3

Frequency selectivity, masking and the critical band

1 INTRODUCTION

This chapter is concerned with the frequency selectivity of the auditory system. Frequency selectivity refers to our ability to resolve the sinusoidal components in a complex sound, and it plays a role in many aspects of auditory perception. However, it is often demonstrated and measured by studying masking. It is a matter of everyday experience that one sound may be obscured, or rendered inaudible, in the presence of other sounds. Thus music from a car radio may mask the sound of the car's engine, provided the music is somewhat more intense. Masking has been defined as:

> (1) The *process* by which the threshold of audibility for one sound is raised by the presence of another (masking) sound. (2) The *amount* by which the threshold of audibility of a sound is raised by the presence of another (masking) sound. The unit customarily used is the decibel. (American Standards Association, 1960)

It has been known for many years that a signal will most easily be masked by a sound having frequency components close to, or the same as, those of the signal (Mayer, 1894; Wegel and Lane, 1924). This led to the idea that our ability to separate the components of a complex sound depends, at least in part, on the frequency-resolving power of the basilar membrane. This idea will be elaborated later in this chapter. It also led to the idea that masking reflects the limits of frequency selectivity: if the selectivity of the ear is insufficient to separate the signal and the masker, then masking will occur. Thus, masking can be used to quantify frequency selectivity. Hence, much of this chapter will be devoted to studies of masking.

As well as having theoretical significance, a knowledge of the rules governing the masking of one sound by another can be very useful in practical situations. For example, one might want to know the extent to which noise from new machinery in a factory will interfere with the ability of workers to hold conversations or to detect warning signals. Masking is also used in the

clinical assessment of hearing. For example, in a patient with one impaired and one normal ear, earphones may be used to present noise to the normal ear when testing the impaired ear, so as to prevent sound that 'leaks' across the head being heard in the normal ear.

An important physical parameter which affects masking is time. Most of this chapter will be devoted to simultaneous masking, in which the signal is presented at the same time as the masker. Later on we will discuss forward masking, in which the signal is masked by a preceding masker, and backward masking, in which the masker follows the signal.

2 THE CRITICAL BAND CONCEPT

A Fletcher's band-widening experiment and the power spectrum model

Fletcher (1940) carried out an experiment which has now become famous and which laid the foundation for the concept of the critical band. He measured the threshold of a sinusoidal signal as a function of the bandwidth of a bandpass noise masker. The noise was always centred 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 (Hamilton, 1957; Greenwood, 1961a; Spiegel, 1981; Schooneveldt and Moore, 1989). An example of the results, taken from Schooneveldt and Moore, is given in Fig. 3.1. The threshold of the signal increases at first as the noise bandwidth increases, but then flattens off, so that further increases in noise bandwidth do not change the signal threshold significantly.

To account for these results, Fletcher (1940) suggested that the peripheral auditory system behaves as if it contained a bank of bandpass filters, with continuously overlapping centre frequencies. These filters are now called the 'auditory filters'. Fletcher thought that the basilar membrane provided the basis for the auditory filters. Each location on the basilar membrane responds to a limited range of frequencies, so that each different point corresponds to a filter with a different centre frequency. Recent data are consistent with this point of view (Moore, 1986).

When trying to detect a signal in a noise background, the listener is assumed to make use of a filter with a centre frequency close to that of the signal. This filter will pass the signal but remove a great deal of the noise. Only the components in the noise which pass through the filter will have any effect in masking the signal. It is usually assumed that the threshold for the signal is

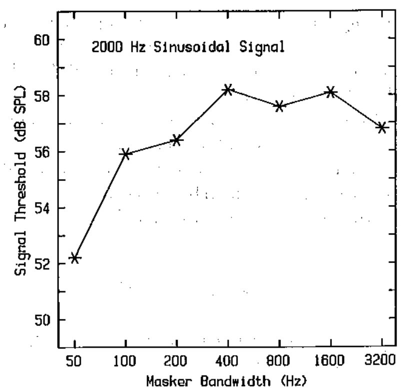

FIG. 3.1 The threshold of a 2000 Hz sinusoidal signal plotted as a function of the bandwidth of a noise masker centred at 2000 Hz. Notice that the threshold of the signal at first increases with increasing masker bandwidth and then remains constant. From Schooneveldt and Moore (1989).

determined by the amount of noise passing through the auditory filter; specifically, threshold is assumed to correspond to a certain signal-to-noise ratio at the output of the filter. This set of assumptions has come to be known as the 'power spectrum model' of masking (Patterson and Moore, 1986), since the stimuli are represented by their long-term power spectra, i.e. the relative phases of the components and the short-term fluctuations in the masker are ignored. We shall see later that the assumptions of this model do not always hold, but the model works well in many situations, and we will accept it for the moment.

In the band-widening experiment described above, increases in noise bandwidth result in more noise passing through the auditory filter, provided 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'.

In analysing the results of his experiment, Fletcher made a simplifying assumption. He assumed that the shape of the auditory filter could be approximated 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. The width of this passband is equal to the critical bandwidth described above. The term 'critical band' is often used to refer to this hypothetical rectangular filter.

Fletcher pointed out that the value of the critical bandwidth could be estimated indirectly, by measuring the threshold of a tone in broadband white noise, given the following hypotheses:

1. Only a narrow band of frequencies surrounding the tone – those falling within the critical band – contributes to the masking of the tone.
2. When the noise just masks the tone, the power of the tone divided by the power of the noise inside the critical band is a constant, K.

As described in Chapter 1, noise power is usually specified in terms of the power in a band of frequencies 1 Hz wide (say from 1000 Hz to 1001 Hz). This is called the noise power density, and is denoted by the symbol N_o. For a white noise N_o is independent of frequency, so that the total noise power falling in a critical band W Hz wide is $N_o \times W$. According to Fletcher's second hypothesis,

$$P/(W \times N_o) = K$$

and

$$W = P/(K \times N_o).$$

By measuring P and N_o, and by estimating K, we can evaluate W.

The first hypothesis follows directly from Fletcher's experiment, although, as we shall see later, it is only an approximation. To estimate the value of the constant K, Fletcher measured the threshold for a tone in a band of noise whose width was less than the estimated critical bandwidth. In this case, K equals the ratio of the power of the tone to the power of the noise, since all of the noise passes through the auditory filter. Fletcher estimated K to equal 1, so that the value of W should be equal to P/N_o. The ratio P/N_o 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 critical bandwidth estimated by more direct methods, such as the band-widening experiment. Also, K varies somewhat with centre frequency, so the critical ratio does not give a correct indication of how the critical bandwidth varies with centre frequency (Patterson and Moore, 1986). The difference between the critical ratio function (critical ratio as a function of frequency) and the generally accepted critical band function is illustrated in Fig. 3.2.

Since Fletcher first described the critical band concept, many different experiments have shown that listeners' responses to complex sounds differ

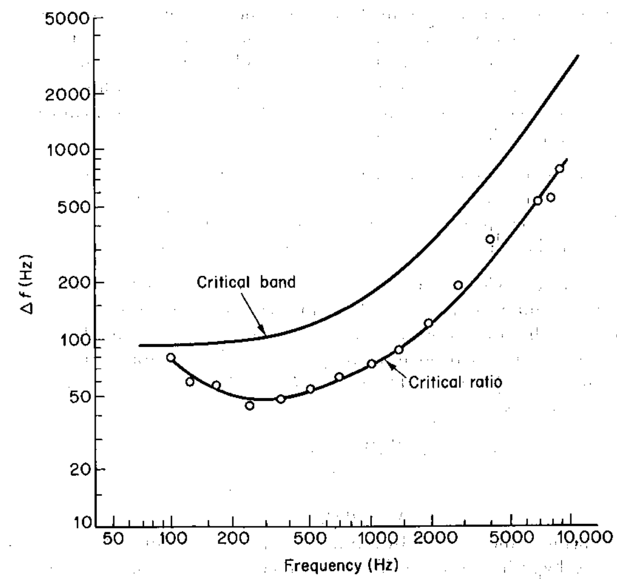

FIG. 3.2 A comparison of the width, Δf, of the critical band as determined by direct measures (upper curve) and as determined from the critical ratio. From Zwicker *et al.* (1957), by permission of the authors and *J. Acoust. Soc. Am.*

according to whether the stimuli are wider or narrower than the critical bandwidth. Further, these different experiments give remarkably similar estimates both of the absolute width of the critical band and of the way the critical bandwidth varies as a function of frequency. Thus the critical band phenomenon pervades and summarizes a great variety of data, and provides a valuable guide in the planning of experiments and the analysis of data. Some of the types of experiment in which critical band phenomena have been observed are described below.

B The loudness of complex sounds

Consider a complex sound of fixed energy (or intensity) having a bandwidth of W. If W is less than the critical bandwidth, then the loudness of the sound is

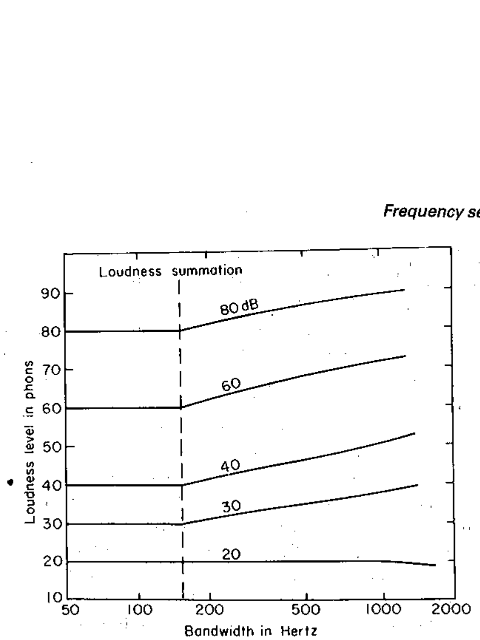

FIG. 3.3 The loudness level in phons of a band of noise centred at 1 kHz, measured as a function of the width of the band. For each of the curves, the overall sound level was constant, and its value, in dB SPL, is indicated in the figure. The dashed line shows that the bandwidth at which loudness begins to increase is roughly the same at all levels tested (except that no increase occurs at the lowest level). From Feldtkeller and Zwicker (1956) by permission of the authors and publisher.

more or less independent of W; the sound is judged to be about as loud as a pure tone of equal intensity lying at the centre frequency of the band. However, if W is increased beyond the critical bandwidth, the loudness of the complex begins to increase. This has been found to be the case for bands of noise (Zwicker *et al.*, 1957) and for complexes consisting of pure tones whose frequency separation is varied (Scharf, 1961, 1970) (see Fig. 3.3). The critical bandwidth for the data in Fig. 3.3 is about 160 Hz for a centre frequency of 1000 Hz. Thus, for a given amount of energy, a complex sound will be louder if the energy is spread over a number of critical bands, than if it is all contained within one critical band. This finding has been incorporated in the models described in Chapter 2, which allow the calculation of the loudness of almost any complex sound (see, for example, Zwicker and Scharf, 1965).

The increase in loudness with increasing bandwidth can be understood if we assume that when the bandwidth of a sound is sufficient to occupy more than one critical band, the loudness in adjacent, but nonoverlapping, bands is summed to give the total loudness. Consider the effect of taking two tones very close together in frequency, so as to occupy one critical band, and increasing their frequency separation so that two critical bands are occupied. The intensity in each band will now be half of what it was, since only one tone is present in each band. According to Steven's Power Law, $L = kI^{0.3}$ (see Chapter 2, section 4). Thus halving intensity is equivalent to a reduction in loudness to 0.81 of its original value. The total loudness in the two bands will be $2 \times 0.81 = 1.62$ times the original value. Thus increasing bandwidth beyond the critical band results in an increase in loudness. The argument can easily be extended to cover multicomponent complex tones and bands of noise (see Moore and Glasberg, 1986).

A corresponding explanation can be offered in terms of the neurophysiological data. When the two tones are very close in frequency they excite essentially the same set of neurones. If the two tones are separated in frequency, they excite essentially independent sets of neurones. Now at moderate and high intensities, the firing rates of individual neurones change relatively slowly as a function of intensity. Thus the decrease in neural firings as a result of halving the effective intensity exciting each set, is more than offset by the doubling of the number of neurones involved. If, then, loudness is related in some way to the total number of neural impulses evoked by the stimulus, it is clear that the loudness of the complex should be greater when the components are widely separated.

At low sensation levels (around 10–20 dB SL) the loudness of a complex sound is roughly independent of bandwidth. This also is easy to explain. At these low levels, neural firing rates change relatively rapidly with intensity, and so does loudness. The loudness of a single critical band changes almost in direct proportion to intensity, so that increasing the spread of energy from one to two critical bands, for example, produces two component bands each half as loud as the original sound (Scharf, 1970). The total loudness is equal to that of the original single band, so that loudness is independent of bandwidth. At very low sensation levels (below 10 dB), if we distribute the energy of a complex sound over a wide range of frequencies, then the energy in each critical band is insufficient to make the sound audible. Accordingly, near threshold, loudness must decrease as the bandwidth of a complex sound is increased from a subcritical value. As a consequence, if the intensity of a complex sound is increased slowly from a subthreshold value, the rate of growth of loudness is greater for a wideband sound than for a narrowband sound.

C The threshold of complex sounds

When two tones with a small frequency separation are presented together, a sound may be heard even when either tone by itself is below threshold.

Gässler (1954) measured the threshold of multitone complexes consisting of evenly spaced sinusoids. The tones 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 tones in a complex was increased, the threshold, specified in terms of total energy, remained constant until the overall spacing of the tones reached the critical bandwidth. Thereafter the threshold increased. The critical bandwidth for a centre frequency of 1000 Hz was estimated to be about 180 Hz. These results suggest 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 critical band. When the components are distributed over more than one critical band, detection is less good.

Unfortunately, more recent data are not in complete agreement with those of Gässler. For example, Spiegel (1981) measured the threshold for a noise signal of variable bandwidth centred at 1000 Hz in a broadband background noise masker. The threshold for the signal as a function of bandwidth did not show a breakpoint corresponding to the critical bandwidth, but increased monotonically as the bandwidth increased beyond 50 Hz. Spiegel suggested that the ear is capable of integration over bandwidths much greater than the critical bandwidth.

D Two-tone masking

Zwicker (1954) measured the threshold for a narrow band of noise, of centre frequency *f*, in the presence of two tones, with frequencies on either side of *f*. He increased the frequency separation of the two tones, starting with a small separation, and found that threshold for the noise signal remained constant until the separation reached a critical value, after which it fell sharply. He took this critical value to be an estimate of the critical bandwidth. Unfortunately the interpretation of this experiment is complicated. One problem is that the lower of the two tones may interact with the noise band to produce combination products; these are frequency components not present in the stimulus applied to the ear, and they appear to result from a nonlinear process in the cochlea (see Chapter 1, section 5D and Chapter 5, section 5A). The listener may detect these combination products even though the signal itself is inaudible. When precautions are taken to mask the distortion products, then the threshold for the signal does not show an abrupt decrease, but decreases smoothly with increasing frequency separation between the two tones (Patterson and Henning, 1977; Glasberg *et al.*, 1984). Nevertheless, the results do clearly indicate the operation of a filtering mechanism in the ear. For an extensive review of results obtained using two-tone maskers, the reader is referred to Rabinowitz *et al.* (1980).

E Sensitivity to phase

The sounds which we encounter in everyday life often change in frequency and amplitude from moment to moment. In the laboratory the perception of such sounds is often studied using either frequency-modulated or amplitude-modulated sine waves. Such waves consist of a carrier frequency (a sine wave) upon which some other signal is impressed. In amplitude modulation (AM) the carrier's amplitude is varied so as to follow the magnitude of a modulating sine wave, while the carrier frequency remains unchanged. In frequency modulation (FM) the carrier's instantaneous frequency is varied in proportion to the modulating signal's magnitude, but the amplitude remains constant. The two types of waveform are illustrated in Fig. 3.4. The expression describing an AM sinewave with carrier frequency f_c and modulating frequency g is

$$(1 + m \sin 2\pi gt) \sin 2\pi f_c t$$

where t is time, and m is a constant determining the amount of modulation; m is referred to as the modulation depth. When $m = 1$, the wave is said to be 100% modulated. The corresponding expression describing an FM sinewave is

$$\sin (2\pi f_c t + \beta \sin 2\pi gt)$$

In this case, β is usually referred to as the modulation index.

These complex waveforms can be analysed into a series of sinusoidal components. For an AM wave the results of the analysis are very simple: the spectrum contains just three frequency components with frequencies $f_c - g$, f_c and $f_c + g$. For an FM wave the spectrum often contains many components, but if m is small, then the FM wave can also be considered as consisting of three components: $f_c - g$, f_c and $f_c + g$. Under some conditions an AM wave and an FM wave may have components which 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) measured one aspect of the perception of such stimuli, namely the just-detectable amounts of amplitude or frequency modulation, for various rates of modulation. He found that for high rates of modulation, where the frequency components are widely spaced, the detectability of FM

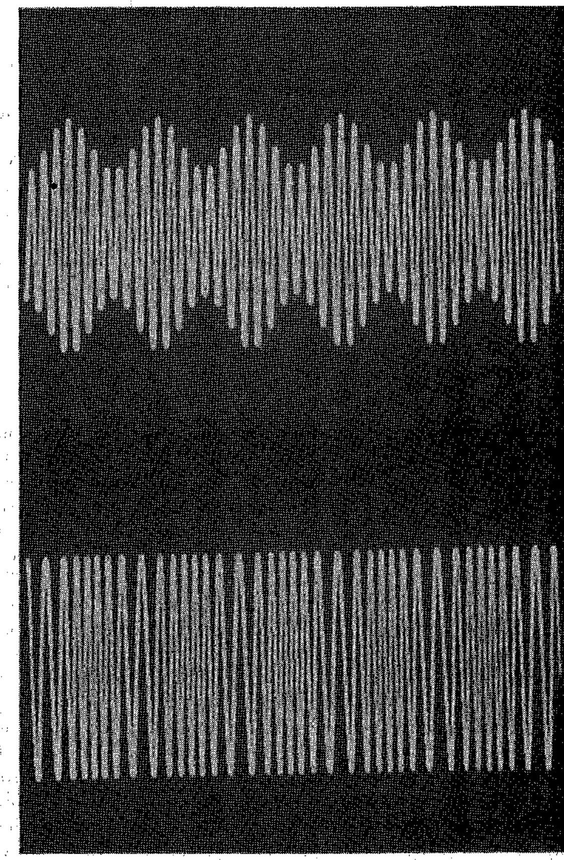

FIG. 3.4 Waveform of an amplitude-modulated wave (upper trace) and a frequency-modulated wave (lower trace).

and AM was equal when the components in each type of wave were of equal amplitude. However, when all three components fell within a critical band, AM was more easily detectable than FM. Thus it appears that we are only sensitive to the relative phase of the components, in the detection of modulation, when those components lie within a critical band. These results have been confirmed by Schorer (1986).

It is not at present clear whether this finding can be generalized to the perception of suprathreshold levels of modulation, or to other aspects of our sensitivity to phase. Indeed there is some evidence to the contrary. For example, it has been shown that subjects can detect phase changes between the components in complex sounds in which the components are separated by considerably more than a critical band (Raiford and Schubert, 1971; Lamore, 1975; Patterson, 1987a). Further work is needed to clarify the nature of these discrepancies, and to determine the applicability of the critical band concept to the phase sensitivity of the ear.

F The discrimination of partials 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 periodic sound. In other words, we can 'hear out' the individual partials. Plomp (1964a) used a complex tone 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 had to judge which of these two tones was a component of the complex. Plomp used two types of complex: a harmonic complex containing harmonics 1 to 12, where the frequencies of the components were integral multiples of that of the fundamental; and a nonharmonic complex, where the frequencies of the components were mistuned from simple frequency ratios. He found that for both kinds of complex only the first five to eight components could be 'heard out'. If it is assumed that a partial will only be distinguished when it is separated from its neighbour by at least one critical bandwidth, then the results can be used to estimate the critical bandwidth. Above 1000 Hz, the estimates obtained in this way coincide with other critical band measures. Below 1000 Hz the estimates are about two-thirds as large. When Plomp repeated the experiment using a two-tone complex, he found that the partials could be distinguished at smaller frequency separations than were found for multitone complexes.

Thus, while the results are roughly in line with other measures of critical bandwidth, there are discrepancies, especially at low frequencies. It is

possible that the analysis of partials from a complex sound depends in part on factors other than pure frequency resolution. Some indication of this is given by the work of Soderquist (1970). He compared musicians and nonmusicians in a task very similar to that of Plomp, and found that the musicians were markedly superior. This result could mean that musicians have narrower critical bands, but this is unlikely if the critical band reflects a basic physiological process, as is generally assumed (Scharf, 1970). It seems more plausible that some other mechanism is involved in this task and that musicians, because of their greater experience, are able to make more efficient use of this mechanism. Haggard (1974) also reports comparisons of different measures of the critical bandwidth, and suggests that discrepancies at low frequencies may be explained by the intervention of mechanisms other than the critical band. This problem is discussed more fully later in this chapter.

G Interim summary

The examples given above show that the phenomenon of the critical band can be revealed in a great variety of different experiments. By and large, the results of the different experiments give reasonably consistent estimates of the value of the critical bandwidth; the critical band function shown in Fig. 3.2 was arrived at by combining the results from many different experiments. However, it is also clear that most of the experiments do not show a distinct breakpoint corresponding to the critical bandwidth. Rather, the pattern of results changes smoothly and continuously as a function of bandwidth.

The idea that there should be distinct breakpoints in the data goes back to Fletcher's approximation of the auditory filter as having the shape of a rectangle. Fletcher was well aware that the filter was not perfectly rectangular. He knew that a tone or narrow band of noise can mask another tone for frequency separations considerably exceeding the critical bandwidth. We are led to consider the critical band as resembling a filter with a rounded top and with sloping edges; the critical bandwidth then becomes some measure of the 'effective' bandwidth of this filter. We will now describe some attempts to measure the characteristics of the auditory filter, in other words, to derive the shape of the auditory filter.

3 ESTIMATING THE SHAPE OF THE AUDITORY FILTER

Most methods for estimating the shape of the auditory filter at a given centre frequency are based on the assumptions of the power spectrum model of masking. The threshold of a signal whose frequency is fixed is measured in the presence of a masker whose spectral content is varied. It is assumed, as a first approximation, that the signal is detected using the single auditory filter which is centred on the frequency of the signal, and that threshold corresponds to a constant signal-to-masker ratio at the output of that filter. The methods described below both use this same basic technique.

A Psychophysical tuning curves

One method involves a procedure which is analogous in many ways to the determination of a neural tuning curve, and the resulting function is often called a psychophysical tuning curve (PTC). To determine a PTC 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. When a sinusoid is used beats occur between the signal and masker, and these can provide a cue as to the presence of the signal. The effectiveness of this cue varies with the frequency separation of the signal and masker, since slow beats (which occur at small frequency separations) are more easily detected than rapid beats (see Chapter 4). This varying sensitivity to beats violates one of the assumptions of the power spectrum model of masking. This problem can be avoided by using a narrowband noise masker, since such a masker has inherent fluctuations in amplitude which prevent beats being detected. Thus noise is generally preferred (Patterson and Moore, 1986).

For each of several masker 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 will tell us the masker level required to produce a fixed output from the auditory filter as a function of frequency. Normally we determine a filter characteristic by plotting the output from the filter for an input varying in frequency and fixed in level (see Chapter 1, section 4 and Fig. 1.4). However, if the filter is linear the two methods will give the same result. Thus, if we assume linearity, the shape of the auditory filter can be obtained simply by inverting the PTC. Examples of some PTCs are given in Fig. 3.5.

The PTCs in Fig. 3.5 are very similar in general form to the neural tuning curves in Fig. 1.13. Remember that the neural tuning curves are obtained by determining the level of a tone required to produce a fixed output from a single neurone, as a function of the tone's frequency. The similarities in the procedures and the results encourage us to believe that the basic frequency selectivity of the auditory system is established at the level of the auditory

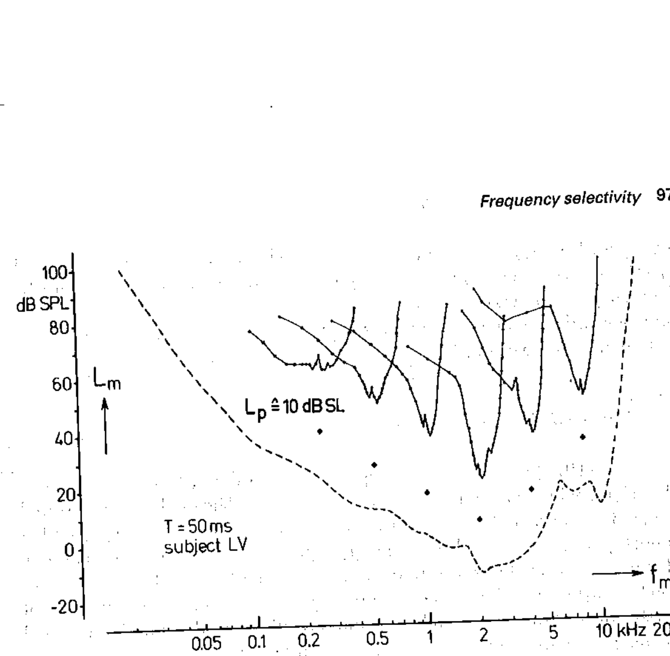

FIG. 3.5 Psychophysical tuning curves (PTCs) determined in simultaneous masking, using sinusoidal signals at 10 dB SL. For each curve, the solid diamond below it indicates the frequency and level of the signal. The masker was a sinusoid which had a fixed starting phase relationship to the brief, 50 ms, signal. The masker level, L_m, required for threshold is plotted as a function of masker frequency, f_m, on a logarithmic scale. The dashed line shows the absolute threshold for the signal. From Vogten (1974), by permission of the author.

nerve, and that the shape of the human auditory filter (or PTC) corresponds to the shape of the neural tuning curve. However, there is a need for caution in reaching this conclusion. In the determination of the neural tuning curve only one tone is present at a time, whereas for the PTC the masker and signal are presented simultaneously. This turns out to be an important point, and we will return to it later.

A second problem is that the neural tuning curve is derived from a single neurone, whereas the PTC inevitably involves activity over a group of neurones with slightly different CFs. Returning to the filter analogy, we have assumed in our analysis that only one auditory filter is involved, but it might be the case that the listener does not attend to just one filter. When the masker frequency is above the signal frequency the listener might do better to attend to a filter centred just below the signal frequency. If the filter has a relatively flat top, and sloping edges, this will considerably attenuate the masker at the filter output, while only slightly attenuating the signal. By using this off-centre filter the listener can improve performance. This is known as 'off-frequency listening', and there is now good evidence that humans do indeed listen 'off-frequency' when it is advantageous to do so. The result 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 and Patterson, 1979; O'Loughlin and Moore, 1981).

B The notched noise method

Patterson (1976) has described an ingenious method of determining auditory filter shape which prevents off-frequency listening. The method is illustrated in Fig. 3.6. The signal is fixed in frequency, and the masker is a noise with a bandstop or notch centred at the signal frequency. The deviation of each edge of the noise from the centre frequency is denoted by Δf. The width of the notch is varied, and the threshold of the signal is determined as a function of notch width. Since the notch is symmetrically placed around the signal frequency, the method cannot reveal asymmetries in the auditory filter, and the analysis assumes that the filter is symmetric on a linear frequency scale. This assumption appears not unreasonable, at least for the top part of the filter and at moderate sound levels since PTCs are quite symmetric around the tips. For a signal symmetrically placed in a bandstop noise, the optimum signal-to-masker ratio at the output of the auditory filter is achieved with a filter centred at the signal frequency, as illustrated in Fig. 3.6. Using a filter not

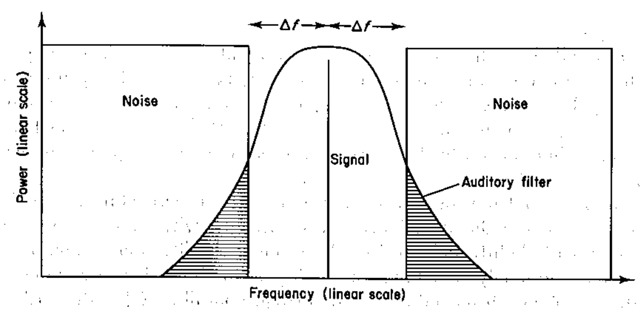

FIG. 3.6 Schematic illustration of the technique used by Patterson (1976) to determine the shape of the auditory filter. The threshold of the sinusoidal signal is measured as a function of the width of a spectral notch in the noise masker. The amount of noise passing through the auditory filter centred at the signal frequency is proportional to the shaded areas.

centred at the signal frequency reduces the amount of noise passing through the filter from one of the noise bands, but this is more than offset by the increase in noise from the other band.

As the width of the spectral notch is increased, less and less noise passes through the auditory filter. Thus the threshold of the signal drops. The amount of noise passing through the auditory filter is proportional to the area under the filter in the frequency range covered by the noise. This is shown as the shaded areas in Fig. 3.6. If we assume that threshold corresponds to a constant signal-to-masker ratio at the output of the filter, then the change in signal threshold with notch width tells us how the area under the filter varies with Δf. The area under a function between certain limits is obtained by integrating the value of the function over those limits. Hence by differentiating the function relating threshold to Δf, the height of the filter is obtained. In other words, the height of the filter for a given deviation, Δf, from the centre frequency is equal to the slope of the function relating signal threshold to notch width, at that value of Δf.

A typical auditory filter derived using this method is shown in Fig. 3.7. It has a rounded top and quite steep skirts. Unlike the simple rectangular filter, a filter with this shape cannot be completely specified with a single number, the critical bandwidth. However, some sort of summary statistic is useful, and one common measure is the bandwidth of the filter at which the response has

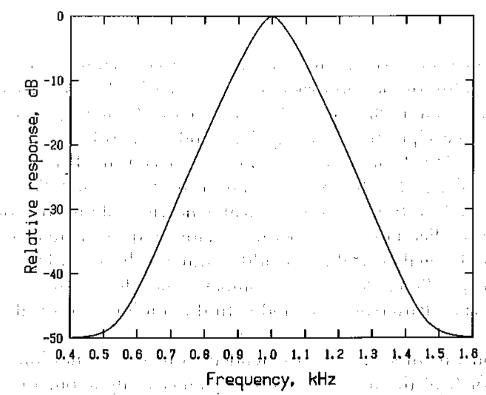

FIG. 3.7 A typical auditory filter shape determined using Patterson's method. The filter is centred at 1 kHz. The relative response of the filter (in dB) is plotted as a function of frequency.

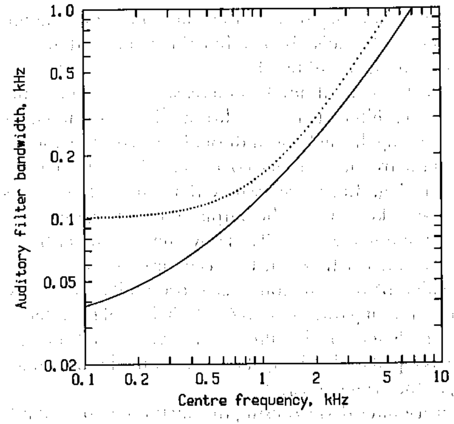

FIG. 3.8 The dotted curve shows the traditional value of the critical bandwidth as a function of frequency (Scharf, 1970). The solid curve shows the value of the ERB of the auditory filter as a function of frequency. The solid curve was obtained by combining the results of several experiments using Patterson's notched noise method of estimating the auditory filter shape. Adapted from Moore and Glasberg (1983a).

fallen by a factor of two in power, i.e. by 3 dB (see Chapter 1, section 4). The 3 dB bandwidths of the auditory filters derived using Patterson's method are typically between 10% and 15% of the centre frequency. An alternative measure is the equivalent rectangular bandwidth (ERB) (see Chapter 1, section 4). The ERBs of the auditory filters derived using Patterson's method are typically between 11% and 17% of the centre frequency. These values are quite close to the estimates of the critical bandwidth obtained in other ways, as described earlier. However, the values at low frequencies tend to be smaller than the traditional critical bandwidth estimates shown in Fig. 3.2 (Moore and Glasberg, 1983a). Figure 3.8 compares the ERB of the auditory filter estimated using Patterson's method with the traditional critical bandwidth function.

Patterson's method has been extended to include conditions where the spectral notch in the noise is placed asymmetrically about the signal frequency. This allows the measurement of any asymmetry in the auditory filter, but the analysis of the results is more difficult, and has to take off-frequency listening into account (Patterson and Nimmo-Smith, 1980). It is beyond the scope of this book to give details of the method of analysis; the interested reader is referred to Patterson and Moore (1986) and Moore and Glasberg (1987). The results show that the auditory filter is reasonably symmetric at moderate sound levels, but becomes increasingly asymmetric at high levels, the low-frequency side becoming shallower than the high-frequency side. The shape of the auditory filter centred at 1 kHz is shown in Fig. 3.9 for a range of sound levels from 20 to 90 dB SPL.

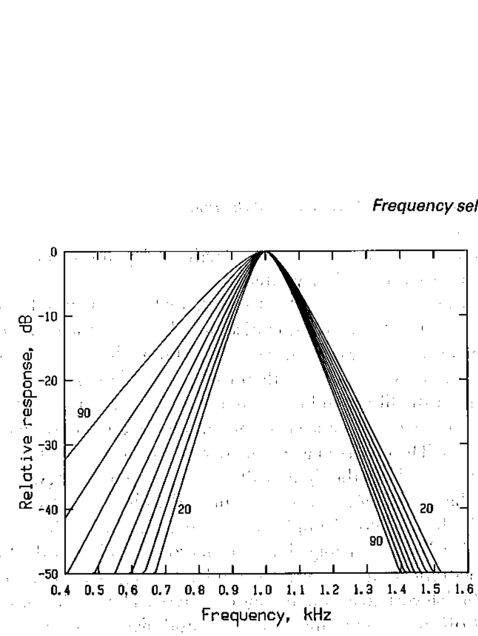

FIG. 3.9 The shape of the auditory filter centred at 1 kHz, plotted for input sound levels ranging from 20 to 90 dB SPL. The attenuation applied by the filter is plotted as a function of frequency. On the low frequency side, the filter becomes progressively less sharply tuned with increasing sound level. On the high frequency side, the sharpness of tuning increases slightly with increasing sound level. At moderate sound levels the filter is approximately symmetric on the linear frequency scale used. Adapted from Moore and Glasberg (1987).

C *Some general observations on auditory filters*

One question we may ask is whether the listener can only attend to one auditory filter at a time. The answer to this is obviously no, since many of the complex signals which we can perceive and recognize occupy more than one critical band; speech is a prime example of this. Indeed we shall see later that the perception of timbre seems to depend, at least in part, on the distribution of activity across different auditory filters. Furthermore, we shall see that the detection of a signal in a masker can sometimes depend on a *comparison* of the outputs of different auditory filters.

In spite of this, it is often possible to predict whether a complex sound will be detected in a given background noise by calculating the thresholds of the most prominent frequency components. If we know the shape of the auditory filter centred on each component, then we can calculate the amount of noise passing through the filter, and the signal-to-noise ratio at the output of the filter. If this ratio exceeds some criterion amount in any filter, then the signal will be detected. The criterion amount corresponds to a signal-to-noise ratio of about 1:2.5 or −4 dB; the signal will be detected if its level is not more than 4 dB below that of the noise at the output of the filter.

This model has practical applications, since it allows the prediction of appropriate levels for warning signals in factories and aircraft. In the past, when no theoretical model was available, the signals were often set at excessively high levels, to err on the side of 'safety'. The result was that when a signal 'went off' it was extremely aversive, and disrupted speech communication. The application of the auditory filter model has shown that sometimes signal levels can be reduced by 20 dB (a factor of 100 in power) and remain clearly audible (Patterson and Milroy, 1980).

Another question which arises is whether there is only a discrete number of critical bands, each one adjacent to its neighbours, or whether there is a continuous series of overlapping critical bands. For convenience, data relating to critical bands have often been presented as though the former were the case. For example, Scharf (1970) presented a table showing critical bandwidths for 24 successive critical bands, the upper cutoff frequency for each band being the same as the lower cutoff for the next highest band. While this method of presentation is convenient, it seems clear that critical bands are continuous rather than discrete; there has been no experimental evidence for any discontinuity or break between different critical bands. Thus we may talk about the critical band around any frequency in the audible range which we care to choose.

4 MASKING PATTERNS AND EXCITATION PATTERNS

So far we have discussed masking experiments in which the frequency of the signal is held constant, while the masker is varied. These experiments are most appropriate for estimating the shape of the auditory filter at a given centre 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) published 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 graph plotting masked threshold as a function of the frequency of the signal 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 (e.g. Egan and Hake, 1950; Greenwood, 1961a) have used a narrow band of noise as either the signal or the masker. Such a noise has 'built in' amplitude and frequency variations and does not produce regular beats when added to a tone.

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 to some extent on the level of the masker. A typical set of results is shown in Fig. 3.10. Notice that on the high-frequency side the slopes of the curves tend to become shallower

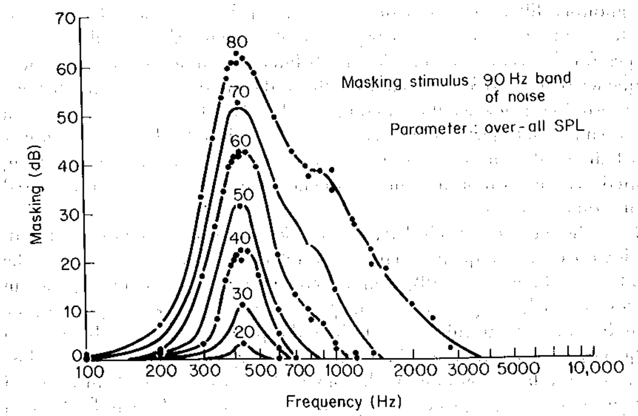

FIG. 3.10 Masking patterns (masked audiograms) for a narrow band of noise centred at 410 Hz. Each curve shows the elevation in threshold of a pure tone signal as a function of signal frequency. The overall noise level for each curve is indicated in the figure. Adapted from Egan and Hake (1950), by permission of the authors and *J. Acoust. Soc. Am.*