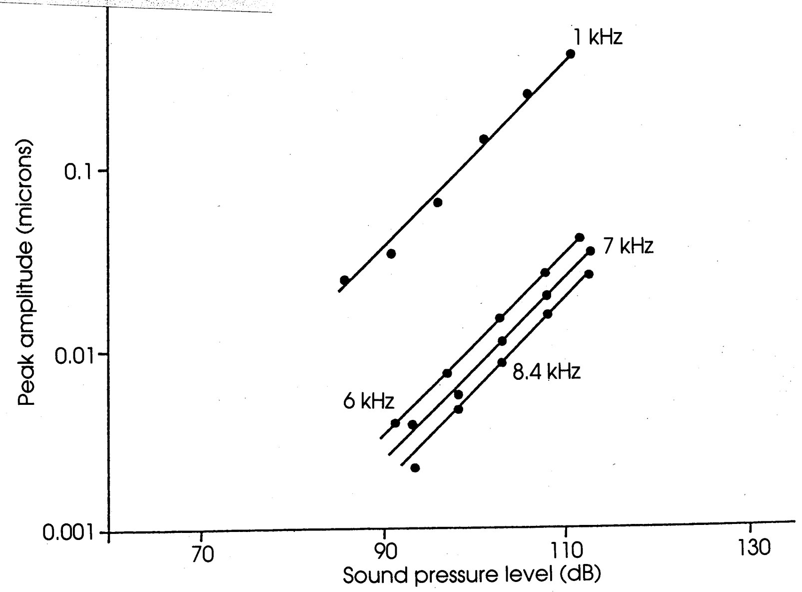
Rosen & Howell
252–257



The causes of the discrepancies between Békésy's and Rhode's results on basilar membrane vibration arise from crucial methodological differences between the two studies. Almost certainly the essential one is that Békésy always used preparations from cadavers. Rhode used live animals and found the nonlinearity to disappear very quickly once the animal had died. Also, Békésy's technique necessitated the use of extremely high signal levels (up to 140–150 dB SPL) in order to make the movements visible, whereas Rhode used levels that would be encountered in everyday life. Nonlinearities in basilar membrane movements have since been found by other groups of experimenters in cats, chinchillas and guinea pigs, and there is now general agreement that the system is nonlinear.

The details of this controversy are, for our purposes, less important than the way in which the concepts of linear systems analysis pervade the entire discussion. LTI systems serve as a benchmark against which other systems can be compared, hence much effort goes into determining the exact nature of the departures from linearity found in a nonlinear system. Therefore, when Rhode claims that basilar membrane motion is nonlinear, he can only do so on the basis of what would be expected of a linear system. This is yet another reason why an understanding of linear systems analysis is crucial to appreciate discussions of even nonlinear systems.

Auditory filters

The usefulness of the concepts of linear systems analysis even for nonlinear systems is well illustrated in our final example, in which we discuss a system that no one has ever supposed to be linear—the entire auditory system of a human being! It has long been known that the frequency analysis exhibited by the basilar membrane is also reflected in human auditory behaviour. Historically, in fact, it was perceptual observations that first led to the suggestion that the basilar membrane operated as a frequency analyser.

This is most easily demonstrated in the context of *masking* experiments. The basic idea is simple. First, a sound is chosen—it can be any sound but usually sinusoids are used. Then, the lowest level at which this sound (known as the *probe*) can be heard is determined. This level would be the *threshold* of the probe under quiet conditions. Then another sound is chosen—known as the *masker*. The threshold of the probe is determined again, only now when it and the masker are presented at the same time. If the presence of the masker makes the probe more difficult to hear—that is, if the level of the probe has to be raised to make it audible—we say that the masker *masks* the probe. The amount of masking is usually expressed as the difference between the level of the just-audible probe under quiet conditions and the level of the just-audible probe in the presence of the masker.

The most striking feature of masking experiments with sinusoidal maskers and probes is that the closer together in frequency the masker and the probe are, the more masking occurs. To put it the other way around, the more different in frequency two sinusoids are, the less they will mask one another. It is as if sinusoids of rather different frequencies are processed by different parts of the auditory system, so they don't interfere with one another. Sinusoids close in frequency seem to go through the same system, and so presenting one affects perception of the other.

Experiments like this led to the concept of the *critical band*, the notion that when spectral components in sounds differ in frequency by less than a certain "critical" amount, they interact. Underlying all such theories is the idea that the auditory system can be regarded as consisting of a large number of *channels* (another word for system) connected up in parallel, the primary difference between the channels being in the frequency range over which each one is responsive. Each one of these channels can then be thought of as a critical band.

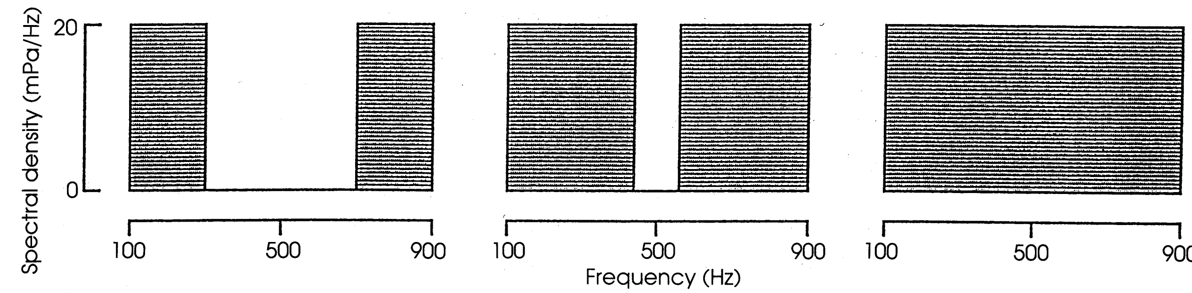
You have already met a system similar to this in Chapter 11—the filter bank. In addition, we've just discussed how the basilar membrane can be thought of as a filter bank. This analogy is probably the single most important concept in the study of auditory perception. The entire auditory system is often characterized as operating like a filter bank in important respects, although all acknowledge that nonlinearities exist.

If we think of the auditory system as being made up of a large number of channels, each of which is preceded by an *auditory filter* which determines what signals get into that channel and what signals don't, what do we need to know about this system to characterize it? Let's assume, as is usually done, that the operations that the channels perform after the filtering are the same in each channel (just as in the spectrograph, where the rectification and smoothing didn't vary). This means that differences in behaviour between the channels must be solely due to the differences in the initial filtering. (Convince yourself that this was true about the spectrograph.) Therefore, a crucial aspect of the functioning of this auditory filter bank will be the properties of the filters—for example, their centre frequencies and bandwidths. In normal listeners there are thousands of filters spread across the frequency range of hearing, so we don't normally need to know their centre frequencies. More informative would be a measure of their bandwidth (or, even better, the shape of their amplitude response) because that would give us a way to predict how the auditory filter bank would respond to particular sounds. We should then be able to, say, predict the degree of masking of one sound on another, or even try to explain the perceptual similarity between different vowel-like sounds.

There have been a number of techniques developed over many years to characterize auditory filters. We will only discuss one of the more recent of these, developed most extensively by Patterson and his colleagues. It has a number of advantages, the main one for our purposes being that it is easy to understand!

At its most basic, Patterson's technique is nothing more than a masking experiment with a standard probe but some specially constructed maskers. First, the frequency of a sinusoidal probe tone is chosen. This will determine the centre frequency of the auditory filter which will be investigated because we assume that the subject is always listening through an auditory filter centred on the probe tone (an assumption that is not always possible to defend).

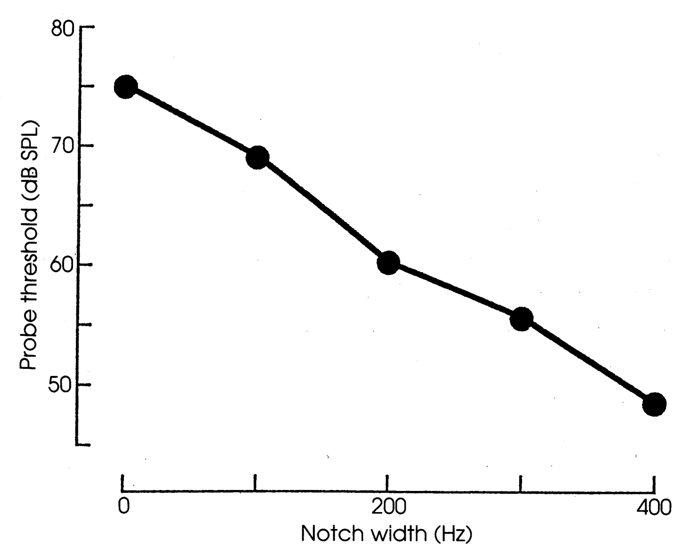
The maskers are a set of fairly wide-band noises, most of which have a spectral notched centred on the frequency of the probe tone. In other words, there is an equal amount of energy in the maskers for a wide range of frequencies except for a region around the probe frequency where there is very little energy. Usually a number of these maskers is created with the width of the spectral notch varying from no notch at all to quite a wide notch. The idealized amplitude spectra of three such noises is shown here (as shaded rectangles), assuming a probe tone of 500 Hz:



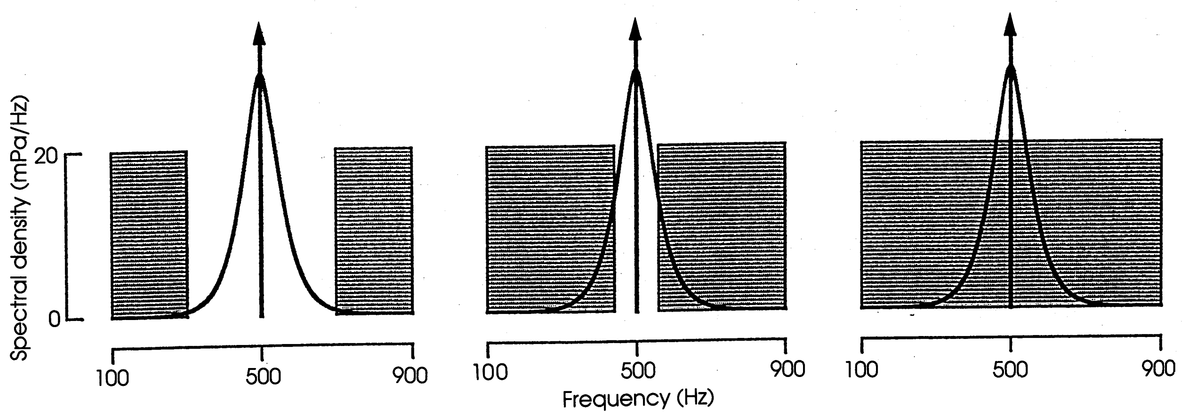
The noise on the left has a wide notch, whereas the middle one has a narrow one. These are known as *notched noise*. The noise on the right has no notch at all; it has a constant spectral density over the frequency range of interest.

There are a number of different ways to create notched noises. The simplest is to pass white noise through a band-stop filter, first described in Chapter 6. Recall that a band-stop filter can be created by adding together the outputs of a high-pass and a low-pass filter as long as the low-pass cutoff frequency is less than the high-pass cutoff.

Once the noises are created, we determine the threshold of the sinusoidal probe as a function of the width of the notch. Here are some actual measurements made on a normal listener. You can see that the threshold for the probe decreases as the notch in the noise is made wider:



How are we to interpret such a result? Let's make a crude model of what's happening. Suppose the subject is listening through only one of his or her auditory channels, the one most sensitive to 500 Hz. We assume that this channel (or critical band) is fed from a filter whose centre frequency is 500 Hz. When the masker puts a lot of energy through the filter, it will become harder for the subject to hear the probe—the masker will interfere with its detection. When the masker only puts a little energy through the filter, the detection of the probe won't be much affected. In short, the more energy in the masker in the pass-band of the auditory filter, the more the detectability of the probe will suffer. Because the wider-notch noises have less energy in the pass-band of the auditory filter than the narrow-notch noises, they should affect the detectability of the probe less. This is easily seen graphically if we superimpose the frequency response of an assumed reasonable auditory filter over the spectra of the probe and masker for the three different maskers illustrated above:

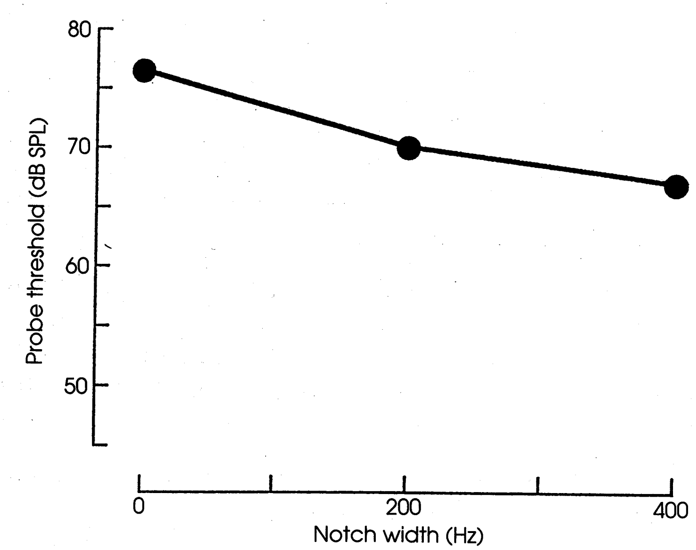


In each of the three examples, the probe is represented as a single component since it is sinusoidal, while the spectrum of the noise is shown lightly shaded. Don't be confused by the arrow at the top of the sinusoid—this is simply to make it stand out in the figure. In all three cases, the probe passes equally well through the auditory filter. For the leftmost case the notch in the noise is wide relative to the filter bandwidth, so little noise gets through the filter. Therefore, little masking occurs and the threshold of the probe is near that found under quiet conditions. For the noise in the middle, however, the notch is fairly narrow compared to the bandwidth of the auditory filter, and so a significant amount of noise gets through to raise the threshold of the probe. Finally, when the noise is white, lots of noise gets through the filter and so the threshold of the probe is at its maximum.

It should be apparent that the narrower the auditory filter, the faster the threshold of the probe will decrease with increasing notch width. Conversely, if the auditory filter is very wide, the threshold of the probe will change fairly gradually with

increasing notch width. It it this principle that is used to determine the width of the auditory filter, and, more generally, its shape. Interestingly, it appears that the auditory filter bandwidth changes with overall level, wider bandwidths being associated with higher levels. The system is thus nonlinear, in a way which is reminiscent of Rhode's results on the basilar membrane.

Such techniques have also come to be used extensively in exploring auditory perception in hearing-impaired listeners. It has been shown that listeners with only a moderate loss in sensitivity can show very large increases in auditory filter bandwidth. This is reflected in the notched-noise experiment with a curve that is rather flat, as shown for one hearing-impaired listener here:



The wider auditory filters found in hearing-impaired listeners contribute to the difficulties they have in understanding speech. Spectral details will tend to be smeared, and there will be more interaction between parts of the speech in different frequency ranges. Outside noises will be more effective maskers. To put it another way, whereas a normal auditory system keeps "each frequency in its place", listeners with widened auditory filters will have everything much more "mixed up". This ties in with the subjective experience of hearing aid users, who have little trouble "hearing" sounds, but can't distinguish them from one another. Current commercial hearing aids can make sounds *louder*, but they will not necessarily make them *clearer*. A thorough understanding of what happens to auditory filtering in the hearing-impaired should make it possible to design hearing aids to do just this.