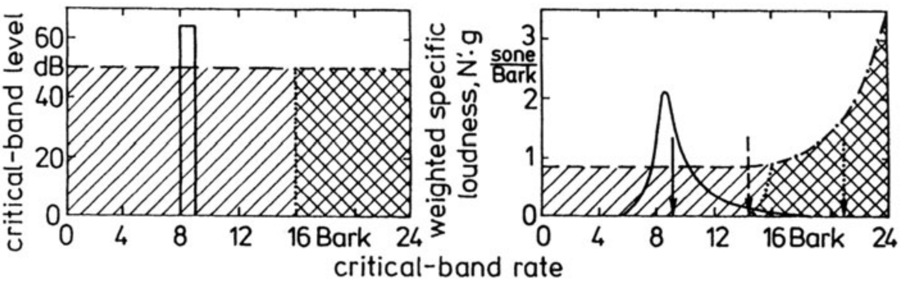

Fig. 9.3. Calculation of sharpness for narrow-band noise centred at 1 kHz (*solid*), uniform exciting noise (*broken* and *hatched* from lower left to upper right) and high-pass noise (*dotted* and *hatched* from upper left to lower right). The left-hand drawing indicates critical-band levels as a function of critical-band rate, while the right-hand drawing indicates the weighted specific loudness, again as a function of critical-band rate. The calculated sharpnesses are indicated by the three vertical arrows

and the noise is changed into a uniform exciting noise, sharpness decreases markedly in agreement with psychoacoustical results. It remains, however, clearly above that of the 1-kHz tone.

9.3 Dependencies of Sensory Pleasantness

Sensory pleasantness is a more complex sensation that is influenced by elementary auditory sensations such as roughness, sharpness, tonality, and loudness. Because of these influences, which make it almost impossible to extract sensory pleasantness as a single elementary sensation, it is necessary to measure the dependence of this sensation in relative values, using the techniques of magnitude estimation with an anchor.

The dependence of sensory pleasantness on sharpness was measured using sinusoidal tones, narrow-band noise of 30-Hz bandwidth and band-pass noise with a 1-kHz bandwidth as a function of centre frequency. Relative sharpness and relative sensory pleasantness were determined psychoacoustically in separate sessions. The data show some scatter. The relationships between sensory pleasantness and other sensations are given as curves in Fig. 9.4. Figure 9.4a shows pleasantness against relative roughness and Fig. 9.4b pleasantness against sharpness for the three sounds. It becomes clear that sensory pleasantness decreases with increasing sharpness. Pure tones already show the largest sensory pleasantness, while the band-pass noise seems to be a sound with low sensory pleasantness.

Similar to the dependence on sharpness, sensory pleasantness depends on roughness, a hearing sensation described in Chap. 11, although the dependence is not as strong. Because the data are again measured using the

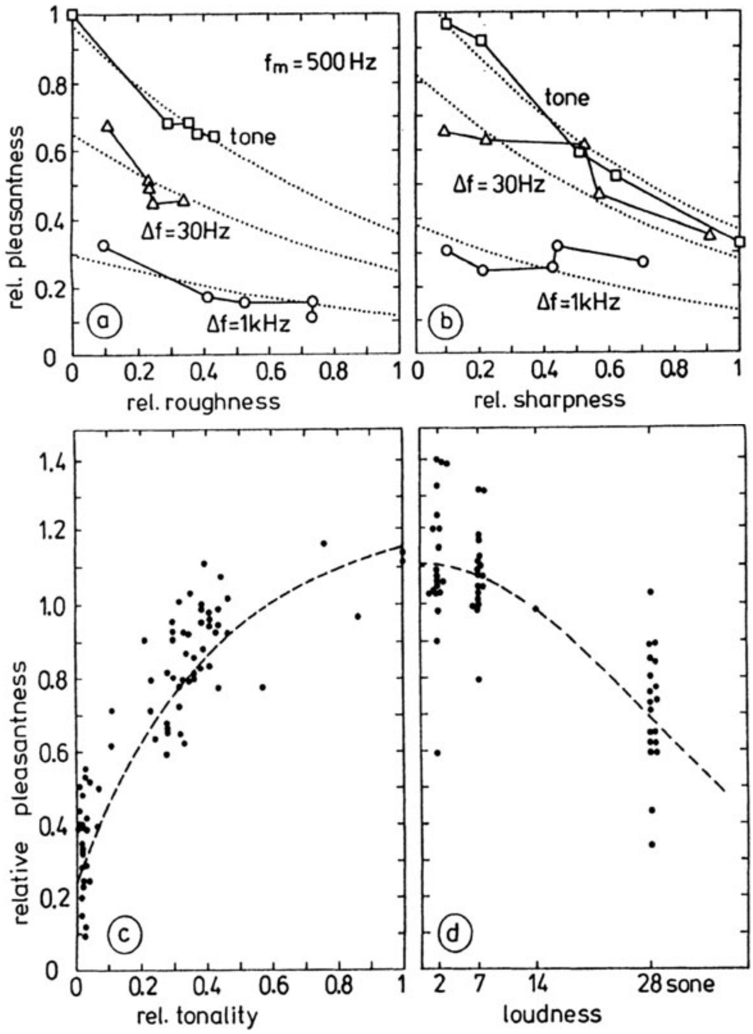

Fig. 9.4a–d. Relative pleasantness as a function of relative roughness with bandwidth as the parameter in (**a**); as a function of relative sharpness in (**b**); as a function of relative tonality in (**c**) and as a function of loudness in (**d**)

method of magnitude estimation with an anchor, only relative values can be given.

The relationship between sensory pleasantness and tonality, i.e. a feature distinguishing noise versus tone quality of sounds, is indicated in Fig. 9.4c. This dependence indicates that sensory pleasantness increases with tonality. Small tonality means small sensory pleasantness. For relative tonality larger than about 0.4, sensory pleasantness does not increase much.

The dependences of sensory pleasantness described so far have been determined using a constant loudness of 14 sone. Sensory pleasantness, however, depends also on loudness in such a way that up to about 20 sone, there is little influence. For values larger than 20 sone, sensory pleasantness decreases. This dependence of sensory pleasantness on loudness cannot be seen in isolation because roughness and sharpness also depend on loudness. If this influence is eliminated, then the relationship between sensory pleasantness and loudness, as given in Fig. 9.4d, remains.

In summary, it can be seen that sensory pleasantness depends mostly on sharpness. The influence of roughness is somewhat smaller and is similar to that of tonality. Loudness, however, influences sensory pleasantness only for values that are larger than the normal loudness of communication between two people in quiet.

9.4 Model of Sensory Pleasantness

Whether a sound is accepted as sounding pleasant or unpleasant depends not only on the physical parameters of the sound but also on the subjective relationship of the listener to the sound. These nonacoustic influences cannot be anticipated and have therefore to be ignored and eliminated if possible. This was done in the experiments, so that a model of sensory pleasantness can only be developed relating the characteristics of the human hearing system on one hand and the physical parameters of the sound on the other.

The relationship between relative values of sensory pleasantness and those of sensation sharpness, roughness, tonality, and loudness can be approximated using equations. A model for the calculation of sharpness has already been given. The sensation of roughness can also be calculated (see Chap. 11). A procedure for calculating the sensation of tonality does not yet exist so that it may be subjectively estimated as a first approximation, while loudness can be calculated relatively precisely. Therefore, it is possible to put the dependencies given in Fig. 9.4 into an equation that contains relative values of sharpness, roughness, tonality, and loudness. The result leads also to a relative value, P/P_0, of sensory pleasantness. The equation, based on relative values of sharpness S, roughness R, tonality T, and loudness N, reads:

$$\frac{P}{P_0} = \mathrm{e}^{-0.7R/R_0}\mathrm{e}^{-1.08S/S_0}\left(1.24 - \mathrm{e}^{-2.43T/T_0}\right)\mathrm{e}^{-(0.023N/N_0)^2} \tag{9.2}$$

Using this equation, it is possible to calculate the sensory pleasantness of any sound, if the sharpness, roughness, and loudness are calculated using the procedures given in Sects. 9.2, 11.2, and 8.7, respectively. Tonality has to be judged subjectively. It has been shown that tonality depends neither on the critical-band rate nor on the loudness. Relative tonality, however, depends on the bandwidth expressed in critical-band-rate spread, such that it decreases with increasing critical-band-rate spread starting from a sinusoidal tone producing relative tonality of unity to about 0.6, 0.3, 0.2, and 0.1 for critical-band-rate spreads of 0.1, 0.2, 0.57, and 1.5 Bark, respectively. Using these values it is possible to estimate the relative sensory pleasantness of different sounds. Results calculated according to (9.2) are indicated in Fig. 9.4 by dotted or dashed lines.

10. Fluctuation Strength

In this chapter, the fluctuation strength of amplitude-modulated broad-band noise, amplitude-modulated pure tones and frequency-modulated pure tones is addressed, and the dependence of fluctuation strength on modulation frequency, sound pressure level, modulation depth, centre frequency, and frequency deviation is assessed. In addition, the fluctuation strength of modulated sounds is compared to the fluctuation strength of narrow-band noises. Finally, a model of fluctuation strength based on the temporal variation of the masking pattern or loudness pattern is proposed.

10.1 Dependencies of Fluctuation Strength

Modulated sounds elicit two different kinds of hearing sensations: at low modulation frequencies up to a modulation frequency of about 20 Hz, the hearing sensation of fluctuation strength is produced. At higher modulation frequencies, the hearing sensation of roughness, discussed in detail in Chap. 11, occurs. For modulation frequencies around 20 Hz, there is a transition between the hearing sensation of fluctuation strength and that of roughness. It is a smooth transition rather than a strong border that exists between the two sensations.

Figure 10.1 shows the dependence of fluctuation strength on modulation frequency. The different panels represent the data for amplitude-modulated broad-band noise (AM BBN), amplitude-modulated pure tone (AM SIN) and frequency-modulated pure tone (FM SIN). In each panel, the fluctuation strength was normalized to the maximum value for that sound (left-hand ordinate). Because fluctuation strength is a sensation which one considers separately from other sensations, both absolute and relative values are useful. A fixed point is therefore defined for a 60-dB, 1-kHz tone 100% amplitude-modulated at 4 Hz , as producing 1 vacil (from vacilare in Latin, or vacillate in English). Using the data of fluctuation strength shown in Fig. 10.7, it is possible to give the absolute values as indicated in the right-hand ordinate scales.

All three panels of Fig. 10.1 clearly show that fluctuation strength shows a band-pass characteristic as a function of modulation frequency, with a maximum around 4 Hz. This means that sounds with a 4-Hz modulation frequency

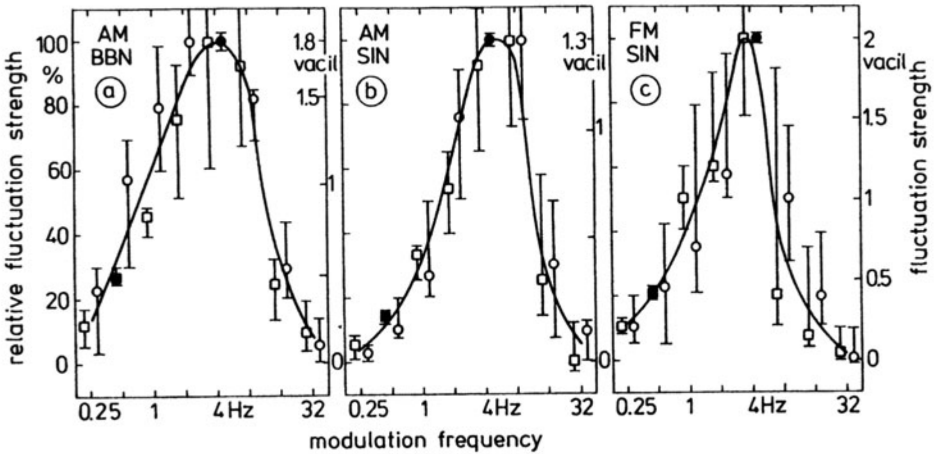

Fig. 10.1a–c. Fluctuation strength of three modulated sounds as a function of modulation frequency. (**a**) Amplitude-modulated broad-band noise of 60-dB SPL and 40-dB modulation depth; (**b**) amplitude-modulated 1-kHz tone of 70-dB SPL and 40-dB modulation depth; (**c**) frequency-modulated pure tone of 70-dB SPL, 1500-Hz centre frequency and $\pm$700-Hz frequency deviation

elicit large fluctuation strength, whether amplitude modulation or frequency modulation is used or whether broad-band or narrow-band sounds are modulated. The maximum fluctuation strength for a modulation frequency of about 4 Hz finds its counterpart in the variation of the temporal envelope of fluent speech: at normal speaking rate, 4 syllables/second are usually produced, leading to a variation of the temporal envelope at a frequency of 4 Hz. This may be seen as an indication of the excellent correlation between speech and hearing system.

The dependence of fluctuation strength on sound pressure level is displayed in Fig. 10.2. Again, the three panels show the results for amplitude-modulated broad-band noise, amplitude-modulated pure tone and frequency-modulated pure tone. In each panel, the fluctuation strength is normalized with respect to the corresponding maximum value on the left ordinate scales, and given in absolute values on the right. With increasing sound pressure level and for all sounds considered, the fluctuation strength increases. For amplitude-modulated sounds (Fig. 10.2a and b), the increase is somewhat more prominent than for the frequency-modulated pure tone (Fig. 10.2c). With an increase of 40 dB in sound pressure level, fluctuation strength of modulated sounds increases on average by a factor of about 2.5 (1.7 to 3).

Figure 10.3 shows the dependence of the fluctuation strength of amplitude-modulated broad-band noise and amplitude-modulated pure tones on both the modulation depth and the modulation factor. In each panel, the fluctuation strength is normalized with respect to the corresponding maximum value on the left ordinate scales and given in absolute values on the right ordinate

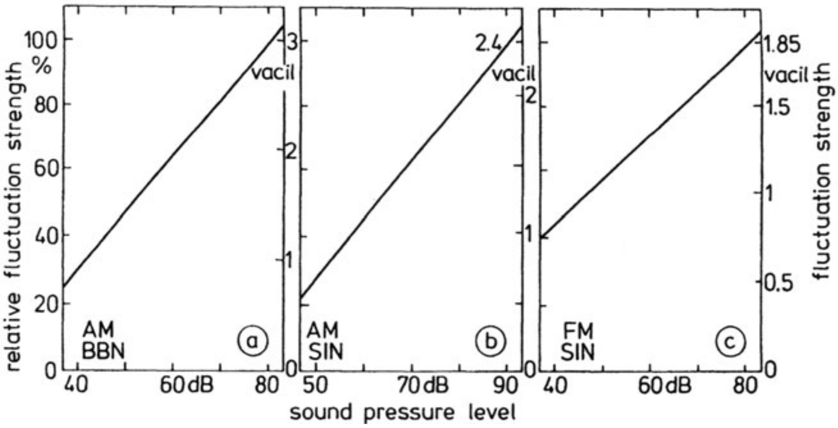

Fig. 10.2a–c. Fluctuation strength of modulated sounds as a function of sound pressure level. Stimulus parameters are the same as in Fig. 10.1, but the modulation frequency is 4 Hz

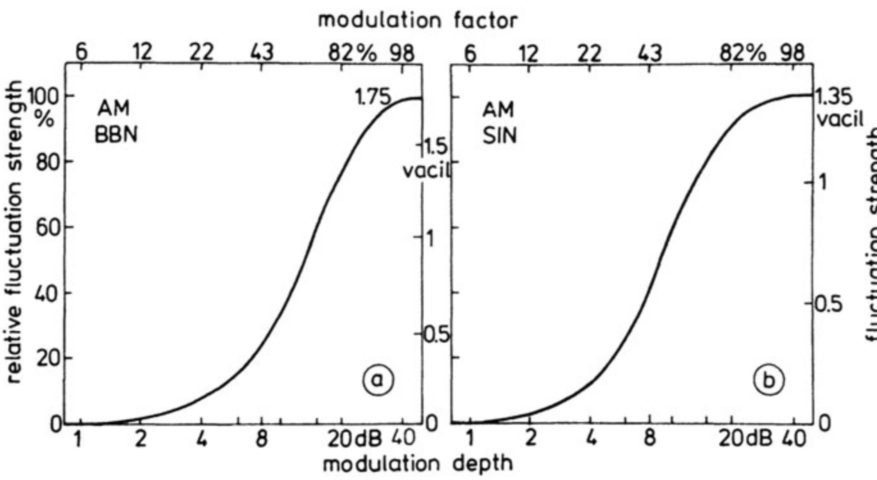

Fig. 10.3a,b. Fluctuation strength of two amplitude-modulated sounds as a function of modulation depth (or modulation factor). (**a**) Amplitude-modulated broad-band noise of 60-dB SPL and 4-Hz modulation frequency; (**b**) amplitude-modulated 1-kHz tone of 70-dB SPL and 4-Hz modulation frequency

scales. According to the results displayed in Fig. 10.3, fluctuation strength is zero until a modulation depth of about 3 dB, after which it increases approximately linearly with the logarithm of modulation depth. To produce the maximum fluctuation strength of either sound, a modulation depth of at least 30 dB (modulation factor 94%) is necessary. Above that modulation depth, fluctuation strength remains constant at its maximal value.

Figure 10.4 shows the dependence of the fluctuation strength produced by modulated pure tones on centre frequency. In the left panel, results for amplitude-modulated pure tones are shown; the right panel indicates data for frequency-modulated pure tones. In each panel, the data are normalized relative to the respective maximal median fluctuation strength, but are also given in absolute values. The data displayed in Fig. 10.4a suggest that the fluctuation strength of amplitude-modulated pure tones depends very little on their centre frequency; despite the large interquartile ranges, the medians indicate the tendency for amplitude-modulated tones at very low (125 Hz) and very high (8 kHz) frequencies to produce less fluctuation strength than AM tones at medium frequencies. However, the results shown in Fig. 10.4b for FM tones indicate a clear dependence of fluctuation strength on centre frequency so that, although the fluctuation strength of FM tones is almost constant up to about 1 kHz, it decreases approximately linearly with the logarithm of the centre frequency towards higher frequencies.

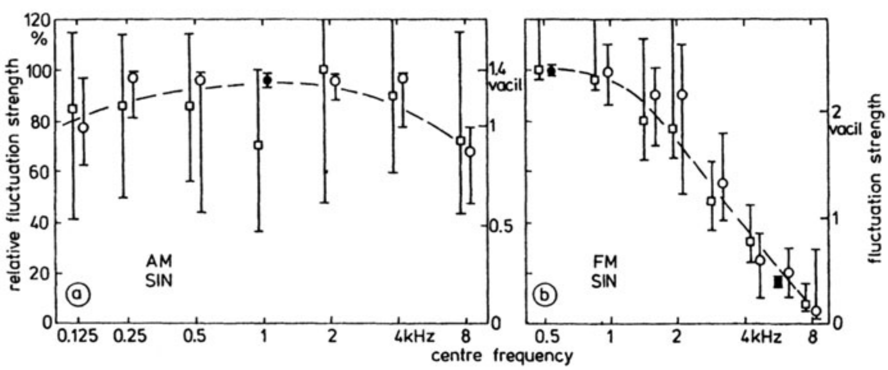

Fig. 10.4a,b. Fluctuation strength of modulated tones as a function of frequency. (**a**) Amplitude-modulated pure tone of 70-dB SPL, 4-Hz modulation frequency and 40-dB modulation depth; (**b**) frequency-modulated pure tone of 70-dB SPL, 4-Hz modulation frequency, and ± 200-Hz frequency deviation

This decrease can be understood in terms of the number of critical bands encompassed at different centre frequencies by FM tones with a constant frequency deviation of 200 Hz. As an example, the FM tone at 0.5 kHz sweeps between 300 and 700 Hz, i.e. between critical-band rates of 3 and 6.5 Bark, respectively. At the 8-kHz centre frequency, the modulation occurs between frequencies of 7.8 and 8.2 kHz, corresponding to 21.1 and 21.3 Bark, respec-

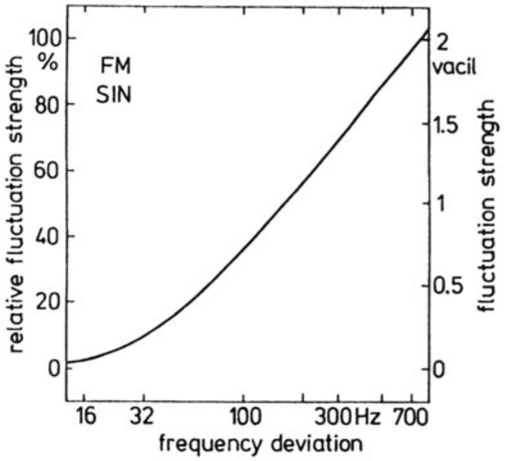

Fig. 10.5. Fluctuation strength of a frequency-modulated tone as a function of frequency deviation. The sound pressure level is 70 dB, the centre frequency 1500 Hz and the modulation frequency 4 Hz

tively. This means that at a centre frequency of 0.5 kHz, the FM tone varies over a critical-band interval of 3.5 Bark, whereas at 8 kHz it varies only over 0.2 Bark. Hence, the critical-band interval at 8 kHz is a factor of 17.5 smaller than the critical-band interval at 0.5 kHz. Regarding Fig. 10.4b, it can be seen that this factor of 17.5 is also found for the difference in the relative fluctuation strength at 0.5 and 8 kHz. This result can be taken as an indication that fluctuation strength of modulated sounds can be described on the basis of the corresponding excitation patterns.

Figure 10.5 shows the dependence of fluctuation strength on frequency deviation of an FM tone at a centre frequency of 1.5 kHz. Fluctuation strength is initially perceived at a frequency deviation of about 20 Hz and increases approximately linearly with the logarithm of frequency deviation. This result applies for an FM tone at 1500 Hz with 70-dB SPL and 4-Hz modulation frequency. For such a tone, the JND for frequency modulation corresponds to about $2\Delta f = 8\,\text{Hz}$ (see Fig. 7.8). Significant values of fluctuation strength (say 10% relative fluctuation strength) are achieved for frequency deviations larger than about 10 times the magnitude of the JNDFM at 4 Hz. This rule seems to apply also for AM sounds: the modulation depth at which about 10% relative fluctuation strength is reached (see Fig. 10.3), is about 10 times larger than the JNDAM at 4 Hz of about 0.4 dB for a 70-dB AM tone, and about 0.7 dB for the AM broad-band noise as displayed in Fig. 7.1.

Not only modulated sounds can elicit the hearing sensation fluctuation strength but also unmodulated narrow noise bands.

Figure 10.6 shows the dependence of relative fluctuation strength of narrow-band noise as a function of its bandwidth as well as effective modulation frequency. According to (1.6) this frequency can be calculated as follows: $f^*_{\text{mod}} = 0.64 \cdot \Delta f$.

The comparison of the data displayed in Figs. 10.1 and 10.6 reveals that – irrespective of periodic or stochastic sound fluctuations – fluctuation strength

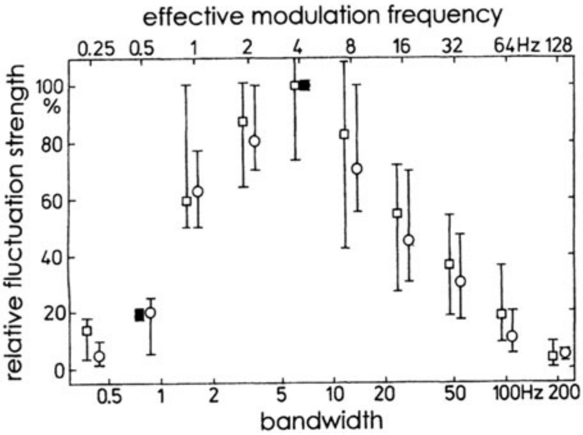

Fig. 10.6. Fluctuation strength of narrow-band noise as a function of its bandwidth. Center frequency 1 kHz, level 70 dB

shows a bandpass characteristic with a maximum around 4 Hz (effective) modulation frequency.

For unmodulated narrow-band noise, fluctuation strength increases with level and shows large values for center frequencies around 1 kHz similar to the data displayed for modulated sounds in Figs. 10.2 and 10.4a.

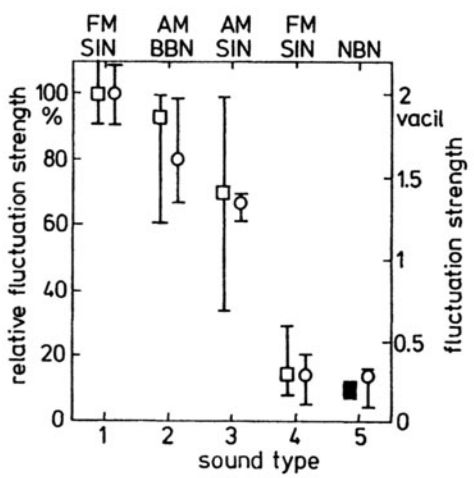

Fig. 10.7. Fluctuation strength of the sounds 1–5 as described in Table 10.1

Figure 10.7 enables a comparison of the fluctuation strength of five different sounds whose characteristics are listed in Table 10.1. The largest fluctuation strength is produced by the 70-dB tone with large frequency deviation. Some 10% less fluctuation strength is elicited by a 60-dB, amplitude-modulated broad-band noise. A 70-dB, amplitude-modulated 2-kHz tone pro-

Table 10.1. Physical data of sounds 1–5

Sound	1	2	3	4	5
Abbreviation	FM SIN	AM BBN	AM SIN	FM SIN	NBN
Frequency [Hz]	1500	-	2000	1500	1000
Level [dB]	70	60	70	70	70
Modulation frequency [Hz]	4	4	4	4	-
Modulation depth [dB]	-	40	40	-	-
Frequency deviation [Hz]	700	-	-	32	-
Bandwidth [Hz]	-	16000	-	-	10

duces a fluctuation strength about 30% down from that of the FM tone. Sound 4, a 70-dB FM tone with small frequency deviation, produces only about 1/10 of the fluctuation strength of sound 1. This result is expected on the basis of the data displayed in Fig. 10.5. Sound 5 represents a narrow-band noise with a bandwidth of 10 Hz. The fluctuation strength elicited by this narrow-band noise can be estimated as follows: as a first approximation, the narrow-band noise can be regarded as an AM tone at 1 kHz with 6.4 Hz modulation frequency (see Sect. 1.1). If an effective modulation factor of 40% for narrow-band noise is assumed, then according to the results displayed in Fig. 10.3, the fluctuation strength of narrow-band noise should be a factor of about 2.5 smaller than the fluctuation strength of AM tones with a 98% modulation factor. A comparison of the relative fluctuation strength of sound 3 and sound 5 in Fig. 10.7, however, reveals that the fluctuation strength of narrow-band noise is smaller by a factor of about 5 than the fluctuation strength of an AM tone. It is apparently the periodic fluctuation of AM tones, in contrast to the random amplitude fluctuations of the noise, that enhances the perceived fluctuation strength of the AM tone.

The large fluctuation strength of amplitude-modulated broad-band noise and frequency-modulated pure tones with large frequency deviation (sounds 2 and 1) can be related to the fact that excitation varies to a large extent along the critical-band rate scale. Therefore, it can be postulated that fluctuation strength is summed up across critical bands. This concept will be explained in more detail in the following section.

10.2 Model of Fluctuation Strength

A model of fluctuation strength based on the temporal variation of the masking pattern can be illustrated by Fig. 10.8, where the temporal masking

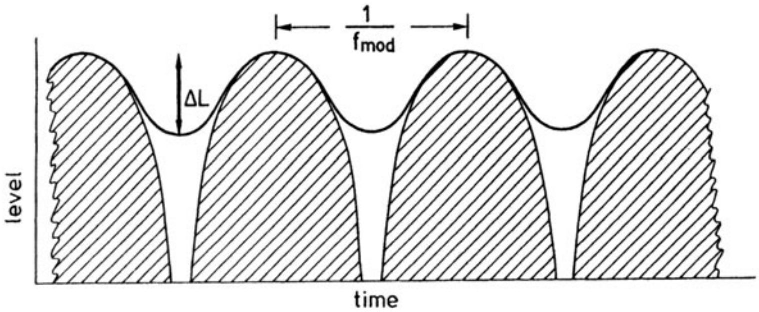

Fig. 10.8. Model of fluctuation strength: temporal masking pattern of sinusoidally amplitude-modulated masker leading to the temporal masking depth ΔL

pattern of a sinusoidally amplitude-modulated masker is schematically indicated by the thick solid line. The hatched areas indicate the envelope of a sinusoidally amplitude-modulated masker plotted in terms of sound pressure level. The interval between two successive maxima of the masker envelope corresponds to the reciprocal of the modulation frequency. The temporal variation of the temporal masking pattern can be described by the magnitude ΔL, which represents the level difference between the maxima and the minima in the temporal masking pattern. This so-called temporal masking depth, ΔL, should not be confused with the modulation depth, d, of the masker's envelope; the masking depth ΔL of the temporal masking pattern is smaller than the modulation depth d of the masker's envelope due to postmasking.

The equation

$$F \sim \frac{\Delta L}{(f_{\mathrm{mod}}/4\mathrm{Hz}) + (4\mathrm{Hz}/f_{\mathrm{mod}})}, \tag{10.1}$$

shows the relationship between fluctuation strength, F, and the masking depth of the temporal masking pattern, ΔL, as well as the relationship between F and modulation frequency f_{mod}. The denominator clearly shows that a modulation frequency of 4 Hz plays an important part in the description of fluctuation strength: for faster modulation frequencies, the ear exhibits the integrative features evidenced in postmasking; for modulation frequencies lower than 4 Hz, effects of short-term memory become important.

For an amplitude-modulated broad-band noise, the magnitude of the masking depth of the temporal masking pattern, ΔL, is largely independent of frequency. For amplitude-modulated pure tones, some frequency dependence occurs because of the nonlinearity of the upper slope in the masking pattern. In addition to those factors, the magnitude of the masking depth shows a strong frequency dependence with frequency-modulated tones. This means that, for describing the fluctuation strength of AM tones and FM tones, the model can be modified as follows: instead of *one* masking depth of the temporal masking pattern ΔL, all of the magnitudes of ΔL occuring

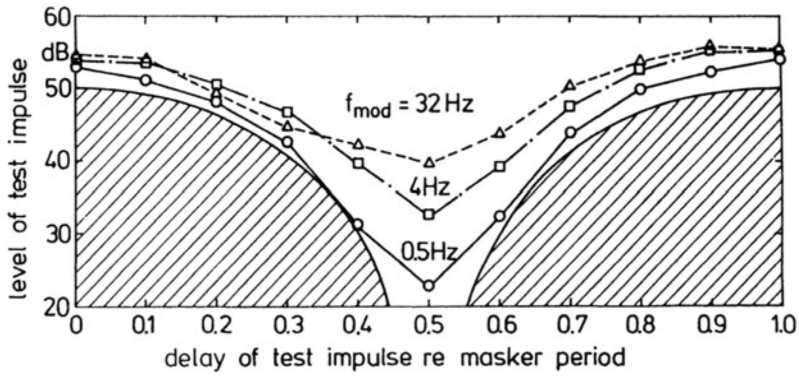

Fig. 10.9. Temporal masking pattern of 100% sinusoidally amplitude-modulated broad-band noise. *Hatched*: temporal envelope of masker; *parameter*: modulation frequency. All data plotted as a function of the temporal location within the period of the modulation of the masker

are integrated along the critical-band rate scale. A more detailed description of this procedure is given in Chap. 11, where the model for roughness is discussed.

A basic feature of the model of fluctuation strength, namely the masking depth of the temporal masking pattern, can be illustrated in Fig. 10.9. The temporal envelope of a sinusoidally amplitude-modulated broad-band noise is indicated by the hatched areas. The different curves represent the temporal masking patterns for the different modulation frequencies indicated. The maximum of the masking pattern shows up at a delay relative to the masker period of 0, the minimum at a delay of 0.5. The difference between maximum and minimum, called the temporal masking depth, clearly decreases with increasing modulation frequency. This means that as a function of modulation frequency, the temporal masking pattern shows a low-pass characteristic, whereas fluctuation strength shows a band-pass characteristic. Equation (10.1) explains how the low-pass characteristic of the temporal envelope is transformed into the band-pass characteristic of fluctuation strength. This bandpass characteristic describes the influence of modulation frequency on fluctuation strength. However, the value ΔL in the formula decreases approximately linearly with increasing frequency of modulation. Taking this into account, a relatively simple formula can be given for the fluctuation strength of sinusoidally amplitude-modulated broad-band noise:

$$F_{\mathrm{BBN}} = \frac{5.8(1.25m - 0.25)[0.05(L_{\mathrm{BBN}}/\mathrm{dB}) - 1]}{(f_{\mathrm{mod}}/5\mathrm{Hz})^2 + (4\mathrm{Hz}/f_{\mathrm{mod}}) + 1.5}\ \mathrm{vacil}\,, \tag{10.2}$$

where m is the modulation factor, L_{BBN} the level of the broad-band noise and f_{mod} the frequency of modulation. For amplitude- or freqency-modulated tones, fluctuation strength may be approximated by integrating the temporal masking depth, ΔL, along the critical-band rate. This leads to the following approximation:

$$F = \frac{0.008 \int_0^{24\mathrm{Bark}} (\Delta L/\mathrm{dB\,Bark}) dz}{(f_{\mathrm{mod}}/4\mathrm{Hz}) + (4\mathrm{Hz}/f_{\mathrm{mod}})} \ \mathrm{vacil}\,, \tag{10.3}$$

where the masking depth, ΔL, may be picked up from the masking patterns described in Chap. 4 for the different critical-band rates of 1 Bark distance. The integral is then transformed into a sum of, at most, 24 terms along the whole range of the critical-band rate.

While for most sounds described in this chapter values of ΔL are available this is usually not the case for sounds typical for practical applications. Therefore, a computer program was developed which uses instead of the ΔL values the corresponding differences in specific loudness. Since the computer models of fluctuation strength and roughness are very similar some more detail is given in Chap. 11 and the related literature.

11. Roughness

Using a 100% amplitude-modulated 1-kHz tone and increasing the modulation frequency from low to high values, three different areas of sensation are traversed. At very low modulation frequencies the loudness changes slowly up and down. The sensation produced is that of fluctuation. This sensation reaches a maximum at modulation frequencies near 4 Hz and decreases for higher modulation frequencies. At about 15 Hz, another type of sensation, roughness, starts to increase. It reaches its maximum near modulation frequencies of 70 Hz and decreases at higher modulation frequencies. As roughness decreases, the sensation of hearing three separately audible tones increases. This sensation is small for modulation frequencies near 150 Hz; it increases strongly, however, for larger modulation frequencies. This behaviour indicates that roughness is created by the relatively quick changes produced by modulation frequencies in the region between about 15 to 300 Hz. There is no need for exact periodical modulation, but the spectrum of the modulating function has to be between 15 and 300 Hz in order to produce roughness. For this reason, most narrow-band noises sound rough even though there is no periodical change in envelope or frequency. Roughness is again a sensation which we can consider while ignoring other sensations.

11.1 Dependencies of Roughness

In order to describe roughness quantitatively, a reference value must be defined. In Latin, the word "asper" characterizes what we call "rough". To define the roughness of 1 asper, we have chosen the 60-dB, 1-kHz tone that is 100% modulated in amplitude at a modulation frequency of 70 Hz. Three parameters are important in determining roughness. For amplitude modulation, the important parameters are the degree of modulation and modulation frequency. For frequency modulation, it is the frequency modulation index and modulation frequency.

Figure 11.1 shows the roughness of a 1-kHz tone at a modulation frequency of 70 Hz, as a function of the degree of modulation. Values of the degree of modulation larger than 1 are not meaningful here. The data, indicated by the solid line, can be approximated by the dashed line. Because the ordinate and abscissa are given in logarithmic scales, the straight dashed line

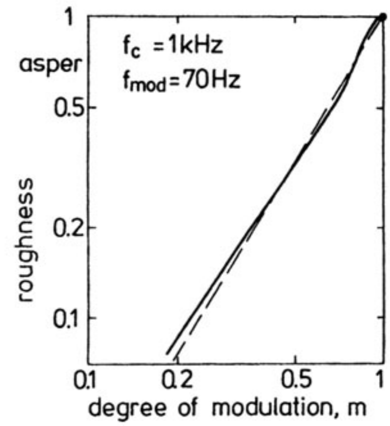

Fig. 11.1. Roughness as a function of the degree of modulation for a 1-kHz tone, amplitude-modulated at a frequency of 70 Hz. The dot in the right upper corner indicates the standard sound, which produces the roughness of 1 asper. The broken line indicates a useful linear approximation

represents a power law. The exponent is near 1.6 so that a roughness of only 0.1 asper is produced for a degree of modulation of 25%. This roughness is quite small and some subjects classify this as "no longer rough".

This dependence of roughness on the degree of modulation holds for tones of other centre frequencies also, although the modulation frequency at which the maximum roughness is reached depends on centre frequency. Figure 11.2 shows the dependence of roughness on modulation frequency at different centre frequencies for 100% modulation. This dependence has a band-pass characteristic. Roughness increases almost linearly from low modulation frequencies, in the double-logarithmic coordinates of Fig. 11.2, before it reaches a maximum. The maximum only depends on carrier frequency below 1 kHz where the maximum is shifted towards lower frequency of modulation with decreasing carrier frequency. The lower slope of this band-pass characteristic remains the same for frequencies of modulation below 1 kHz, even though the maximum decreases with decreasing centre frequency. For centre frequencies

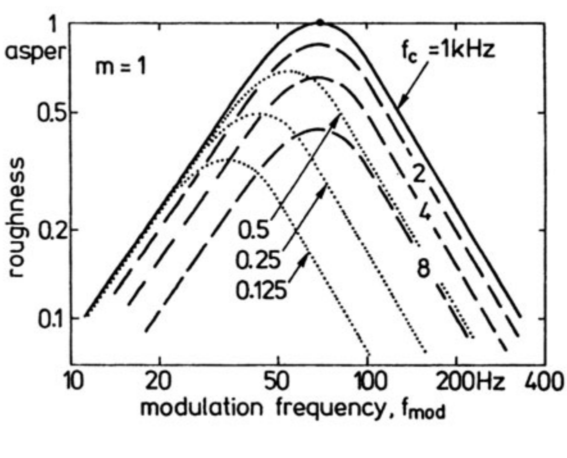

Fig. 11.2. Roughness of 100% amplitude-modulated tones of the given centre frequency as a function of the frequency of modulation

above 1 kHz, the height of the maximum is reduced although the frequency of modulation at which this maximum is reached remains unchanged. This means that above 1-kHz centre frequency there is a parallel shift downwards of the characteristic with increasing centre frequency.

The upper part of the band-pass characteristic can again be approximated by a straight line, with a relatively quick decay of roughness with increasing modulation frequency. It seems that the width of the critical band at lower centre frequencies plays an important role. At 250 Hz, the critical bandwidth is only 100 Hz. For a modulation frequency of 50 Hz, the two sidebands already have a separation of 100 Hz. For even higher modulation frequencies, the two sidebands fall into different critical bands. For centre frequencies above about 1 kHz, all the dependencies on frequency of modulation have the same shape. Here, the maximal roughness seems to be limited by the temporal resolution of our hearing system. Thus, two characteristics of the ear seem to influence the sensation of roughness: at low centre frequencies, it is frequency selectivity; at high centre frequencies, it is the limited temporal resolution.

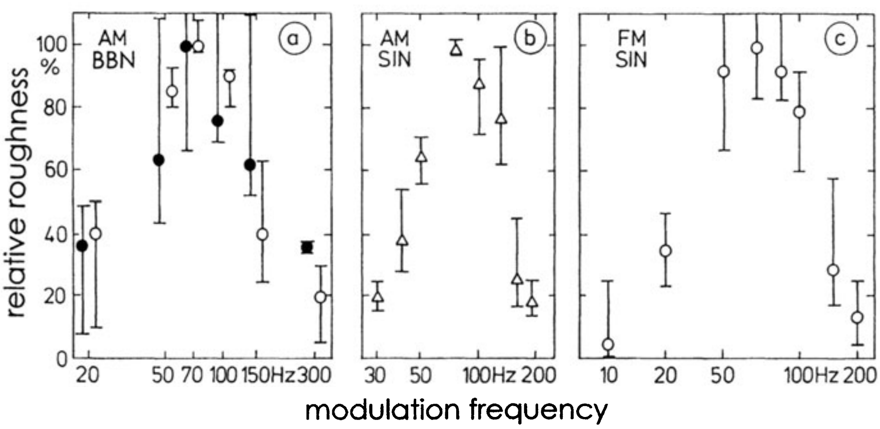

Fig. 11.3a–c. Roughness of three modulated sounds as a function of modulation frequency. (**a**) Amplitude modulated broadband noise of 60 dB SPL and 40 dB modulation depth; (**b**) amplitude modulated 1 kHz tone of 70 dB SPL and 40 dB modulation depth; (**c**) frequency modulated pure tone of 70 dB SPL, 1500 Hz center frequency, and ±700 Hz frequency deviation

Figure 11.3 shows the dependence of relative roughness on modulation frequency for amplitude modulated broadband noise, amplitude modulated pure tones, and frequency modulated pure tones. In line with the data presented in Fig. 11.2 the maximum of roughness occurs near a modulation frequency of 70 Hz irrespective of bandwidth or type of modulation. As with amplitude modulated pure tones also for amplitude modulated broadband noise or frequency modulated pure tones roughness vanishes for modulation frequencies above about 300 Hz.

Bands of noise often sound rough, although there is no additional amplitude modulation. This is because the envelope of the noise changes randomly. These changes become audible especially for bandwidths in the neighbourhood of 100 Hz, where the average rate of envelope change is 64 Hz (see Sect. 1.1). Therefore, roughness is particularly large at such bandwidths. For increasing bandwidth, the creation of roughness is limited by frequency selectivity. Nonetheless, at very high centre frequencies noises can be produced which are still within the critical band but have a bandwidth of 1 kHz. Such a noise, although it is randomly amplitude modulated, sounds relatively steady and produces only a very small sensation of roughness.

Figure 11.4 shows the dependence of roughness on sound pressure level. Data are given for amplitude-modulated broad-band noise, amplitude-modulated pure tone, and frequency modulated pure tone.

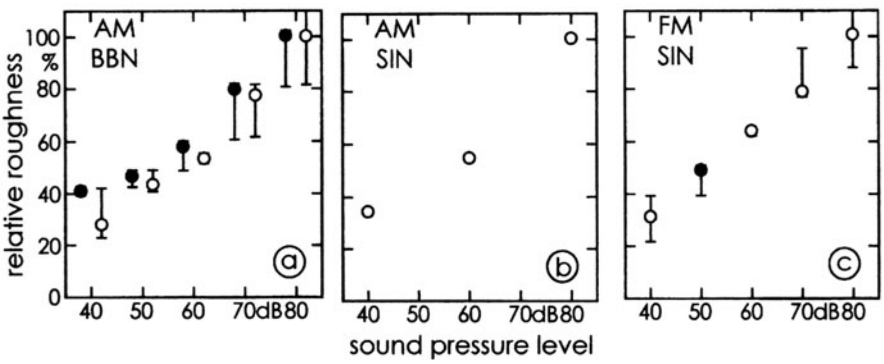

Fig. 11.4a–c. Roughness of modulated sounds as a function of sound pressure level. Modulation frequency is 70 Hz. (**a**) Amplitude-modulated broad-band noise with 40 dB modulation depth, (**b**) amplitude-modulated 1 kHz-tone with 40 dB modulation depth, (**c**) frequency-modulated pure tone at 1500 Hz with ±700 Hz frequency deviation

For an increase in sound pressure level by 40 dB roughness increases by a factor of about 3. This dependence of roughness on level is similar to the increase of fluctuation strength with level as displayed in Fig. 10.2.

An increment of roughness becomes audible for an increment in the degree of modulation of about 10%, which corresponds to an increment of about 17% in roughness. For amplitude-modulated 1-kHz tones and a modulation frequency of 70 Hz, a threshold of roughness is reached for values close to 0.07 asper. One asper is close to the maximum roughness for amplitude-modulated tones; there are thus only about 20 audible roughness steps throughout the total range of roughnesses.

Frequency modulation can produce much larger roughness than amplitude modulation. A strong frequency modulation over almost the whole frequency range of hearing produces a roughness close to 6 asper. Only amplitude modulation of broad-band noises is able to produce such a large roughness.

Figure 11.5 shows the dependence of roughness of an FM-tone on frequency deviation. An FM-tone centered at 1500 Hz with a level of 70 dB was modulated by a modulation frequency of 70 Hz.

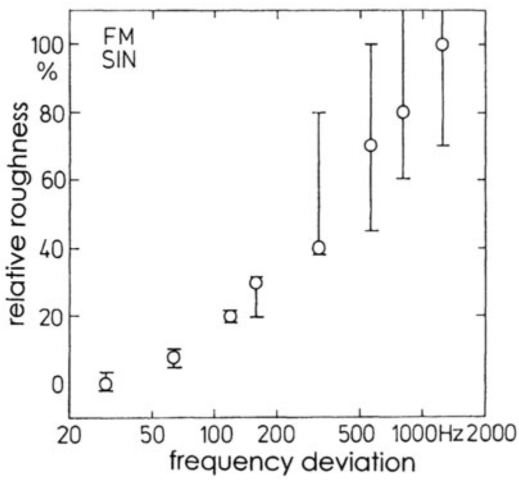

Fig. 11.5. Roughness of a sinusoidally frequency modulated pure tone as a function of frequency deviation. Center frequency 1500 Hz, sound pressure level 70 dB, modulation frequency 70 Hz

The results displayed in Fig. 11.5 indicate that for frequency deviations up to 50 Hz only negligible values of roughness show up. For larger values of the frequency deviation Δf roughness increases almost linearly with the logarithm of Δf.

11.2 Model of Roughness

As mentioned above, there are two main factors that influence roughness. These are frequency resolution and temporal resolution of our hearing system. Frequency resolution is modelled by the excitation pattern or by specific-loudness versus critical-band rate pattern.

It is assumed that our hearing system is not able to detect frequency as such, and is only able to process changes in excitation level or in specific loudness at all places along the critical-band rate scale; thus the model for roughness should be based on the differences in excitation level that are produced by the modulation. Starting with amplitude modulation, we can refer to data that describe the masking effect produced by strongly temporally varying maskers. Because masking level is an effective measure for determining excitation, it can be used to estimate the excitation-level differences produced by amplitude modulation. This procedure incorporates the two main effects already discussed, i.e. frequency and temporal resolution.

The temporal masking patterns outlined in Figs. 4.24, 4.25, and 4.27 show the temporal effects and indicate values of ΔL that can be used as a measure to estimate differences between the maximum and the minimum of the temporal masking pattern. This temporal masking depth, ΔL, becomes larger for lower modulation frequency. If roughnesses were determined only by this masking depth, then one would expect the largest roughness at the lowest modulation frequencies. This is not the case, indicating that roughness is a sensation produced by temporal changes. A very slow change does not produce roughness; a quick periodic change does, however. This means that roughness is proportional to the speed of change, i.e. it is proportional to the frequency of modulation. Together with the value of ΔL, this leads to the following approximation:

$$R \sim f_{\text{mod}} \Delta L \,. \tag{11.1}$$

For very small frequencies of modulation, roughness remains small although ΔL is large. In this case f_{mod} is small, so that the size of the product remains small. For medium frequencies of modulation around 70 Hz, the value of ΔL is smaller than at low modulation frequencies. However, f_{mod} is much larger in this case, so that the product of the two values reaches a maximum. At high frequencies of modulation, f_{mod} is a large value but ΔL becomes small because of the restricted temporal resolution of our hearing system. Thus the product diminishes again. In this context, it should be realized that a modulation frequency of 250 Hz corresponds to a period of 4 ms, and the effective duration of the valley is only about 2 ms. In this case, the temporal masking depth, ΔL, becomes almost zero. Consequently, the product of ΔL and f_{mod} becomes very small and the roughness disappears.

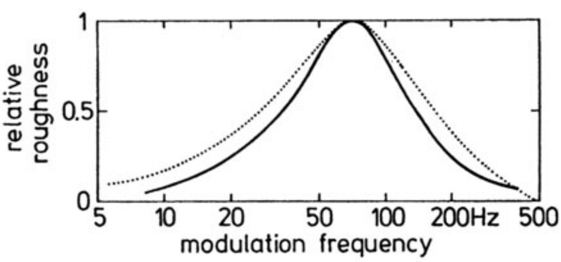

Fig. 11.6. Relative roughness as a function of the frequency of modulation as subjectively measured (*solid*) and calculated (*dotted*)

This way, roughness can be approximated and compared with the measured data. This is done in Fig. 11.6 by using the maximum roughness at 70 Hz as a reference. Roughness relative to the maximal value reached is plotted as a function of modulation frequency. The solid line corresponds to the data shown in Fig. 11.2 for centre frequencies above 1 kHz. The dotted line corresponds to the calculated values based on the assumption that roughness is proportional to the product of modulation frequency and masking depth. The calculated dependence agrees well with the subjectively measured one, although there are some differences especially at lower frequencies of modulation. There, subjectively measured roughness disappears more quickly than calculated roughness. In this region, subjects have difficulties differentiating between sensations of roughness and fluctuation strength, and concentrate mostly on one or the other.

For more precise calculations, it has to be realized that the value ΔL depends on the critical-band rate. The nonlinear rise of the upper masking slope outlined in Fig. 4.9 produces $\Delta L's$ in the region of the upper slope that are much larger than those corresponding to the main excitation. These effects can also be seen in Fig. 4.24. Comparing ΔL produced at a test-tone frequency identical to the frequency of the masker (1 kHz), with that produced on the upper slope near 1.6 kHz, it is evident that ΔL on the slope is much larger. To account for this effect, the given approximation is changed to:

$$R \sim f_{\mathrm{mod}} \int_0^{24\mathrm{Bark}} \Delta L_{\mathrm{E}}(z)\, dz \; . \tag{11.2}$$

Using the boundary condition that a 1-kHz tone at 60 dB and 100%, 70 Hz amplitude-modulated, produces the roughness of 1 asper, the roughness R of any sound can be calculated using the equation

$$R = 0.3 \frac{f_{\mathrm{mod}}}{\mathrm{kHz}} \int_0^{24\mathrm{Bark}} \frac{\Delta L_{\mathrm{E}}(z)\, dz}{\mathrm{dB/Bark}} \;\mathrm{asper}\; . \tag{11.3}$$

Unfortunately, we do not have data for ΔL as a function of critical-band rate that are as numerous as the data for excitation level or specific loudness. Therefore, the calculations are somewhat limited. Using the data available, however, we have been able to demonstrate that roughness can be calculated precisely as a function of the degree of modulation. In this case, the calculated value is influenced mainly by the dependence of ΔL on the degree of modulation, but not by the frequency of modulation. For modulated tones, the nonlinear rise of the upper slope of the masking versus critical-band rate pattern, creates larger contributions to roughness than those produced at the main excitation. This leads to the prediction – in agreement with psychoacoustical data – that the slope of the relationship between roughness and degree of modulation (both on logarithmic scales) is larger for sinusoidal tones (1.6) than for broad-band noises (1.3).

Some data concerning the value ΔL with frequency-modulated sounds are available. Approximations based on the equation given follow qualitatively and in many cases even quantitatively the psychoacoustically measured dependencies.

It is of advantage to transfer the ΔL values necessary for the calculation of roughness into the corresponding variations of specific loudness. As input to the model, the specific loudness-time function in each channel of a loudness-meter as illustrated in Fig. 8.26 is necessary. Moreover, the correlation between signals in neighboring channels has to be taken into account. On this basis, a computer program was developed which nicely accounts for the measured psychoacoustic data and can also describe quantitatively the roughness of noise emissions.

A variant of this program which essentially is based on the same features can also quantitatively assess the dependencies of fluctuation strength on relevant stimulus parameters (cf. Chap. 10).

12. Subjective Duration

When we talk about duration, we normally think of objective duration, i.e. physical duration measured in seconds, milliseconds or minutes. This is so, although we often check durations by listening to them in music, for example, or by giving a talk, where a short silence can add emphasis. If such durations can be measured by listening, they cannot be objective durations but must be subjective because they correspond to sensations. Subjective duration is not drastically different from objective duration if the durations of long-lasting sound bursts are compared. Therefore, it is often assumed that subjective duration and objective duration are almost equal. This is not so, however, when the duration of sound bursts is compared with the duration of sound pauses. In this case, drastic differences appear which indicate the need to consider subjective duration as a separate sensation.

12.1 Dependencies of Subjective Duration

The scale of subjective duration can be quantified by fixing a reference value and a unit. We have chosen as the unit the "dura" and have determined that a 1-kHz tone of 60 dB sound pressure level and 1 s physical duration, produces a subjective duration of 1 dura. By halving and doubling, we can produce the relationship between subjective duration and physical duration for 1-kHz tone bursts. The result is plotted in Fig. 12.1 in double logarithmic scales. The subjective duration, D, is the ordinate, and physical duration, T_i, is the abscissa. The reference point is marked by an open circle. Proportionality between the two values would be indicated by the broken 45^o line.

This proportionality is effective over a wide range, starting at large durations of 3 s down to a duration of about 100 ms. Below that physical duration, subjective duration deviates from this proportionality with the tendency that subjective duration decreases less than physical duration. However, these results may be influenced by the fact that reducing the duration of the 1-kHz tone produces a different spectrum. Therefore, white noise was used to produce shorter sound bursts without changing the spectrum. For large durations, white-noise bursts show the same proportionality between physical and subjective duration as found for 1-kHz tones. The physical duration of

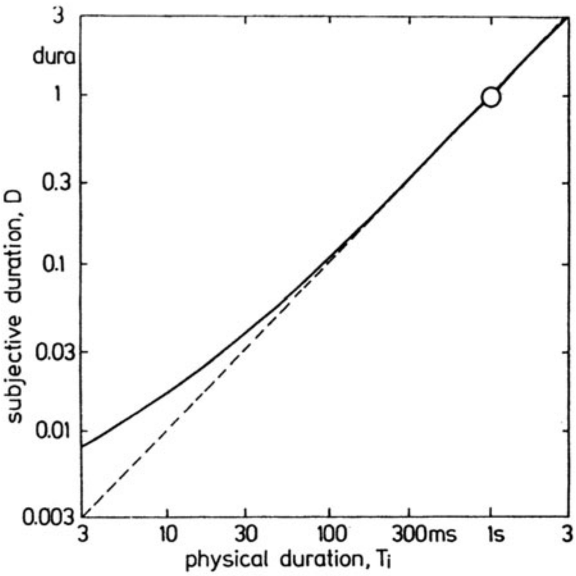

Fig. 12.1. Subjective duration as a function of the physical duration of 1-kHz tones at 60 dB SPL (*solid line*). The broken line indicates equality of physical and subjective duration. The open circle marks the standard sound producing a subjective duration of 1 dura

white noise can be reduced to 0.3 ms without much influence on the spectral shape. The results in this short duration area confirmed the tendency of 1-kHz tones to show little difference in subjective durations when physical duration is changed from 1 ms to 0.5 ms. From these measurements, it can be concluded that the effect shown in Fig. 12.1 is not a side effect based on the spectral broadening of the shorter 1-kHz tone bursts, but an effect that is based on the behaviour of human hearing. This means that one cannot expect subjective duration and objective duration to be equal for durations shorter than 100 ms.

This finding is somewhat astonishing; however, the results produced by comparing subjective durations produced by pauses with those produced by tone bursts are even more surprising. In this case, the method of comparison is used and pauses are changed to sound as long as tone bursts, and vice versa. The inset in Fig. 12.2 shows the temporal sequence of the test sounds to be compared. A sound burst of duration T_i is followed by a pause that lasts at least 1 s. After this, the tone is switched on again for 0.8 s, followed by a pause of duration T_p, switched on again for another 0.8 s and followed by a pause lasting at least 1 s. The subject compares the perceived duration of the tone burst with the perceived duration of the pause. In one experiment, the physical duration of the tone burst is changed, in the next the physical duration of the pause is changed by the subject to produce equal subjective durations. The use of the method of adjustment produces the same results as the use of the method of constant stimuli. The results plotted in Fig. 12.2 show curves of equal subjective duration plotted in the plane with physical duration of the pause as the ordinate and physical duration of the burst as

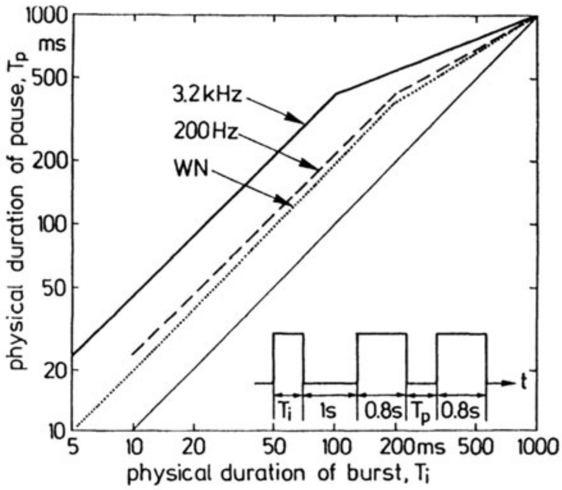

Fig. 12.2. Comparison of subjective durations produced by pauses and those produced by sound bursts. The physical duration of pauses (*ordinate*) producing the same subjective durations as a sound burst of given physical duration (*abscissa*) are shown for white noise (WN), 200-Hz and 3.2-kHz tones. The inset indicates the sequence of the sounds

the abscissa, both on logarithmic scales. The parameter is the type of sound used.

Tones of 3.2 kHz produce the largest effects while low-frequency tones of 200 Hz, or white noise, produce less pronounced effects. Both show, however, strong deviations from the 45^o line that indicates equality of the durations of the burst and the pause. The expected result, that subjective duration of a pause and subjective duration of a burst are equal, is true for physical durations larger than 1 s but is certainly not true for smaller physical durations. Using 3.2-kHz tones as the sound from which the bursts and the pauses are extracted, it can be seen from Fig. 12.2 that a burst duration of 100 ms produces the same subjective duration as a pause which actually lasts as long as 400 ms. The difference in this case is a factor of four! For 200 Hz or white noise, the effect is smaller but still about a factor of two. The difference between physical burst duration and physical pause duration necessary to produce equal subjective durations, is true for burst durations below 100 ms down to durations as short as 5 ms.

The effect, which must play a strong role in music and speech perception, is not too dependent on loudness level if loudness levels larger than 30 phon are used.

12.2 Model of Subjective Duration

There are few psychoacoustical data available with regard to subjective duration. Thus, it is not possible to give an extensive model. The results so far, however, allow us to start with a model that is relatively simple, and makes use of the effect of premasking and postmasking. In most cases, the effect of