Suggested formulae for calculating auditory-filter bandwidths and excitation patterns

Brian C. J. Moore and Brian R. Glasberg
Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, England

(Received 3 December 1982; accepted for publication 22 April 1983)

Recent estimates of auditory-filter shape are used to derive a simple formula relating the equivalent rectangular bandwidth (ERB) of the auditory filter to center frequency. The value of the auditory-filter bandwidth continues to decrease as center frequency decreases below 500 Hz. A formula is also given relating ERB-rate to frequency. Finally, a method is described for calculating excitation patterns from filter shapes.

PACS numbers: 43.66.Ba, 43.66.Dc [JH]

INTRODUCTION

The peripheral auditory system has often been modeled as a bank of overlapping bandpass filters (Fletcher, 1940; see Plomp, 1976 and Moore, 1982 for reviews). This concept is closely related to the concept of the critical band (Zwicker *et al.*, 1957), and many workers would consider that the critical bandwidth as measured in a variety of experiments (Scharf, 1970), is related to the bandwidth of the auditory filter at a given center frequency (e.g., Plomp, 1976, p. 10; Moore, 1982). However, the critical bandwidth may also be considered as a purely empirical phenomenon, for example as "that bandwidth at which subjective responses rather abruptly change" (Scharf, 1970, p. 159). Classical estimates of the value of the critical bandwidth as a function of frequency (the critical-band function) indicate that its value is approximately constant below 500 Hz, but increases above that frequency roughly in proportion with center frequency (Zwicker *et al.*, 1957; Scharf, 1970; Zwicker and Terhardt, 1980). However, there have been comparatively few direct estimates of critical bandwidth below 500 Hz, and indirect estimates based on the critical ratio may be in error, since the efficiency of the detector mechanism following the auditory filter appears to change with center frequency (Patterson *et al.*, 1982; Fidell *et al.*, 1983).

Recently several groups of workers have estimated the shape of the auditory filter, using the notch-noise method described by Patterson (1976) or the rippled-noise method of Houtgast (1977). The filter shape has been estimated using a linear model of masking in which it is assumed that the threshold corresponds to a constant signal-to-masker ratio at the output of the auditory filter. In this paper we present a summary of these results which shows that the equivalent rectangular bandwidth (ERB) of the auditory filter continues to decrease below 500 Hz. If the critical bandwidth is related to the ERB of the auditory filter, this may indicate a need to revise the classical critical-band function.

I. ESTIMATES OF AUDITORY FILTER SHAPE AND ERB

A. The notch-noise method

In this method the threshold of a sinusoidal signal is measured as a function of the width of a spectral notch in a noise masker; the notch is arithmetically centered at the signal frequency (Patterson, 1976). This method has been used by Patterson (1976), Weber (1977), Moore and Glasberg (1981), Patterson *et al.* (1982), Fidell *et al.* (1983), Glasberg *et al.* (1983), and Shailer and Moore (1983). Patterson *et al.* (1982) suggested a method of deriving the filter shape which assumed that the filter had the form

$$W(g) = (1 - r)(1 + pg)e^{-pg} + r, \qquad (1)$$

where g is the deviation in frequency from the filter center frequency divided by the center frequency, and p and r are adjustable parameters. The parameter p determines the shape of the passband of the filter, while r determines the point at which the passband gives way to the shallower "tails" of the filter. The filter shape is derived by fitting the integral of Eq. (1) to the data relating threshold to notch width. For full details see Patterson *et al.* (1982). For this filter shape the ERB expressed as a proportion of center frequency is approximately equal to $4/p$.

We have derived values of p and r, and hence estimates of ERB, for all studies where more than one frequency was used, and more than one subject was tested. Results are shown in Fig. 1. All the data presented in the figure were obtained at a noise spectrum level of 40 dB except those of Fidell *et al.* (1983) which were obtained at a spectrum level of 60 dB. At middle to high frequencies the ERB tends to increase when the spectrum level is increased above a certain level (about 40 dB; see Weber, 1977), but we do not know whether this happens at low frequencies, or at what noise spectrum level it occurs; the level would probably be higher than at high frequencies, owing to the transmission characteristics of the middle ear. The continuing decrease in the ERB as frequency decreases below 500 Hz would, if anything, be more marked if data at low spectrum levels were available. The data shown in Fig. 1 cover a range of frequencies from 0.124 to 6.5 kHz.

B. The rippled-noise method

This method is fully described by Houtgast (1977). Glasberg *et al.* (1983), in a direct comparison of the notch-noise method and the rippled-noise method, found that ERBs estimated from the latter (using Houtgast's method of

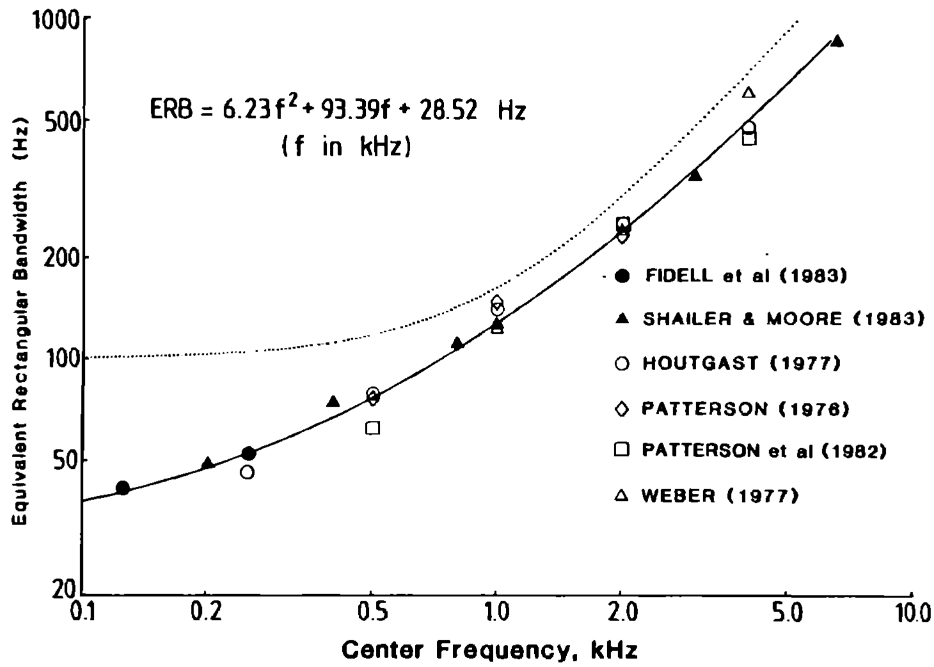

FIG. 1. The symbols indicate estimates of the equivalent rectangular bandwidth (ERB) of the auditory filter at various center frequencies, taken from the results of the workers indicated. The curve fitted to the data is specified by the equation in the figure. The dotted line is the classical critical-band function.

analysis, or using the method described by Pick, 1980) were about 35% greater than those estimated from the former. The discrepancy appears to arise largely from the different methods used to derive filter shapes from the data. The methods used by Houtgast and Pick can lead to overestimates of the filter bandwidth for certain filter shapes, including the rounded-exponential filter shape described by Eq. (1). The discrepancy is much reduced if the data are fitted assuming a filter shape such as the rounded exponential; see Glasberg *et al.* (1983) for details. For the purpose of comparison, Houtgast's estimates of ERBs have been divided by 1.35. His data were obtained at moderate noise levels (below 35 dB). The results are also shown in Fig. 1. After the "adjustment" Houtgast's results closely matched those derived using notch noise, in terms of their variation with center frequency. This adjustment should not be taken to imply that Houtgast's results are wrong; indeed we feel that the filter shape in Eq. (1) may be slightly too sharp around the tip of the filter, and may lead to underestimates of the ERB (see Glasberg *et al.*, 1983). However, our main interest here is in the variation of ERB with center frequency.

C. A simple expression relating ERB to center frequency

The data shown in Fig. 1 have been fitted with an equation of the form

$$\mathrm{ERB} = Af^2 + Bf + C, \qquad (2)$$

where f is center frequency in kHz and A, B, and C are adjustable parameters. The fit was obtained by expressing deviations from the fitted function as a proportion of center frequency, and minimizing the mean-squared deviation of the data from the fit. The resulting function, shown as the solid line in Fig. 1, is

$$\mathrm{ERB} = 6.23f^2 + 93.39f + 28.52. \qquad (3)$$

For comparison, the dotted line shows the classical critical-band function according to a formula given by Zwicker and Terhardt (1980).

II. ERB RATE

Zwicker and his co-workers (Zwicker and Feldtkeller, 1957; Zwicker and Scharf, 1965; Zwicker and Terhardt, 1980) have argued that it is often instructive and convenient to plot psychoacoustical data not as a function of frequency, but rather in terms of units of the critical bandwidth, or Barks. Zwicker called the scale relating frequency to Barks the critical-band-rate scale, or z scale. This scale can be obtained by integrating the reciprocal of the critical-band function (Zwicker and Feldtkeller, 1967). A comparable func-

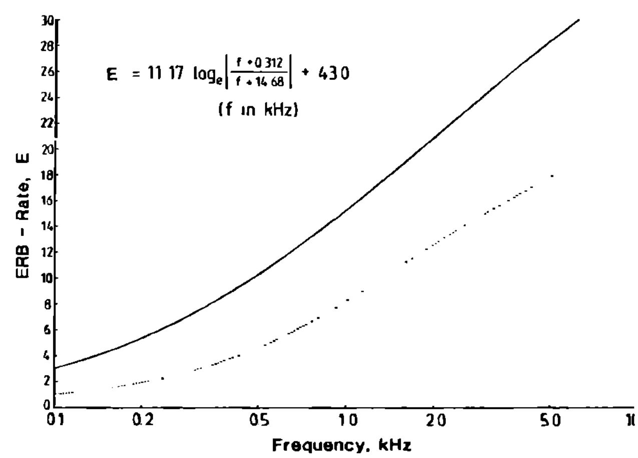

FIG. 2. The function relating ERB-rate to center frequency. The function is specified by the equation in the figure. The dotted line is the classical critical-band-rate scale.

tion, ERB-rate, can be derived from Eq. (2)

$$\text{ERB-rate} = \int \frac{1}{Af^2 + Bf + C}\,df$$
$$= \frac{1}{R}\log_e\left|\frac{f+(B-R)/2A}{f+(B+R)/2A}\right| + \text{const}, \quad (4)$$

where $R = (B^2 - 4AC)^{1/2}$, and $B^2 > 4AC$. Substituting the values of A, B, and C from Eq. (3), we get

$$\text{ERB-rate} = 11.17 \log_e \left|\frac{f+0.312}{f+14.675}\right| + 43.0. \quad (5)$$

This ERB-rate function is plotted in Fig. 2. The constant of integration, 43.0, has been chosen to make the ERB-rate $= 0$ when $f = 0$. However, the expression should only be considered valid for values of f between 0.1 and 6.5 kHz. For comparison, the critical-band-rate scale suggested by Zwicker and Terhardt (1980) is shown as the dotted line.

III. DERIVATION OF EXCITATION PATTERNS FROM FILTER SHAPES

The excitation pattern evoked by a given sound is the distribution of internal excitation as a function of some internal variable related to frequency. In terms of the filter bank analogy, the excitation pattern may be conceived as the output of each filter as a function of filter center frequency. For calculating excitation patterns from filter shapes, it is convenient to use a simple one-parameter approximation to the auditory filter shape, by dropping the parameter r from Eq. (1). The filter shape then becomes

$$W(g) = (1 + pg)e^{-pg}. \quad (6)$$

This approximation appears reasonable for the main passband of the filter (within about 25–30 dB of the tip), but it does not adequately characterize the shallower tails of the filter (Patterson *et al.*, 1982). For this simplified filter shape, the value of ERB/f is exactly $4/p$. Hence Eqs. (3) and (6) together give a method of estimating the filter shape at any center frequency between 0.1 and 6.5 kHz. It is then a simple matter to calculate the filter output as a function of center frequency for any given input stimulus. Figure 3 illustrates this graphically for a 1-kHz sinusoid. At the top are shown simplified filter shapes at several center frequencies, with the form given by Eq. (6), and with ERBs given by Eq. (3). For the filter with the lowest center frequency shown, the relative output of the filter in response to the 1-kHz tone is about -40 dB, indicated by the point a. In the lower half of the figure this gives rise to the point a on the excitation pattern; the point has an ordinate of -40 dB, and it is positioned on the abscissa at a frequency corresponding to the center of the lowest filter illustrated. The relative outputs of the other filters shown are indicated, in order of increasing center frequency, by the points b–e. These each give corresponding points on the excitation pattern. The entire excitation pattern was derived by calculating the filter output for filter center frequencies spaced at 10-Hz intervals. Note that although the filter shape is symmetric on a linear frequency scale, the derived excitation pattern is asymmetric.

If desired, the excitation patterns can be plotted as a function of ERB-rate, using Eq. (5). When plotted in this way the shapes of the excitation patterns are independent of stimulus frequency. They resemble the excitation patterns derived from the masked audiograms of narrow-band noise for moderate masker levels (Zwicker and Feldtkeller, 1967). In particular, the low-frequency slope is close to the value suggested by Maiwald (1967).

The method is easily extended to derive the excitation patterns for complex stimuli. The patterns for the individual sinusoidal components are calculated, and are then

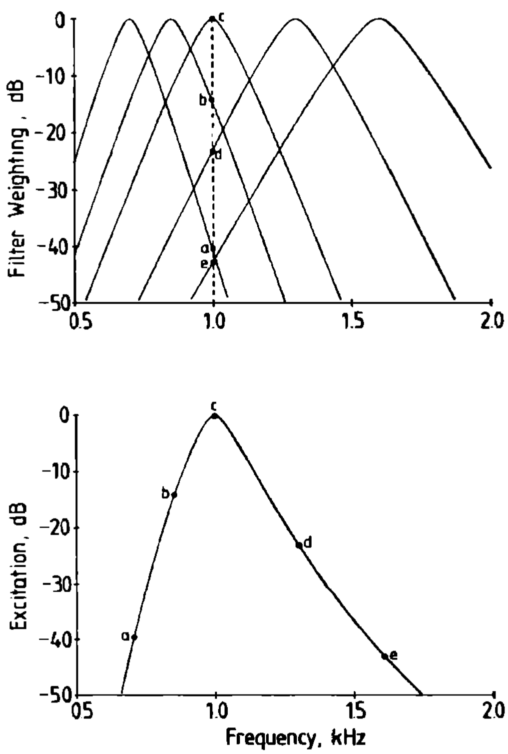

FIG. 3. Illustration of the derivation of an excitation pattern from filter shapes. At the top are shown simplified filter shapes at several center frequencies, with the form given by Eq. (6). The filter shape is assumed symmetrical, but the ERB increases with frequency in accord with Eq. (3). The excitation pattern for a 1-kHz tone (dashed line) is obtained by calculating the output of each filter as a function of filter center frequency.

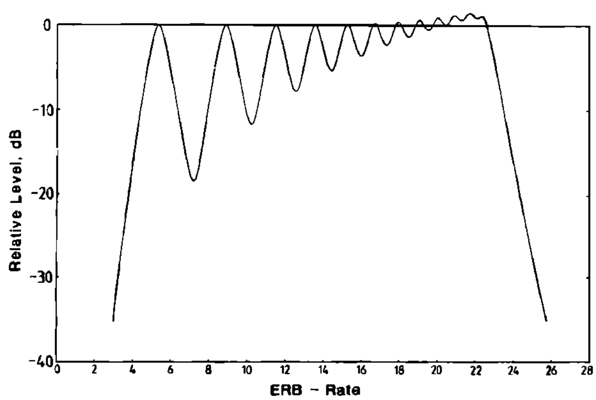

FIG. 4. An excitation pattern calculated for a complex tone containing the first 12 harmonics of a 200-Hz fundamental, all harmonics having equal amplitude. The pattern was calculated by summing the responses to individual components, the summation being done in linear power terms. The abscissa is an ERB-rate scale.

summed. Notice that the summation must be done in linear power, not in dB. An example is given in Fig. 4 for a complex tone consisting of the first 12 harmonics of a 200-Hz fundamental, all components having equal amplitude. The derivation could be made more accurate by taking into account the transmission characteristics of the outer and middle ear.

IV. LIMITATIONS

The ERB values estimated here apply to young listeners at moderate sound levels. The ERB tends to increase with age (Patterson *et al.*, 1982) and with increasing sound level (Weber, 1977; Pick, 1980). The values will also vary somewhat from one listener to another. The assumption of filter symmetry is probably only valid at moderate sound levels (Patterson and Nimmo-Smith, 1980). With increasing sound level the low-frequency slope of the filter becomes shallower and the high-frequency slope becomes somewhat steeper. Note that these changes are the opposite way round from those in excitation patterns. [The reason for this can be illustrated by Fig. 3; the high-frequency side of the excitation pattern (points d and e in the lower half) is determined by the low-frequency slopes of filters with center frequencies above that of the stimulus.] Further work is needed to quantify these changes before excitation patterns can be derived over a wide range of levels. The method of deriving excitation patterns from filter shapes, unlike the more "direct" methods using masked audiograms, is not affected by combination tone detection (Greenwood, 1971) or by off-frequency listening (Patterson, 1976; Verschuure, 1978).

All of the estimates of filter shape and ERB described in this paper were derived for simultaneous masking. Thus the results do not reflect the influence of suppression (Houtgast, 1974). Suppression will have the effect of sharpening the filter shape derived in nonsimultaneous masking (Houtgast, 1977; Moore and Glasberg, 1981), and correspondingly of sharpening excitation patterns (Moore and Glasberg, 1982). The degree of sharpening may vary with stimulus type, and is not predictable on the basis of linear filtering (Glasberg *et al.*, 1983).

ACKNOWLEDGMENTS

We are grateful to Roy Patterson and Ian Nimmo-Smith for many helpful discussions, and for the use of their computer programs. Roy Patterson provided constructive criticism of an earlier version of this paper. We thank Lorraine Evans for typing the manuscript and Roy Hammans for help with the figures. This work was supported by the Medical Research Council.

Fidell, S., Horonjeff, R., Teffeteller, S., and Green, D. M. (1983). "Effective masking bandwidths at low frequencies," J. Acoust. Soc. Am. 73, 628–638 (1983).

Fletcher, H. (1940). "Auditory patterns," Rev. Mod. Phys. 12, 47–65.

Glasberg, B. R., Moore, B. C. J., and Nimmo-Smith, I. (1983). "Comparison of auditory filter shapes derived with three different maskers," J. Acoust. Soc. Am. (submitted).

Greenwood, D. D. (1971). "Aural combination tones and auditory masking," J. Acoust. Soc. Am. 50, 502–543.

Houtgast, T. (1974). "Lateral suppression in hearing," thesis, Free University of Amsterdam (Academische Pers, B. V., Amsterdam).

Houtgast, T. (1977). "Auditory-filter characteristics derived from direct-masking data and pulsation-threshold data with a rippled-noise masker," J. Acoust. Soc. Am. 62, 409–415.

Maiwald, D. (1967). "Die Berechnung von Modulationsschwellen mit Hilfe eines Funktionsschemas," Acustica 18, 193–207.

Moore, B. C. J (1982). *An Introduction to the Psychology of Hearing*, 2nd. ed. (Academic, London).

Moore, B. C. J., and Glasberg, B. R. (1981). "Auditory filter shapes derived in simultaneous and forward masking," J. Acoust. Soc. Am. 70, 1003–1014.

Moore, B. C. J., and Glasberg, B. R. (1982). "Interpreting the role of suppression in psychophysical tuning curves," J. Acoust. Soc. Am. 72, 1375–1379.

Patterson, R. D. (1976). "Auditory filter shapes derived with noise stimuli," J. Acoust. Soc. Am. 59, 640–654.

Patterson, R. D., and Nimmo-Smith, I. (1980). "Off-frequency listening and auditory-filter asymmetry," J. Acoust. Soc. Am. 67, 229–245.

Patterson, R. D., Nimmo-Smith, I., Weber, D. L., and Milroy, R. (1982). "The deterioration of hearing with age: Frequency selectivity, the critical ratio, the audiogram and speech threshold," J. Acoust. Soc. Am. 72, 1788–1803.

Pick, G. F. (1980). "Level dependence of psychophysical frequency resolution and auditory filter shape," J. Acoust. Soc. Am. 68, 1085–1095.

Plomp, R. (1976). *Aspects of Tone Sensation* (Academic, London).

Scharf, B. (1970). "Critical bands," in *Foundations of Modern Auditory Theory, Vol. I*, edited by J. V. Tobias (Academic, New York).

Shailer, M. J., and Moore, B. C. J. (1983). "Gap detection as a function of frequency, bandwidth, and level," J. Acoust. Soc. Am. 74, 467–473 (1983).

Verschuure, J. (1978). "Auditory excitation patterns," thesis, Erasmus University, Rotterdam (W. D. Meinema, B. V., Delft).

Weber, D. L. (1977). "Growth of masking and the auditory filter," J. Acoust. Soc. Am. 62, 424–429.

Zwicker, E., and Feldtkeller, R. (1967). *Das Ohr als Nachrichtenempfanger* (Hirzel, Stuttgart).

Zwicker, E., Flottorp, G., and Stevens, S. S. (1957). "Critical bandwidth in loudness summation," J. Acoust. Soc. Am. 29, 548–557.

Zwicker, E., and Scharf, B. (1965). "A model of loudness summation," Psych. Rev. 72, 3–26.

Zwicker, E., and Terhardt, E. (1980). "Analytical expressions for critical-band rate and critical bandwidth as a function of frequency," J. Acoust. Soc. Am. 68, 1523–1525.