A Nonlinear Psychoacoustic Model Applied to the ISO MPEG Layer 3 Coder

Frank Baumgarte, Charalampos Ferekidis, Hendrik Fuchs

Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung
Universität Hannover, Germany

Abstract

A psychoacoustic model which approximates the masked threshold evoked by complex sounds is presented. It features nonlinear superposition of masking components in order to generate masked thresholds which closely match known psychoacoustic data. First results obtained with the psychoacoustic model for controlling the quantizers of the ISO MPEG Layer 3 coder are discussed.

1 Introduction

Significant improvements of high quality audio bit rate reduction have been achieved by considering the properties of the human auditory perception. This is generally realized by the introduction of a psychoacoustic model which generates the masked threshold evoked by a sound signal and which controls the quantizers of a coding system. The masked threshold for quantization errors is defined as the maximum level of quantization noise which is just non audible in the presence of a masking sound. Therefore, the quantization noise will only become audible if the level exceeds the masked threshold.

Bit rate reduction is achieved by exploitation of statistical redundancy and perceptual irrelevance defined by the masked threshold. The reduction of irrelevance apart from redundancy is obtained by adapting the spectral and temporal shape of the quantization noise to the fluctuations of the masked threshold.

The generation of the masked threshold by psychoacoustic models used so far in coding systems is carried out by two steps. In a first step the masking sound spectrum is decomposed into simple masker components which are superposed in a second step to result in the overall masked threshold. The superposition of threshold components used in models proposed by the ISO MPEG standard [1] and others ([2],[3]) is based on linear addition. From psychoacoustic measurements ([4],[5]) it is known that linear addition of masker components often results in a much lower overall threshold than determined experimentally. Thus a nonlinear superposition was proposed by Lutfi [7] which closely matches the measured threshold. It is expected that the incorporation of a generalized nonlinear superposition into a psychoacoustic model offers an improved approximation of the masked threshold evoked by complex sounds and an improved reduction of irrelevance.

The developed nonlinear model is described in chapter 2 emphasizing the properties of the nonlinear superposition. A comparison of the masked thresholds resulting from a linear model of the ISO MPEG Layer 3 coder and the nonlinear model applied to this coding system is presented in chapter 3.

2 Nonlinear Psychoacoustic Model

Psychoacoustic models are based on psychoacoustic measurements of the masked threshold. Measurements are carried out for well defined combinations of maskers and test signals to adjust the perceptual threshold for the test signal in the presence of the masker during a subjective listening test. Due to this test conditions the masked threshold can only be determined for simple combinations of maskers and test signals, for example a narrow band noise masker and a test tone. In contrast the determination of the masked threshold of arbitrary complex sounds by psychoacoustic measurements is impracticable. So the results from psychoacoustics are only applicable if the complex sound is represented by a combination of simpler maskers with a known threshold. The overall masked threshold can than be approximated by a superposition of the particular masked thresholds of the masker components. Given an analysis algorithm which successful divides a complex sound into masker components, the properties of the superposition of the masked thresholds are to be determined. In a first approach to this problem a linear behavior of perception was assumed, yielding linear addition of threshold component intensities [4]. Several psychoacoustic models ([1],[2],[3]) and sound quality measurement systems [6] are based on linear superposition of masked threshold components. Further results from psychoacoustics concerning the additivity of masking proved that a linear model fails in most cases of spectral overlapping threshold components ([4],[5],[7]). Thus a nonlinear model was introduced to account for the significant higher thresholds resulting from the experiments compared to the results of a linear model [8]. Such a nonlinear model of additivity is successfully used with a sound quality measurement system [9]. The psychoacoustic model presented here incorporates this nonlinear superposition as main part. An earlier version of the model is described in [10].

Differences of the masked thresholds resulting from a linear and a nonlinear superposition are discussed later for some special masker configurations. The results indicate considerable deviations of the approximated thresholds showing that significant improvements are possible from a nonlinear model.

The suggested nonlinear psychoacoustic model is described in the following paragraphs according to the functional block diagram in figure 1. Considering the model as a system approximating the masked threshold of complex sounds it is independent of any underlying coding scheme. The only assumption concerning the intended application consist in noiselike disturbances resulting from quantization noise. Binaural masking effects are not permitted by the model so that in case of stereo signals it is independently applied to both channels.

2.1 Spectral Analysis

As a first step in determining the masked threshold for noise masked by a sound a spectral representation of the signal similar to the sound analysis in the inner ear must be obtained. This representation is approximated by a short time FFT using a 1024 point Hann window. The FFT is calculated in time intervals of 12 ms at 48 kHz sampling frequency. The uniformly distributed frequency samples of the FFT are mapped to the critical band scale [11]. This scale (unit Bark) corresponds to a perceptual pitch scale and offers the advantage of an approximately invariant masking behavior in contrast to the frequency scale. The mapping is carried out by averaging the squared frequency samples $X(l)$ located in each critical band interval z_k which results in sound intensities on a critical band scale [12]

$$I_M^*(z_k) = \frac{1}{b_{k+1} - b_k} \sum_{l=b_k}^{b_{k+1}-1} |X(l)|^2. \tag{1}$$

In equ. (1) the boundaries b_k indicate the lower index of the frequency samples located in the critical band interval k which has the width Δz.

$$b_k = \frac{f(z_k - \frac{1}{2}\Delta z)}{\Delta f} \tag{2}$$

The function $f(z)$ denotes the critical band to frequency mapping. This nonlinear relation of frequency and critical band rate is shown in figure 2. The frequency resolution Δf is determined by the FFT length and the sampling rate. At a sampling rate of 48 kHz and a 1024 point FFT it amounts to $\Delta f = 47$ Hz. The resolution Δz is determined by psychoacoustic considerations and will be discussed later.

The frequency mapping shows a dependency of the obtained intensity level I_M^* from the signal bandwidth in each critical band interval. Assuming for example a single nonzero frequency sample $X(l)$, the level of the sample is attenuated according to the critical band width referred to the frequency scale. A constant critical band width of Δz corresponds to a nonlinear growing bandwidth on the frequency scale. The attenuation of a single nonzero frequency sample is determined by the factor $1/(b_{k+1} - b_k)$ of equ. (1) where b_k is the lower boundary of the critical band interval which contains the frequency sample. This is in contrast for a white spectrum $X(l)$ because there is no critical band rate dependent attenuation. The negative attenuation referred to as gain $g_z(z_k)$ is shown in figure 3 for both cases. The gain is given by the ratio of intensities in the critical band domain and the frequency domain

$$g_z(z_k) = 10 \log \frac{I_M^*(z_k)}{|X(b_k)|^2}. \tag{3}$$

For this figure a higher resolution Δf and Δz is used and it is assumed that Δz is equal to Δf at the lowest critical band rate. The lower line of figure 3 is obtained by assuming exactly one nonzero frequency sample in each critical band interval. From this consideration it can be stated that sound signals with narrow band spectra which are smaller than the their corresponding critical band interval are attenuated up to 15 dB at the upper critical band limit while there is no attenuation appearing on the lower critical band limit. This property of the frequency to critical band mapping models the critical band width summation of sound intensities by the auditory system. In general it is desirable to use a finer resolution than the critical band width $\Delta z = 1$. In figure 3 the shape of the two lines will remain the same by changing the resolution Δz but there will be a vertical shift of the lower line according to the ratio of Δf and Δz at a critical band rate of $z = 0$.

2.2 Prefiltering

The sound intensities $I_M^*(z)$ obtained by the frequency mapping are interpreted as individual maskers with the corresponding level $L_M^*(z_k)$ [dB]. Previous to the determination of masked thresholds the individual maskers are weighted according to their loudness. This is performed by a prefilter which approximates an equal loudness function [13]. The different weighting of masker components is applied before the superposition of the threshold components in order to consider the critical band rate dependent masker effectivity. After the superposition the inverse filter is applied in order to remove the prefilter characteristic from the resulting overall threshold. The prefilter in conjunction with the inverse filter only influences the relative weighting of masker components to each other. This concept considers different masking properties with respect to the loudness of maskers. For example two maskers of equal level and different critical band rate will only produce the same amount of masking if the maskers provide equal loudness. In case of different perceived loudness the masked threshold of the louder masker has to be amplified relative to the other masker. The effect of the filtering will be discussed in more detail in conjunction with the threshold generation.

2.3 Determination of Masked Threshold Components

The masked thresholds known from psychoacoustics [14] are applied to the individual maskers. Because of the underlying spectral analysis in the critical band domain the individual masked thresholds $L_{T,i}$ can easily be determined using a spreading function. As seen in figure 4 this function is described by three parameters. The attenuation a_v corresponds to the difference of masker level and the maximum of the spreading function. The slopes s_l and s_u correspond to the lower and upper slope respectively in units of dB/Bark. While positive values of s_l indicate rising characteristics of the lower slope positive values of s_u indicate falling characteristics of the upper slope. The mathematical representation of the spreading function belonging to a masker component $L_M(z_i)$ at the critical band rate z_i is given by equ. (4).

$$L_{T,i}(z_k) = \begin{cases} L_M(z_i) - a_V - s_l \cdot (z_i - z_k)\ ; & z_k < z_i \\ L_M(z_i) - a_V - s_u \cdot (z_k - z_i)\ ; & z_k \geq z_i \end{cases} \tag{4}$$

Except s_u the parameters are constant for different masker levels and critical band rates. The upper slope is adapted level dependent according to equ. (5).

$$s_u = \left(22\mathrm{dB} - 0.2\ L_M\right)/\mathrm{Bark} \tag{5}$$

For the model calculations a discrimination of the critical band rate is performed. With the assumption of a constant resolution of Δz the discrete Bark values are determined by the index k with the relation $z_k = k\Delta z$.

2.4 Nonlinear Superposition

The calculation of the overall masked threshold from the individual masker components is performed using a "power–law model" proposed by Lutfi [7]. This model of masking additivity was verified for measurements of several authors [8]. The temporal and spectral boundaries for the application of the model are discussed in ([15],[16],[17],[18]). Contrary the linear superposition proposed by the ISO MPEG standard the nonlinear model uses a compressive exponential characteristic prior to the addition of masker components. The expansion of the sum is performed afterwards according to equ. (6).

$$I_T(z_i) = \left[\sum_k \left[I_{T,k}(z_i)\right]^{\alpha}\right]^{1/\alpha} \tag{6}$$

It should be noted that the nonlinear addition is applied to sound intensities which are calculated from levels by

$$I_{T,k}(z_i) = 10^{L_{T,k}(z_i)/10}. \tag{7}$$

Figure 5 shows the result of the nonlinear addition for two masker components $L_M(z_1)$ and $L_M(z_2)$ evoking the masked thresholds $L_{T,1}$ and $L_{T,2}$. The nonlinear addition of the intensities results in the overall threshold L_T. The inscribed term ΔL_T referred to as "additional masking" is defined as the minimum difference of the overall threshold and the threshold components:

$$\Delta L_T(z_i) = L_T(z_i) - \max_k\left(L_{T,k}(z_i)\right) \tag{8}$$

Additional masking is introduced because it is suitable for the description of the masking differences occurring in case of complex maskers compared to single maskers. According to [8] a parameter of $\alpha = 0.3$ permits additional masking in agreement with psychoacoustic data. This setting yields a maximum additional masking of 10 dB in the presence of two maskers. In case of $\alpha = 1.0$ the model degenerates to a linear model which corresponds to a linear addition of intensities. The linear addition results in a maximum additional masking of only 3 dB.

Increasing the number of masker components so that their critical band distance is smaller than 1 Bark leads to even more elevated thresholds because of the higher number of components which add up. Assuming white noise as the sound signal and a critical band resolution of 1/4 Bark the additional masking amounts to an average of 30 dB as shown in figure 6. In contrast the linear addition remains nearly unchanged at a 3 dB additional masking. Compared to psychoacoustic measurements the elevated masking for wide band noise has its counterpart in the different masking properties of noise and tone. Differences of threshold for noiselike and tonal maskers in the order of 20 dB were reported by [19] which are in agreement with results obtained by the nonlinear model. But the model fails in discriminating between tonal and narrow band noise maskers because of the limited frequency resolution. In this case the model always assumes a tonal masker, ensuring that the determined masked threshold does not exceed the true threshold for both maskers.

The different results for tonal and noiselike maskers are overlaid by the different gains g_z of the frequency to critical band mapping for these signal types. As shown in figure 3 the gain of a single nonzero frequency sample at high critical band rates amounts up to –15 dB. This results in an increased masking difference for noiselike and tonal signals in the higher critical band range.

Considering the behavior of the nonlinear superposition the exponent α and the resolution Δz of the model are of great importance. Because the parameters cannot be specified independently the following strategy seems reasonable. First the exponent is adjusted according to psychoacoustic data concerning additional masking. Second the critical band resolution is adjusted to match the 20 dB increment of threshold for noise maskers compared to tonal maskers at low critical band rates. Both conditions are fulfilled with the chosen parameters $\alpha = 0.3$ and $\Delta z = 0.25$. In figure 7 the masking increment resulting from the critical band resolution for a wide band noise masker compared to a tonal masker is shown. For a doubling of resolution it approximately amounts to 6 dB in case of $\Delta z < 1$.

2.5 Inverse Filtering

The inverse filter exhibits the inverted frequency response of the prefilter. As aforementioned the purpose of the filtering is a relative weighting of maskers relative to each other. So the prefilter characteristic must be compensated to avoid an overall threshold shift resulting from the prefilter. The remaining effect is shown in figure 8. The masked thresholds for single tones of equal level obtained from the model obviously show varying slopes according to the response of the filter. Flatter slopes reflect a greater influence of the belonging masker on neighboring maskers. At the boundaries of the perceptible frequency range the flat slopes indicate the considerable influence of the threshold in quiet on the shape of the masked thresholds. The threshold in quiet is not yet considered by the model. In audio coding applications the sound level of the reproduction cannot be controlled so that the ratio of the sound level and the threshold in quiet cannot be precisely determined.

An additional effect is obtained in conjunction with the nonlinear superposition. The nonlinearity additionally amplifies the masked threshold in the range of a falling prefilter characteristic because of the asymmetry of the underlying spreading functions. In the range of rising characteristics the converse is true. In case of white noise the amplification originating from the prefilter amounts to 5 dB above the average 30 dB additional masking, as shown in figure 6. A rising threshold is also observed in psychoacoustic measurements using white noise maskers [20]. Because sinusoidal test tones were used for these masking experiments in contrast to the noiselike test signals assumed here, the masking increment is considerable higher and reaches a maximum of 15 dB increment at the upper critical band boundary.

3 Results of the Application to a Layer 3 Coder

The audio part of the established ISO MPEG standard [1] offers a framework of three layers, each containing a coding scheme for different tradeoffs between complexity and achieved quality at a given bit rate. The Layer 3 coder currently reaches the best ratios of quality over bit rate in applications requiring high sound quality. At a bit rate of 2×128 kbit/s the quality is comparable to CD.

The Layer 3 coder applies a psychoacoustic model to approximately adjust the introduced quantization noise according to the masked threshold. A uniform hybrid filterbank for the decomposition into spectral components is used offering a spectral resolution of 576 bands. An improved temporal resolution can be obtained by switching to shorter filters with a reduced spectral resolution of 192 bands. A nonuniform division of the sound spectrum according to perceptual properties is provided by the concept of scalefactor bands. The spectral components located in a scalefactor band are grouped and quantized together using a common scalefactor. In each scalefactor band noise shaping according to the sound spectrum is provided due to the non equal step size of the quantizers used in each scalefactor band.

The scalefactor bands allow individual adjustment of the introduced quantization noise according to a resolution of approximately critical bandwidth (1 Bark). Because of the finer resolution of the masked threshold, the maximum allowed noise level of a subband is determined by the minimum threshold value in that band. In figure 9 the scalefactor band noise levels resulting from the masked threshold of five maskers and nonlinear superposition are given. For comparison the threshold generated by the nonlinear model using $\alpha = 1$ which yields linear addition of intensities is also shown. Compared to linear superposition the allowable noise levels for nonlinear superposition are considerable higher especially near the minima of the masked threshold curve. For this graph the influence of a possible noise shaping has been ignored.

A first implementation of the nonlinear psychoacoustic model in a Layer 3 coder permits a masked threshold generation in case of the standard temporal resolution. If the coder switches to short filters gaining a better temporal resolution a constant signal–to–mask ratio is assumed.

Figure 10 shows typical proportions of the approximated masked threshold in conjunction with the short time spectrum of one block of a clarinet recording. The threshold obtained from the nonlinear model obviously shows a smoothing effect compared to that from the ISO–model. The consequence is a higher allowable average noise level resulting from the raised minima of the threshold. Another difference between the generated thresholds is the deviation which increases towards the lower and upper frequency bounds. This deviation is occurring systematically for all sequences tested. For low frequencies it emerges from the binaural masking level difference (BMLD) considered by the ISO–model implementation which is realized as a minimum signal–to–mask ratio ranging up to 24 dB for the lower frequency boundary. The nonlinear model considers no BMLD since this perceptual property can only be demonstrated for special binaural signal configurations which are not likely to occur in natural sounds. At low frequencies the ISO–model generally does not exploit masking in full extend which results in a higher bit need for the coding of the lower subbands.

The lifted threshold of the ISO–model compared to the nonlinear model in the high frequency range follows from the assumption that maskers in this range are noiselike. Consequently the ISO–model fails in case of high frequency tonal maskers determining a considerable higher masked threshold than expected. Quantization noise in this frequency range may be audible especially for tonal maskers. Therefore, the reduced bit need for the coding of the higher subbands can lead to a quality degradation.

Regarding the Layer 3 coder the achieved sound quality is not only determined by the approximated masked threshold but also by the bit allocation algorithm that controls the quantization noise level. For critical test signals the target bit rate is generally not sufficient for keeping the quantization noise below the masked threshold. In this situation the masked threshold must be approximated by the quantization noise as good as possible. If the resulting noise level still exceeds the threshold by a certain amount, a reduction in coder bandwidth gaining higher noise–to–mask ratios may be subjectively less annoying. These considerations show that the bit allocation algorithm plays an important part if the quantization noise exceeds the masked threshold because of an insufficient bit rate.

4 Conclusions

The developed nonlinear psychoacoustic model for the approximation of the masked threshold for arbitrary sounds features several important properties also found in psychoacoustic masking experiments. The kernel of the model consisting of a nonlinear superposition of masker components leads to a more realistic threshold compared to earlier approaches using a linear superposition especially for complex maskers. The nonlinear superposition adapted from [7] yields considerable higher thresholds in case of overlapping masked threshold components which correspond to psychoacoustic measurements. For instance, two overlapping threshold components result in an up to 7 dB higher overall masked threshold using the nonlinear superposition compared to a linear superposition.

In addition the different masking properties of tonal and noiselike sounds are taken into account by the nonlinear psychoacoustic model. The different masking properties are obtained from three basic elements of the model. The nonlinear superposition results in a lower threshold for tonal sounds with masker components of different amplitudes than for noiselike sounds with masker components of almost constant amplitude. The masked threshold for a noiselike sound can be 30 dB above that of a tonal sound due to the nonlinear superposition. The introduction of critical band rate instead of frequency contributes a damping of up to 15 dB for high critical band rates of a tonal sound. The filtering contributes an even smaller amount by damping the thresholds of noiselike maskers at low critical band rates and amplifying the thresholds at high critical band rates. The results from the nonlinear model are in agreement with the measured masking properties of noiselike and tonal sounds so that a tonality estimation as demanded by the psychoacoustic model proposed by the ISO MPEG standard is not needed.

Compared to the psychoacoustic model of ISO MPEG the nonlinear model presented here shows an improved masked threshold approximation in accordance with psychoacoustic measurements. The application of the nonlinear model to an ISO MPEG Layer 3 coder offers the possibility of an optimized quantization noise allocation with respect to the masking properties. Thus the Layer 3 coder is expected to yield an improved subjective quality if also the bit allocation algorithm is optimized according to the demands of the nonlinear psychoacoustic model.

References

[1] ISO/IEC. "*Information technology – Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s –*". Part 3: Audio. ISO/IEC 11172–3 International Standard, 1993.

[2] C. Colomes et al. "*A Perceptual Model Applied to Audio Bit–Rate Reduction*". J. Audio Eng. Soc., Vol. 43, No. 4, April 1995.

[3] J. D. Johnston. "*Estimation of Perceptual Entropy Using Noise Masking Criteria*". ICASSP 1988, pp. 2524–2527.

[4] D. M. Green. "*Additivity of Masking*". J. Acoust. Soc. Am., 41(6), Jan. 1967.

[5] E. Zwicker, S. Herla. "*Über die Addition von Verdeckungseffekten*". Acustica Vol. 34, pp. 89–97, 1975.

[6] T. Sporer et al. "*Evaluating a Measurement System*". J. Audio Eng. Soc., Vol. 43, No. 5, May 1995.

[7] R. A. Lutfi. "*Additivity of simultaneous masking*". J. Acoust. Soc. Am. 73, pp. 262–267, 1983.

[8] R. A. Lutfi. "*A Power–Law Transformation Predicting Masking by Sounds with Complex Spectra*". J. Acoust. Soc. Am. 77 (6), June 1985.

[9] J. G. Beerends, J. A. Stemerdink. "*A Perceptual Audio Quality Measure Based on a Psychoacoustic Sound Representation*". J. Audio Eng. Soc., Vol. 40, No.12, Dec. 1992.

[10] C. Ferekidis: "*Entwicklung eines Modells der Verdeckungswirkung des menschlichen Gehörs zur Irrelevanzreduktion von Audiosignalen*" (German). "Studienarbeit". Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung. April 1993.

[11] H. Fletcher. "*Auditory Patterns*". Reviews of Modern Physics, Vol. 12, pp. 47–65, Jan. 1940.

[12] E. Zwicker, E. Terhardt. "*Analytical Expressions for Critical–Band Rate and Critical Bandwith as a Function of Frequency*". J. Acoust. Soc. Am., 68(5), Nov. 1980.

[13] E. Zwicker, R. Feldtkeller. "*Das Ohr als Nachrichtenempfänger*" (German). Hirzel Verlag Stuttgart, Germany 1967.

[14] E. Terhardt. "*Calculating Virtual Pitch*". Hearing Research, Vol. 1, pp. 262 – 267, 1979.

[15] L. E. Humes, W. Jesteadt. "*Models of the additivity of masking*". J. Acoust. Soc. Am., Vol. 85(3), pp. 1285–1294, March 1989.

[16] L. E. Humes, L. W. Lee. "*Two experiments on the spectral boundary conditions for nonlinear additivity of simultaneous masking*". J. Acoust. Soc. Am., Vol. 92(5), pp. 2598–2606, Nov. 1992.

[17] C. G. Cokely, L. E. Humes. "*Two experiments on the temporal boundaries for the non-linear additivity of masking*". J. Acoust. Soc. Am., Vol. 94(5), pp. 2553–2559, Nov. 1993.

[18] B. C. J. Moore. "*Additivity of simultaneous masking, revisited*". J. Acoust. Soc. Am. 78(2), pp. 488–494, Aug. 1985.

[19] R. P. Hellman. "*Asymmetry of Masking between Noise and Tone*". Perception & Psychophysics Vol. 11 (3), pp. 241–246, 1972.

[20] E. Zwicker, H. Fastl. "*Psychoacoustics. Facts and Models*". Springer–Verlag, Berlin, 1990.

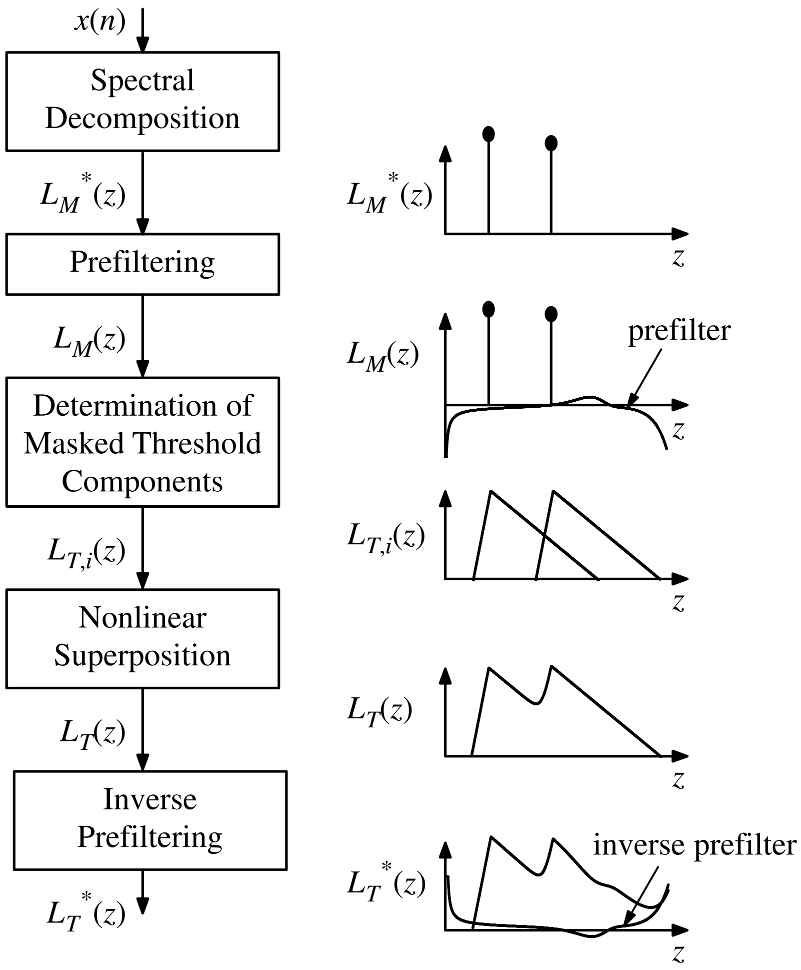

Figure 1 Overview of the nonlinear psychoacoustic model. Sound signal samples are input and overall masked threshold is the output of the model. Block diagram (left side) and associated signal levels over critical band rate (right side) are shown for an example with two masker components.

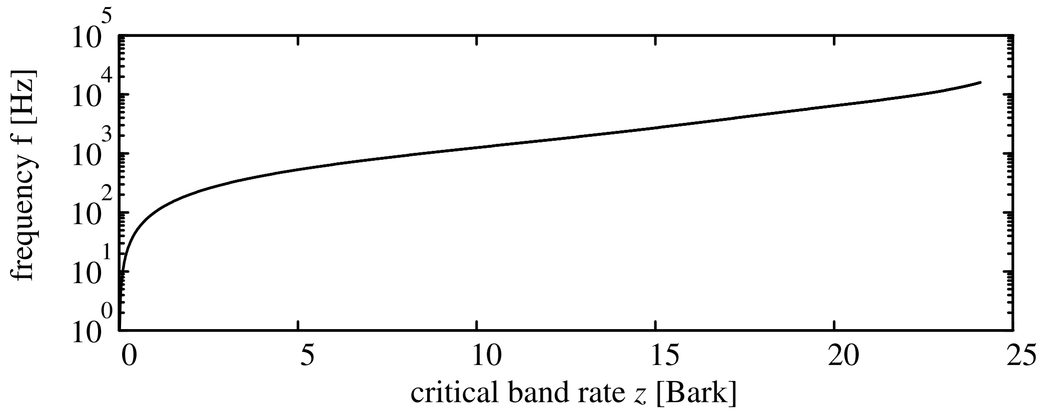

Figure 2 Relation of frequency and critical band rate

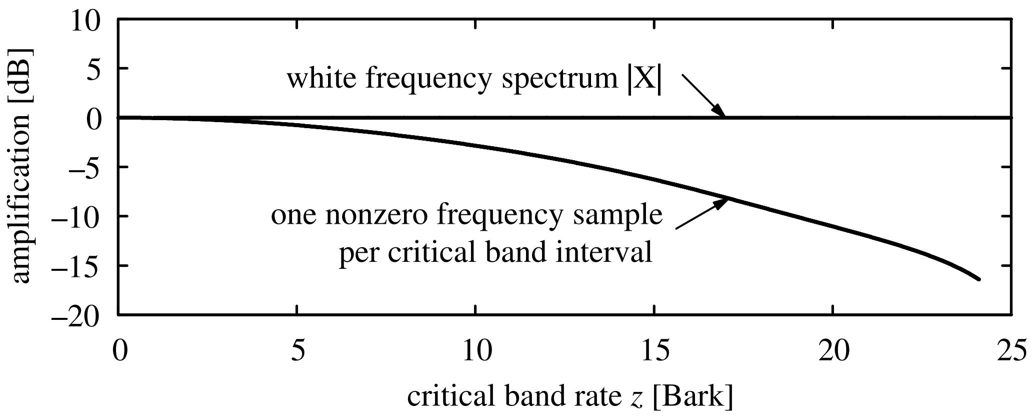

Figure 3 Amplification $g_z(z)$ resulting from the mapping of frequency to critical band rate. Assuming equal resolutions Δf=Δz at the lower critical band rate boundary.

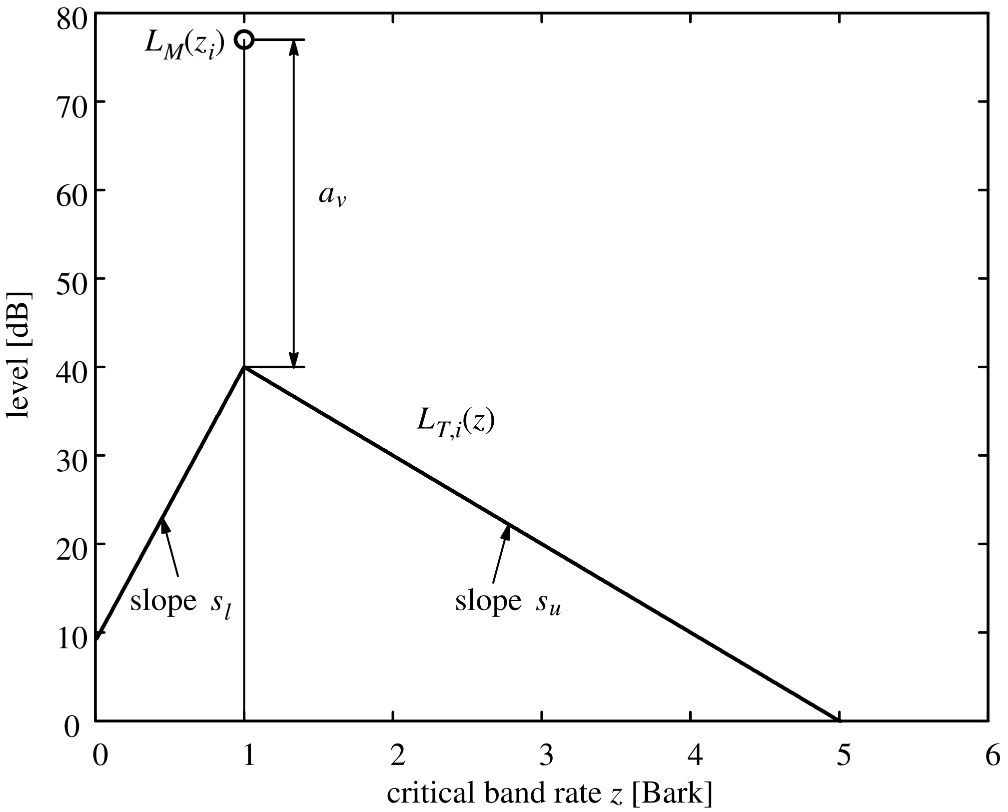

Figure 4 Spreading function $L_{T,i}$ of one masker component L_M at the critical band rate z_i. The lower and upper slopes of the spreading function are indicated as s_l and s_u. The attenuation of the maximum relative to the masker level is denoted with a_v.

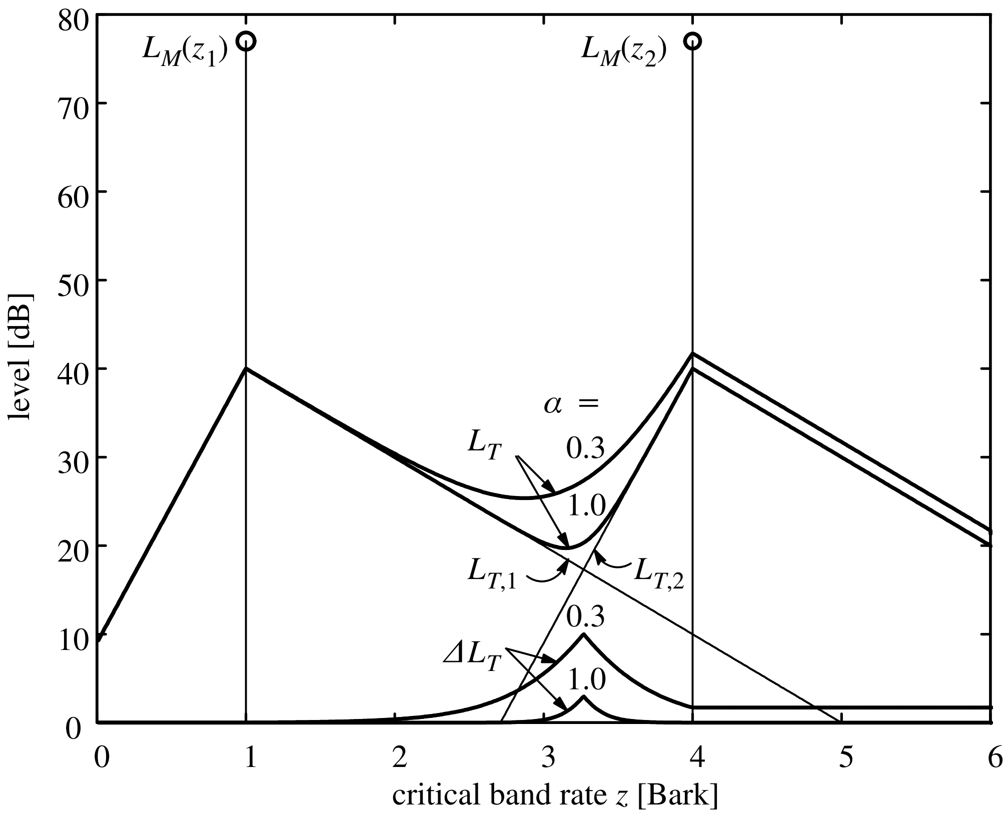

Figure 5 Superposition of two masked threshold components $L_{T,1}$ and $L_{T,2}$. The resulting overall threshold L_T is shown for different parameters α. The additional masking ΔL_T is also shown.

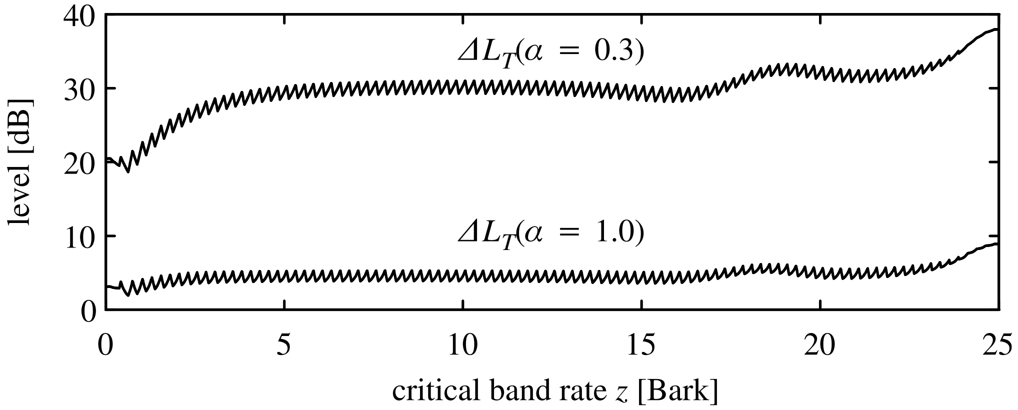

Figure 6 Additional masking of a white noise masker at a resolution of Δz = 1/4 Bark obtained from the nonlinear model for different parameters α.

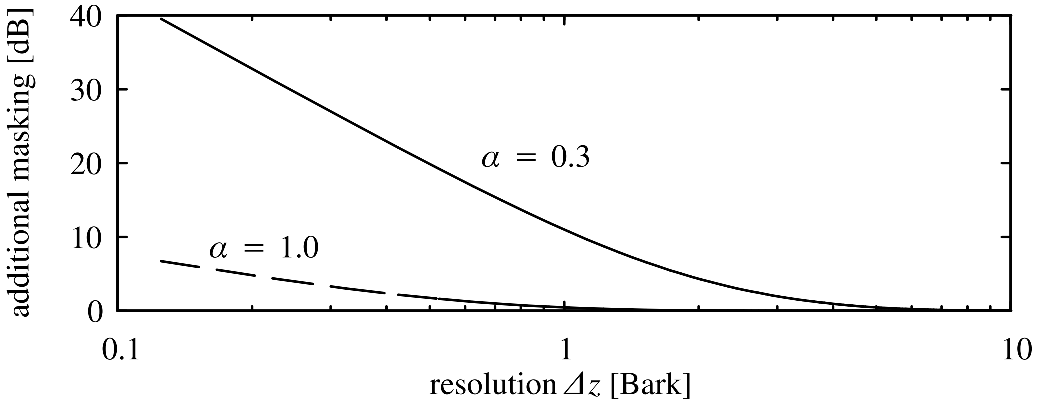

Figure 7 Additional masking ΔL_T over critical band resolution Δz for wide band maskers compared to one tonal masker. The parameter α determines the exponent used for the superposition.

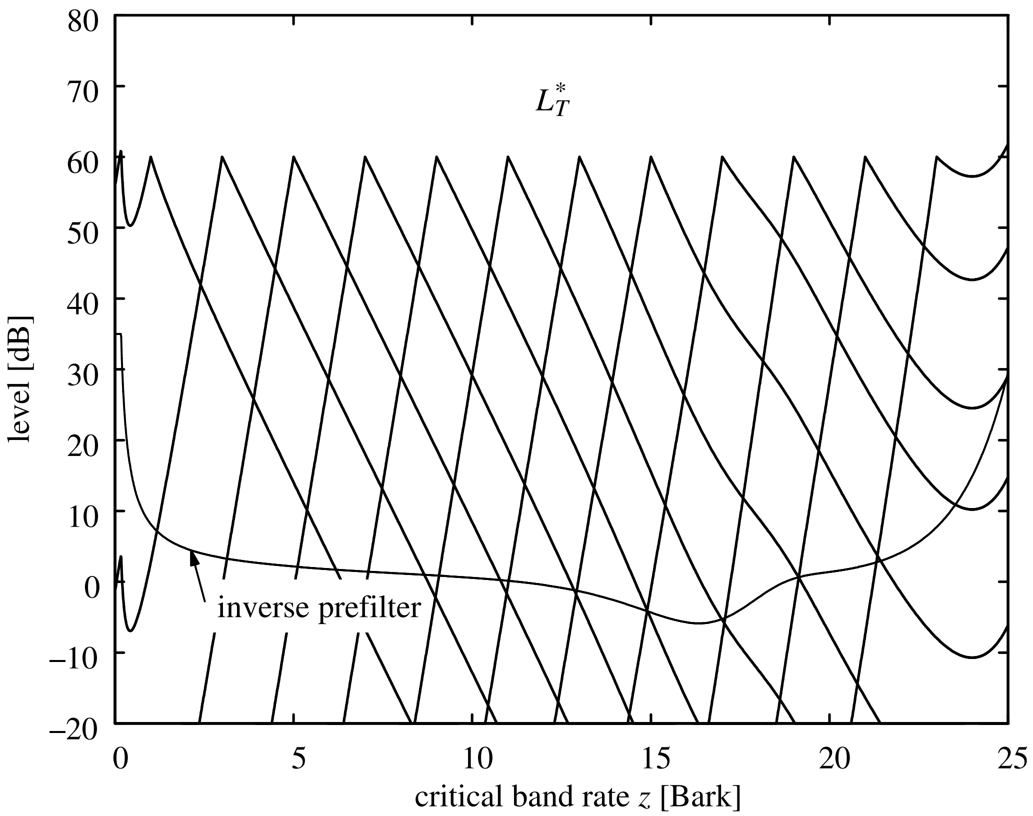

Figure 8 Masked thresholds for single tones at different critical band rates adjusted at equal maximum level. The inverse prefilter characteristic is shown for comparison.

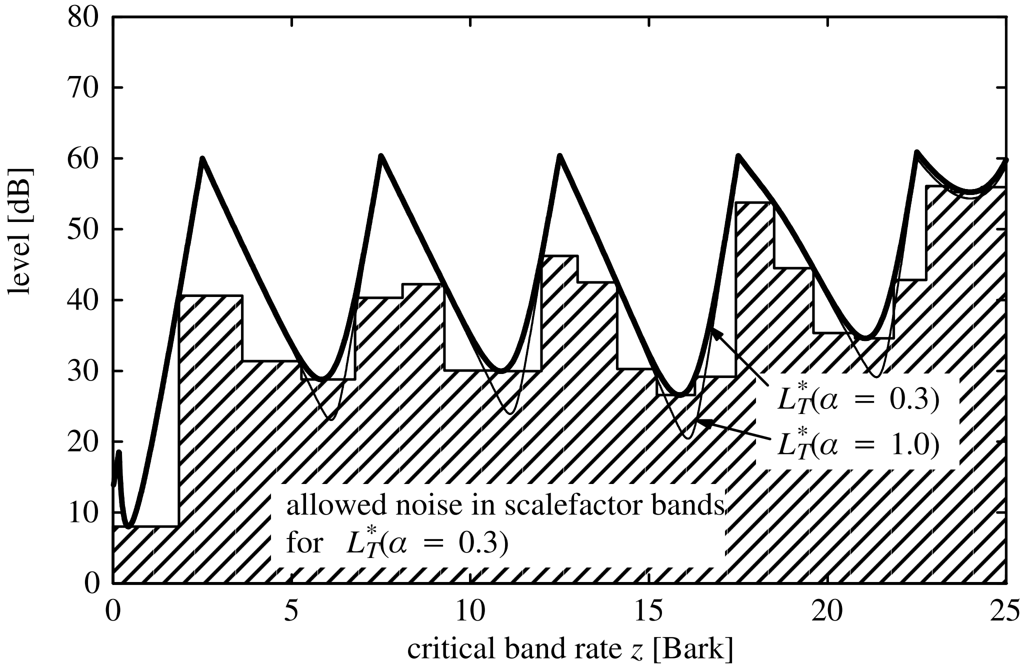

Figure 9 Overall masked threshold resulting from the superposition of five masked threshold components for different parameters α. The allowed noise levels in the scalefactor bands of a Layer 3 coder obtained from one masked threshold are given by the hatched area.

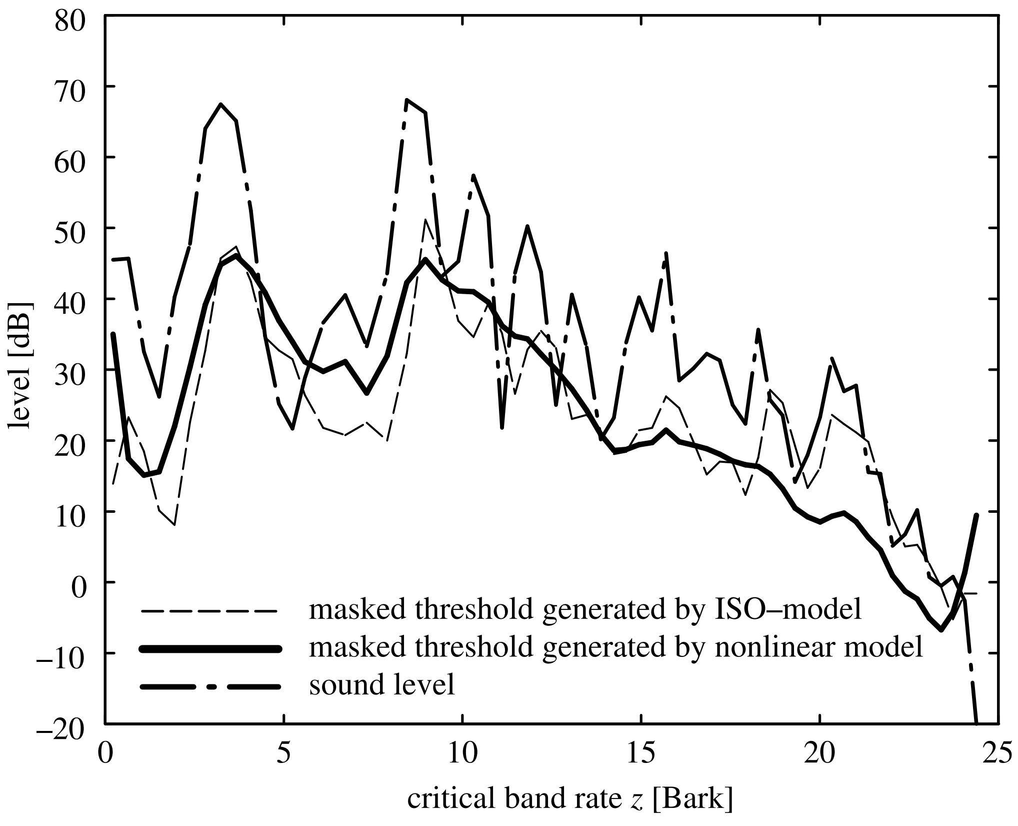

Figure 10 Generated masked thresholds obtained from the nonlinear model and the ISO–model for one block (12ms) of a clarinet recording. For comparison the sound level is also shown.