevens 1966.)

Figure 3. Patterns of excitation for equally loud tones spaced 250 mels apart. When plotted in sensation coordinates, sones per mel versus mels, the patterns all have the same form. The area under a curve is proportional to the perceived magnitude. At the lower perceived level, 2.5 sones or 44 PLdB, there is no overlap of the patterns, and the loudnesses add. For the upper curves, for 25 sones or 74 PLdB, the overlap is extensive and the total perceived loudness is less than the sum of the separate components. (After 15, Howes 1950.)

Figure 4. Showing how the fraction F depends upon level. The value F stands for the fractional loudness contributed by each 1/3 octave band. The value of F used in the calculation of perceived level is determined by the value in sones of the loudness or noisiness of the 1/3 octave band that produces the maximum perceived magnitude. Rule for octave bands: subtract 4.9 dB from the loudest 1/3 octave band; determine from Fig.6 the sone value with which to read the F value from the graph; then double the F value. Values of F are tabled in Stevens (1, 1972).

Figure 5. The results of averaging 25 separate contours of equal perceived magnitude. The decibel means are shown by the circles, the medians by the triangles. The 5-section weighting function of Mark VII has been fitted to both kinds of averages. (From 1 Stevens,1972).

Figure 5. Contours of equal perceived magnitude in sones. This family of contours can be used for the calculation of loudness and noisiness. The sound is presumed to be measured in octave or 1/3 octave bands in decibels re 20 $\mu N/m^2$. Values for these contours are tabled in Stevens (1, 1972).

Figure 6. Calculated response of a network for a sound level meter designed to give a suitable approximation to the "ear weighting". Such a meter could be calibrated to read dB(E). For may common noises dB(E) would approximate the perceived level in PLdB. (From 1, Stevens 1972).

REFERENCES

1. S.S.Stevens 1972 Journal of the Acoustical Society of America 51, 575 - 601. Perceived level of noise by Mark VII and dB(E).

2. Lord Rayleigh 1877 (reprinted by Dover, New York). The theory of sound.

3. G. v. Bekesy 1960 McGraw-Hill, New York. Experiments in hearing.

4. B.A.Kingsbury 1927 Physical Review 29, 588-600. A direct comparison of the loudness of pure tones.

5. B.G.Churcher, A.J.King, and H.Davies 1934 Journal of the Institution of Electrical Engineers 75,401-446. The measurement of noise, with special reference to engineering noise problems.

6. H.Fletcher and W.A.Munson 1933 Journal of the Acoustical Society of America 5, 82-108. Loudness, its definition, measurement, and calculation.

7. S.S.Stevens 1936 Psychological Review 43, 405-416. A scale for the measurement of a psychological magnitude, loudness.

8. S.S.Stevens 1953 Science 118, 576. On the brightness of lights and the loudness of sounds.

9. J.C.Stevens and L.E.Marks 1965 Proceedings of the National Academy of Sciences 54, 407-411. Cross-modality matching of brightness and loudness.

10. S.S.Stevens 1959 Journal of Experimental Psychology 57, 201-209. Cross-modality validation of subjective scales for loudness, vibration, and electric shock.

11. S.S.Stevens 1969 Perception and Psychophysics 6, 251-256. On predicting exponents for cross-modality matches.

12. S.S.Stevens 1966 Perception and Psychophysics 1, 5-8. Matching functions between loudness and ten other continua.

13. C.H.Graham (editor) 1965 Wiley, New York. Vision and visual perception.

14. A.Chapanis and R.M.Halsey 1955 Journal of the Optical Society of America 45, 1-6. Luminance of equally bright colours.

15. D.H.Howes 1950 American Journal of Psychology 63, 1-30. The loudness of multi-component tones.

16. S.S.Stevens and J.Volkmann 1940 American Journal of Psychology 53, 329 -353. The relation of pitch to frequency: a revised scale.

17. H.Fletcher 1953 D.Van Nostrand, New York. Speech and hearing in communication.

18. S.S.Stevens 1956 Journal of the Acoustical Society of America 28, 807 -832. Calculation of the loudness of complex noise.

19. S.S.Stevens 1961 Journal of the Acoustical Society of America 33, 1577 -1585. Procedure for calculating Loudness: Mark VI.

20. International Organisation for Standardisation 1967. Procedure for calculating loudness level, ISO/R-532.

21. International Organisation for Standardisation 1967. Procedure for describing aircraft noise around an airport. ISO/R 507.

22. S.S.Stevens (editor) 1951 Wiley, New York. Handbook of experimental Psychology.