SOUND QUALITY SYMPOSIUM 98
Ypsilanti, Michigan
1998 April 9

CALCULATING ROUGHNESS USING TIME-VARYING SPECIFIC LOUDNESS SPECTRA

Ulrich Widmann[1)] and Hugo Fastl[2)]

1) Müller-BBM, Schalltechnisches Beratungsbüro, D-82152 Planegg n. München, Germany
2) Lehrstuhl für Mensch-Maschine-Kommunikation, TU München, Germany

INTRODUCTION

Modulated sounds elicit besides other hearing sensations also roughness and fluctuation strength [1]. In technical applications roughness and fluctuation strength play an important role. Aures proposed a method to calculate roughness [2]. Vogel described a common way to model loudness and roughness [3]. Related to this basic concepts some later models we a calculation procedure for roughness is proposed. The concept is based on a model for roughness developped by Fastl [4]. Fastl's model is based on psychoacoustically measured temporal masking patterns. In the following section the characteristics of such patterns is simulated by means of the time-dependant course of specific loudness.

PSYCHOACOUSTIC MODEL OF ROUGHNESS

For sinusoidal tones, FM-tones as well as for modulated broadband noises the hearing sensation roughness depends on the modulation frequency (f_{mod}) [1]. Fig. 1a shows the rel. roughness of AM-modulated broadband sounds according to Fastl [4] as function of f_{mod}. For modulated sinusoidal tones with low carrier frequencies there is a special influence of the bandwith of the critical bands [1].

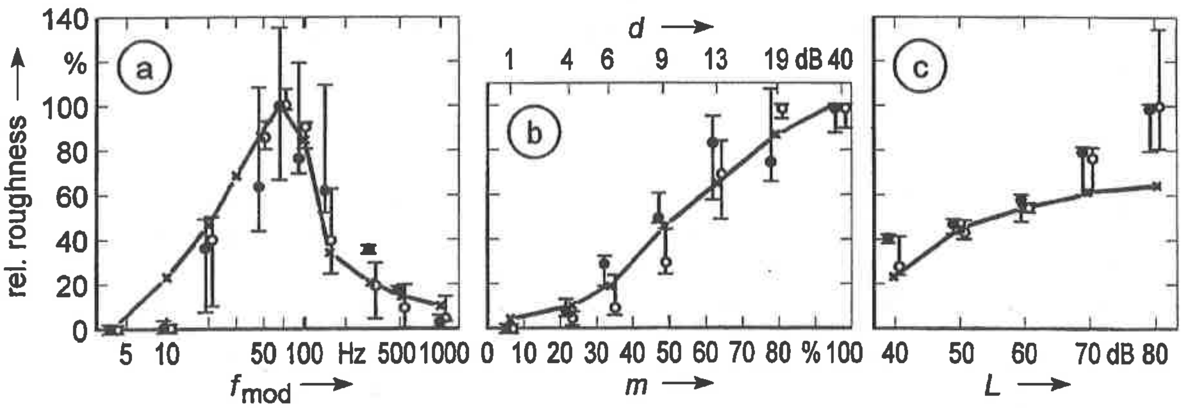

Fig. 1: a) rel. roughness of AM-modulated broadband noise as a function of f_{mod}, b) rel. roughness as a function of the degree of modulation (m) or the modulation depth (d) of an AM-modulated broadband noise for f_{mod} = 70 Hz, c) rel. roughness as a function of the level (L) of AM-modulated broadband noise for f_{mod} = 70 Hz; crosses and line: simulation results.

Figure 1b displays how roughness is related to the degree of modulation (m) or the modulation depth (d). Roughness increases almost proportionally to the square of the degree of modulation (m). Figure 1c depicts the increase of roughness with increasing level (L) of the noise. The increase is relatively small. For a level difference of 40 dB an increase in roughness by a factor of approx. 3 occurs.
Fastl develloped a model for roughness of modulated broadband noises on the basis of psychoacoustically measured temporal masking patterns [4]. Figure 2 displays a schematic sketch of a temporal masking pattern for an amplitude modulated broadband noise for different modulation frequencies. The hatched area indicates the temporal envelope of the masker. It can be seen that the human listening system is able to follow the envelope of the masker for slow modulation frequencies. For higher modulation frequencies the quantity ΔL is decreasing, due to temporal masking effects [4]. The quantity ΔL represents the difference between the maximum and the minimum of the level at threshold of the test tone when measured during one period of the modulation.

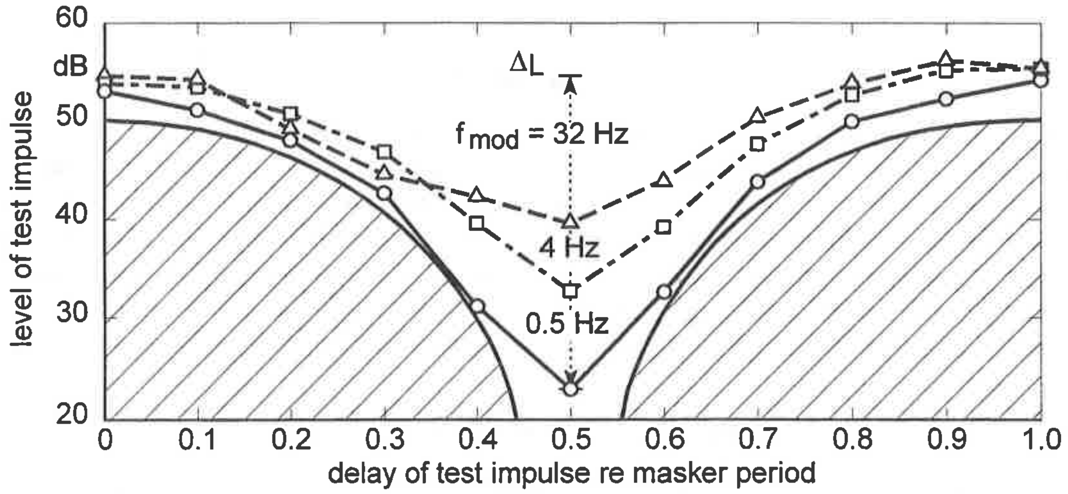

Fig. 2: Temporal masking patterns for an amplitude modulated broadband noise; hatched area: masker envelope; the symbols indicate the threshold of short test tone impulses, triangles: f_{mod} = 32 Hz, circles: f_{mod} = 0,5 Hz, quadrangles: f_{mod} = 4Hz.

Fastl concluded that the main characteristics of roughness can be simulated with model using the quantity ΔL, derived from the measured temporal masking patterns, and the modulation frequency f_{mod}. This can be described by the equation

$$R \approx f_{mod} \bullet \Delta L \,. \quad \text{(eq. 1)}$$

For low modulation frequencies roughness is small although ΔL is big. For increasing modulation frequencies the value of ΔL is decreasing but f_{mod} is increasing. The product reaches a maximum for a modulation frequency of about 70 Hz. At modulation frequencies above 70 Hz the value of ΔL is decreasing rapidly and therefore the product aproaches zero, and roughness disappears [1, 2].

The model states that not directly frequency is detected, but that the temporal changes in the patterns are analyzed [4].

Measured data of temporal masking patterns are very rare and therefore in this paper a method is described simulating their behaviour by means of temporally variable specific loudness patterns as they occur in loudness meters [5]. For broadband noise the value ΔL is almost independant of the critical band rate (z). For narrow-band sounds ΔL depends on z. The nonlinear spread of masking may cause a ΔL at the upper slope which is larger than at the main excited band of the narrow-band sound [1]. This can be included in the model by expanding eq. 1 to

$$R = c \cdot f_{mod} \cdot \int_0^{24\,\mathrm{Bark}} \Delta L_E(z)\, dz = c \cdot f_{mod} \cdot \Delta \quad \text{asper} \,. \qquad \text{(eq. 2)}$$

The constant c is used to standardize the calculated roughness values on the reference sound. This is an 100% amplitude modulated 1kHz tone with a level of 60dB and a modulation frequency of 70 Hz. The reference sound elicits a roughness of 1 asper (vox aspera (lat.): rough voice).

SIMULATING THE MODEL USING SPECIFIC LOUDNESS PATTERNS

From temporal masking patterns a time-variable course of the quantity $L'_E(z, t)$ can be derived. This quantity is very similiar to the quantity $N'(z,t)$ used in a loudness meter. In loudness meters following equations are used [6]:

$$N'(z) = c_1 \cdot ((1\text{-}s + s \cdot 10^{0.1(L'_E(z,t)-L_{HS}(z))/dB})^{0.25} - 1) \ \frac{\text{sone}}{\text{Bark}} \qquad \text{(eq. 3)}$$

with $c_1 = 0.0635 \cdot 10^{0.025\, L_{HS}(z)/dB}$.

Is eq. 3 rearranged, neglecting the term 1-s, the excitation level $L'_E(z, t)$ can be written as

$$L'_E(z,t)/dB = 40\log\left(\frac{N'(z,t)/\text{sone/Bark}}{c_1} + 1\right) - 10\log(s) + L_{HS}(z)/dB. \qquad \text{(eq. 4)}$$

The value ΔL can be derived from the difference

$$\Delta L'_E(z) = L'_E(z,t_{max}) - L'_E(z,t_{min}) =$$

$$= 40 \cdot \left(\log\left(\frac{N'(z,t_{max})/\text{sone/Bark}}{c_1} + 1\right) - \log\left(\frac{N'(z,t_{min})/\text{sone/Bark}}{c_1} + 1\right)\right) \text{dB} \qquad \text{(eq. 5)}$$

where t_{max} states the time where N' is maximal and t_{min} where N' is minimal during one period of modulation. The time-variable values of $N'(z, t)$ can be derived from a loudness meter [5]. This offers the opportunity that well-known procedures for the consideration of spectral and temporal masking can be applied. The N'-values must be calculated every 2ms to allow the detection of modulation frequencies up to approx. 200 Hz. For calculation purposes the integral in eq. 2 can be replaced by a sum of 240 partial level differences, every 0.1 Bark, along the critical band rate scale. The processing of 240 discrete values is successfully used in loudness calculation systems [5, 6].
If big ratios of the term N'/c_1 are assumed eq. 5 can be simplyfied to

$$\Delta L'_E(z) = 40 \cdot \lg \frac{N'(z, t_{max})}{N'(z, t_{min})} \ \text{dB}. \qquad \text{(eq. 6)}$$

Eq. 5 shows that the excitation level difference is related to a specific loudness ratio. This approximation can be used instead of eq. 6, but it gives somewhat worse results with respect to the level dependance of roughness for lower levels. Therefore in the following chapter eq. 5 is used.

CALCULATION PROCEDURE

Figure 3a-e shows the main steps for the roughness calculation procedure. For each of the 21 channels of the loudness meter the modulation frequency and the values N'_{max} and N'_{min} for a detected modulation period are calculated (Fig. 3a).
From the 21 values N'_{max} a loudness pattern with 240 values is derived taking spectral masking into

account (Fig. 3b). The same procedure is applied to the N'_{min} values. From the two patterns, in an intermediate step, the quantity ΔL according to eq. 5 is calculated (Fig. 3c).
Following psychoacoustic results presented by Vogel in 1975 [3] the level difference is set to be maximal 30dB. This feature is especially important at very low modulation frequencies and high degrees of modulation [3].
The quantity is in the next step weighted with a fequency dependant factor.

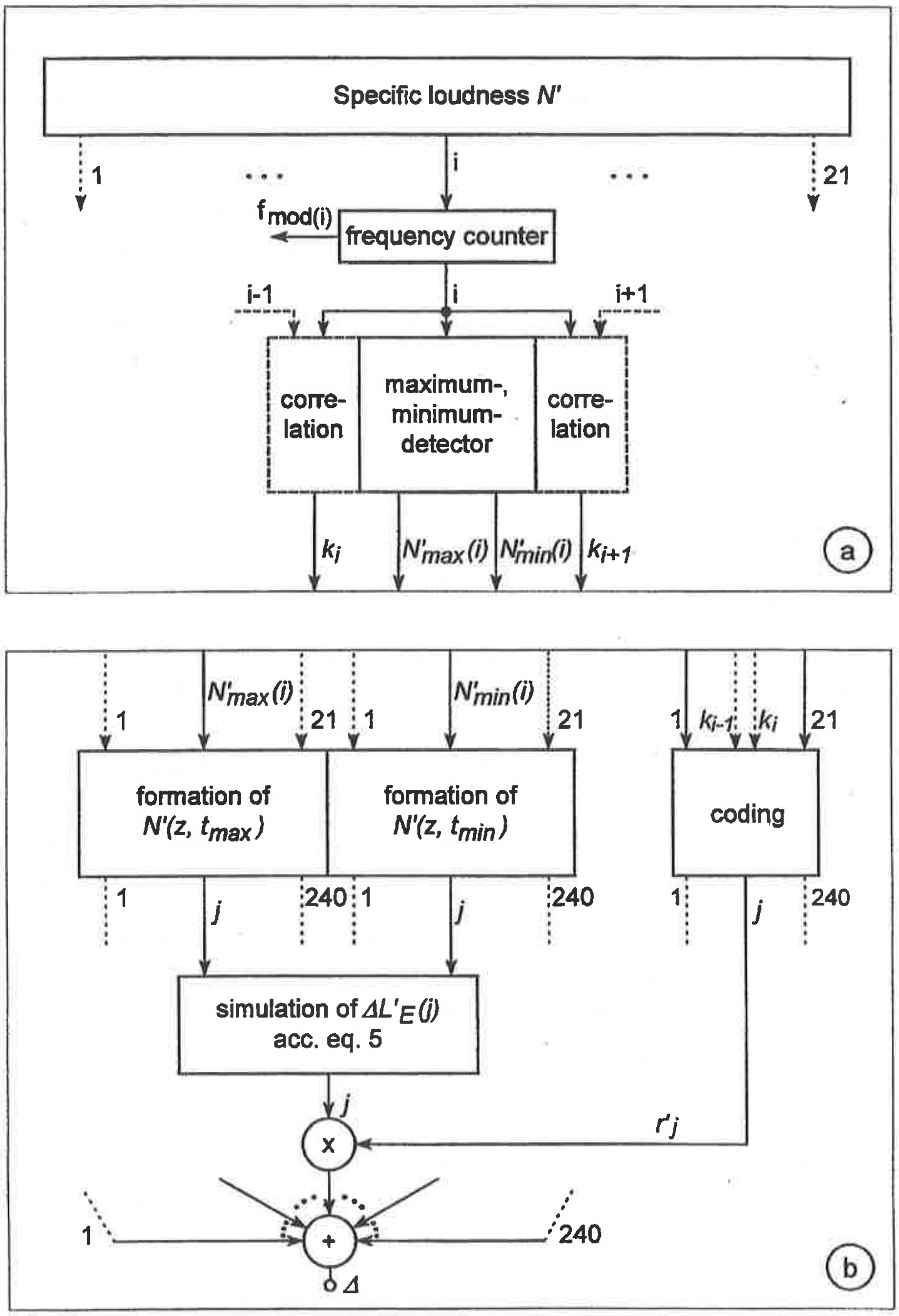

Fig. 3a,b: Calculation procedure for roughness: a) extraction of f_{mod} and maximal and minimal values of specific loudness values N´ in each channel of a loudness meter, b) derivation of the quantity Δ (details see text).

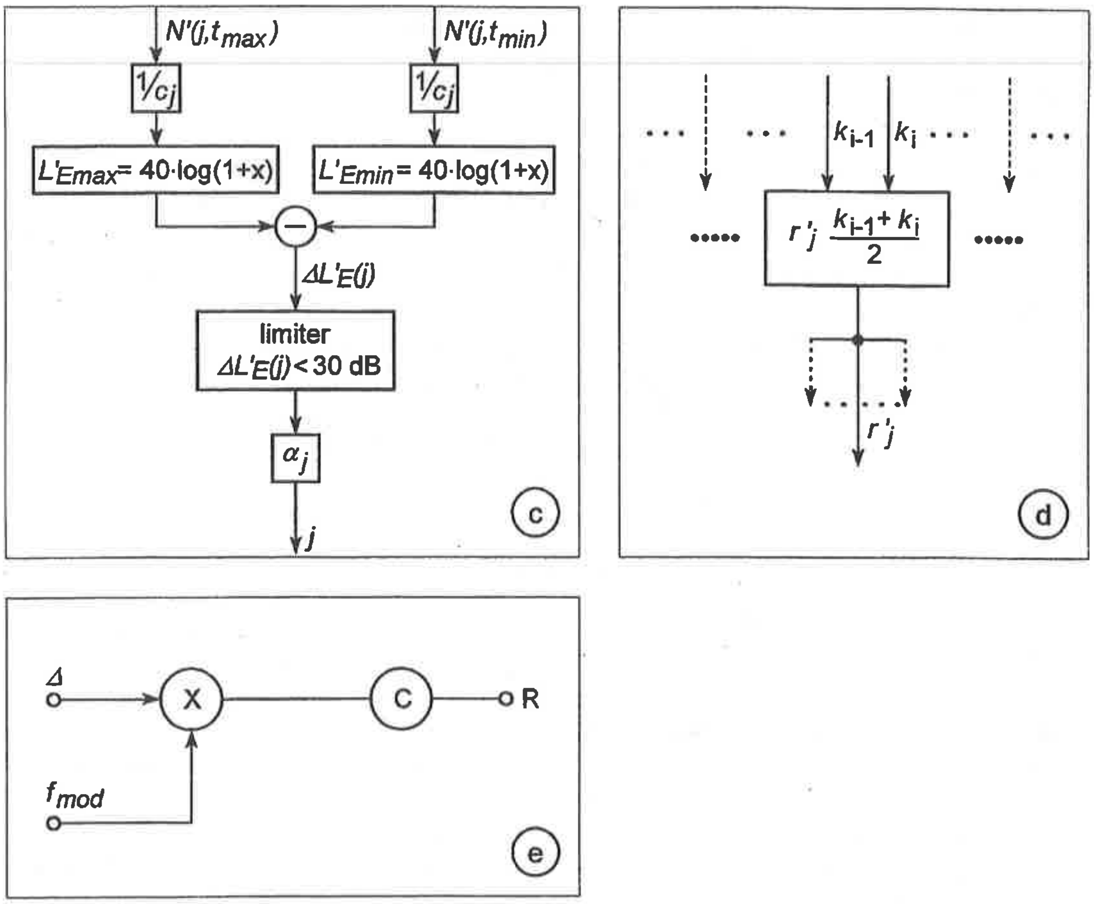

Fig. 3c-e: Calculation procedure for roughness: c) calculation of weighted ΔL, d) crosscorrelation of specific loudness channels, e) schematic sketch of the overall procedure (details see text).

This is necessary to consider results of an investigation of Suchowerskyj, found in 1977 [8]. His studies showed that the hearing system is more sensitive for changes in the high frequency area than in the low frequency area.

A frequency dependant characteristic is also found in measurements of masking patterns [1]. The logarithmic threshold factor s described by Zwicker and Fastl [1] follows approximately

$$\Delta L_S = 10 \cdot \log(1+s) = 2.2 \text{ dB} - 0.05 \text{ dB/Bark} \cdot \text{z/Bark} \qquad \text{(eq. 7)}$$

The weighting factors α(j) in Fig. 3c are therefore empirically set to be

$$\alpha(j) = \frac{1}{10^{\Delta L_s/10\text{dB}} - 1} = \frac{1}{s} \qquad \text{(eq. 8)}$$

Changes in excitation at higher frequencies are therefore weighted 2.5 times higher than at lower frequencies.

From psychoacoustic experiments it is known that unmodulated broadband noise does not produce roughness. Loudness meters divide the measured noise inside a bandpass filter bank into several critical bands having a bandwith Δf(z). Due to this step statistical modulation frequencies of narrow-band noise (f_{mod} = 0.64* Δf) are produced.

To have therefore an additional measure for the correlation of the time course of N´-values in neighboured bands the crosscorrelation is analyzed in parallel (Fig. 3d). This is necessary to reduce the calculated roughness of unmodulated broadband noise. For unmodulated broadband noise the modulations found in neighboured bands are uncorrelated and therefore calculated roughness values can be reduced [for details see e.g. 1, 2].

The weighted quantity $\Delta L(j)$ is later on multiplied with the above mentioned correlation coefficients. As a result the quantity Δ, which simulates the sum of all ΔL´s along the critical band rate scale is derived.
The total roughness R is calculated from the quantity Δ and the analyzed modulation frequency as shown in Fig. 3e.
The described method allows a quantitative simulation of parameters of the hearing sensation roughness. The experimental data for roughness of broadband noise investigated by Fastl [4] can be compared with calculated results of the computer model. The simulation results are indicated in Fig. 1a-c by crosses and solid lines. The calculated values and the subjective data agree very well. Only for the level dependance of roughness some deviations show up with increasing level.

CONCLUSIONS

The above described procedure allows the calculation of roughness of modulated sounds inline with the model proposed by Fastl [4]. From temporally variable specific loudness spectra a quantity can be derived simulating the level difference in the time-course of temporal masking patterns. Besides this quantity also the detected modulation frequency has to be considered.
The model shows that the concept of specific loudness provides an aurally adequate representation of sounds. From temporally variable specific loudness patterns not only loudness, but also other time-variable hearing sensations can be simulated as well, if the appropriate time-constants were applied [7]. Fastl proposed also a psychoacoustic model for the hearing sensation fluctuation strength [9]. With the described procedure also the main characteristics of this hearing sensation could be simulated with specific loudness spectra. Only the combination of f_{mod} and Δ has to be changed [1, 9].

REFERENCES

[1] Zwicker, E., Fastl, H., Psychoacoustics - facts and models, Springer Verlag, Berlin, Heidelberg, New York, 1990.

[2] Aures, W., Ein Berechnungsverfahren für die Rauhigkeit, Acustica 58, 268-281.

[3] Vogel, A., Ein gemeinsames Funktionsschema zur Beschreibung der Lautheit und der Rauhigkeit , Biol. Cybernetics 18, 31-40.

[4] Fastl, H. , Roughness and temporal masking patterns of sinusoidally amplitude modulated broadband noise. In: psychophysics and physiology of hearing (E.F. Evans & J.P. Wilson, eds.), Academic Press, London, 403-414.

[5] Zwicker, E., Deuter, K., Peisl, W., Loudness meters based on ISO 532 B with large dynamic range. Proc. Internoise ´85, Vol. II, 1119-1122.

[6] Zwicker, E., Fastl, H., Widmann, U., Kurakata,K., Kuwano, S., Namba, S., Program for Calculating Loudness According to DIN 45631 (ISO532B). J. Acoust. Jpn (E) **12**, 39-42.

[7] Widmann, U., Lippold, R., Fastl, H., A computer program simulating post-masking for applications in sound analysis systems, In: proc. of Noise Con´98.

[8] Suchowerskyj, W. (1977b): Funktionsschema zur Beurteilung der subjektiven Bewertung von Schalländerungen. Biol. Cybernetics 26,169-124.

[9] Fastl, H., Fluctuation strength and temporal masking patterns of amplitude modulated broadband noise. Hearing Research **8**, 59-69.