I. INTRODUCTION

*Frequency analysis* refers to the ability of the auditory system to separate or resolve (to a certain extent) the components in a complex sound. For example, if two tuning forks, each tuned to a different frequency, are struck simultaneously, two different tones can usually be heard, one corresponding to each frequency. This ability is also known as *frequency selectivity* and *frequency resolution*; these terms will be used interchangeably in this chapter.

It seems likely that frequency analysis depends to a large extent on the filtering that takes place in the cochlea (see Chapters 2 and 3, this volume). Thus, any complex sound, such as a note produced by a musical instrument or a vowel sound produced by the human voice, undergoes such an analysis at an early stage of auditory processing; the sinusoidal components of the sound are separated, and coded independently in the auditory nerve, provided that their frequency separation is sufficiently large. Furthermore, this stage of analysis cannot be bypassed; all sounds are subject to frequency analysis within the cochlea. Hence, the percept of such sounds as a coherent whole depends upon the representations of the individual components being "reassembled" at some later stage in the auditory system (see Chapters 8, 11 and 12).
Frequency analysis is most often demonstrated and quantified by studying masking. Masking may be regarded as reflecting the limits of frequency analysis. If a sound of a given frequency is masked by another sound with a different frequency, then the auditory system has failed to resolve the two sounds. Hence, by measuring when one sound is just masked by another, it is possible to characterize the frequency analysis capabilities of the auditory system.

II. THE POWER SPECTRUM MODEL AND THE CONCEPT OF THE CRITICAL BAND

Fletcher (1940) measured the threshold for detecting a sinusoidal signal as a function of the bandwidth of a bandpass noise masker. The noise was always centered at the signal frequency, and the noise power density was held constant. Thus, the total noise power increased as the bandwidth increased. This experiment has been repeated several times since then, with similar results (Hamilton, 1957; Greenwood, 1961a; Spiegel, 1981; Schooneveldt & Moore, 1989; Bernstein & Raab, 1990; Moore, Shailer, Hall, & Schooneveldt, 1993). An example of the results, taken from Moore et al. (1993), is given in Figure 1. The threshold of the signal increases at first as the noise bandwidth increases, but then flattens off; further increases in noise bandwidth do not change the signal threshold significantly.

To account for this pattern of results, Fletcher (1940) suggested that the peripheral auditory system behaves as if it contained a bank of bandpass filters, with overlapping passbands. These filters are now called the auditory filters. Fletcher suggested that the signal was detected by attending to the output of the auditory filter centered on the signal frequency. Increases in noise bandwidth result in more noise passing through that filter, as long as the noise bandwidth is less than the filter bandwidth. However, once the noise bandwidth exceeds the filter bandwidth, further increases in noise bandwidth will not increase the noise passing through the filter. Fletcher called the bandwidth at which the signal threshold ceased to increase the critical bandwidth (CB). It is usually assumed that this bandwidth is closely related to the bandwidth of the auditory filter at the same center frequency.

Traditionally, the value of the CB has been estimated by fitting the data with two straight lines, a horizontal line for large bandwidths where thresholds are roughly constant, and a sloping line for smaller bandwidths. However, this approach has two problems. First, the data often show no distinct "break point" at which the slope abruptly decreases to 0. Rather, the slope gradually decreases as the bandwidth increases. Second, small errors of measurement can lead to rather large errors in the estimated CB. Thus, Fletcher's band-widening experiment does not provide a precise way of estimating the bandwidth of the auditory filter (Patterson & Moore, 1986).
Nevertheless, the experiment is important for the concepts to which it gave rise.

Fletcher's experiment led to a model of masking known as the *power-spectrum model*, which is based on the following assumptions:

1. The peripheral auditory system contains an array of linear overlapping bandpass filters.
2. When trying to detect a signal in a noise background, the listener is assumed to use just one filter with a center frequency close to that of the signal. Usually, it is assumed that the filter used is the one that has the highest signal-to-masker ratio at its output.
3. Only the components in the noise that pass through the filter have any effect in masking the signal.
4. The threshold for detecting the signal is determined by the amount of noise passing through the auditory filter; specifically, the threshold is assumed to correspond to a certain signal-to-noise ratio, $K$, at the output of the filter. The stimuli are represented by their long-term power spectra, that is, the relative phases of the components and the short-term fluctuations in the masker are ignored.

We now know that none of these assumptions is strictly correct: The filters are not linear, but are level dependent (Moore & Glasberg, 1987b); listeners can combine information from more than one filter to enhance signal detection (Spiegel, 1981; Buus, Schorer, Florentine, & Zwicker, 1986; see also
Chapter 4); noise falling outside the passband of the auditory filter centered at
the signal frequency can affect the detection of that signal (Hall, Haggard, &
Fernandes, 1984; see also Chapter 7); and fluctuations in the masker can play a
strong role (Patterson & Henning, 1977; Kohlrausch, 1988; Moore, 1988).

These failures of the model do not mean that the basic concept of the
auditory filter is wrong. Indeed, the concept is widely accepted and has
proven to be very useful. Although the assumptions of the model do some-
times fail, it works well in many situations. Nevertheless, it should be
remembered that simplifying assumptions are often made in attempts to
characterize and model the auditory filter.

In analyzing the results of his experiment, Fletcher made a simplifying
assumption. He assumed that the shape of the auditory filter could be ap-
proximated as a simple rectangle, with a flat top and vertical edges. For such
a filter all components within the passband of the filter are passed equally,
and all components outside the passband are removed totally. The width of
the passband of this hypothetical filter would be equal to the CB described
previously. However, it should be emphasized that the auditory filter is not
rectangular and that the data rarely show a distinct break point correspon-
ding to the CB. It is surprising how, even today, many researchers talk about
the critical band as if the underlying filter were rectangular.

Fletcher pointed out that the value of the CB could be estimated indi-
directly, by measuring the power of a sinusoidal signal \( P_r \) required for the
signal to be detected in broadband white noise, given the assumptions of the
power-spectrum model. For a white noise with power density \( N_0 \), the total
noise power falling within the CB is \( N_0 \times \text{CB} \). According to assumption 4,

\[
P_r/(\text{CB} \times N_0) = K
\]

and

\[
\text{CB} = P_r/(K \times N_0).
\]

By measuring \( P_r \) and \( N_0 \), and by estimating \( K \), the value of the CB can be
evaluated.

Fletcher estimated that \( K \) was equal to 1, indicating that the value of the
CB should be equal to \( P_r/N_0 \). The ratio \( P_r/N_0 \) is now usually known as the
critical ratio. Unfortunately, Fletcher's estimate of \( K \) has turned out not to be
accurate. More recent experiments show that \( K \) is typically about 0.4
(Scharf, 1970). Thus, at most frequencies the critical ratio is about 0.4
times the value of the CB estimated by more direct methods, such as the band-
widening experiment. Also, \( K \) varies with center frequency, increasing
markedly at low frequencies, so the critical ratio does not give a correct
indication of how the CB varies with center frequency (Patterson & Moore,

One other aspect of the data in Figure 1 should be noted. If the assump-
tions of the power spectrum model were correct and if the auditory filter were rectangular, then for subcritical bandwidths the signal threshold should increase by 3 dB per doubling of bandwidth; each doubling of bandwidth should lead to a doubling of the noise power passing through the filter, which corresponds to a 3 dB increase in level. In fact, the rate of change is markedly less than this. The exact slope of the function varies from study to study, but it has often been found to be less than the theoretical 3 dB per doubling of bandwidth (Bernstein & Raab, 1990). The deviation from the theoretical value can probably be explained by two factors: the filter is not actually rectangular, but has a rounded top and sloping edges; and for narrow noise bandwidths, the slow fluctuations in the noise have a deleterious effect on detection (Bos & de Boer, 1966; Patterson & Henning, 1977).

III. ESTIMATING THE SHAPE OF THE AUDITORY FILTER

Most methods for estimating the shape of the auditory filter at a given center frequency are based on the assumptions of the power-spectrum model of masking. If the masker is represented by its long-term power spectrum, \( N(f) \), and the weighting function or shape of the auditory filter is \( W(f) \), then the power-spectrum model is expressed by

\[
P_s = K \int_{0}^{\infty} W(f) N(f) \, df.
\]

(3)

where \( P_s \) is the power of the signal at threshold. By manipulating the masker spectrum, \( N(f) \), and measuring the corresponding changes in \( P_s \), it is possible to derive the filter shape, \( W(f) \).

The masker chosen to measure the auditory filter shape should be such that the assumptions of the power-spectrum model are not strongly violated. A number of factors affect this choice. If the masker is composed of one or more sinusoids, beats between the signal and masker (see Chapter 1) may provide a cue to the presence of the signal. This makes sinusoids unsuitable as maskers for estimating the auditory filter shape, since the salience of beats changes as the masker frequency is altered; this violates the assumption of the power-spectrum model that threshold corresponds to a constant signal-to-masker ratio at the output of the auditory filter.

In general, noise maskers are more suitable than sinusoids for estimating the auditory filter shape, because noises have inherent amplitude fluctuations that make beats much less effective as a cue. However, for narrowband noises, which have relatively slow fluctuations, temporal interactions between the signal and masker may still be audible. In addition, the slow fluctuations may strongly influence the detectability of the signal in a way that depends on the difference between the center frequency of the masker
and the frequency of the signal (Buus, 1985; Moore & Glasberg, 1987a). For these reasons, the assumptions of the power-spectrum model are best satisfied using reasonably broadband noise maskers.

A second important consideration in choosing a noise masker for measuring auditory filter shapes is that the filter giving the highest signal-to-masker ratio is not necessarily centered at the signal frequency. For example, if the signal has a frequency of 1 kHz, and the masker spectrum consists entirely of frequencies above 1 kHz, the highest signal-to-masker ratio may occur for a filter centered below 1 kHz. The process of detecting the signal through a filter that is not centered at the signal frequency is called off-frequency listening. In this context, the center frequency of the filter is “off frequency.” Furthermore, if the masker spectrum is concentrated primarily above or below the signal frequency, there may be a range of filter center frequencies over which the signal-to-masker ratio is sufficiently high to give useful information. Under these conditions, the observer may combine information over several auditory filters, rather than listening through a single filter as assumed by the power-spectrum model (Patterson & Moore, 1966; Moore, Glasberg, & Simpson, 1992; for a similar concept applied to intensity discrimination, see Chapter 4).

A. Psychophysical Tuning Curves

The measurement of psychophysical tuning curves (PTCs) involves a procedure that is analogous in many ways to the determination of a neural tuning curve (Chistovich, 1957; Small, 1959); see Chapter 3. The signal is fixed in level, usually at a very low level, say, 10 dB SL. The masker can be either a sinusoid or a narrow band of noise, but a noise is generally preferred, for the reasons given earlier.

For each of several masker center frequencies, the level of the masker needed just to mask the signal is determined. Because the signal is at a low level it is assumed that it will produce activity primarily in one auditory filter. It is assumed further that, at threshold, the masker produces a constant output from that filter, in order to mask the fixed signal. Thus the PTC indicates the masker level required to produce a fixed output from the auditory filter as a function of frequency. Normally, a filter characteristic is determined by plotting the output from the filter for an input varying in frequency and fixed in level. However, if the filter is linear, the same result can be obtained by plotting the input required to give a fixed output. Thus, if linearity is assumed, the shape of the auditory filter can be obtained simply by inverting the PTC. Examples of some PTCs are given in Figure 2; the data are taken from Vogten (1974).

It has been assumed so far that only one auditory filter is involved in the determination of a PTC. However, there is now good evidence that off-
frequency listening can influence PTCs. When the masker frequency is above the signal frequency, the highest signal-to-masker ratio occurs for a filter centered below the signal frequency. Conversely, when the masker frequency is below the signal frequency, the highest signal-to-masker ratio occurs for a filter centered above the signal frequency. In both these cases, the masker level required for threshold is higher than would be the case if off-frequency listening did not occur. When the masker frequency equals the signal frequency, the signal-to-masker ratio is similar for all auditory filters that are excited and off-frequency listening is not advantageous. The overall effect of off-frequency listening is that the PTC has a sharper tip than would be obtained if only one auditory filter were involved (Johnson-Davies & Patterson, 1979; O’Loughlin & Moore, 1981a, 1981b).

One way to limit off-frequency listening is to add to the masker a fixed, low-level noise with a spectral notch centered at the signal frequency (O’Loughlin & Moore, 1981a; Moore, Glasberg, & Roberts, 1984; Patterson & Moore, 1986). Such a masker should make it disadvantageous to use an auditory filter whose center frequency is shifted much from the signal frequency. The effect of using such a noise, in addition to the variable narrowband masker, is illustrated in Figure 3. The main effect is to broaden the tip of the PTC; the slopes of the skirts are relatively unaffected.

A final difficulty in using PTCs as a measure of frequency selectivity is connected with the nonlinearity of the auditory filter. Evidence will be
FIGURE 3  Comparison of PTCs where off-frequency listening is not restricted (triangles) and where it is restricted using a low-level notched noise centered at the signal frequency (squares). (Data from Moore et al., 1984.)

presented later indicating that the auditory filter is not strictly linear, but changes its shape with level. The shape seems to depend more on the level at the input to the filter than on the level at the output (although this is still a matter of some debate, as will be discussed later in this chapter). However, in determining a PTC, the input is varied while the output is held (roughly) constant. Thus, effectively, the underlying filter shape changes as the masker frequency is altered. This can give a misleading impression of the shape of the auditory filter; in particular, it leads to an underestimation of the slope of the lower skirt of the filter and an overestimation of the slope of the upper skirt (Verschuure, 1981a, 1981b; Moore & O'Loughlin, 1986).

B. The Notched-Noise Method

To satisfy the assumptions of the power-spectrum model, it is necessary to use a masker that limits the amount by which the center frequency of the filter can be shifted (off-frequency listening) and that limits the range of filter center frequencies over which the signal-to-masker ratio is sufficiently high to be useful. This can be achieved using a noise masker with a spectral
notch around the signal frequency. For such a masker, the highest signal-to-masker ratio occurs for a filter that is centered reasonably close to the signal frequency, and performance is not improved (or is improved very little) by combining information over filters covering a range of center frequencies (Patterson, 1976; Patterson & Moore, 1986; Moore et al., 1992). The filter shape can then be estimated by measuring signal threshold as a function of the width of the notch.

For moderate noise levels, the auditory filter is almost symmetrical on a linear frequency scale (Patterson, 1974, 1976; Patterson & Nimmo-Smith, 1980; Moore & Glasberg, 1987b). Hence, the auditory filter shape can be estimated using a notched-noise masker with the notch placed symmetrically about the signal frequency. The method is illustrated in Figure 4. For a masker with a notch width of $2\Delta f$, and a center frequency $f_c$, Eq. (3) becomes

$$P_s = KN_0 \int_{f_c-\Delta f}^{f_c+\Delta f} W(f) \, df + KN_0 \int_{f_c+\Delta f}^{\infty} W(f) \, df,$$

(4)

where $N_0$ is the power spectral density of the noise in its passbands. The two integrals on the right-hand side of Eq. (4) represent the respective areas in Figure 4 where the lower and upper noise bands overlap the filter. Because both the filter and the masker are symmetrical about the signal frequency, these two areas are equal. Thus, the function relating $P_s$ to the width of the notch provides a measure of the integral of the auditory filter. Hence, the value of $W(f)$ at a given deviation $\Delta f$ from the center frequency is given by the slope of the threshold function at a notch width of $2\Delta f$.

**FIGURE 4** Schematic illustration of the technique used by Patterson (1976) to determine the shape of the auditory filter. The threshold of the sinusoidal signal is measured as a function of the width of a spectral notch in the noise masker. The amount of noise passing through the auditory filter centered at the signal frequency is proportional to the shaded areas.
When the auditory filter is asymmetric, as it is at high masker levels (see later), the filter shape can still be measured using a notched-noise masker if some reasonable assumptions are made and if the range of measurements is extended to include conditions where the notch is placed asymmetrically about the signal frequency. It is necessary first to assume that the auditory filter shape can be approximated by a simple mathematical expression with a small number of free parameters. Patterson, Nimmo-Smith, Weber, and Milroy (1982) suggested a family of such expressions, all having the form of an exponential with a rounded top, called roex for brevity. The simplest of these expressions was called the roex\((p)\) filter shape. It is convenient to measure frequency in terms of the absolute value of the deviation from the center frequency of the filter, \(f_c\), and to normalize this frequency variable by dividing by the center frequency of the filter. The new frequency variable, \(g\), is

\[
g = |f - f_c|/f_c.
\]

The roex\((p)\) filter shape is then given by

\[
W(g) = (1 + pg) \exp(-pg),
\]

where \(p\) is a parameter that determines both the bandwidth and the slope of the skirts of the auditory filter. The higher the value of \(p\), the more sharply tuned is the filter. The equivalent rectangular bandwidth (ERB) is equal to \(4f_c/p\) (see Chapter 1 for a definition of the ERB). When the filter is assumed to be asymmetric, then \(p\) is allowed to have different values on the two sides of the filter: \(p_l\) for the lower branch and \(p_u\) for the upper branch. The ERB in this case is \(2f_c/p_l + 2f_c/p_u\).

Having assumed this general form for the auditory filter shape, the values of \(p_l\) and \(p_u\) for a particular experiment can be determined by rewriting Eq. (4) in terms of the variable \(g\) and substituting the preceding expression for \(W\); the value of \(p_l\) is used for the first integral, and the value of \(p_u\) for the second. The equation can then be solved analytically; for full details see Patterson et al. (1982) and Glasberg, Moore, and Nimmo-Smith (1984a). Starting values of \(p_l\) and \(p_u\) are assumed, and the equation is used to predict the threshold for each condition (for notches placed both symmetrically and asymmetrically about the signal frequency). The center frequency of the filter is allowed to shift for each condition so as to find the center frequency giving the highest signal-to-masker ratio; this center frequency is assumed in making the prediction for that condition. Standard least-squares minimization procedures are then used to find the values of \(p_l\) and \(p_u\) that minimize the mean-squared deviation between the obtained and predicted values. The minimization is done with the thresholds expressed in decibels. Full details are given in Patterson and Nimmo-Smith (1980), Glasberg et al. (1984a), Patterson and Moore (1986) and Glasberg and Moore (1990).
The roex\( (p) \) filter shape is usually quite successful in predicting the data from notched-noise experiments, except when the thresholds cover a wide range of levels or when the masked thresholds approach the absolute threshold. In such cases there is a decrease in the slope of the function relating threshold to notch width, a decrease that is not predicted by the roex\( (p) \) filter shape. This can be accommodated in two ways. The first involves limiting the dynamic range of the filter, using a second parameter, \( r \). This gives the roex\( (p,r) \) filter shape of Patterson et al. (1982):

\[
W(g) = (1 - r)(1 + pg) \exp(-pg) + r. \tag{7}
\]

As before, \( p \) can have different values for the upper and lower branches of the filter. However, the data can generally be well predicted using the same value of \( r \) for the two sides of the filter (Tyler, Hall, Glasberg, Moore, & Patterson, 1984; Glasberg & Moore, 1986). The method of deriving filter shapes using this expression is exactly analogous to that described earlier.

When the noise level used is relatively high and a large range of notch widths is used, there may be systematic deviations of the data from values predicted by the roex\( (p,r) \) model. In such cases, a better fit to the data can be obtained using a model in which the slope of the filter is assumed to decrease once its attenuation exceeds a certain value. This is achieved using the roex\( (p,w,t) \) model suggested by Patterson et al. (1982). The filter is assumed to be the sum of two exponentials, both of which are rounded:

\[
W(g) = (1 - w)(1 + pg) \exp(-pg) + w(1 + tg) \exp(-tg) \tag{8}
\]

The parameter \( t \) determines the slope of the filter at large deviations from the center frequency, and the parameter \( w \) determines the point at which the shallower "tail" takes over from the steeper central passband. In principle, all three parameters, \( p, w, \) and \( t \), could be different for the two sides of the filter, giving six free parameters all together. In practice, it has been found that the results can be well fitted by assuming that one side of the filter is a "stretched" version of the other side (Patterson & Nimmo-Smith, 1980; Glasberg, Moore, Patterson, & Nimmo-Smith, 1984b). In this case, \( w \) is assumed to be the same for the two sides of the filter, and the ratio \( p/t \) is assumed to be the same for the two sides of the filter. This reduces the number of free parameters to four.

One limitation of the notched-noise method occurs when the auditory filter is markedly asymmetric, as it is, for example, at high sound levels. In such cases, the method does not define the sharper side of the filter very well. As a rule of thumb, when the value of \( p \) for one side of the filter is more than twice that for the other, the slope of the steeper side is very poorly determined.

A second potential problem with the method is that components within the upper band of noise may interact to produce combination products
whose frequencies lie within the notch in the noise. Such combination products are produced by nonlinear processes within the cochlea, and they occur even when the input is at low to moderate sound levels (Greenwood, 1971; Smoorenburg, 1972a, 1972b). The effect of this is that the upper band of noise may produce more masking than would be the case if no combination products were present. This can result in a derived filter shape with a shallower upper skirt. However, the effect on the derived filter shape is usually small (Moore, Glasberg, van der Heijden, Houtsma, & Kohlrausch, 1995).

C. The Rippled-Noise Method

Several researchers have estimated auditory filter shapes using rippled noise, sometimes also called comb-filtered noise, as a masker. This is produced by adding white noise to a copy of itself that has been delayed by T seconds. The resulting spectrum has peaks spaced at 1/T Hz, with minima in between. When the delayed version of the noise is added to the original in phase, the first peak in the spectrum of the noise occurs at 0 Hz; this noise is referred to as cosine+. When the polarity of the delayed noise is reversed, the first peak is at 0.5/T Hz; this is referred to as cosine−. The sinusoidal signal is usually fixed in frequency, and the values of T are chosen so that the signal falls at either a maximum or minimum in the masker spectrum; the signal threshold is measured for both cosine+ and cosine− noise for various ripple densities (different values of T).

The auditory filter shape can be derived from the data either by approximating the auditory filter as a Fourier series (Houtgast, 1977; Pick, 1980) or by a method similar to that described for the notched-noise method (Glasberg et al., 1984a; Patterson & Moore, 1986). The filter shapes obtained in this way are generally similar to those obtained using the notched-noise method, although they tend to have a slightly broader and flatter top (Glasberg et al., 1984a). The method seems to be quite good for defining the shape of the tip of the auditory filter, but it does not allow the auditory filter shape to be measured over a wide dynamic range.

D. Allowing for the Transfer Function of the Outer and Middle Ear

The transfer function of the outer and middle ear varies markedly with frequency, particularly at very low and high frequencies. Clearly this can have a significant influence on measures of frequency selectivity. For example, if one of the bands of noise in a notched-noise experiment is very low or high in center frequency, it will be strongly attenuated by the middle ear and so will not do much masking. It is possible to conceive of the auditory filter
shape as resulting from the overall response properties of the outer and middle ear and the cochlea. However, it is theoretically more appealing to conceive of the auditory filter as resulting from processes occurring after the outer and middle ear. The effect of the outer and middle ear can be thought of as a fixed frequency-dependent attenuation applied to all stimuli before auditory filtering takes place.

If this is the case, then the frequency-dependent attenuation should be taken into account in the fitting procedure for deriving filter shapes. Essentially, the spectra of the stimuli at the input to the cochlea have to be calculated by assuming a certain form for the frequency-dependent transfer. The fitting procedure then has to work on the basis of these "corrected" spectra. In practice, this means that the integral in Eq. (3) cannot be solved analytically, but has to be evaluated numerically. Glasberg and Moore (1990) have considered several possible types of "correction." One is appropriate for stimuli presented in a free field (e.g., via a loudspeaker in an anechoic chamber) or via earphones designed to have a free-field response, such as the Sennheiser HD414 or the Etymotic Research ER4. Another is appropriate for earphones designed to give a flat response at the eardrum, such as the Etymotic Research ER2. In both cases, the "correction" may be modified to take into account the specific properties of the transducers used. Glasberg and Moore (1990) list a computer program for deriving auditory filter shapes from notched-noise data that includes the option of using "corrections" to allow for the transfer function of the outer and middle ear.

E. An Example of Measurement of the Auditory Filter Shape

Figure 5 shows an example of data obtained using the notched-noise method, and of the filter shape obtained. The data are for a normally hearing subject and a signal frequency of 200 Hz. In the top panel, signal thresholds are plotted as a function of the width of the spectral notch in the noise masker. Asterisks indicate conditions where the spectral notch was placed symmetrically about the signal frequency; the notch width, Δ, is specified as the deviation of each edge of the notch from the signal frequency, divided by the signal frequency. The left-pointing arrows indicate conditions where the lower edge of the notch was 0.2 units farther from the signal frequency than the upper edge. The right-pointing arrows indicate conditions where the upper edge of the notch was 0.2 units farther from the signal frequency than the lower edge. Moving the lower edge of the notch farther from the signal frequency has a greater effect than moving the upper edge farther from the signal frequency.

The lines in the top panel are the fitted values derived from the roex(p,r) model, as described by Glasberg and Moore (1990). The model fits the data well. The derived filter shape is shown in the bottom panel. The filter is
FIGURE 5  The top panel shows thresholds for a 200 Hz signal as a function of the width of a notch in a noise masker. The value on the abscissa is the deviation of the nearer edge of the notch from the signal frequency, divided by the signal frequency, represented by the symbol $\Delta$. Asterisks (*) indicate conditions where the notch was placed symmetrically about the signal frequency. Right-pointing arrows indicate conditions where the upper edge of the notch was
somewhat asymmetric, with a shallower lower branch. The ERB is 48 Hz, which is typical at this center frequency.

IV. SUMMARY OF THE CHARACTERISTICS OF THE AUDITORY FILTER

A. Variation with Center Frequency

Moore and Glasberg (1983b) presented a summary of experiments measuring auditory filter shapes using symmetric notched-noise maskers. All of the data were obtained at moderate noise levels and were analyzed using the roex(p,r) filter shape. Glasberg and Moore (1990) updated that summary, including results that extend the frequency range of the measurements and data from experiments using asymmetric notches. The ERBs of the filters derived from the data available in 1983 are shown as asterisks in Figure 6. The dashed line shows the equation fitted to the data in 1983. Other symbols show ERBs estimated in more recent experiments, as indicated in the figure.

The solid line in Figure 6 provides a good fit to the ERB values over the whole frequency range tested. It is described by the following equation:

\[ \text{ERB} = 24.7(4.37F + 1), \]

where \( F \) is center frequency in kHz. This equation is a modification of one originally suggested by Greenwood (1961b) to describe the variation of the CB with center frequency. He based it on the assumption that each CB corresponds to a constant distance along the basilar membrane. Although the constants in Eq. (9) differ from those given by Greenwood, the form of the equation is the same as his. Each ERB corresponds to a distance of about 0.89 mm on the basilar membrane.

It should be noted that the function specified by Eq. (9) differs somewhat from the "traditional" critical band function (Zwicker, 1961), which flattens off below 500 Hz at a value of about 100 Hz. The traditional function was obtained by combining data from a variety of experiments. However, the data were sparse at low frequencies, and the form of the function was strongly influenced by measures of the critical ratio. As described earlier, the critical ratio does not provide a good estimate of the CB, particularly at low frequencies. It seems clear that the CB does continue to decrease below 500 Hz.

---

0.2 units farther away from the signal frequency than the lower edge. Left-pointing arrows indicate conditions where the lower edge of the notch was 0.2 farther away than the upper edge. The fact that the left-pointing arrows are markedly below the right-pointing arrows indicates that the filter is asymmetric. The bottom panel shows the auditory filter shape derived from the data. (From Moore et al., 1990.)
FIGURE 6  Estimates of the auditory filter bandwidth from a variety of experiments, plotted as a function of center frequency. The dashed line represents the equation suggested by Moore and Glasberg (1983b). The solid line represents the equation suggested by Glasberg and Moore (1990). (Adapted from Glasberg & Moore, 1990.)

It is sometimes useful to plot experimental data and theoretical functions on a frequency-related scale based on units of the CB or ERB of the auditory filter. A traditional scale of this type is the Bark scale (Zwicker & Terhardt, 1980) where the number of Barks is indicated by the symbol \( z \). A good approximation to the traditional Bark scale is

\[
z = \left[ 26.8 / (1 + 1.96 / F) \right] - 0.53
\]

(Traunmuller, 1990). A scale based on the ERB of the auditory filter, derived from Eq. (9), is

\[
\text{Number of ERBs, } E = 21.4 \log_{10}(4.37F + 1)
\]

Auditory filter bandwidths for young, normally hearing subjects vary relatively little across subjects; the standard deviation of the ERB is typically about 10% of its mean value (Moore, 1987; Moore et al., 1990). However, the variability tends to increase at very low frequencies (Moore et al., 1990) and at very high frequencies (Patterson et al., 1982; Shailer, Moore, Glasberg, Watson, & Harris, 1990).
B. Variation with Level

If the auditory filter were linear, then its shape would not vary with the level of the noise used to measure it. Unfortunately, this is not the case. Moore and Glasberg (1987b) presented a summary of measurements of the auditory filter shape using maskers with notches placed asymmetrically about the signal frequency. They concluded that the lower skirt of the filter becomes less sharp with increasing level, while the higher skirt becomes slightly steeper. Glasberg and Moore (1990) reanalyzed the data from the studies summarized in that paper, but using a modified fitting procedure including "corrections" for the transfer function of the middle ear. They also examined the data presented in Moore et al. (1990) and Shailer et al. (1990). The reanalysis led to the following conclusions:

1. The auditory filter for a center frequency of 1 kHz is roughly symmetric on a linear frequency scale when the level of the noise is approximately 51 dB/ERB. This corresponds to a noise spectrum level of about 30 dB. The auditory filters at other center frequencies are approximately symmetric when the effective input levels to the filters are equivalent to the level of 51 dB/ERB at 1 kHz (after making allowance for changes in relative level produced by passage of the sound through the outer and middle ear).

2. The low-frequency skirt of the auditory filter becomes less sharp with increasing level. The variation can be described in terms of the parameter $p_I$. Let $X$ denote the effective input level in dB/ERB. Let $p_{I(X)}$ denote the value of $p_I$ at level $X$. Then,

$$p_{I(X)} = p_{I(51)} - 0.38 \left( \frac{p_{I(51)}}{p_{I(51,1kHz)}} \right) (X - 51),$$

where $p_{I(51)}$ is the value of $p$ at that center frequency for an effective input noise level of 51 dB/ERB and $p_{I(51,1kHz)}$ is the value of $p_I$ at 1 kHz for an input level of 51 dB/ERB.

3. Changes in slope of the high-frequency skirt of the filter with level are less consistent. At medium center frequencies (1–4 kHz) there is a trend for the slope to increase with increasing level, but at low center frequencies there is no clear trend with level, and the filters at high center frequencies show a slight decrease in slope with increasing level.

These statements are based on the assumption that, although the auditory filter is not linear, it may be considered as approximately linear at any given noise level. Furthermore, the sharpness of the filter is assumed to depend on the input level to the filter, not the output level. This issue is considered further later. Figure 7 illustrates how the shape of the auditory filter varies with input level for a center frequency of 1 kHz.
As mentioned earlier, the notched-noise method does not give a precise estimate of the slope of the steeper side of the filter when the filter is markedly asymmetric. This is a particular problem at high sound levels, where the lower branch becomes very shallow. Thus, at high levels, there may well be significant errors in the estimates of the sharpness of the high-frequency side of the filter.

V. MASKING PATTERNS AND EXCITATION PATTERNS

In the experiments described so far, the frequency of the signal was held constant, while the masker was varied. These experiments are most appropriate for estimating the shape of the auditory filter at a given center frequency. However, many of the early experiments on masking did the opposite; the signal frequency was varied while the masker was held constant.

Wegel and Lane (1924) reported the first systematic investigation of the
masking of one pure tone by another. They determined the threshold of a signal with adjustable frequency in the presence of a masker with fixed frequency and intensity. The function relating masked threshold to the signal frequency is known as a masking pattern, or sometimes as a masked audiogram. The results of Wegel and Lane were complicated by the occurrence of beats when the signal and masker were close together in frequency. To avoid this problem, later experimenters (Egan & Hake, 1950; Fastl, 1976a) have used a narrow band of noise as either the signal or the masker.

The masking patterns obtained in these experiments show steep slopes on the low-frequency side, of between 80 and 240 dB/octave for pure tone masking and 55–190 dB/octave for narrowband noise masking. The slopes on the high-frequency side are less steep and depend on the level of the masker. A typical set of results is shown in Figure 8. Notice that on the high-frequency side the slopes of the curves tend to become shallower at high levels. Thus, if the level of a low-frequency masker is increased by, say, 10 dB, the masked threshold of a high-frequency signal is elevated by more than 10 dB; the amount of masking grows nonlinearly on the high-frequency side. This has been called the upward spread of masking.

The masking patterns do not reflect the use of a single auditory filter. Rather, for each signal frequency the listener uses a filter centered close to the signal frequency. Thus the auditory filter is shifted as the signal frequency is altered. One way of interpreting the masking pattern is as a crude

---

**FIGURE 8** Masking patterns (masked audiograms) for a narrow band of noise centered at 410 Hz. Each curve shows the elevation in threshold of a sinusoidal signal as a function of signal frequency. The overall noise level for each curve is indicated in the figure. (Data from Egan & Hake, 1950.)
indicator of the excitation pattern of the masker. The excitation pattern of a sound is a representation of the activity or excitation evoked by that sound as a function of characteristic frequency (Zwicker, 1970). In the case of a masking pattern, one might assume that the signal is detected when the excitation it produces is some constant proportion of the excitation produced by the masker in the frequency region of the signal. Thus the threshold of the signal as a function of frequency is proportional to the masker's excitation level. The masking pattern should be parallel to the excitation pattern of the masker, but shifted vertically by a small amount. In practice, the situation is not so straightforward, since the shape of the masking pattern is influenced by factors such as off-frequency listening and the detection of combination tones produced by the interaction of the signal and the masker (Greenwood, 1971).

A. Relationship of the Auditory Filter to the Excitation Pattern

Moore and Glasberg (1983b) have described a way of deriving the shapes of excitation patterns using the concept of the auditory filter. They suggested that the excitation pattern of a given sound can be thought of as the output of the auditory filters as a function of their center frequency. This idea is illustrated in Figure 9. The upper portion of the figure shows auditory filter shapes for five center frequencies. Each filter is symmetrical on the linear frequency scale used, but the bandwidths of the filters increase with increasing center frequency, as illustrated in Figure 6. The dashed line represents a 1 kHz sinusoidal signal whose excitation pattern is to be derived. The lower panel shows the output from each filter in response to the 1 kHz signal, plotted as a function of the center frequency of each filter; this is the desired excitation pattern.

To see how this pattern is derived, consider the output from the filter with the lowest center frequency. This has a relative output in response to the 1 kHz tone of about −40 dB, as indicated by point a in the upper panel. In the lower panel, this gives rise to the point a on the excitation pattern; the point has an ordinate value of −40 dB and is positioned on the abscissa at a frequency corresponding to the center frequency of the lowest filter illustrated. The relative outputs of the other filters are indicated, in order of increasing center frequency, by points b to e, and each leads to a corresponding point on the excitation pattern. The complete excitation pattern was actually derived by calculating the filter outputs for filters spaced at 10 Hz intervals. In deriving the excitation pattern, excitation levels were expressed relative to the level at the tip of the pattern, which was arbitrarily labeled 0 dB. To calculate the excitation pattern for a 1 kHz tone with a level of, say, 60 dB, the level at the tip would be labeled 60 dB, and all other excitation levels would correspondingly be increased by 60 dB.
FIGURE 9  An illustration of how the excitation pattern of a 1 kHz sinusoid can be derived by calculating the outputs of the auditory filters as a function of their center frequency. The top half shows five auditory filters, centered at different frequencies, and the bottom half shows the calculated excitation pattern. (From Moore & Glasberg, 1983b.)

Note that, although the auditory filters were assumed to be symmetric on a linear frequency scale, the derived excitation pattern is asymmetric. This happens because the bandwidth of the auditory filter increases with increasing center frequency. As pointed out by Patterson (1974), the increase in auditory filter bandwidth with frequency can also explain why masking patterns are asymmetric when the auditory filter itself is roughly symmetric.

B. Changes in Excitation Patterns with Level

One problem in calculating excitation patterns from filter shapes is how to deal with the level dependence of the auditory filter. It seems clear that the shape of the auditory filter does change with level, the major change being a decrease in sharpness of the low-frequency side with increasing level. However, to determine the effect of this on excitation patterns, it is necessary to decide exactly what aspect of level determines the filter shape.

As one approach to this problem, Moore and Glasberg (1987b) consid-
quired whether the shape of the auditory filter depends primarily on the level of the input to the filter or on the level of the output of the filter. This way of posing the problem may well be over-simplistic, especially when the input is a complex sound. However, the question may be a reasonable one for a simple stimulus such as a sinusoid.

To examine this question, Moore and Glasberg (1987b) calculated excitation patterns for sinusoids assuming that the shape of the auditory filter depended either on the input level to the filter or the output level from the filter. They assumed that the filter had the form of the \( \text{roex}(p) \) filter described earlier. An example of the results is shown in Figure 10, for a 1 kHz sinusoid at levels ranging from 20 to 90 \( \text{dB} \) SPL. For the left-hand panels the output level of each filter was assumed to determine its shape. For the right-hand panels the input level was assumed to determine the shape.

As described previously, the shapes of excitation patterns for narrowband stimuli as a function of level can be determined approximately from their masking patterns. The patterns shown in the right-hand panels of Figure 10 closely resemble masked audiograms at similar masker levels, whereas those in the left-hand panels are very different in form and do not show the classic “upward spread of masking.” Moore and Glasberg (1987b) concluded that the critical variable determining the auditory filter shape is the input level to the filter.

Rosen, Baker, and Kramer (1992) have taken the opposite viewpoint, arguing that the sharpness and asymmetry of the auditory filter are determined by the level at the output of the auditory filter. Their argument is based on an analysis of data from a notched-noise experiment conducted using three fixed masker spectrum levels and three fixed signal levels, with a signal frequency of 2000 Hz. Although their data can be fitted well on the assumption that the shape of the auditory filter is determined by its output level, this assumption leads to problems with other types of stimuli. One such problem was noted earlier: excitation patterns for narrowband stimuli calculated using this assumption have the “wrong” shape. In addition, the equations given by Rosen et al. to define the variation of the filter parameters (such as \( p_l \) and \( p_u \)) with level lead to substantial errors in predicting the masking produced by sounds such as low-frequency narrow bands of noise. Thus, even though their conclusion is consistent with the restricted data set analyzed, it does not seem to be generally applicable.

At present, it appears that the data from a range of experiments are best accounted for on the assumption that the shape of the auditory filter is controlled by the level at its input, rather than the level at its output. Unfortunately, the situation cannot be as simple as this. For sounds with complex broadband spectra, it seems likely that only components that produce a significant output from a given auditory filter have any influence in determining the shape of that filter. Moore and Glasberg (1987b) suggested that
the shape of the auditory filter may be determined primarily by the input level of the component that produces the greatest output from the filter. When the input spectrum is continuous or contains closely spaced components, they suggested that the power of the components should be summed with a range of 1 ERB (around the frequencies of the components in question) to determine the effective input level.

Further analysis of these ideas, by Brian Glasberg and myself (unpublished results), reveals a significant problem. When the input is composed of a few discrete sinusoidal components with different levels, excitation patterns calculated in this way can show discontinuities. This happens because the component producing the greatest output from the auditory filters changes as the center frequency of the filters changes. At center frequencies, where a change in the dominant component occurs, there is an abrupt jump in the calculated sharpness of the filter, and this leads to a discontinuity in the
excitation pattern. It seems very unlikely that such discontinuities would occur in the auditory system.

The program for calculating excitation patterns published by Glasberg and Moore (1990) is based on a different set of assumptions, although these assumptions were not explicitly stated in that paper. The assumptions are as follows. Each component (or group of components if several lie within 1 ERB) gives rise to an excitation pattern whose spread is determined by the level of the component (or group). Thus, the extent to which a given auditory filter is excited by the component is determined by the input level of the component. This idea is similar to that proposed by Zwicker (Zwicker & Feldtkeller, 1967; Zwicker, 1970). The simplest way to calculate the effective level at a given center frequency is to sum the powers of components within ±0.5 ERB of that frequency. Zwicker used a similar approach and referred to the resulting quantity as *psychoacoustical incitation*. This summed power determines the spread of excitation from that frequency. The assumption that component powers are summed within a rectangular ERB (around the frequencies of the components) is unrealistic, and it is probably more satisfactory to perform the summation using a rounded-exponential weighting function.

In cases where the stimulus is complex, with components spread over several ERBs, it is assumed that excitation patterns arising from components lying in different ERBs are summed in terms of linear power. Again, Zwicker (Zwicker & Feldtkeller, 1967; Zwicker, 1970) made a similar assumption, and the assumption is implicit in some models for calculating loudness and for predicting intensity discrimination performance (see Chapter 4).

There is a consequence of these assumptions that may, at first sight, appear paradoxical. A given auditory filter may have several different sharpnesses (for example, several values of \( p_i \)) simultaneously. The value of \( p_i \) is calculated separately for each group of components (summed within an ERB around the frequencies of the components). However, this is not so strange in terms of excitation patterns. The program by Glasberg and Moore (1990) works by calculating excitation patterns from filter shapes, but, for the auditory system, the excitation pattern may be "primary" in some sense. If the spread of the excitation pattern produced by a given component is determined by the level of that component, then the effective shapes of the auditory filters excited by that component have to vary depending on the level of the component. And if two components are present simultaneously (separated by more than 1 ERB), each will give rise to an excitation pattern whose spread is determined by the level of the respective component.

Many of these problems arise from the fact that auditory filtering is inherently nonlinear, but the models used are based on quasi-linear filtering.
More appropriate models may lead to a better understanding of the factors controlling the selectivity of the auditory filters.

VI. THE ADDITIVITY OF MASKING AND EXCESS MASKING

Many years ago, Green (1967) measured the masking of a gated sinusoidal signal produced by a continuous sinusoidal masker of the same frequency and, separately, by a broadband continuous noise. He adjusted the levels of the two maskers so that they produced equal amounts of masking. He then measured the amount of masking produced by combining the two equally effective maskers. If the threshold of the signal were determined simply by the power of the masker at the output of the auditory filter centered at the signal frequency, the combined maskers should produce 3 dB more masking than either masker alone. In fact, the amount of extra masking produced by the combined masker was usually markedly greater than the expected 3 dB. The amount of masking above 3 dB is sometimes referred to as excess masking.

In the last two decades, many cases of excess masking have been reported; the amount of masking produced by two maskers that are equally effective when presented individually is often more than 3 dB greater than the masking produced by each masker alone (Lutfi, 1983, 1985; Humes & Jesteadt, 1989). Two general approaches have been taken to explain the excess masking.

In one approach, it is assumed that the detection cues used by the subject differ for the two individual maskers (Bilger, 1959; Green, 1967; Moore, 1985). For example, if one masker is a continuous broadband noise, as in Green’s (1967) experiment, the detection cue may be the additional energy produced by the signal at the output of the auditory filter centered on the signal frequency. If the other masker is a continuous sinusoid, the detection cue may be a fluctuation in the envelope of the auditory filter output. This cue is very effective, so when the two maskers produce equal amounts of masking, the sinusoid produces a considerably greater output from the auditory filter than the noise. When the two maskers are combined, the noise introduces random fluctuations in amplitude that make it difficult to use the fluctuation cue previously employed with the sinusoidal masker. However, the energy cue previously used with the noise masker is also less effective, because the sinusoidal masker considerably increases the energy at the filter output. Hence, considerable excess masking occurs. According to this type of explanation, excess masking occurs when the detection processes or cues following the auditory filter are different for the two maskers used and when each masker renders less effective the cue used with the other masker.

A second type of explanation has been proposed by Lutfi (1983, 1985) and Humes and Jesteadt (1989). These researchers have suggested that the
effects of the two maskers are summed after each has undergone a compressive nonlinear transformation. A model of this type was originally proposed by Penner and Shiffrin (1980) to account for the excess masking obtained with pairs of maskers that do not overlap in time (see Chapter 6, Section VIII of this chapter, and Oxenham & Moore, 1994), but Lutti and Humes and Jesteadt have argued that a similar model can be applied to maskers that do overlap in time.

In the model of Humes and Jesteadt (1989), the compressed internal effect of each masker is given by the following transform:

\[ i_{mt} = (10^{(mt/10)})^\alpha - (10^{(qt/10)})^\alpha \]  

(13)

where \( i_{mt} \) reflects the internal effect of the masker, \( mt \) is the masked threshold of the signal, \( qt \) is the absolute threshold of the signal, and \( \alpha \) is a parameter that is adjusted to fit the data. The value of \( \alpha \) reflects the amount of compression: the smaller the value of \( \alpha \), the greater is the compression. The effects of combined maskers are assumed to be summed after the individual maskers have been subjected to the nonlinear compressive transform of Eq. (13). An inverse transform is then applied to the sum to predict the signal threshold for the combined maskers. If \( \alpha < 1 \), excess masking is predicted. If \( \alpha = 1 \), there is no compression, and no excess masking is predicted.

The basic concept behind this type of model is that the excess masking arises from a fundamental physiological property of the auditory system: it is assumed that all stimuli are subject to a compressive nonlinearity in the peripheral auditory system. The amount of excess masking should be determined by the characteristics of this compressive nonlinearity, and excess masking should always occur.

Although the model proposed by Humes and Jesteadt can account for a large body of experimental data, several problems can be identified with this approach. First, it is assumed that the two maskers are compressed independently. It is hard to see how this could happen for two maskers that are presented simultaneously. Consider Green’s (1967) experiment described earlier. Presumably, the masking produced by both the noise masker and the sinusoidal masker depends on the extent to which those maskers produce an output from the auditory filter centered at the signal frequency. If the two maskers are presented simultaneously and excite the same auditory filter, how can the effects of the two maskers be subjected to independent compressive nonlinearities?

A second problem arises from the assumption that the amount of excess masking is determined by the characteristics of the compressive nonlinearity, and that this nonlinearity reflects a physiological property of the peripheral auditory system. If this were the case, then the form of the nonlinearity (the value of \( \alpha \)) needed to fit the data should be similar regardless of the specific combination of maskers used. In fact, the required value
of α varies markedly across different data sets (Humes & Jesteadt, 1989). This is not consistent with the idea that the excess masking arises from a fixed physiological property of the auditory system.

A third problem is raised by a study of Oxenham and Moore (1995) on excess masking in subjects with cochlear hearing loss. It is known that the compressive nonlinearity on the basilar membrane is reduced in such subjects (see Chapter 2). Hence, if the model of Humes and Jesteadt (1989) were correct, excess masking should be reduced. In an experiment similar to that of Green (1967), Oxenham and Moore showed that excess masking in simultaneous masking was similar for normally hearing subjects and subjects with cochlear hearing loss. However, excess masking produced by combined forward and backward masking was absent in subjects with cochlear hearing loss, while it was marked in normally hearing subjects. These results suggest that compressive nonlinearities in the peripheral auditory system can account for excess masking in nonsimultaneous masking, but not in simultaneous masking.

A fourth problem is that the model predicts that excess masking will always occur. This is not the case. For example, Moore (1985) determined masking functions (signal threshold versus masker level) separately for two narrowband noise maskers, one centered at 1.4 kHz and the other at 1.6 kHz. The signal was always a 2 kHz sinusoid. The masking functions were used to select pairs of maskers (i.e., the two bands of noise presented together) that would be equally effective if presented individually. When the two bands of noise had independent envelopes, excess masking occurred, as predicted by the model of Humes and Jesteadt. However, when the two bands of noise had the same envelope (i.e., the bands were comodulated; see Chapter 7), no excess masking occurred; the combined maskers produced 3 dB more masking than each masker individually. This is not consistent with the model of Humes and Jesteadt. Green (1967) and Bilger (1959) also reported cases where no excess masking occurred. In my opinion, these results are sufficient to demonstrate that the model cannot be correct.

Moore (1985) explained the pattern of his results in the following way. When a single narrowband noise masker is used, subjects exploit the envelope fluctuations in the masker to improve signal detection. They do this by detecting the signal in minima of the masker envelope and by comparing temporal patterns of modulation across different auditory filters (Buus, 1985; Moore & Glasberg, 1987a). When the signal is absent, the pattern of modulation is similar at the outputs of all auditory filters. When the signal is present, the pattern of modulation at the output of filters tuned close to the signal frequency differs from that in the remaining filters. This across-filter disparity provides a detection cue; see Chapter 7 for further discussion of this topic. When two uncorrelated narrowband noise maskers are combined, this cue is disrupted and excess masking occurs. When two narrow-
band maskers with the same envelope are combined, the cue is preserved, and the combined masking is correctly predicted by a linear power summation of the effects of the two maskers.

In summary, it seems that most cases of excess masking in simultaneous masking can be explained by a detailed consideration of the detection cues available to and used by the subjects for the individual maskers and for the combined maskers. Alternative models, assuming that the effects of the individual maskers are subject to a compressive nonlinearity before the effects are combined, lead to a number of conceptual difficulties, and cannot account for cases where excess masking does not occur.

Finally, a general problem in the study of excess masking should be noted. It is very hard to formulate a set of rules defining whether a given stimulus should be described as one masker or as two (or even more). For example, should a band of noise extending from 500 to 1000 Hz be described as a single masker or as two maskers, one extending from 500 to 750 Hz, and the other from 750 to 1000 Hz? In many cases, the definition of what constitutes one masker or two maskers appears completely arbitrary. This problem is discussed by Humes and Jesteadt (1989) and by Humes, Lee, and Jesteadt (1992). The latter proposed that maskers that do not overlap within the critical band centered on the signal frequency should be treated as two separate maskers; in such cases they predicted that excess masking should occur. On the other hand, they suggested that for simultaneous maskers with spectral overlap within the critical band centered on the signal frequency there is effectively only one masker, and no excess masking should occur. While this rule is consistent with the data presented by Humes et al. (1992), it clearly does not always work. For example, Green (1967) showed that excess masking could occur for spectrally overlapping maskers; and Moore (1985) showed that linear additivity of masking could occur for pairs of maskers that did not overlap spectrally either with each other or with the signal.

VII. PHENOMENA REFLECTING THE INFLUENCE OF AUDITORY FILTERING

Many aspects of auditory perception are affected by auditory filtering. Some of these are considered in other chapters, especially Chapters 4 and 8. This section describes just a few examples of these phenomena.

A. The Threshold of Complex Sounds

Gässler (1954) measured the threshold for detecting multicomponent complexes consisting of evenly spaced sinusoids. The complexes were presented both in quiet and in a special background noise, chosen to give the same
masked threshold for each component in the signal. As the number of components in a complex was increased, the threshold, specified in terms of total energy, remained constant until the overall spacing of the tones reached a certain bandwidth, the CB for threshold. Thereafter the threshold increased by about 3 dB per doubling of bandwidth. The CB for a center frequency of 1 kHz was estimated to be about 180 Hz. These results were interpreted as indicating that the energies of the individual components in a complex sound will sum, in the detection of that sound, provided the components lie within a CB. When the components are distributed over more than one CB, detection is based on the single band giving the highest detectability.

Other data are not in complete agreement with those of Gässler. Indeed, most subsequent experiments have failed to replicate Gässler’s results. For example, Spiegel (1981) measured the threshold for a noise signal of variable bandwidth centered at 1 kHz in a broadband background noise masker. The threshold for the signal as a function of bandwidth did not show a break point corresponding to the CB, but increased monotonically as the bandwidth increased beyond 50 Hz. The slope beyond the CB was close to 1.5 dB per doubling of bandwidth. Higgins and Turner (1990) have suggested that the discrepancy may be explained by the fact that Gässler widened the bandwidth, keeping the upper edge of the complex fixed in frequency, while Spiegel used stimuli with a fixed center frequency. However, other results clearly show that the ear is capable of combining information over bandwidths much greater than the CB (Buus et al., 1986; Langhans & Kohlrausch, 1992). For example, Buus et al. (1986) showed that multiple widely spaced sinusoidal components were more detectable than any of the individual components.

These results should not be interpreted as evidence against the concept of the auditory filter. They do indicate, however, that detection of complex signals may not be based on the output of a single auditory filter. Rather, information can be combined across filters to improve performance.

B. Sensitivity to the Relative Phase

An amplitude modulated (AM) sinewave with modulation index $m$ and a frequency modulated (FM) sinewave with modulation index $\beta$ may each be considered as composed of three sinusoidal components, corresponding to the carrier frequency and two sidebands (an FM wave actually contains many components but for small modulation indices only the first two sidebands are important); see Chapter 1. When the modulation indices are numerically equal ($m = \beta$) and the carrier frequencies and modulation frequencies are the same, the components of an AM wave and an FM wave are identical in frequency and amplitude, the only difference between them
being in the relative phase of the components. If, then, the two types of wave are perceived differently, the difference is likely to arise from a sensitivity to the relative phase of the components.

Zwicker (1952), Schorer (1986), and Sek (1994) have measured one aspect of the perception of such stimuli, namely, the just-detectable amounts of amplitude or frequency modulation, for various rates of modulation. They found that, for high rates of modulation, where the frequency components were widely spaced, the detectability of FM and AM was equal when the components in each type of wave were of equal amplitude \( m = \beta \). However, for low rates of modulation, when all three components fell within a narrow frequency range, AM could be detected when the relative levels of the sidebands were lower than for a wave with a just-detectable amount of FM \( m < \beta \). This is illustrated in the upper panel of Figure 11. Thus, for small frequency separations of the components, subjects appear to be sensitive to the relative phases of the components, while for wide frequency separations they are not.

If the threshold for detecting modulation is expressed in terms of the modulation index, \( m \) or \( \beta \), the ratio \( \beta/m \) decreases as the modulation frequency increases and approaches an asymptotic value of unity. This is illustrated in the lower panel of Figure 11. The modulation frequency at which the ratio first becomes unity is called the critical modulation frequency (CMF). Zwicker (1952) and Schorer (1986) suggested that the CMF corresponded to half the value of the CB; essentially, the CMF was assumed to be reached when the overall stimulus bandwidth reached the CB. If this is correct, then the CMF may be regarded as providing an estimate of the CB at the carrier frequency.

Further analysis suggests that this interpretation of the results may not be completely correct. The CMF appears to correspond to the point where one of the sidebands in the spectrum first becomes detectable; usually the lower sideband is more detectable than the upper one (Hartmann & Hnath, 1982; Moore & Sek, 1992). The threshold for detecting a sideband depends more on the selectivity of auditory filters centered close to the frequency of the sideband than on the selectivity of the auditory filter centered on the carrier frequency. Furthermore, for low carrier frequencies, the upper sideband may be more detectable than the lower sideband (Sek & Moore, 1994). The change in the most detectable sideband with carrier frequency can account for the finding that the ERB decreases more with decreasing center frequency than does the CMF. It also makes the CMF unsuitable as a direct measure of the CB. Additionally, it should be noted that the detectability of a sideband may be influenced by factors not connected with frequency selectivity, such as the efficiency of the detection process following auditory filtering. This efficiency may well vary with center frequency, just as it does for the detection of tones in notched noise. Thus, like the critical ratio described earlier, the CMF does not provide a direct measure of the CB.
FIGURE 11 The upper panel shows thresholds for detecting sinusoidal amplitude modulation (squares) or frequency modulation (circles) of a 1 kHz carrier, plotted as a function of modulation rate. The thresholds are expressed in terms of the modulation indices, $m$ and $\beta$, respectively (the indices are multiplied by 100 to give convenient numbers). The lower panel shows the ratio $\beta/m$, plotted on a logarithmic scale as a function of modulation rate. (Data from Sek, 1994, with permission of the author.)

It also appears to be incorrect to assume that changes in the relative phase of the components in a complex sound are detectable only when those components lie within a CB. In cases where all components are well above threshold, subjects can detect phase changes between the components in complex sounds in which the components are separated by considerably more than a CB (Craig & Jeffress, 1962; Blauert & Laws, 1978; Patterson, 1987). The detection of these phase changes may depend partly on the ability to compare the time patterns at the outputs of different auditory filters; see Chapter 6 for further information on this topic.
C. The Audibility of Partial in Complex Tones

According to Ohm's (1843) Acoustical Law, the ear is able to hear pitches corresponding to the individual sinusoidal components in a complex sound. In other words, we can "hear out" the individual partials. For periodic complex sounds, the most prominent pitch is usually the residue pitch or virtual pitch associated with the sound as a whole, and we are not normally aware of hearing pitches corresponding to individual partials; see Chapter 8. Nevertheless, such pitches can be heard if attention is directed appropriately (Helmholtz, 1863).

Plomp (1964) and Plomp and Mimpen (1968) used complex tones with 12 sinusoidal components to investigate the limits of this ability. The listener was presented with two comparison tones, one of which was of the same frequency as a partial in the complex; the other lay halfway between that frequency and the frequency of the adjacent higher or lower partial. The listener was allowed to switch freely between the complex tone and the comparison tones and was required to decide which of the two comparison tones coincided with the partial in the complex tone. The score (varying between 50 and 100%) was used as an index of how well the partial could be heard out from the complex tone. For harmonic complex tones, only about the first five to seven harmonics could be heard out.

Plomp and Mimpen (1968) suggested that a component can be heard out only when its frequency separation from adjacent components exceeds the CB. The spacing of the components in a harmonic complex is uniform on a linear frequency scale, but, relative to the CB, the upper harmonics are more closely spaced than the lower harmonics. Harmonics above about the eighth are separated by less than a CB and cannot be heard out. Results for two subjects using complex tones where the frequency ratios between components were "compressed" relative to a harmonic complex gave basically the same result; only the lower components could be heard out.

The concept that the CB was the main factor limiting the audibility of partials in complex tones was questioned by Soderquist (1970). He used a task similar to that of Plomp (1964), but compared the results for musicians and nonmusicians. He found that musicians performed markedly better than nonmusicians. A possible explanation for this is that musicians have narrower CBs than nonmusicians. However, Fine and Moore (1993) estimated auditory filter bandwidths in musicians and nonmusicians, using the notched-noise method, and found that ERBs did not differ for the two groups. An alternative possibility is that performance depends on some factor or factors other than the CB.

Some aspects of the data of Plomp and Mimpen (1968) also suggest the involvement of factors other than the CB. The frequency difference between adjacent harmonics required to hear them separately was somewhat
greater than traditional CB values (Zwicker & Terhardt, 1980) above 1000 Hz and was distinctly smaller below 1000 Hz. Recent estimates of the ERB of the auditory filter, described in Section IV.A, are smaller than traditional CB values at low frequencies and more consistent with the data of Plomp and Mimpen. Nevertheless, some discrepancy remains. The relative value of the CB or the ERB (i.e., bandwidth divided by center frequency) increases as the center frequency decreases below about 1000 Hz. As a consequence, the number of resolvable harmonics in a harmonic complex tone would be expected to decrease at low fundamental frequencies. In fact, the data of Plomp and Mimpen show that the number of resolvable harmonics increases as the fundamental frequency decreases below 250 Hz.

Moore and Ohgushi (1993) examined the ability of musically trained subjects to hear out individual partials in complex tones with partials uniformly spaced on a scale related to the ERB of the auditory filter. ERB spacings of 0.75, 1.0, 1.25, 1.5, and 2 were used, and the central component always had a frequency of 1000 Hz. On each trial, subjects heard a pure tone (the “probe”) followed by a complex tone. The probe was close in frequency to one of the partials in the complex, but was mistuned downward by 4.5% on half the trials (at random) and mistuned upward by 4.5% on the other half. The task of the subject was to indicate whether the probe was higher or lower in frequency than the nearest partial in the complex. The partial that was “probed” varied randomly from trial to trial. If auditory filtering were the only factor affecting performance on this task, then scores for a given ERB spacing should be similar for each component in the complex sound.

Scores for the highest and lowest components in the complexes were generally high for all components spacings, although they worsened somewhat for ERB spacings of 0.75 and 1.0. Scores for the inner components were close to chance level at 0.75 ERB spacing, and improved progressively as the ERB spacing was increased from 1 to 2 ERBs. For ERB spacings of 1.25 or less, the scores did not change smoothly with component frequency; marked irregularities were observed, as well as systematic errors. Moore and Ohgushi suggested that these resulted from irregularities in the transmission of sound through the middle ear; such irregularities could change the relative levels of the components, making some components more prominent than others and therefore easier to hear out.

Performance for the inner components tended to be worse for component frequencies above 1000 Hz than below 1000 Hz. This is consistent with the pattern of results found by Plomp and Mimpen (1968) and indicates that some factor other than auditory filtering influences the audibility of partials in complex tones.

Moore and Ohgushi suggested that the pitches of individual components may be partly coded in the time patterns of neural activity (phase locking)
in the auditory nerve, as has also been suggested by previous researchers (Ohgushi, 1978, 1983; Srulovicz & Goldstein, 1983; Moore, Glasberg, & Shailer, 1984; Moore, 1989; Hartmann, McAdams, & Smith, 1990; Moore & Glasberg, 1990; see Chapter 3). Phase locking is more precise below 1000 Hz than above 1000 Hz. This can explain why, for partials uniformly spaced on an ERB scale, the identification of partials was better for components with frequencies below 1000 Hz than above 1000 Hz.

The influence of phase locking can also explain the superior identification of the lowest and highest components in the complex tones. Generally, neurons with characteristic frequencies (CFs) close to the frequency of a given partial will phase lock to that partial, provided the frequency separation of partials is sufficient and the frequency of the partial is not too high. For an inner partial in a complex tone, the pattern of phase locking in neurons with CFs close to the frequency of the partial will be disturbed by the partials on either side, making it more difficult to extract the pitch of that partial, especially when the components are closely spaced. In contrast, neurons tuned just below the lower edge frequency or just above the higher edge frequency will show a pattern of phase locking that is less disturbed by the other components. A similar explanation can be offered for Plomp's (1964) finding that the partials in a two-tone complex could be heard out for smaller frequency separations than were found for multitone complexes.

In summary, it seems clear that auditory filtering plays a strong role in limiting the ability to hear out partials in complex tones. However, it is probably not the only factor involved. Specifically, the pitches of individual partials may partly be coded in the patterns of phase locking of neurons in the auditory nerve. This coding is more accurate at low frequencies than at high.

VIII. NONSIMULTANEOUS MASKING

Simultaneous masking describes situations where the masker is present for the whole time that the signal occurs. Masking can also occur when a brief signal is presented just before or after the masker; this is called nonsimultaneous masking. Two basic types of nonsimultaneous masking can be distinguished: (1) backward masking, in which the signal precedes the masker (also known as prestimulatory masking); and (2) forward masking, in which the signal follows the masker (also known as poststimulatory masking).

Although many studies of backward masking have been published, the phenomenon is poorly understood. The amount of backward masking obtained depends strongly on how much practice the subjects have received, and practiced subjects often show little or no backward masking (Miyazaki & Sasaki, 1984; Oxenham & Moore, 1994, 1995). The larger masking ef-
ffects found for unpracticed subjects may reflect some sort of "confusion" of the signal with the masker. In contrast, forward masking can be substantial even in highly practiced subjects. The main properties of forward masking are as follows:

1. Forward masking is greater the nearer in time to the masker that the signal occurs. This is illustrated in the left panel of Figure 12. When the delay $D$ of the signal after the end of the masker is plotted on a logarithmic scale, the data fall roughly on a straight line. In other words, the amount of forward masking, in dB, is a linear function of $\log(D)$.

2. The rate of recovery from forward masking is greater for higher masker levels. Thus, regardless of the initial amount of forward masking, the masking decays to 0 after 100–200 ms.

3. Increments in masker level do not produce equal increments in amount of forward masking. For example, if the masker level is increased by 10 dB, the masked threshold may increase by only 3 dB. This contrasts with simultaneous masking, where, at least for wideband maskers, the threshold usually corresponds to a constant signal-to-masker ratio. This effect can be quantified by plotting the signal threshold as a function of masker level. The resulting function is called a growth of masking function.

![Graph](image)

**FIGURE 12** The left panel shows the amount of forward masking of a brief 2 kHz signal, plotted as a function of the time delay of the signal after the end of the noise masker. Each curve shows results for a different noise spectrum level (10–50 dB). The results for each spectrum level fall on a straight line when the signal delay is plotted on a logarithmic scale, as here. The right panel shows the same thresholds plotted as a function of masker spectrum level. Each curve shows results for a different signal delay time (17.5, 27.5, or 37.5 ms). Note that the slopes of these growth of masking functions decrease with increasing signal delay. (Adapted from Moore and Glasberg, 1983a.)
Several such functions are shown in the right panel of Figure 12. In simultaneous masking such functions would have slopes close to 1. In forward masking the slopes are less than 1, and the slopes decrease as the value of $D$ increases.

4. The amount of forward masking increases with increasing masker duration for durations up to at least 20 ms. The results for greater masker durations vary somewhat across studies. Some studies show an effect of masker duration for durations up to 200 ms (Kidd & Feth, 1982), while others show little effect for durations beyond 50 ms (Fastl, 1976b).

The mechanisms underlying forward masking are not clear. It could be explained in terms of a reduction in sensitivity of recently stimulated neurons or in terms of a persistence in the pattern of neural activity evoked by the masker. Both points of view can be found in the literature. In addition, the response of the basilar membrane to the masker takes a certain time to decay, and for small intervals between the signal and the masker this may result in forward masking (Duifhuis, 1973); see Chapter 6 for further discussion of these issues.

IX. EVIDENCE FOR LATERAL SUPPRESSION FROM NONSIMULTANEOUS MASKING

Measurements of basilar membrane motion (Ruggero, 1992), or in single neurons (Arthur, Pfeiffer, & Suga, 1971), show that the response to a tone of a given frequency can sometimes be suppressed by a tone with a different frequency, a phenomenon known as two-tone suppression; see Chapters 2 and 3. For other complex signals, similar phenomena occur and are given the general name lateral suppression or suppression. This can be characterized in the following way. Strong activity at a given characteristic frequency can suppress weaker activity at adjacent CFs. In this way, peaks in the excitation pattern are enhanced relative to adjacent dips. The question now arises as to why the effects of suppression are not usually seen in experiments on simultaneous masking.

Houtgast (1972) has argued that simultaneous masking is not an appropriate tool for detecting the effects of suppression. In simultaneous masking, the masking stimulus and the signal are processed simultaneously in the same channel (the same auditory filter). Thus any suppression in that channel will affect the neural activity caused by both the signal and the masker. In other words, the signal-to-masker ratio in a given frequency region will be unaffected by suppression, and thus the threshold of the signal will remain unaltered.

Houtgast suggested that this difficulty could be overcome by presenting the masker and the signal successively, for example, by using forwar
masking. If suppression does occur, then its effects will be seen in forward masking provided (1) in the chain of levels of neural processing, the level at which the suppression occurs is not later than the level at which most of the forward masking effect arises; and (2) the suppression built up by the masker has decayed by the time that the signal is presented (otherwise the problems described for simultaneous masking will be encountered).

Following the pioneering work of Houtgast (1972, 1973, 1974), many workers have reported that there are systematic differences between the results obtained using simultaneous and nonsimultaneous masking techniques. An extensive review is provided by Moore and O'Loughlin (1986). One major difference is that nonsimultaneous masking reveals effects that can be directly attributed to suppression. A good demonstration of this involves a psychophysical analog of neural two-tone suppression. Houtgast (1973, 1974) measured the threshold for a 1 kHz signal and a 1 kHz nonsimultaneous masker. He then added a second tone to the masker and measured the threshold again. He found that sometimes the addition of this second tone produced a reduction in the threshold, and he attributed this to a suppression of the 1 kHz component in the masker by the second component. If the 1 kHz component is suppressed, then there will be less activity in the frequency region around 1 kHz, producing a drop in the threshold for detecting the signal. The second tone was most effective as a "suppressor" when it was somewhat more intense than the 1 kHz component and above it in frequency. Similar results have been obtained by Shannon (1976).

Under some circumstances, the reduction in threshold (unmasking) produced by adding one or more extra components to a masker can be partly explained in terms of additional cues provided by the added components, rather than in terms of suppression. Specifically, in forward masking the added components may reduce "confusion" of the signal with the masker by indicating exactly when the masker ends and the signal begins (Moore, 1980, 1981; Moore & Glasberg, 1982a, 1985; Neff, 1985; Moore & O'Loughlin, 1986). This may have led some researchers to overestimate the magnitude of suppression as indicated in nonsimultaneous masking experiments. However, it seems clear that not all unmasking can be explained in this way.

X. THE ENHANCEMENT OF FREQUENCY SELECTIVITY REVEALED IN NONSIMULTANEOUS MASKING

A second major difference between simultaneous and nonsimultaneous masking is that the frequency selectivity revealed in nonsimultaneous masking is greater than that revealed in simultaneous masking. A well-studied example of this is the psychophysical tuning curve. PTCs determined in forward masking are typically sharper than those obtained in simultaneous masking (Moore, 1978). An example is given in Figure 13. The difference is
FIGURE 13 Comparison of psychophysical tuning curves determined in simultaneous masking (triangles) and forward masking (squares). The masker frequency is plotted as deviation from the center frequency divided by the center frequency ($\Delta f/f$). The center frequency is indicated in kHz in each panel. A low-level notched noise was gated with the masker to provide a consistent detection cue in forward masking and to restrict off-frequency listening. (From Moore et al., 1984.)

particularly marked on the high-frequency side of the tuning curve. According to Houtgast (1974) this difference arises because the internal representation of the masker (its excitation pattern) is sharpened by a suppression process, with the greatest sharpening occurring on the low-frequency side. In simultaneous masking, the effects of suppression are not seen, because any reduction of the masker activity in the frequency region of the signal is accompanied by a similar reduction in signal-evoked activity. In other words, the signal-to-masker ratio in the frequency region of the signal is unaffected by the suppression. In forward masking, on the other hand, the suppression does not affect the signal. For maskers with frequencies above that of the signal, the effect of suppression is to sharpen the excitation pattern of the masker, resulting in an increase of the masker level required to mask the signal. Thus the suppression is revealed as an increase in the slopes of the PTC.

An alternative explanation is that, in simultaneous masking, the low-level signal may be suppressed by the masker, so that it falls below absolute
threshold. The neural data indicate that tones falling outside of the region bounded by the neural tuning curve can produce suppression. Thus, the PTC in simultaneous masking might map out the boundaries of the more broadly tuned suppression region (Delgutte, 1988).

It remains unclear which of these two explanations is correct. Moore and Glasberg (1982b) concluded on the basis of a psychophysical experiment that the first explanation was correct, and a physiological experiment by Pickles (1984) supported this view. However, Delgutte (1990) has presented physiological evidence suggesting that simultaneous masking by intense low-frequency tones (upward spread of masking) is due largely to suppression rather than spread of excitation.

Several other methods of estimating frequency selectivity have indicated sharper tuning in nonsimultaneous masking than in simultaneous masking. For example, auditory filter shapes estimated in forward masking using a notched-noise masker, have smaller bandwidths and greater slopes than those estimated in simultaneous masking (Moore & Glasberg, 1981; Moore, Poon, Bacon, & Glasberg, 1987). This encourages the belief that a general consequence of suppression is an enhancement of frequency selectivity.

XI. SUMMARY

The peripheral auditory system contains a bank of bandpass filters, the auditory filters, with center frequencies spanning the audible range. The basilar membrane appears to provide the initial basis of the filtering process. The auditory filter can be thought of as a weighting function that characterizes frequency selectivity at a particular center frequency. The shape of the auditory filter at a given center frequency can be estimated using the notched-noise masking technique and the assumptions of the power-spectrum model. Its bandwidth for frequencies above 1 kHz is about 10–17% of the center frequency. At moderate sound levels the auditory filter is roughly symmetric on a linear frequency scale. At high sound levels the low-frequency side of the filter becomes less steep than the high-frequency side. The shape of the auditory filter appears to depend mainly on the level at the input to the filter.

When two maskers are combined, the resulting masking is sometimes greater than predicted from linear summation of the individual effects of the maskers. One explanation for this excess masking is that the individual maskers are subject to a compressive nonlinearity before their effects are combined. However, a more plausible explanation can be given in terms of the detection cues used with the individual maskers and the combined maskers.

The excitation pattern of a given sound represents the distribution of activity evoked by that sound as a function of the characteristic frequency of the neurons stimulated. In psychophysical terms, the excitation pattern can
be defined as the output of each auditory filter as a function of its center frequency. The shapes of excitation patterns for sinusoids or narrowband noises are similar to the masking patterns of narrowband noises.

The critical bandwidth is related to the bandwidth of the auditory filter. It is revealed in experiments on masking, loudness, absolute threshold, phase sensitivity, and the audibility of partials in complex tones. However, factors other than auditory filtering play a role in many of these experiments, so that they often do not provide a "direct" measure of the bandwidth of the auditory filter.

Houtgast and others have shown that nonsimultaneous masking reveals suppression effects similar to the suppression observed in primary auditory neurons. This suppression is not revealed in simultaneous masking, possibly because suppression at a given CF does not affect the signal-to-masker ratio at that CF. One result of suppression is an enhancement in frequency selectivity.

Acknowledgments

I thank Joseph Alcántara, Tom Baer, Brian Glasberg, Andrew Oxenham, Roy Patterson, Robert Peters, Aleksander Sek, Michael Shailer, and Michael Stone for helpful comments on an earlier version of this chapter. I also thank Aleksander Sek for producing Figure 11.

References


Moore, B. C. J., & Glasberg, B. R. (1987b). Formulae describing frequency selectivity as a
function of frequency and level and their use in calculating excitation patterns. Hearing
Research, 28, 209–225.
overlapping and non-overlapping harmonics. Journal of the Acoustical Society of America, 87,
2163–2177.
Moore, B. C. J., Glasberg, B. R., & Roberts, B. (1984). Refining the measurement of psycho-
limens for harmonics within complex tones. Journal of the Acoustical Society of America, 75,
550–561.
Moore, B. C. J., & O'Loughlin, B. J. (1986). The use of nonsimultaneous masking to measure
frequency selectivity and suppression. In B. C. J. Moore (Ed.), Frequency selectivity in
of masking and the auditory filter shape. Journal of the Acoustical Society of America, 81,
1873–1880.
masking release in subjects with unilateral and bilateral cochlear hearing impairment.
Ohgushi, K. (1978). On the role of spatial and temporal cues in the perception of the pitch of
Ohgushi, K. (1983). The origin of tonality and a possible explanation of the octave enlarge-
Ohm, G. S. (1843). Uber die Definition des Tones, nebst daran geknüpfter Theorie der Sirene
Hearing Research, 5, 343–346.
O'Loughlin, B. J., & Moore, B. C. J. (1981b). Off-frequency listening: Effects on psycho-
acoustical tuning curves obtained in simultaneous and forward masking. Journal of the
Acoustical Society of America, 69, 1119–1125.
masking. Hearing Research, 80, 105–118.
802–809.
tical Society of America, 59, 640–654.


