NOISE-CON 98
Ypsilanti, Michigan
1998 April 5-8

A COMPUTER PROGRAM SIMULATING POST-MASKING FOR APPLICATIONS IN SOUND ANALYSIS SYSTEMS

U. Widmann[1], R. Lippold[1], H. Fastl[2]

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

INTRODUCTION

A model and a procedure for calculating loudness of steady-state sounds from measured third-octave band levels was published by Zwicker. The procedure is based on the distribution of the specific loudness along the critical band scale. The procedure was adopted in ISO 532B. Some years ago DIN 45631 was revised including a computer program for calculating loudness level in BASIC which runs on IBM compatile or japanese PC's [1]. Also a computer program in ANSI-C was published [2].
If loudness of temporally variable sounds is considered in a first step the specific loudness values have to be treated as time-dependant values. In each critical band the effects of temporal masking must be considerd. For of a complete loudness model also in a second step a network summing up all specific loudness values along the critical bands must be realized in the instrument. Such a network should consider that a loudness meter shows for impulses of tones with decreasing duration total loudness values N_{max} decreasing as well ([3]). This network is not discussed here.
In the following section the first step towards an aurally-adequate time-variable loudness calculation procedure is developped using data of post-masking as measured by Zwicker [4]. A computer program for post-masking is derived from the data. In modern digital sound-analysis systems this procedure can not only be applied for calculation procedures of total loudness, but also for other temporally variable hearing sensations [5].

DATA ON POSTMASKING

Post (forward-) masking has been investigated in several studies in which the dependance of post-masking on masker duration has been measured. One effect seems to be generally accepted: post-masking becomes smaller for masker durations shorter than about 100 ms. Zwicker [4] published some data that show that the form of the decay decreases more rapidly for shorter masker impulses. The masking patterns can be used as a guide for the decay of the ears internal excitatition and to the corresponding specific loudness.
Fig. 1 summarizes the calculated specific loudness-time function $N'_{th}(t)$ (the index 'th' stands for simulated post-masking threshold) as given by Zwicker [4]. An exponential decay in this figure corresponds to a straight line as indicated in the inset of Fig. 1. All curves decay from the same starting point, regardless of duration, much more rapidly for shorter masker durations than for a long lasting impulse. For masker durations of T_M = 5, 10, 30, 200 ms the time constants for the decay as proposed by Zwicker are T_N = 4, 4.5, 6.5 and 11 ms. At longer delays the specific loudness tends to decrease more slowly. A time constant of about 15 ms can be derived from the data.

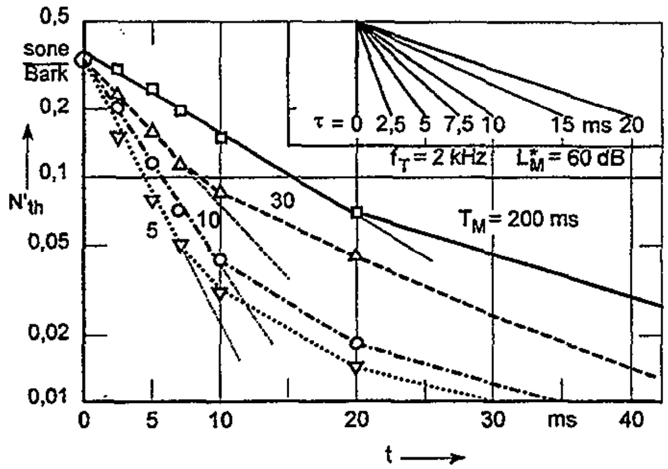

Fig. 1: Post-masking curves according to [4] transformed into specific loudness N'_{th} as a function of masker duration T_M. Inset: Several time constants for exponential decay.

The described behaviour could be simulated with an anolog non-linear model. Fig. 2 shows an analog network already proposed by Zwicker [4].

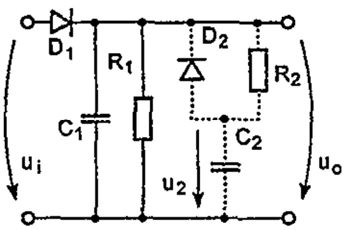

Fig. 2: Network simulating a duration dependant decay

CONTINOUS TREATMENT OF THE NETWORK

The behaviour of the network shown above is determined by the time-constants $\tau_{short}=R_1 \cdot C_1$, $\tau_{long}=R_1 \cdot (C_1+C_2)$ and $\tau_{var} = R_2 \cdot C_2$.
Thinking with ideal diodes we have to distinguish several cases. Because of the Diode D_1 the voltage u_o at the output of the network is never lower than the input voltage u_i. If u_i becomes lower than u_o (case 1) C_1 is discharged. Depending on the voltage u_2 of capacitor C_2 there are

- case 1.1 with $u_o > u_2$: C_1 is discharged via R_1 and via R_2 in C_2 and from the differential equations

$$\dot{u}_o = -\frac{1}{C_2}\left(\frac{u_o}{R_1}+\frac{u_o-u_2}{R_2}\right) \tag{1}$$

$$\dot{u}_2 = \frac{u_o-u_2}{C_2 \cdot R_2} \tag{2}$$

we get

$$u_o(t) = \left(\tau_{var} \cdot \lambda_1 + 1\right) K_1 \cdot e^{\lambda_1 t} + \left(\tau_{var} \cdot \lambda_2 + 1\right) K_2 \cdot e^{\lambda_2 t} \tag{3}$$

with:
t indicating the time since u_2 became lower than u_o and the abbreviations

$$\lambda_{1,2} = \frac{1}{2\cdot\tau_{var}\cdot\tau_{short}}\left(-\tau_{var}-\tau_{long}\pm\sqrt{\left(\tau_{var}+\tau_{long}\right)^2-4\cdot\tau_{short}\cdot\tau_{var}}\right).$$

$$K_1 = \frac{u_o(t=0)-u_2(t=0)\cdot\left(\tau_{var}\lambda_2+1\right)}{\tau_{var}\left(\lambda_1-\lambda_2\right)}, \qquad K_2 = \frac{u_o(t=0)-u_2(t=0)\cdot\left(\tau_{var}\lambda_1+1\right)}{\tau_{var}\left(\lambda_2-\lambda_1\right)}$$

and

- case 1.2 with $u_o=u_2$: C_1+C_2 are discharged via R_1, so

$$\dot{u}_o = \frac{-u_o}{(C_1+C_2)\cdot R_1} \tag{4}$$

is valid and we get

$$u_o(t) = u_o(t=0)\cdot e^{-t/\tau_{long}}, \tag{5}$$

where t now indicates the time, since $u_o=u_2$ holds.

Diode D_2 prevents u_2 from becoming higher than u_o.
Besides case 1 it is possible that $u_o=u_i$ (for instance if u_i has risen and remained afterwards constant). Under this condition C_2 is charged (without discharging C_1), if $u_o>u_2$ (case 2.1). The valid differential equation in case 2.1 is

$$\dot{u}_2 = \frac{1}{\tau_{var}}(u_o-u_2) \tag{6}$$

and u_2 follows

$$u_2(t) = \left(u_2(t=0)-u_o\right)\cdot e^{-t/\tau_{var}}+u_o, \tag{7}$$

t is the time since u_o became equal u_i.
It is important to know $u_2(t)$ because it is one initial value, necessary for determining u_o in case 1 (see eqs. 3 and 5). If C_2 is already charged on the same voltage like C_1 (case 2.2) u_2 remains unchanged. As mentioned above $u_o<u_2$ can never occur.
Fig. 3 shows an example for the time-dependant decay of the voltages involved in the decay process. The output signal follows immediately an increase of the input voltage. After the input signal is switched of, the decay of the output signal starts. The slope of the decrease follows the short time constant. After the condenser C_2 is charged the slope changes to the slow decay rate.

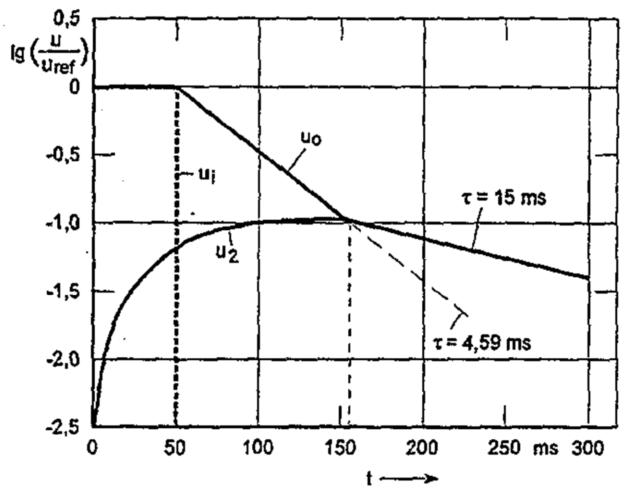

Fig. 3: Time history of the voltages of the network for an input impulse with a duration of 50 ms.

DISCRETE TREATMENT OF THE NETWORK

To realize the transfer characteristic of the above discussed network on sampled time data (sampling frequency $f_s = 1/\Delta t$) we can use the time dependent functions of eqs. (3), (5) and (7) to determine the new value $u_o(t)$ from the input value $u_i(t)$ and the known values $u_o(t-\Delta t)$ and $u_2(t-\Delta t)$. Indicating

- $u_o(t) \rightarrow u_o$,
- $u_o(t-\Delta t) \rightarrow u_{o0}$,
- $u_2(t) \rightarrow u_2$,
- $u_2(t-\Delta t) \rightarrow u_{20}$ and
- $u_i(t) \rightarrow u_i$

we get:

- case 1.1: ($u_i < u_{o0}$ and $u_{o0} > u_{20}$)
  $$u_2 = u_{o0} \cdot B_0 - u_{20} \cdot B_1, \tag{8}$$
  $$u_o = u_{o0} \cdot B_2 - u_{20} \cdot B_3, \tag{9}$$
- case 1.2: ($u_i < u_{o0}$ and $u_{o0} = u_{20}$)
  $$u_o = u_2 = u_{o0} \cdot B_4, \tag{10}$$
- case 2.1: ($u_i = u_{o0}$ and $u_{o0} > u_{20}$)
  $$u_o = u_i \tag{11}$$
  $$u_2 = (u_{20} - u_i) \cdot B_5 + u_i \tag{12}$$

with the constants

$$B_0 = \frac{1}{\tau_{var}(\lambda_1 - \lambda_2)}\left(e^{\lambda_1 \Delta t} - e^{\lambda_2 \Delta t}\right),$$

$$B_1 = \frac{1}{\tau_{var}(\lambda_1 - \lambda_2)}\left((\tau_{var}\lambda_2 + 1)e^{\lambda_1 \Delta t} - (\tau_{var}\lambda_1 + 1)e^{\lambda_2 \Delta t}\right),$$

$$B_2 = \frac{1}{\tau_{var}(\lambda_1 - \lambda_2)}\left((\tau_{var}\lambda_1 + 1)e^{\lambda_1 \Delta t} - (\tau_{var}\lambda_2 + 1)e^{\lambda_2 \Delta t}\right),$$

$$B_3 = \frac{(\tau_{var}\lambda_1 + 1)(\tau_{var}\lambda_2 + 1)}{\tau_{var}(\lambda_1 - \lambda_2)}\left(e^{\lambda_1 \Delta t} - e^{\lambda_2 \Delta t}\right),$$

$$B_4 = e^{-\Delta t/\tau_{long}},$$

$$B_5 = e^{-\Delta t/\tau_{var}}.$$

In the discrete realization of the network we only can check the conditions distinguishing between several cases at discrete times. So, solving the problem in finite steps of length Δt, it can happen that u_i became larger than u_{o0}. In this new case (case 3) u_o follows u_i whereas u_2 obeys eq. (12). Furthermore we also have to prevent that u_o becomes lower than u_2 during one step (case 1.1.1). It is not exact to set $u_2 = u_o$, but the error using this approach is negligible.
The following C-program shows the complete algorithm for a discrete realization of the network. τ_{short} and τ_{long} were chosen to be 5 ms and 15 ms. τ_{var} is 75 ms, which seems to be an appropriate setting for loudness analysis purposes (see Fig. 4).
First we have to initialize the constants B_i depending on the sampling interval Δt and to set a definite state of the capacitors by calling routine *init_n_lp* and afterwards function *nl_lp* calculates for the next value u_i the corresponding u_o.

REALIZATION AS C-PROGRAM

Figure 4 describes the source code of a C program realizing the digital structure described above.

```
#include <math.h>
#include <stdio.h>

#define t_short 0.005
#define t_long  0.015
#define t_var   0.075

float B[6],
      u_o_0,  /* u_o(t-delta_t) */
      u_2_0;  /* u_2(t-delta_t) */

void init_nl_lp(float delta_t)
/* initializes constants B and states
   of capacitors C1 and C2,
   1/delta_t = sampling frequency */
{
  float lambda_1, lambda_2,
        p, q, den,
        e1, e2;
  int i;

  p = (t_var + t_long)/
      (t_var*t_short);
  q = 1/(t_short*t_var);
  lambda_1 = -p/2+sqrt(p*p/4 - q);
  lambda_2 = -p/2-sqrt(p*p/4 - q);
  den = t_var*(lambda_1 - lambda_2);
  e1 = exp(lambda_1*delta_t);
  e2 = exp(lambda_2*delta_t);

  B[0] = (e1-e2)/den;
  B[1] = ((t_var*lambda_2+1)*e1 -
         (t_var*lambda_1+1)*e2)/den;
  B[2] = ((t_var*lambda_1+1)*e1 -
         (t_var*lambda_2+1)*e2)/den;
  B[3] = t_var*lambda_1+1)*
         (t_var*lambda_2+1)*
         (e1-e2)/den;
  B[4] = exp(-delta_t/t_long);
  B[5] = exp(-delta_t/t_var);

  u_o_0 = 0; /* at beginning
    capacitors C1 and C2 are
    discharged */
  u_2_0 = 0;
}
```

```
float nl_lp(float u_i)
/* calculates u_o(t) from u_i(t) using
   u_o(t-delta_t) and u_2(t-delta_t)*/
{
  float u_o, u_2;

  if (u_i < u_o_0)             /* case 1 */
    if (u_o_0 > u_2_0){ /* case 1.1 */
      u_2 = u_o_0*B[0] - u_2_0*B[1];
      u_o = u_o_0*B[2] - u_2_0*B[3];
      if (u_i > u_o)
        u_o = u_i; /* u_o can't become
          lower than u_i */
      if (u_2 > u_o)    /* case 1.1.1 */
        u_2 = u_o; /* u_2 can't become
          higher than u_o */
    }
    else{                    /* case 1.2 */
      u_o = u_o_0*B[4];
      if (u_i > u_o){
        u_o = u_i; /* u_o can't become
          lower than u_i */
      }
      u_2 = u_o;
    }
  else{
    if (u_i == u_o_0){     /* case 2 */
      u_o = u_i;
      if (u_o > u_2_0)   /* case 2.1 */
        u_2 = (u_2_0 - u_i)*B[5] +
              u_i;
      else                 /* case 2.2 */
        u_2 = u_i;
    }
    else{                    /* case 3 */
      u_o = u_i;
      u_2 = (u_2_0 - u_i)*B[5] + u_i;
    }
  }

  u_o_0 = u_o; /* preparation for next
    step */
  u_2_0 = u_2;

  return(u_o);
}
```

Fig. 4: C-source code describing the duration dependant decay

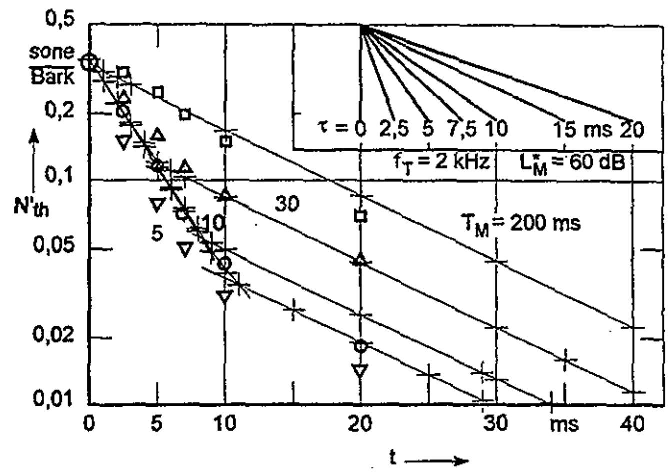

Fig. 5: comparison of the original data (solid lines) and the simulated post-masking data (symbols): T_M = 200 ms (quadrangles), T_M = 30 ms (triangles), T_M = 10 ms (circles), T_M = 5 ms (tilted triangles).

Fig. 5 compares the calculated decays with the algoritm for masker durations of 5, 10, 30 and 200 ms and the measured data (see Fig. 1). For longer masker durations the calculated data and the values N'_{th} agree very well. For short masker durations some deviations occur with the chosen time-constants.

CONCLUSIONS

If loudness of temporally variable sounds is considered the specific loudness values have to be treated as time-dependant values. A temporal resolution of 2 ms can be chosen for practical applications. In digital sound analysis systems post-masking effects must be taken into account. Post-masking depends strongly on duration. With the above described digital realization a procedure is available to account for this duration dependant decay of post-masking. For durations of 10 ms to 200 ms the behaviour of the algorithm fits the published data of post-masking sufficiently. Therefore this algorithm may be applied in loudness meters or other pre-processing devices.

REFERENCES

[1] Zwicker, E., Fastl, H., Widmann, U., Kurakata,K., Kuwano, S., Namba, S. (1991): Program for Calculating Loudness According to DIN 45631 (ISO532B). J. Acoust. Jpn (E) **12**, 39-42.

[2] Widmann, U. (1993): Krach gemessen - Gehörbezogene Geräuschbewertung, c't 4/93 (including a C-listing for loudness acc. DIN 45631),

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

[4] Zwicker, E. (1984): Dependance of post-masking on masker duration and ist relation to temporal effects in loudness. J. Acoust.Soc.Am. 75.

[5] Widmann, U., Fastl, H., Calculating roughness using time-varying specific loudness spectra. In: Proc. Sound Quality Symposium '98 together with Noise Con'98.