ARCHIVES OF ACOUSTICS
8, 2, 139-154 (1983)

EVALUATION OF ELECTROACOUSTIC DEVICES BY THE EQUIVALENT SCALE METHODS

TOMASZ ŁĘTOWSKI, JACEK SMURZYŃSKI

Laboratory of Music Acoustics, Chopin Academy of Music
(00-368 Warsaw, ul. Okólnik 2)

This paper presents the preliminary results of experiments on the possibility of using methods based on the mutual masking of signals compared, in auditory evaluation. An attempt was made to objectivize the measure of the similarity of the signals, or the transmission channel quality, and to evaluate the degree of differentiation of objects investigated on an equivalent physical scale. In the preliminary experiments a number of experimental procedures were examined, two of which, characterized by the relatively highest differentiation sensitivity, underwent more detailed evaluation. The results indicated the possibility of objectivizing the measure of the transmission circuit quality in the range of frequency response variation over a third-octave.

1. Introduction

The ultimate criterion of quality (usefulness) of devices for broadly understood sound generation and transmission is auditory evaluation of sound signals generated or transmitted by these devices. This evaluation can also extend to the sound signals themselves, considered apart from the forming devices used. This can occur, for example, in the evaluation of the naturalness of sound effects or the efficience of the performance of different warning signal types.

The purpose of both the evaluation of sound signals themselves (direct auditory evaluation) and the evaluation of sound generation and transmission devices (indirect auditory evaluation) which is based on the former, may be to determine the degree of mutual similarity between objects evaluated and the degree to which one of the objects dominates over the other. In such cases the basic kinds of auditory evaluation are differentiation and classification of evaluated objects, respectively. A discussion of the applications of the two kinds of evaluation and a review of the different forms of auditory evaluation in practice is given by ŁĘTOWSKI in paper [3].

The basic disadvantage of the above two kinds of auditory evaluation is the lack of a stable and unambiguously defined general measure of the quality (similarity) of objects which would permit the reliable reference of the results of different auditory experiments to one another and unambiguous determination of the degree of differentiation between given evaluated objects. Although the categorization of evaluated objects on interval scales characterized by the patterns (anchors) of some selected categories permits the comparison of the results of different (independent) auditory experiments, it is, however, a very imprecise tool and covers a rather limited range of practical application. A greater opportunity for the evaluation of objects is provided by the anchored absolute (ratio) scales, but in the case of evaluating most sensations and emotions the use of scales of this type raises justifield objections [6, 7].

In view of this situation, for some time attempts have been made to use in the field of auditory evaluation the procedures of differentiation and classification of objects, with a simultaneous representation of the results on some different types of equivalent substitute continua. They can be abstract (the auxiliary scale of evaluation reliability degrees), graphic (a linear distance scale) or expressed in physical units. Attempts have also been made to use for this purpose the procedures of cross modality scaling introduced into experimental psychology by STEVENS [9].

MUNSON and KARLIN [4] were the first to develop, in 1962, an equivalent quality scale of physical character in the field of auditory impressions. In their investigations of the evaluation of devices for speech transmission they introduced as the "objective" measure of the quality difference between two speech signals the difference in the sound intensity level (dB) between them which is necessary for the achievement of the same subjective quality of the two sounds. The paired comparison method was used as the scaling procedure. The signal at the output of the transmission channel investigated is compared with the original (input) signal masked by some additional random noise level. The difference between the levels of the output and input signals which were recognized as qualitatively equal with a specific random noise level interfering with the input signal is recognized as the measure of the quality of a given transmission channel. This method was later developed, among others, by ROTHAUSER *et al.* [8] and NAKATANI and DUKES [5].

Another method of physical equivalent scales which has been proposed for description of (electroacoustic) transmission channels consists in a controlled introduction of a given type of deformation into the input signal. The measure of the quality of the transmission channel is the degree of deformation which can be introduced into the input signal without causing a noticeable change in the output signal. Such qualitative scales were used by BORDONE-SACERDOTE and MODENA [1] and KULESZA [2], for example.

The aim of this paper is to investigate the possibility of using in the field of auditory evaluation methods based on the mutual interference (masking) between signals compared with each other and to evaluate the degree of differentiation between objects investigated on a corresponding physical equivalent scale. The methods proposed are based on the assumption that the sound intensity level of the reference signal X, or some additional stimulus M, which is necessary for a given perception threshold of the signal Y evaluated to be achieved, can be the "objectivized" measure of the similarity of the signal Y to the signal X, or the measure of the quality of the transmission channel under study.

2. General characteristic of the investigations

The previous experience of the authors in the field of auditory evaluation and investigations in the fundamentals of hearing led to the hypothesis that the hearing threshold for a given signal under given conditions of simultaneous or succesive masking can be the measure of the deviation of this signal from some "norm" (reference signal) assumed. In order to verify this hypothesis five experimental procedures based on the use of the masking effect were investigated. These procedures were named in the following way:

a. the method of alternative sequences I,
b. the method of alternative sequences II,
c. the method of alternative sequences III,
d. the pulsation method
e. the method of succesive masking.

The time paradigms of the particular procedures are shown schematically in Fig. 1.

The method of alternative sequences I consists in the presentation, against a continuous reference signal, of a sequence of pulses of a signal to be compared. The listeners' task is to find such a low level of the compared signal, or such a high level of the reference signal, at which a characteristic pulsation (modulation) of the resultant signal occurs. The measure of similarity between X and Y is the difference in level which occurs between these signals for the controlled signal level defined by the listener.

The method of alternative sequences II consists in the presentation, against a given continuous masking signal, of alternating sequences of equal-level pulses of the reference and comparative signals. The listeners' task is to select such a low level of the masking signal at which the difference between the sequences of the pulses becomes inaudible. The measure of similarity between the signals X and Y is the minimum level of the signal M defined by the listener according to the present instruction.

The method of alternative sequences III consists in the alternating presentation, against pulses of the masking signals, of pulses of the reference or comparative signals. The listeners' task is to find such a high level of the masking signal at which differences between the pulses begin to be audible. The measure of similarity between the signals X and Y is the level of the masking signal, defined by the listener, with a constant level of the reference and comparative signals.

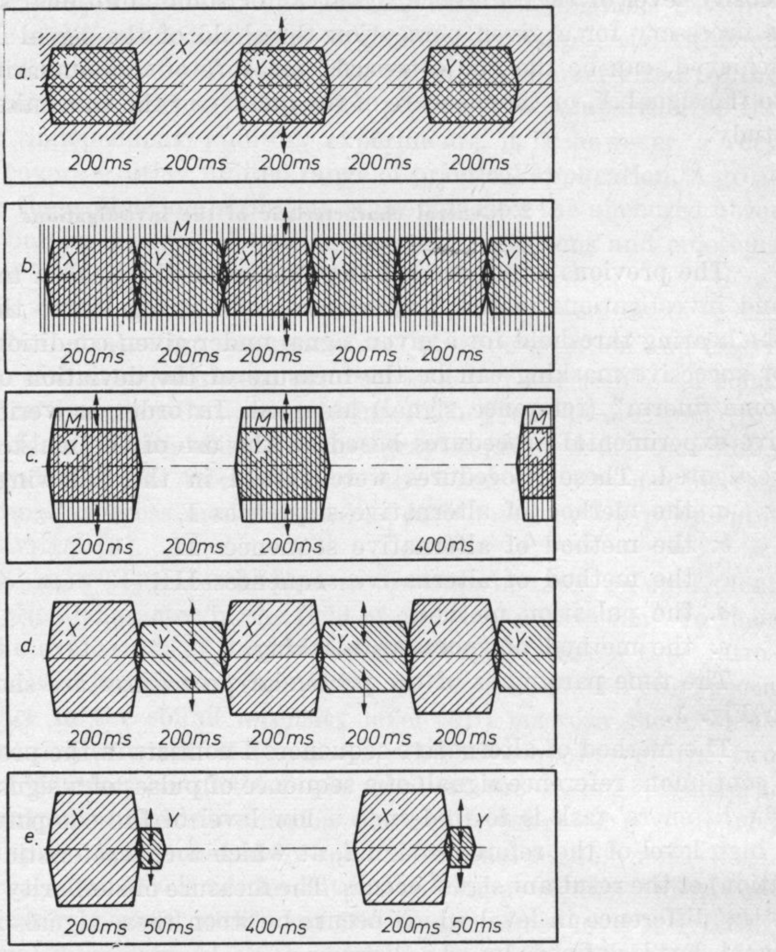

Fig. 1. Time paradigms of the measurement procedures: a. method of alternative sequences I, b. method of alternative sequences II, c. method of alternative sequences III, d. pulsation method, e. method of successive masking

The pulsation method consists in the alternative presentation of pulses of the reference signal X and the comparative signal Y. The listener's task is to adjust the level of the signal Y so as to obtain the sensation of continuity of this signal against the background of pulses of the signal X. The measure of difference between the signals X and Y is the difference between their levels when the pulsation threshold occurs.

The method of succesive masking consists in the presentation of sequences of pulses of the reference signal X, followed by short pulses of the comparative signal Y. The listener increases the level of the signal Y until pulses of this signal begin to be audible when a pulse of the signal X ends. The subjective measure of difference between the two signals is the difference between the levels of the signals X and Y selected for the above purpose by the listener.

Table 1. The physical signals in particular investigation methods

Investigation method	Variant	Reference signal	Comparative signal	Masking signal
method of alternative sequences I	*A*	*SL*	*FSL*	–
	B	*SR*	*FSR*	–
	C	*SRM*	*FSRM*	–
method of alternative sequences II	*D*	*SL*	*FSL*	*SL*
	E	*SL*	*FSL*	*SR*
	F	*SL*	*FSL*	*SRM*
method of alternative sequences III	*G*	*SL*	*FSL*	*SL*
	H	*SL*	*FSL*	*SR*
	I	*SL*	*FSL*	*SRM*
pulsation method	*J*	*SL*	*FSL*	–
method of succesive masking	*K*	*SL*	*FSL*	

SL – random noise (white noise)
SR – pink noise (with power density spectrum decreasing as a function of frequency at a rate of 3 dB/oct.)
SRM – uniformly masking noise (with power density spectrum stable up to the frequency $f = 500$ Hz and decreasing at a rate of 3 dB/oct. above this frequency)
FSL – filtered random noise
FSR – filtered pink noise
FSRM – filtered uniformly masking noise

Table 1 shows physical signals used by the authors in the implementation of particular investigation procedures. In all cases the transmission channel system investigated was simulated with a controlled Brüel and Kjaer 5587 spectrum shaper. The (output) comparative signals corresponded to four high-pass and four low-pass responses of the equaliser with the following cut-off frequencies (−3 dB):

1. the low-pass system: a. 7 kHz, b. 8.9 kHz, c. 11.2 kHz, d. 14.1 kHz;
2. the high-pass system: e. 179 Hz, f. 224 Hz, g. 282 Hz, h. 355 Hz.

The selection of such a “laboratory” transmission system was caused by the necessity of unambiguous description at this stage of investigations of physical transmission properties of the systems compared.

The investigations presented in this paper were carried out in two stages. In the first stage, the authors examined all the 8 "laboratory" transmission systems for all the 11 stimuli configurations ($A-K$) shown in Table 1. The investigations also included some "laboratory" systems with narrow-band responses (single 1/3 octave and octave bands). The loudness level of the monitoring of stable signals, i.e. those not adjusted by the listener, was in all cases 50-60 phons. In individual investigations both loudspeaker monitoring (loudspeaker GK 124) and earphone monitoring (earphones Peerless PMB-6) were used.

It was found in these preliminary investigations that among all the psychoacoustic procedures examined, only the version J of the method of alternative sequences III, called the method of alternative sequences below, and the pulsation method seem to provide the purpose-desired sensitivity in differentiating evaluated objects and the repeatability of the investigation results. Therefore, these methods underwent more formalized experiments whose results are the object of the present communication.

3. Pulsation method

The measurement system used in the pulsation method is shown schematically in Fig. 2. The reference signal X (channel I) was white noise recorded on magnetic tape (tape recorder Revox A77). The comparative signal Y (channel II) was a white noise signal (from Brüel and Kjaer 1024 generator) filtered by a set of 36 parallel 1/3 octave filters (Spectrum Shaper 5587). Both signals

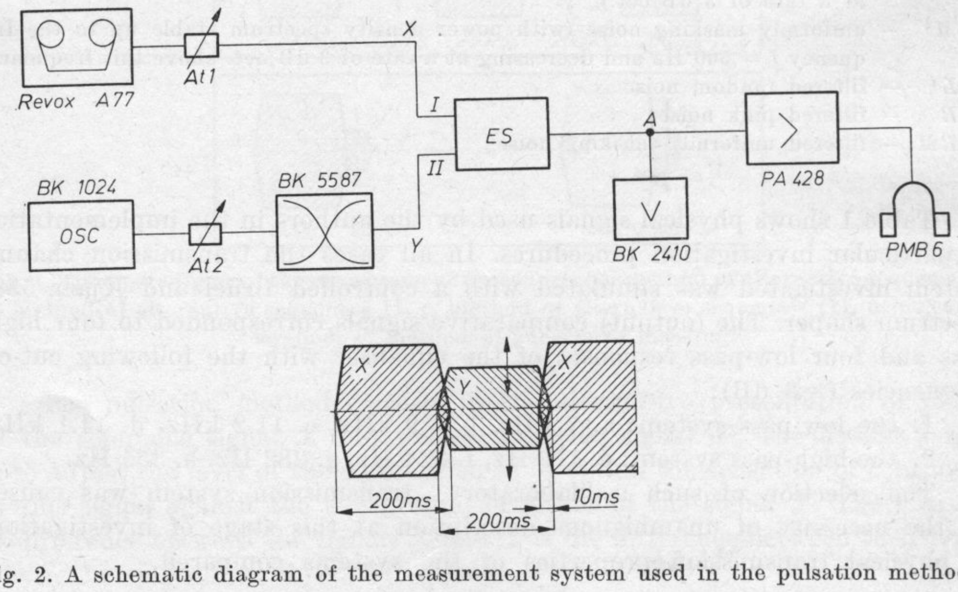

Fig. 2. A schematic diagram of the measurement system used in the pulsation method

were fed to the corresponding inputs of a multi-channel electronic switch (*ES*). At the output of the switch the two signals were presented in the form of alternate pulses. The duration of each pulse was 200 ms (the rise time being 10 ms). When appropriately amplified (by the amplifier Fonica PA 428), the signals were heard by the listener through the orthodynamic headphones PMB-6. In the course of the experiment the listener adjusted with a decade attenuator At2 the level of the comparative signal *Y* so that is was heard as continuous. The corresponding frequency responses of channel II were created, analogously to the preliminary experiments, by means of a set of 1/3 octave filters. 3 listeners, workers of the Sound Engineering Department, took part in the investigations. After 3 trial series each of the listeners made 5 measurements for each frequency response of channel II. The level of the reference signal at point *A* was 290 mV, while its loudness level at the output of the headphones was 60 phons. The level of the comparative signal at which the listener set the pulsation threshold was read from a Brüel and Kjaer 2410 voltmeter. It should be noted that the pilot experiments showed that the use of pulses with shorter duration than the assumed one (higher switch frequency) and loudspeaker monitoring decreases the reliability of evaluation.

4. Method of alternative sequences

The measurement system used in the method of alternative sequences is shown in Fig. 3. White noise from a Brüel and Kjaer 1024 generator was supplied to two inputs (II and III) of the electronic switch through two separate

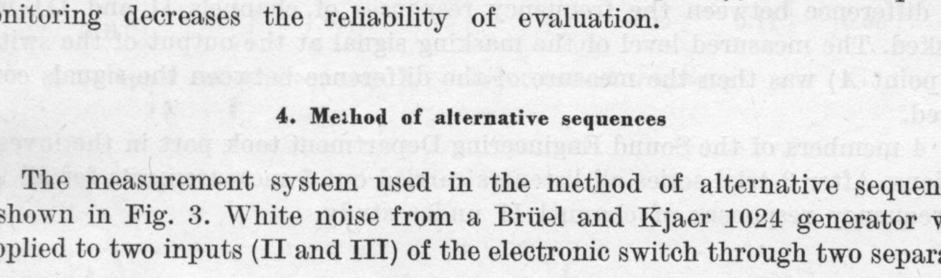

Fig. 3. A schematic diagram of the measurement system used in the method of alternative sequences

channels. One of the channels (III) had a flat frequency response in the whole acoustic band. The other channel (II) included an adjustable spectrum shaper. The pulses of the signals X and Y (with 200 ms duration) thus obtained were presented alternatively at the output of the switch. The interstimulus interval was 200 ms. The next sequence was repeated after 400 ms. The rise and decay times of the signals were 10 ms. A signal of uniformly masking noise (M), reproduced from magnetic tape, was added to both signals. The signal M was supplied to the input of the electronic switch. Summed-up pulses with the same duration were presented to listeners through the PMB-6 headphones. The voltage of the reference signal X (channel III) at the output of the switch was kept at a constant level of 80 m V (point A). Before the signal M was switched on, the loudness level of the pulses X and Y was adjusted subjectively and was about 60 phons.

A system of filters was used to control the frequency response of channel II, as in the preliminary experiments. In the course of the experiment the listener, using the attenuator $At1$, increased the level of the masking signal M until he recognized the timbre of the pulses as the same. This signified that the difference between the frequency responses of channels II and III was masked. The measured level of the masking signal at the output of the switch (at point A) was then the measure of the difference between the signals compared.

4 members of the Sound Engineering Department took part in the investigations. After 3 trial series all listeners carried out 5 measurements for all the 8 frequency responses of channel II under study.

5. Discussion of the results

The results obtained in the investigations by the pulsation method are given in Table 2 and Fig. 4. Particular rows of Table 2 and Fig. 4 show the mean values of the voltage level of the comparative signal at which the listener set the pulsation threshold.

The results were obtained for 3 listeners and 8 frequency responses. Table 2 also gives the values of the corresponding standard deviations (dB). For these results the significance of differences among the mean values was examined by the t-test for uncorrelated data

$$t = \frac{Y_1 - Y_2}{\sqrt{\frac{\sigma_1^2 + \sigma_2^2}{n-1}}},$$

where Y_1 and Y_2 are the mean values of the voltage of the comparative signal for the pulsation threshold determined by the listener for two different frequency responses of channel II (a, b, c, d and e, f, g, h), σ_1^2 and σ_2^2 are the values

Table 2. The results for the pulsation method. The mean values of the values of the level of the pulsation threshold and the values of the corresponding standard deviations

Listener		Cut-off frequency of channel II f_g [kHz]				f_d [Hz]			
		7 kHz	8.9 kHz	11.2 kHz	14.1 kHz	179 Hz	224 Hz	282 Hz	355 Hz
		a	*b*	*c*	*d*	*e*	*f*	*g*	*h*
1	*Y* [dB]	−14.52	−13.43	−12.4	−11.4	−10.23	−10.03	−9.9	−9.79
	σ [dB]	0.71	0.63	0.21	0.74	0.21	0.12	0.17	0.13
2	*Y* [dB]	−14.52	−13.19	−12.22	−10.99	−10.12	− 9.95	−9.87	−9.84
	σ [dB]	0.27	0.54	1.0	0.18	0.11	0.11	0.05	0.11
3	*Y* [dB]	−15.29	−13.64	−12.92	−12.04	−10.31	−10.26	−10.2	−10.28
	σ [dB]	0.77	1.02	0.15	0.34	0.09	0.11	0.06	0.16

f_g — upper cut-off frequency, f_d — lower cut-off frequency

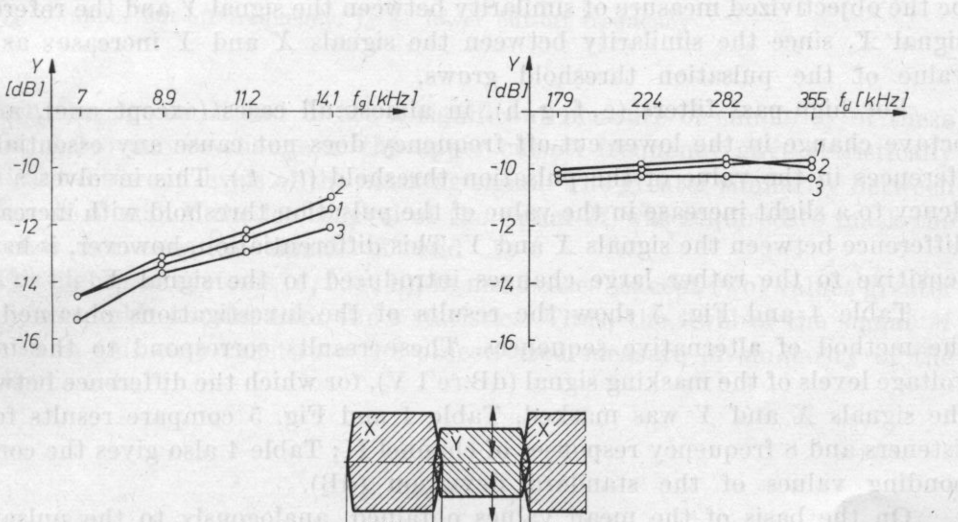

Fig. 4. The mean values of the level of the signal *Y* corresponding to the pulsation threshold. Results for 3 listeners as a function of the cut-off frequencies of the frequency response of channel II

of the corresponding variances and n is the number of results for a measurement point ($n = 5$). Table 3 shows the results of the t – test obtained for neighbouring measurement points. For $2n-2 = 8$ degrees of freedom and the significance level $a = 0.1$ the critical value of the t-test is $t_a = 1.86$.

Table 3. The values of the t statistics in comparison of different frequency responses of channel II in the pulsation method

Comparison of frequency responses of channel II	Listener 1	2	3
$a-b$	8.84	13.86	11.11
$b-c$	12.37	7.84	6.57
$c-d$	9.86	12.33	12.83
$e-f$	1.61	2.12	0.78
$f-g$	1.30	1.34	0.89
$g-h$	1.0	0.45	0.97

For low-pass systems (a, b, c, d) the results – $t \geqslant t_a$ – indicate the significance of differences between the values of pulsation thresholds with a 1/3 octave change in the cut-off frequency. The value of the voltage level of the comparative signal Y at which the pulsation threshold was set can in this case be the objectivized measure of similarity between the signal Y and the reference signal X, since the similarity between the signals X and Y increases as the value of the pulsation threshold grows.

For high-pass filters (e, f, g, h), in almost all cases (except one), a 1/3 octave change in the lower cut-off frequency does not cause any essential differences in the value of the pulsation threshold ($t < t_a$). This involves a tendency to a slight increase in the value of the pulsation threshold with increasing difference between the signals X and Y. This differentiation, however, is hardly sensitive to the rather large changes introduced to the signal Y.

Table 4 and Fig. 5 show the results of the investigations obtained for the method of alternative sequences. These results correspond to the mean voltage levels of the masking signal (dB re 1 V), for which the difference between the signals X and Y was masked. Table 4 and Fig. 5 compare results for 4 listeners and 8 frequency responses of channel II; Table 4 also gives the corresponding values of the standard deviation (dB).

On the basis of the mean values obtained, analogously to the pulsation method, the investigation results were verified statistically. Table 5 shows the values of the t test obtained.

The results obtained ($t \geqslant t_a$) for low-pass systems (a, b, c, d) indicate that the value of the voltage of the signal M required to mask the differences

Table 4. The results for the method of alternative sequences. The mean values of the voltage level of the masking signal and the values of the corresponding standard deviations

Listener		Cut-off frequency of channel II f_g [kHz] 7 kHz *a*	8.9 kHz *b*	11.2 kHz *c*	14.1 kHz *d*	f_d [Hz] 179 Hz *e*	224 Hz *f*	282 Hz *g*	355 Hz *h*
1	*Y* [dB]	−16.8	−18.6	−20.9	−24.7	−23.7	−23.2	−20.6	−18.9
	σ [dB]	0.29	0.58	0.64	0.28	0.60	0.52	0.54	0.71
2	*Y* [dB]	−15.1	−13.8	−15.3	−19.6	−24.4	−22.6	−18.6	−15.8
	σ [dB]	0.29	0.19	0.29	0.96	1.15	0.27	0.49	0.51
3	*Y* [dB]	−13.2	−16.7	−21.6	−26.1	−26.9	−22.5	−18.1	−14.3
	σ [dB]	0.08	0.36	0.41	0.57	0.15	0.67	0.65	0.33
4	*Y* [dB]	−15.4	−19.7	−21.2	−25.5	−18.0	−15.3	−13.8	−13.1
	σ [dB]	0.60	0.47	0.36	0.65	1.10	0.41	0.66	0.42

f_g — upper cut-off frequency, f_d — lower cut-off frequency

between the signals *X* and *Y* are a significant measure of similarity of these signals. A 1/3 octave change in the upper cut-off frequency gave statistically significant different levels of the masking signal. The greater similarity between the signals *X* and *Y* the lower level of the signal *M* was required to mask the difference between the signals *X* and *Y*.

For high-pass systems (e, f, g, h), in most cases (except two) values greater than t_a were also obtained for the *t* statistics. Thus, the level of the signal *M* can also in this case constitute the objectivized measure of similarity of the signals *X* and *Y*.

6. Conclusions

1. On the basis of experiments, two measurement procedures characterized by the highest sensitivity of differentiating the evaluated objects were chosen of 11 procedures. For the method of alternative sequences II (variants *G*, *H* and *J*), the best results in the preliminary investigations were obtained

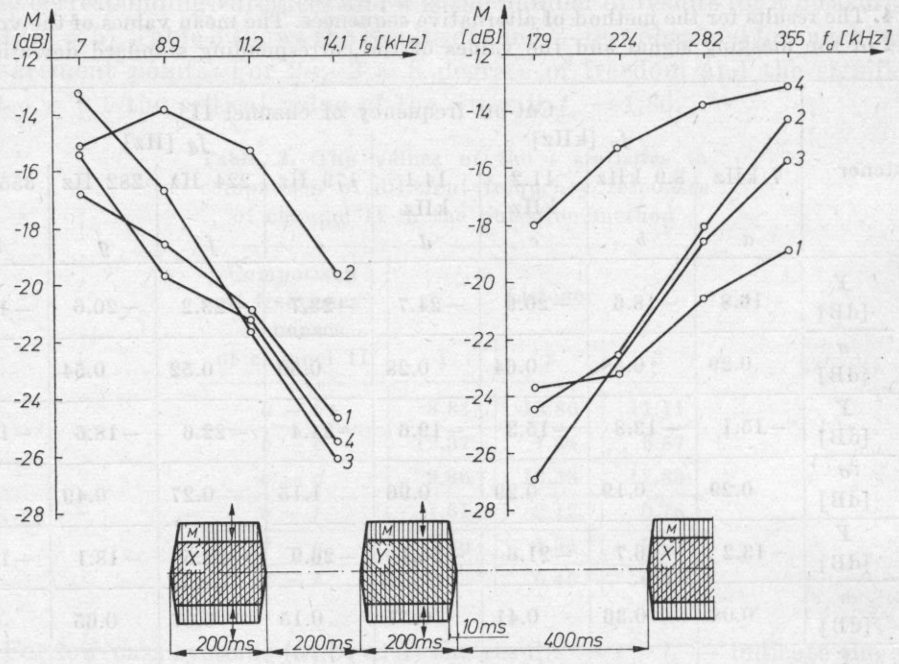

Fig. 5. The results of the method of alternative sequences. The mean values of the level of the signal M as a function of the cut-off frequencies of the frequency response of channel II obtained for 4 listeners

Table 5. The values of the t statistics in comparison of different frequency responses of channel II in the method of alternative sequences

Comparison of frequency responses of channel II	Listener 1	2	3	4
$a-b$	5.69	10.95	22.14	9.74
$b-c$	5.15	9.25	17.28	4.98
$c-d$	8.88	9.62	12.83	12.27
$e-f$	1.33	3.07	9.77	4.52
$f-g$	6.49	11.97	8.54	3.63
$g-h$	3.47	7.72	11.10	1.54

using uniformly masking noise as the masking noise. The use of pink or white noise decreased the sensitivity of differentiating objects.

2. In the case of the pulsation method the characteristics of changes in the pulsation threshold as a function of the cut-off frequency of channel II are very close to one another for all the three listeners. The differences between the values of the pulsation threshold obtained for particular listeners do not exceed 1.1 dB.

3. In the case of the pulsation method a 1/3 octave change in the upper cut-off frequency causes an essential change in the value of the sensation threshold. This method can thus serve to differentiate the frequency responses of the electroacoustic channel over the high frequency range. The determination of the sensitivity of the method requires further investigations.

4. Fig. 6 shows the mean values of the pulsation threshold calculated on the basis of the results for 3 listeners. The response of the pulsation threshold obtained as a function of frequency was approximated by linear regression

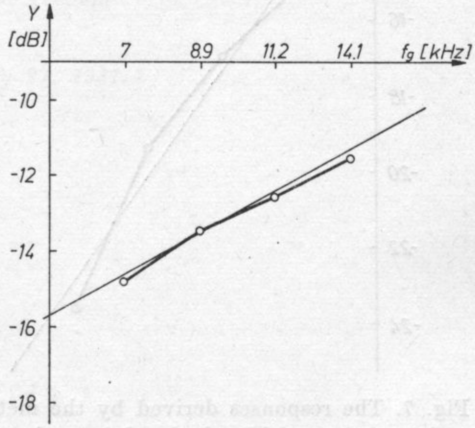

Fig. 6. The response of the pulsation threshold as a function of the cut-off frequency f_g of channel II, calculated from the results for 3 listeners (thick line) and its linear approximation (thin line)

(with the correlation coefficient $r = 0.996$). The directional coefficient of the straight line is 1.081. The approximation shown in Fig. 6 is based on the results obtained for 4 values of f_g of channel II. However, the determination of the frequency response of the pulsation threshold requires further investigation with denser measurement points (with a change in t_g by values less than 1/3 octave).

5. The values of the pulsation threshold obtained with adjustment of the lower cut-off frequency indicate a low differentiating sensitivity. The results obtained are characterized by very high repeatability (low variance), which in turn indicates the sharp character of the occurrence of the pulsation threshold.

6. In the case of alternative sequences III a 1/3 octave change in the cut-off frequency response of channel II causes an essential change in the value of the signal M required to mask the differences between the signals X and Y. The responses obtained for particular listeners (Fig. 5) are however, much more differentiated than those obtained using the pulsation method. It follows from Fig. 5 that with upper limiting of the transmission band the curves derived for listeners 1, 3 and 4 are close. In turn, the curve for listener 2, which is different from them, seems to indicate his sharper differentiation of timbre changes over the high frequency range. Analogously, with lower limiting of the signal, listener 4 shows a better than average differentiation of timbre.

7. Fig. 7 shows the mean responses obtained by the method of alternative sequences. With lower limiting of the band, linear regression (with the correlation coefficient $r = 0.999$) gives the directional coefficient of 2.181. With upper limiting, linear regression (with the correlation coefficient $r = 0.982$) gives

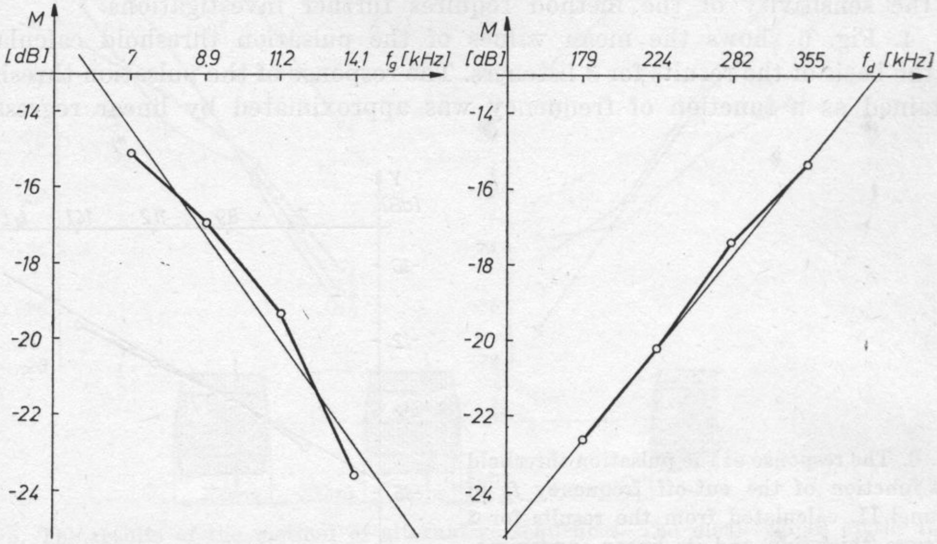

Fig. 7. The responses derived by the method of alternative sequences as a function of f_g and f_d of channel II, calculated from the results for 4 listeners (thick lines) and their linear approximation (thin lines)

the directional coefficient of − 2.802. The determination of the exact behaviour of the frequency responses obtained by the method of alternative sequences requires further investigation

8. The results obtained suggest that the method of alternative sequences assures better sensitivity of auditory differentiation of spectrum changes than the pulsation method does.

9. On the basis of the investigations carried out, it seems that the method of alternative sequences can be used to differentiate the frequency responses of electroacoustic devices.

References

[1] C. Bordene-Sacerdote, C. Modena, *Prove d'ascolto su sistemi di altoparlanti*, Elettronica a Telecomunicazioni, **6**, 1-20 (1971).

[2] B. W. Kulesza, *Method of auditory evaluation of radio sets in industrial conditions* (in Polish), Proc. XXII Open Seminar on Acoustics, Świeradów 1975.

[3] T. Łętowski, *Auditory evaluation of electroacoustic devices* (in Polish), Zesz. Nauk. PRiTV, 27 (1976).

[4] W. A. MUNSON, J. W. KARLIN, *Isopreference method for evaluating speech transmission circuits,* J. Acoust. Soc. Am., **34**, 762-774 (1962).

[5] L. H. NAKATANI, K. D. DUKES, *A sensitive test of speech communication quality,* J. Acoust. Soc. Am., **53**, 1083-1092 (1973).

[6] L. C. W. POLS, L. J. T. van der KAMP, R. PLOMP, *Perceptual and physical space of vowel sounds,* J. Acoust. Soc. Am., **46**, 458-467 (1969).

[7] D. L. RICHARDS, *Design and analysis of subjective acoustical experiments which involve a quantal response,* Acoustica, **2**, 83 (1952).

[8] E. M. ROTHAUSER, G. E. URBANEK, W. P. PACHL, *Isopreference method for speech evaluation,* J. Acoust. Soc. Am., **44**, 408-418 (1968).

[9] S. G. STEVENS, *Cross modality validation of subjective scales for loudness vibration and electric shock,* J. Exp. Psychol., **57**, 201-209 (1959).

Received on May 21, 1981.