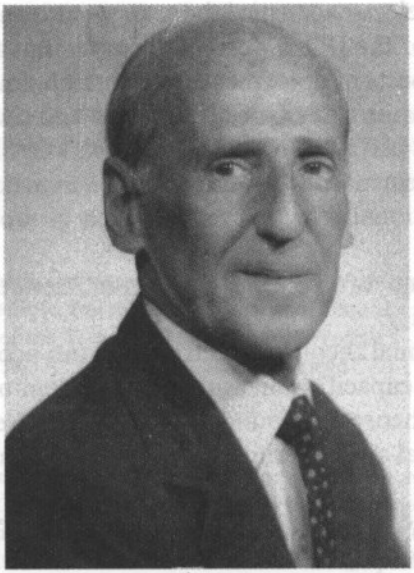
ARCHIVES OF ACOUSTICS
19, 1, 3–4 (1994)



PROFESSOR DR HAB. JANUSZ KACPROWSKI

Polish acousticians suffered a great loss on January 23, 1993, the date of death of Professor Janusz Kacprowski, who passed away after a lengthy and grave disease at the age of 76. The entire professional life of prof. Kacprowski was connected with acoustics: a student of the Electrical Department of the Warsaw Technological University before the war in 1939, he continued his education at the underground courses during the German occupation in Warsaw to obtain the M.Sc. degree right after the war in 1945.

He started his professional career in the Telecommunication Institute in Warsaw doing the technological and scientific research on electroacoustics; he made considerable contributions to the telephonometric speech measurement technique by preparing original telephone set quality standards and the "artificial ear" standard, henceforth generally accepted by the Polish industry. He was also the constructor of a talking clock. At the same time he was teaching electroacoustics at the Warsaw Technological University an adviced numerous students preparing their M.Sc.-dissertations. His two books on electroacoustics werefor several years the fundamental handbooks for students specializing in acoustics. In that period prof. Kacprowski defended his Ph.D-thesis in the field of electroacoustics and was nominated a "docent" after the usual habilitation procedure.

In 1952 moved to the Polish Academy of Sciences, first to the Vibration Research Department and then, at the beginning of 1953, he became one of the organizers of the newly created Institute of Fundamental Technological Research of the P.A.S. Under his guidance, the electroacoustic laboratory was transformed to the Cybernetic Acoustics Department of the IFTR dealing mainly with various problems of speech acoustics and certain aspects of noise measurement techniques. Remarkable results of his research brought him the title of a full professor and several scientific awards and state decorations. In the period of more than 30 years of his activity in the Institute, concluding with his retirement in 1987, he succeeded in making his department one of the leading and internationally recognized research groups in the field of acoustics and electroacoustics.

Let us mention some of the most important results of prof. Kacprowski's extensive scientific activity:

- Original four-terminal theory of electro-mechano-acoustics transducers serving as a basis for improved capacitive measurement microphons;
- Extension of the theoretical foundations of the noise level measurements;
- Complex theoretical and experimental investigations on the speech analysis and synthesis aimed at

a) Establishing of the acoustic structure of Polish speech,
b) Optimization of wave-guide transmission lines;
c) Development of the man-machine communication systems;
d) Digital coding of the speech signal.

Prof. Kacprowski's results concerning the speech analysis and synthesis were also used in medicine for improving the methods of phonetic diagnosis and rehabilitation methods.

Prof. Kacprowski was also active in the field of organization of Polish science: he was one of the founders of the Polish Acoustical Society also Archives of Acoustics, thus contributing to the general development of Polish acoustics. We are sure that a large number of his students and co-workers will continue and develop his original ideas and concepts.

His death is a particularly painful event for all of us — the large group of people who used his friendly and valuable advice and collaborated with him in the Polish Academy of Sciences during the last four decades.

Ignacy Malecki

ARCHIVES OF ACOUSTICS
19, 1, 5–35 (1994)

SPEECH SIGNAL TRANSMISSION RATE COMPRESSION USING OF THE TIME PARAMETERS CODING METHOD

CZ. BASZTURA

Acoustic signals analysis and processing division
Institute Telecommunications and Acoustics
of the Technical University of Wrocław
(50-370 Wrocław, ul. Janiszewskiego 7/9)

A new method of speech signal encoding and decoding is the subject of the presented research project. This method may find applications in the field of telecommunications (for telephone or radio transmission) as well as in the field of speech synthesis. The concept of the method is based on the extraction and transmission of such time parameters as the intervals between the subsequent speech signal zero-crossings and the amplitude of the signal in these intervals as well as on the subsequent speech signal's reconstruction based on the given parameters and knowledge derived from the analysis of speech. The system is composed of two main processes: the extraction and encoding of the transmission parameters and the original speech signal reconstruction (resynthesis). The method makes it possible ot decrease the transmission rate about 10 times, as compared to the original speech signal. The results prove that the synthesized speech signal quality, when the new method is used, may be better than the one obtained by the use of other vocoder methods. The method's low cost and a relatively simple hardware system developed for the parameters' extraction and for the reconstruction of the speech signal are the most important advantages of the described method.

1. Introduction

1.1. Presentation

The switching to digital processing, encoding and transmission circuits and systems began as early as in the beginning of the sixties when the first system of the encoded information transmission was presented. The information was passed with the decreased error rate as well as the transmission rate.

The interest and demand of efficient coding methods and speech signal transmission rate compression as well as the transmission quality estimation grew along with the progress and development of the new speech signal coding for teletransmission purposes [1, 2, 3, 5, 6, 7, 11, 12, 13, 18, 29]. As compared to the analog information transmission, the digital method has a number of

advantages including the possibilities of detecting and correcting errors and also better distortion protection.

Digital speech signal transmission systems are divided into two categories:

1) Waveform encoding systems,
2) Particular speech signal parameters encoding systems.

The latter are called vocoders or source coders [23].

1.2. The concept of the speech signal transmission rate compression method based on time parameters coding

The two categories of coding described in Section 1.1. were applied in digital teletransmission systemes. Both methods are included in the international recommendation CCITT, the pulsecode modulation (PCM) at the bit rate of 64 Kbit/s as well as the adapted differential PCM at the rate of 32 Kbit/s.

Intensive research is being carried out in order to develop relatively little complicated vocoder systems which have a transmission rate less than 16 Kbit/s. At the same time the aim is to preserve the processed signal's quality as similar to the one achieved by using PCM or ADPCM methods. The solution to this problem is being searched in the parametric encoding method. The parametric encoding method bound with the speech signal information volume (transmission rate) compression consists in isolating from speech signal certain parameters that describe its amplitude variations and the component frequencies of the spectrum as well as the transmission of these parameters actual values to the receiver by a telecommunication channel instead of the original speech signal. The parameters isolated from the speech signal usually form slow changing courses; hence for their transmission a channel of a much lower bit rate may be used than the one which transmits the original signal. This is the method that is used in encoding, transmission and decoding in vocoder technology [1, 2, 8, 9, 25].

The known methods of parametric coding, i.e., speech signal transmission rate compression based on harmonic spectral-band formant and linear predicted coding methods, make it possible to achieve a speech signal transmission rate of the range from 4.8 Kbit/s to 1 Kbit/s. The concept of a new speech signal transmission rate compression is based on the zero crossing analysis (ZCA). There are two main premises for this concept:

1) Theoretical, described, among others, in the publication by Logan [15] and Basztura [4] as well as in publications by other authors concerning the possibility of reconstructing (with a certain inaccuracy) a limited bandwidth signal on the basis of knowledge about zero crossings,
2) Practical, resulting from the fact that the parameters based on ZCA and describing linguistic differences of linguistic units as well as their particular features in automatic voice and speech recognition processs are the parameters of similar discrimination power to other — more complex parameters as spectral and predictive ones [4, 10].

The substance of the time vocoder's concept, i.e., speech signal transmission rate compression based on time parameters encoding, is the asynchronical isolation of such basic parameters as lengths of subsequent intervals between the zero-crossings I_n and amplitudes of courses in these intervals A_n form the digital signal. These two parameters are transmitted by a telecommunication link to a receiver where the signal is resynthesized (reconstructed) basing on the received sequences $\{I_n, A_n\}$. They may be additionally quantized, what makes it possible to diminish the information volume. These data are supplemented by the information obtained during preliminary research and analyses residing in the resynthesis circuit's memory. This allows to recreate the input (to the receiver) speech signal U_n as the signal U'_n corresponding very well with the input signal. The method efficiency estimation criterion is, above all, the relation between signal transmission rate compression rate and the output signal quality estimation U'_n. The resynthesized signal estimation comprises a subjective estimation (including intelligibility, the transmission of individual features and the quality of sound) and objective estimation [2, 11, 12, 13, 28].

The planned research had the following purposes:

1) the statistical (for stationary segments of speech) and deterministic (for each of the Polish speech phonemes) data elaboration regarding the values of: time intervals resolution between zero-crossings of the speech signal I_n, the values and probabilities of occurrence of A_n amplitudes in these intervals,
2) the determination of the possibility of scalar or vector quantization for $\{A_n, I_n\}$ allowing further transmission rate compression,
3) the design of the digital simulation model of a vocoder,
4) carrying out research concerning subjective distinctness and intelligibility in order to set a selection of optimal parameters and the estimation of the vocoder's quality,
5) carrying out tests on objective compatibilities of the input signal U_n and the resynthesized signal U'_n,
6) determining the vocoder's primary parameters and presenting the concept of the method's implementation possibilities.

2. Methods and systems of digital speech signal encoding

2.1. Presentation

All wave from encoding devices and digital vocoders have one feature in common, i.e., analog-to-digital and digital-to-analog processing application. It has to be admitted that all imperfections introduced by these converters may reduce significantly the benefits of digital speech signal processing. The importance of selecting the converters' properties is motivated by a few factors:

- too high quality of the converter increases costs and may result in useless engaging of the computing capacity during further steps of signal processing,

- the measured decrease (increase) of quality (objective) might to correspond with the decrease (increase) of the individually perceived quality.

2.2. *Waveform encoding*

DM and PCM are the most common methods of waveform encoding. Pulse-code modulation is based on the fact that each sample is transmitted by means of the code created form the group of impulses (Fig. 1.). The advantage of this encoding method is a high distortion resistance rate, its disadvantage, though, is a relatively big information volume of the encoded form of the speech signal. The sample frequency $f_{pr} \geqslant 2\, f_g$, where f_g — maximum transmission frequency. Usually $f_g = 4$ kHz is assumed according to the telephone transmission bandwidth.

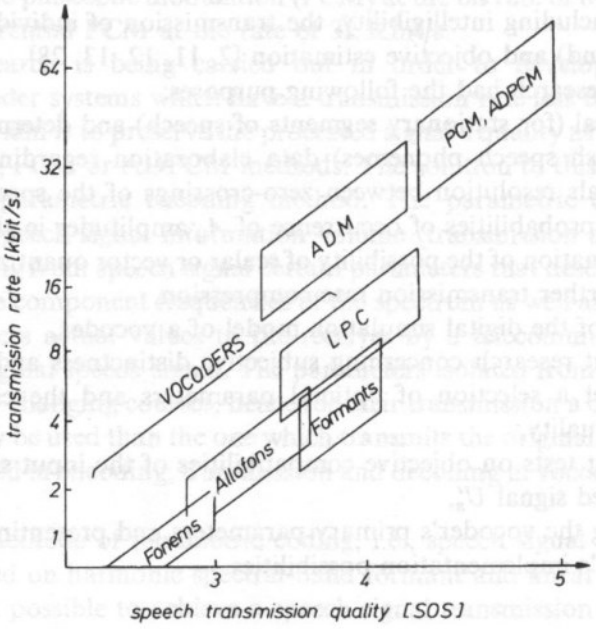


Fig. 1. The different parameters of speech coding methods.

The Delta Modulation (DM) is a way of encoding where the difference of levels between two samples is expressed digitally. The main advantage of this way of processing is an extremely simple circuit realisation but PCM converters do not have to have logical ciruits generating subsequent approximations, amplifying or sampled-data-memorizing filters, introductory amplifiers and other high precision elements [20].

In differential pulse-code modulation (DPCM) the signal increments during Δt intervals between samples as well as their signs are given. In adapted delta modulation the increment's value increases or decreases depending on the signal's changes.

Table 1. The parameters of the selected digital speech signal encoding systems

Speech signal encoding method	Transmission rate [Kbit/.s]	Signal/noise ratio [dB]
Waveform encoding		
Linear pulse-code modulation (PCM)	128	43
Logarithmic pulse-code modulation	56	29
Linear-logarithmic pulse-code modulation	32	29
Linear Delta Modulation (DM)	56	21
Adapted Delta modulation (ADM)	24–16	18
Parametric encoding		
Encoding method	Transmission rate [Kbit/s]	word intelligibility [%]
Spectral-band encoding	4.8–4.2	92
Orthogonal (harmonic) encoding	2.4	86
formant encoding	1.2	80
predicted encoding	4.8–1.2	88

2.3. Parametric coding — vocoders

Parametric encoding is bound with speech signal information companding. This consists in isolating from the speech signal certain parameters describing its amplitude changes as well as the spectrum component frequencies and the transmission of their instantaneous values (in an analog or digital way) by a telecommunication channel instead of a proper speech signal. The parameters isolated from the signal form slow-changing courses; therefore, for their transmission a channel of information volume less than for the proper speech signal transmission may be used. The speech signal at the receiver is subject to reconstruction to the shape approximate to the primary one, therefore it may be assumed that its volume was subject to compression and then — to expansion. At present the following methods of speech signal parametric encoding are known: band spectral (or: band-channel) and harmonic encoding, formant predictive encoding and a few mixed methods developed from them like PARCOR [14, 21]. All types of vocoders mentioned above are characterized by the following disadvantages limiting so far their range of application.

These are:

a) the complexity of encoding and decoding procedures which causes the high cost of emitting and receiving devices,
b) usually non-natural quality of sound,
c) lack of transmission of the individual voice features which results in the fact that the collocutor cannot be recognized by his voice.

Figure 1 shows the dependence of the transmission rate and the speech transmission quality expressed in the listeners' scale of five grades (SOS) from the encoding methods [23]. At present the most developed and preferred vocoders are those based on linear predictive coding. As compared to others, they have many advantages, nevertheless they require special complex processors LPC working in the real time. In spite of the differences, all of the parametric coding methods show many common features. For example, all of them are based on the phonetic-acoustic speech signal microstructure analysis and the isolation of those parameters which describe univocally this microstructure.

The time vocoder has the advantages of waveform encoding methods such as PCM (pulse-code modulation), ADPCM (adapted pulse-code modulation), DM (delta modulation) or ADDM (adapted delta modulation) and at the same time it can ensure a signal information volume compression comparable to those of LPC vocoders. On the ground of the time structured information transmission, the individual pronounciation character is preserved, which is an important element in human intercommunication.

3. Transmission rate compression based on the analysis of time parameters

3.1. Presentation

The concept of the new parametric encoding method, outlined in Section 1.2. of this publication can, be illustrated schematically as it is in the Fig. 2 where the block

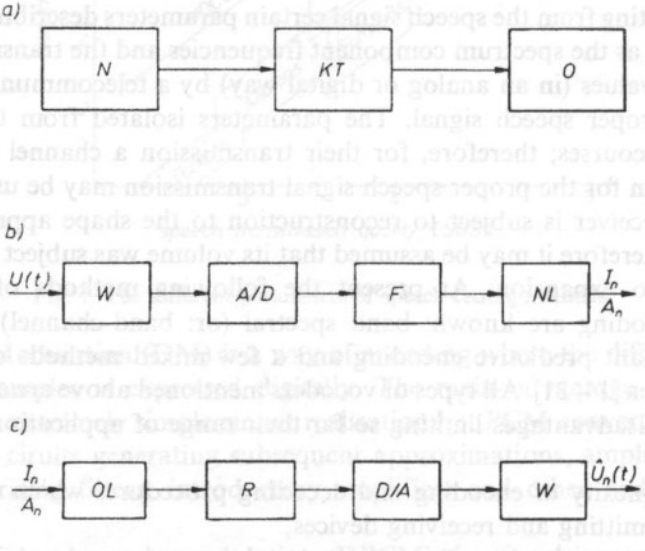


Fig. 2. The block scheme illustrating the conception of the speech signal transmission rate compression method based on coding of the time parameters. *W* — amplifier, *F* — filter, *NL* — line driver, *OL* — line receiver, *R* — resynthesis.

scheme of the vocoder system based on the analysis of time parameters shown. The system encloses the transmitter unit (N) and the receiver (O) unit linked by the digital telecommunication line (KT).

The isolation of the numerical values' sequence of two quantities, i.e., the time intervals' length between the subsequent speech signal's zero-crossings (I_n) and the amplitudes in these intervals (A_n) from the time course is made in the transmitter unit (the analyzer). The pair of these numbers: the interval (I_n) and the amplitude (A_n) are transmitted to the receiver (the synthesizer) where to each interval (I_n) and amplitude (A_n) the course of the actual values is assigned (Fig. 3). The information of the course's shape comes from the signal's statistical and deterministic analyses of Polish speech.

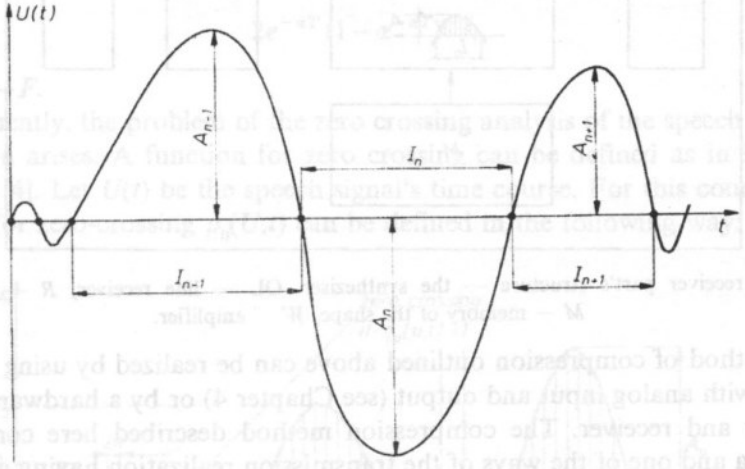


Fig. 3. The basic speech signal's time parameters I_n, A_n.

3.1.1. The transmitter unit — the analyzer

The analyzer (Fig. 4) comprises: an amplifier, analog-to-digital converter with sampling parameters set to $f_{pr} = 10$ kHz and dynamics D (D described on 10 – 12 bits per sample). These are the minimal typical processing parameters A/D for the speech signal as well as for the time intervals extractor, the amplitude and the line transmitter.

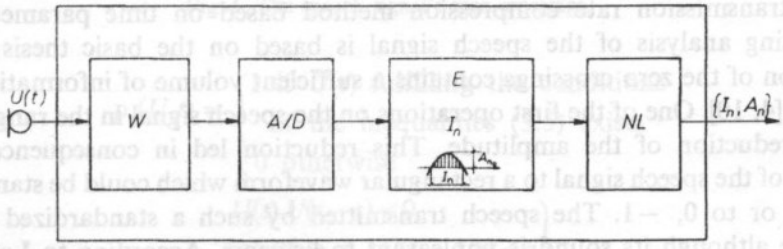


Fig. 4. The transmitter part's structure — the analyzer. W — amplifier, E — extractor, NL — line driver.

3.1.2. The receiver unit — the synthesizer

The synthesizer (Fig. 5) comprises: the line receiver, the resynthesizer system, the actual courses shapes' memory (or the processor computing current values of $U_n(t)$ — the rule of synthesis) as well as the D/A converter the parameters of which are compatible with the A/D converter in the analyzer plus the output system.

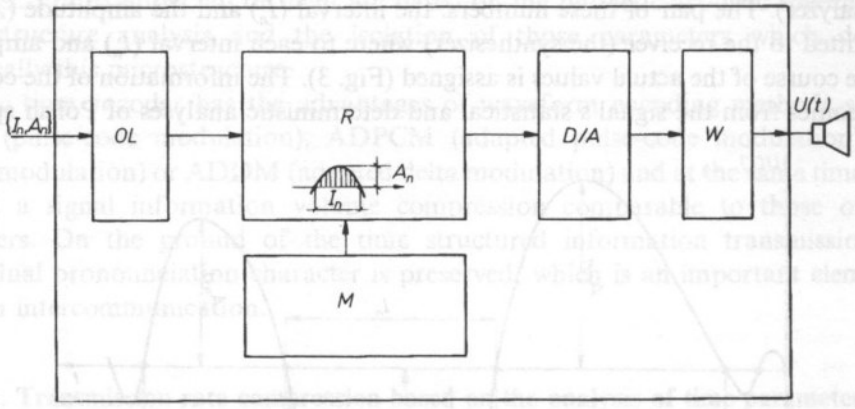


Fig. 5. The receiver part's structure — the syntheziser. OL — line receiver, R — resynthesizer, M — memory of the shape, W — amplifier.

The method of compression outlined above can be realized by using a universal computer with analog input and output (see Chapter 4) or by a hardware system of transmitter and receiver. The compression method described here comprises the general idea and one of the ways of the transmission realization having a significant transmission rate compression. There are some possible modifications to this method that should be examined, e.g., the possibility of transmitting the interval's values I_n alone from the transmitter to the receiver. The A_n values would then be assigned automatically in the receiver on the basis of previously gained data and rules obtained from statistical and deterministic analyses (the correlation $A_n = f(I_n)$ etc.).

3.2. Zero-crossing analysis

The premises resulting from zero-crossing analysis form the theoretical base for the new transmission rate compression method based on time parameters. The zero-crossing analysis of the speech signal is based on the basic thesis that the distribution of the zero crossings contains a sufficient volume of information about the signal [4, 10]. One of the first operations on the speech signal in the range of time was the reduction of the amplitude. This reduction led in consequence, to the reduction of the speech signal to a rectangular waveform which could be standardized to $+1$, 0 or to 0, -1. The speech transmitted by such a standardized signal is intelligible although its sound is unpleasant to listeners. According to LOGAN [15], a signal

$$U(t)=z(t)+\cos Ft, \tag{3.1}$$

where $z(t)$ is a bandlimited signal $(-f_g, f)$ and $0<f<F<\infty$ and at the same time it fulfils the conditions of the inequality (3.2), can be reproduced if

$$(-1)^j U(kT/F)>0, \tag{3.2}$$

and

$$|z(t)|<1.$$

The zeros within the time interval $t-T$ and $t+T$ make it possible to reproduce $U(t)$ with a relative error less than

$$2e^{-\alpha T}(1-\alpha^{-\alpha T})^{-2}, \tag{3.3}$$

where $\lambda=f-F$.

Consequently, the problem of the zero crossing analysis of the speech signal and of extraction arises. A function for zero crossing can be defined as in BASZTURA's publication [4]. Let $U(t)$ be the speech signal's time course. For this course (Fig. 6) a function for zero-crossing $\rho_0(U,t)$ can be defined in the following way:

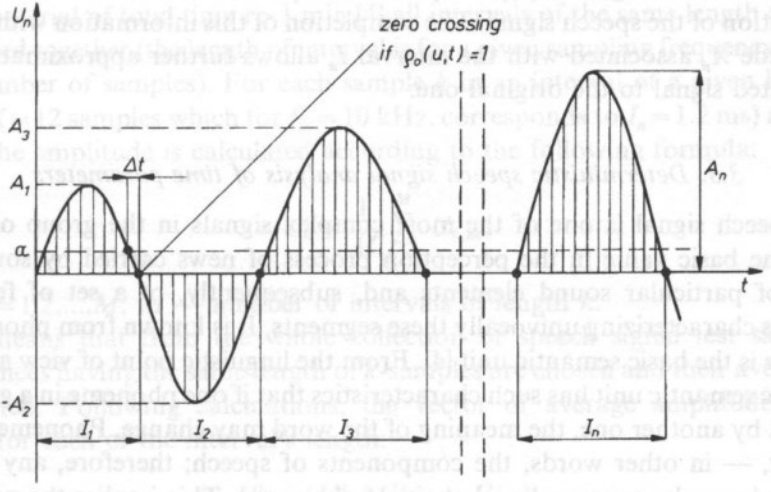


Fig. 6. The speech signal's zero crossings.

$$\rho_0(U,t)=\begin{cases}1 \text{ if } U(t) \text{ fulfilling the conditions}\\ \quad \text{of the inequalities (3.5) exists}\\ 0 \text{ otherwise}\end{cases} \tag{3.4}$$

$$\left.\begin{array}{l}U(t)\,U(t-\tau)<0\\ |U(t)|\geqslant\alpha,\ |U(t-\Delta t)|<\alpha\\ U(t)<\alpha, \text{ for } t<t<t+\Delta t\end{array}\right\} \tag{3.5}$$

where:

α — describes a certain threshold value ($\alpha < 0$) protecting from occurrence of any additional zero-crossing caused by distortions, Δt — the displacement of a zero-crossing position resulting from the fact that $\alpha \neq 0$.

Zero-crossing analysis makes it possible to isolate a few parameters characterizing the speech signal with defined precision. These are, among others, the current zero crossing density $\rho_0(t)$ and the average zero-crossing density within the segment T:

$$\rho_0 = \frac{I}{T}, \tag{3.6}$$

where: I — number of points within the segment T with $\rho_0(U,t) = 1$ or the time intervals distribution between subsequent zero crossings $R(t)$ [4].

$$R(t) = \sum_{j=1}^{I} \delta(t - t_j), \tag{3.7}$$

where: $\delta(t)$ — delta function $j = 1,2,...,I$ — the point on the axis of time with $\rho_0(U,t) = 1$.

The time intervals (in practice they are surveyed as a number of samples of the discrete form of the signal $U(n)$) carry the most part of information that allow the reconstruction of the speech signal. The completion of this information with the value of amplitude A_n associated with the interval I_n allows further approximation of the reconstructed signal to the original one.

3.3. Deterministic speech signal analysis of time parameters

The speech signal is one of the most complex signals in the group of acoustic signals. The basic value in the perception process of news carried by sound is the function of particular sound elements and, subsequently, of a set of features or parameters characterizing univocally these segments. It is known from phonetics that a phoneme is the basic semantic unit [4]. From the linguistic point of view a phoneme as the basic semantic unit has such characteristics that if one phoneme in a given word is replaced by another one, the meaning of the word may change. Phonemes are "the atoms" or, — in other words, the components of speech; therefore, any utterance pronounced may be expressed as a chain of phonemes. This implies the necessity of examining and analysing these individual components of Polish speech. It has to be observed that in spite of the fact that phonemes make up 37 classes in Polish speech, they show a significant differentiation within each of those classes. The reasons for that are the following:

a) the differences between the class subjects,

b) the differences inside the subjects,

c) the context differentiation (phonemes' context representations, i.e., allophones).

These differences constitute a significant problem when creating a relatively simple model of signal parametrization not only in vocoder technology but also in

broadly viewed research in the range of speech analysis and recognition [4]. The results of observation quasi-stationary causes of phonemes make it possible to assume the elementary dependences $A_n = f(I_n)$ (possibly corrected on the basis of statistical research). They also allow for both the choice of shapes of time envelopes of the amplitudes for the intervals U'_n and the introductory quantization of the values A_n, I_n (and possibly the shapes) in order to minimize the transmission rate.

3.4. Statistical analysis of the signal for time parameters

In order to settle the best coding parameters I_n and A_n as well as the synthesiser's analytical data, a statistical analysis of time intervals between subsequent zero-crossings I_n was carried out. The dependences between the amplitudes A_n and the length of interval I_n were settled expressing the probability function $P(A_n)$ for the given interval I_n. The average course of U_n was estimated (the amplitude's time envelope) for the n^{th} interval along with the standard deviation for each particular time sample. The data obtained from a statistically representative sound material sample for the Polish speech made it possible to settle the most probable amplitudes and provided information on the distributions of the shape of statistical envelopes (standard deviations for samples). The analysis of statistical dependences runs according to the following algorithm: for a sound material of total time $t > 1$ min [4] all intervals of the same length are chosen and grouped together (the length of interval is for a given sampling frequency measured by the number of samples). For each sample k in an interval of a given length (for example $K = 12$ samples which for $f_{pr} = 10$ kHz, corresponds to $I_n = 1.2$ ms) an average value of the amplitude is calculated according to the following formula:

$$U'_k = \frac{1}{M} \sum_{m=1}^{M} U_{k,m}, \tag{3.8}$$

where: $m = 1,2,...,M$; M — number of intervals of length k.

This means that from the whole collection of speech signal test samples the sub-sequences having the same length of k-samples are chosen and their average value is calculated. Following calculations, the vector of average amplitude values is obtained for each of the interval's length:

$$U'_k = \{U'_{k,1}, U'_{k,2}, \ldots, U'_{k,K}\}, \tag{3.9}$$

where:

$$U'_{k,K} = \frac{1}{M} \sum_{m=1}^{M} U_{k,m}. \tag{3.10}$$

The standard deviation indicates the statistical dispersion of the values $U_{k,K}$ resulting from the differences mentioned above:

$$\delta_{k,K} = \frac{1}{M-1} \left[\sum_{m=1}^{M} U_{k,K} - M U'^{2}_{k} \right]. \tag{3.11}$$

3.6. The speech signal's resynthesis

The speech signal's resynthesis realized in a contractual "receiver" of the time vocoder (see Fig. 5) runs simultaneously with the sequel of samples from a generator set to the analyzer's sampling frequency (f_{pr}). The line receiver with a memory buffer with compensates the asynchroneity resulting from the fact that the portions of data $\{I_n, A_n\}$ or the number indicating the vector representing the pair $\{I_n, A_n\}$ measured in the analyzer is transmitted in constant time distances while their usage in time (either the numbers or the vector) depends on the value of I_n (the interval's length). This makes it necessary to store the data (A_n, I_n) and allow for a slight delay of transmissoin, which should not exceed 100 ms and should not have any influence on the quality of transmission. The synthesiser's memory contains the information, the envelope shape calculation program which would also make up these values alternating their sign as well as the reconstructed signal's output through the channel of the analog-to-digital converter and the output system. Depending on the accepted option of resynthesis, the following data may be present in the memory:

a) the data concerning the most probable values of amplitudes (in this case it is possible to transmit only the values of I_n),

b) the shape or a few settled shapes of U_n related to the ranges of length of the intervals I_n.

4. Estimation of the parameters of the time vocoder

4.1. Presentation

The concept of asynchronic coding of time parameters and their analysis became the base for the functional design of the model of the functional vocoder based on an all-purpose computer hardware with acoustic input and output and specially designed software realizing not only the estimation of the time parameters but also the time vocoder's simulation of operation. Therefore the universal option was accepted; this allowed for the realization of the analysis research program and for the time vocoder's simulation of operation along with the elements of its testing and transmission quality estimation.

4.2. Research hardware

For research purposes computer hardware was utilized making it possible, owing to the software, to settle the data for analysis and the parameters of the time vocoder. The indispensable completion to the hardware were the following (Fig. 7):

a) an analog acoustic input/output system (a microphone, amplifiers, low pass filters, a loudspeaker) ordered at ZAIPSA ITA, the workshop of the Technical University of Wrocław,

b) analog-to-digital and digital-to-analog converters' card CONVERT of the TAD-01 type along with the software for the preliminary working registration and for the speech signal's output,

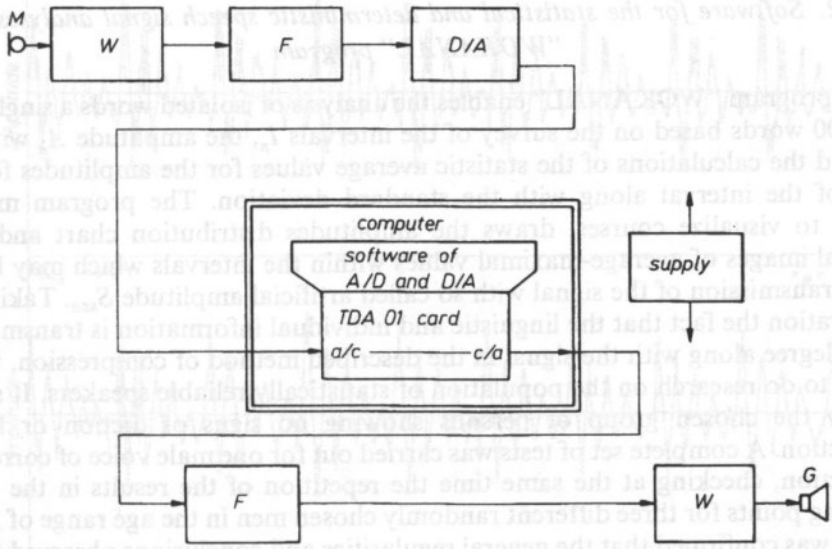


Fig. 7. The block scheme of the acoustic input and output. W — amplifier, F — filter.

c) the software for the algorithm of the analysis "WOKANAL",
d) the software realizing functions of the time vocoder "WOKCZAS".

4.3. Acoustic input/output system

The acoustic input/output system designed and developed as a part of the research program along with the A/D and D/A card is an indispensable linker between the analog and digital speech signal's structure. A block scheme of input and output systems cooperating with the A/D and D/A converter's card type TAD-01 placed on the computer's main card is shown in Fig. 7. The input system is composed of a directional microphone (M), an input amplifier, a low pass filter (F), an amplitude limiter, a registration measuring machine with an indicator. The output system is composed of a low pass filter, a tape recording amplifier, an output power amplifier with a loudspeaker. Since the sampling frequency $f_{pr} = 10$ kHz was assumed a priori, the analyzed speech signal baseband practically does not exceed 4.0 kHz. The developed systems ensure the speech signal's transmission without any distortion wsithin the frequency band limited by the low pass filter (F) from the microphone to the A/D cord and from the A/D to the loudspeaker.

4.3.1. The A/D *and* D/A *converter's card and its software*

The analog-to-digital A/D and digital-to-analog D/A converters card cooperates with both: the acoustic input and output system as well as — owing to the library functions for the Turbo Pascal compiler — the computer's compiler.

4.3.2. Software for the statistical and deterministic speech signal analyses — the "WOKANAL" program

The program "WOKANAL" enables the analysis of isolated words a single list of up to 100 words based on the survey of the intervals I_n, the amplitude A_n with their signs and the calculations of the statistic average values for the amplitudes for each length of the interval along with the standard deviation. The program makes it possible to visualize courses, draws the amplitudes distribution chart and traces statistical images of average-maximal values within the intervals which may be used for the transmission of the signal with so called artificial amplitude S_{sza}. Taking into consideration the fact that the linguistic and individual information is transmitted in a great degree along with the signal in the described method of compression, there is no need to do research on the population of statistically reliable speakers. It suffices to study the chosen group of persons showing no signs of diction or hearing imperfection. A complete set of tests was carried out for one male voice of correct and clear diction, checking at the same time the repetition of the results in the chosen examining points for three different randomly chosen men in the age range of 25–45 years. It was confirmed that the general regularities and conclusions observed for one voice may be the transmission only of the intervals I_n, i.e., the signal with the artificial amplitude. It must be considered that the statistics which allow to determine artificial amplitudes are the closest to the speaker whose utterances were analized. For other speakers this part of the observations will not be reliable.

4.3.3. Software for the process of the speech signal's analysis and resynthesis — the "WOKCZAS" program

The vocoder's software (the "WOKCZAS" program) comprises signal analysis in time, extraction of the intervals I_n and of the amplitudes A_n, coding, recording, reading and the signal's resynthesis. In the speech signal analysis part (the emitter), the intervals I_n are isolated and their length is stored. The values and signs of the amplitudes A_n for subsequent intervals are measured analogically. The program for analysis also comprises the option of extraction of the interval I_n alone (This was marked as the operation with the artificial amplitude S_{sza}). In order to ensure the same phase for the "artificial amplitude" signal and the original signal as well as to assign an amplitude with the correct sign to a given interval the appropriate sign is encoded using the same bits which are used to store the interval's length. The program encloses many options that make it possible to examine the signals in different scales, to compare the time courses, the signal's performance through the loudspeaker; the signals may also be written onto a disk or a tape recorder.

4.4. Deterministic dependences — length of time intervals' between zero-crossings and amplitude

The program "WOKANAL" makes it possible to estimate the deterministic parameters. The research was carried out on the experimental material of 37 Polish speech phonemes pronounced in the following way:

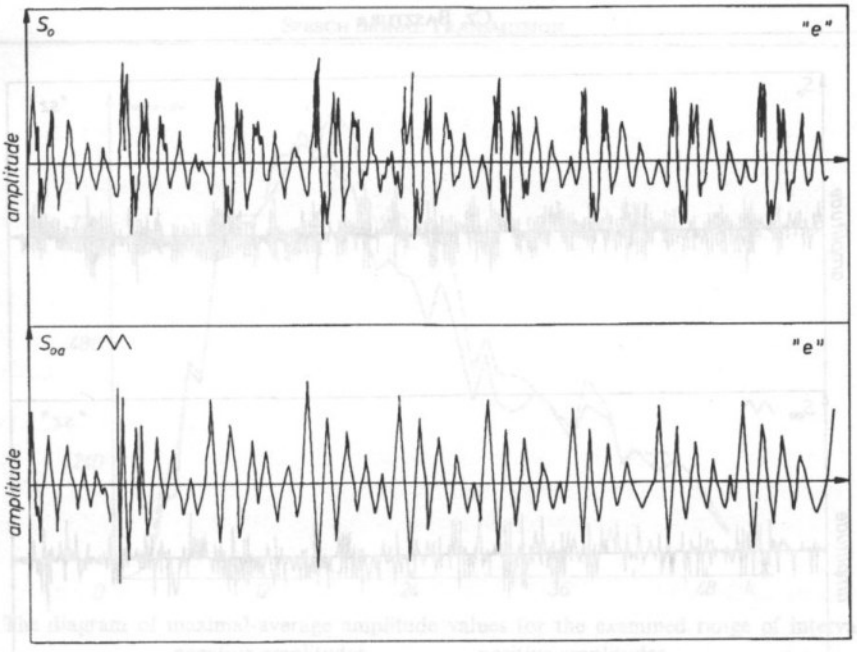


Fig. 8. The original (top) and the reconstructed (bottom) of the "e" wovel — triangular function approximation.

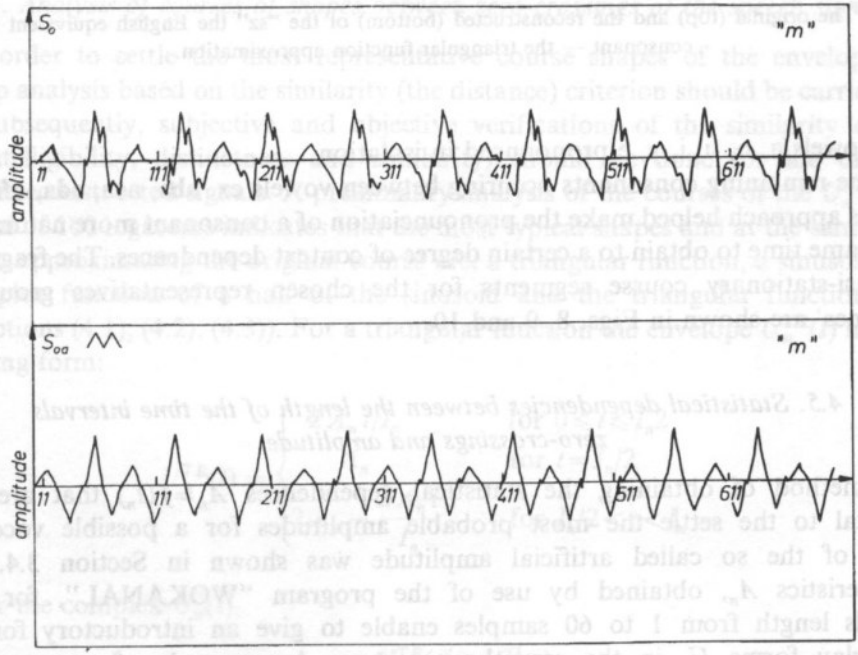


Fig. 9. The original (top) and the reconstructed (bottom) of the "m" consonant — triangular function approximation.

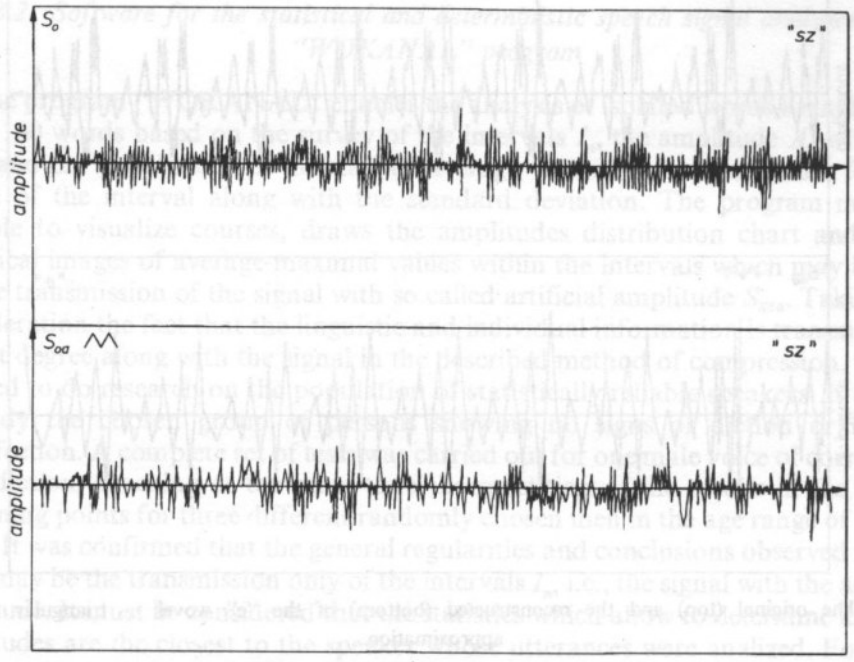


Fig. 10. The original (top) and the reconstructed (bottom) of the "sz" the English equivalent in "sh" consonant — the triangular function approximation.

- vowels/a, o, u, i, y, e/pronounced in isolation,
- the remaining consonants occuring between vowels ex. **aba, aca, ada, afa,** etc.

This approach helped make the pronounciation of a consonant more natural and at the same time to obtain to a certain degree of context dependences. The fragments of quasi-stationary course segments for the chosen representatives groups of phonemes' are shown in Figs. 8, 9 and 10.

4.5. Statistical dependencies between the length of the time intervals zero-crossings and amplitude

A method of obtaining the statistical dependences $A_n=f(I_n)$ that are fundamental to the settle the most probable amplitudes for a possible vocoder's option of the so called artificial amplitude was shown in Section 3.4. The characteristics A_n, obtained by use of the program "WOKANAL", for each intervals length from 1 to 60 samples enable to give an introductory forecast of overlay forms U_n in the resynthesis system. An example of aggregate distribution of maximum-average amplitude values for a list of 100 words is shown in Fig. 11.

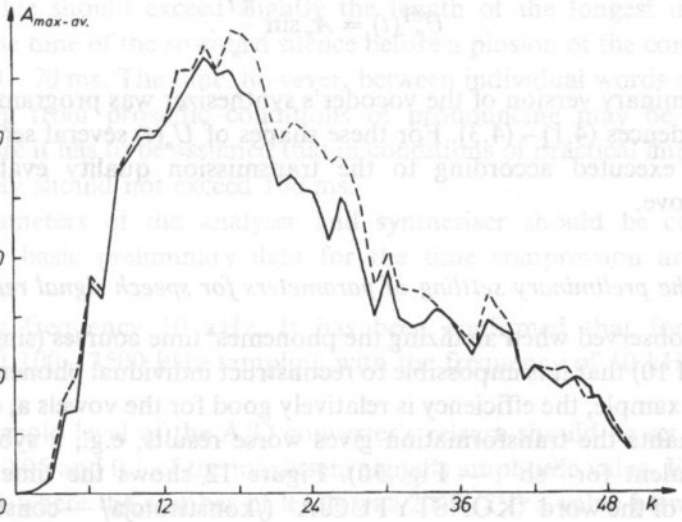


Fig. 11. The diagram of maximal-average amplitude values for the examined range of intervals. ..... negative amplitudes, _____ positive amplitudes.

4.6. Analysis of courses of shapes between zero-crossings of the speech signal

In order to settle the most representative course shapes of the envelope U_n, a group analysis based on the similarity (the distance) criterion should be carried out and, subsequently, subjective and objective verifications of the similarity degree (the intelligibility, distinctness and similarity) should be done for the original and the reconstructed signals. A preliminary analysis of the courses of the U_n of the material of 100 logatoms indicates that the most typical shapes and at the same time the best approximating the original course are: a triangular function, a sinusoid and a complex function of a half of the sinusoid and the triangular function (the assumptions (4.1), (4.2), (4.3)). For a triangular function the envelope $U_n^{TR}(t)$ has the following form:

$$U_n^{TR}(t) = \begin{cases} 2A_n t/I_n & \text{for } 0 \leqslant t \leqslant I_n 2 \\ A_n & \text{for } t = I_n/2 \\ 2A_n - \dfrac{2A_n t}{I_n} & \text{for } I_n/2 < t < I_n \end{cases} \tag{4.1}$$

and for the complex $U_n^Z(t)$

$$U_n^Z(t) = \frac{U_n^{TR}(t) + U_n^{SI}(t)}{2}, \tag{4.2}$$

where: $U_n^{TR}(kt)$ as in Eq. (4.1) dependency and

$$U_n^{SI}(t) = A_n \sin \frac{\pi t}{I_n}. \tag{4.3}$$

The preliminary version of the vocoder's synthesizer was programmed according to the dependences (4.1)–(4.3). For these shapes of $U_n(t)$ several series of measurements were executed according to the transmission quality evaluation criteria described above.

4.7. The preliminary settling of parameters for speech signal resynthesis

It has be observed when analizing the phonemes' time courses (single examples in Figs. 8, 9 and 10) that it is impossible to reconstruct individual phonemes to the same degree. For example, the efficiency is relatively good for the vowels a, e, i, y, o, u. For some consonants the transformation gives worse results, e.g., a sybilant "sz" (the Polish equivalent for "sh" — Fig. 10). Figure 12 shows the time course of the pronouncing of the word "KONSTYTUCJA" (/konstitutɕja/ — constitution). When transmitting a continuous speech or whole sentences pronounced spontaneously, delay distortions of a certain kind may arise. These distortions are caused by the necessity of a certain length of the signal buffering in the transmitter (or in the receiver) as a result of asynchronous extraction of the sequence $\{I_n, A_n\}$ connected with the variability of I_n and, therefore, with the transmitted information's matter.

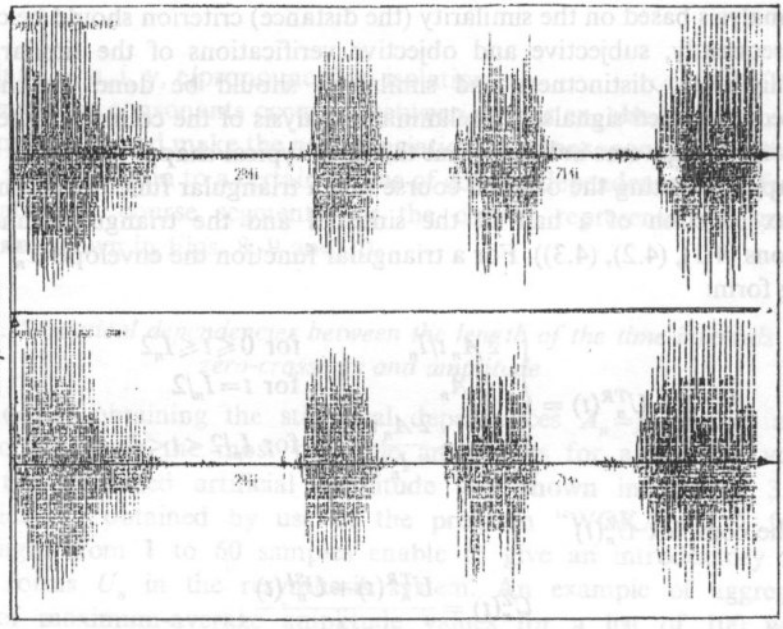


Fig. 12. The time courses for the pronounciation of "konstytucja" (constitution) /konstitutɕja/, the first window from the top — S_o, the second one — S_{oa}.

The delay's value should exceed slightly the length of the longest interval. It is assumed that the time of the so-called silence before a plosion of the consonants p, t, k, g is about 30 – 70 ms. The time, however, between individual words and so-called pauses resulting from prosodic conditions of pronouncing may be significantly longer. Therefore it has to be assumed that in conditions of practical implementation the time of delay should not exceed 100 ms.

Which parameters of the analyser and synthesiser should be considered as presettled? The basic preliminary data for the time compression are given and described below:

1. Sampling frequency 10 kHz. It has been confirmed that for the speech baseband 100 – 3500 kHz sampling with the frequency of 10 kHz is satisfactory.
2. The threshold level of the A/D converter's release should be set to the value between 0.05 and 0.1 of the maximum signal's amplitude value. For the 12-bit converter where the number of levels was 2048 (2^{11}) a value between 100 and 200 was assumed.
3. The signal's beginning retrieval was based on the verification as to whether a certain number of m samples exceed the threshold level α. It was assumed that if $m=10$ samples exceed the threshold level, this point is the beginning of the signal's analysis. The value of m and the mode of indication of the beginning as well as the end is an arbitrary matter from the point of view of the method and was not subject to any analysis.
4. Three shapes constituted of shape analysis of the envelope of the signal $U(n)$ assigned to the intervals in the synthesiser irrespective of the value of the amplitude S_{oa}:
   - a triangle,
   - a sinusoid,
   - average superposition of a sinusoid and a triangle.

   The dependences for the shapes mentioned above were described by the patterns (4.1), (4.2) and (4.3). The results of formal and informal audio monitoring research showed that "statistically" the best quality of sound for the listeners as well as the best distinctness and intelligibility are ensured by the complex function of a sinusoid and a triangle. Examples of the approximation of the vowels "a" and "o" by the three shapes: a sinusoid, a triangle and a complex function are shown in Figs. 13, 14 and 15.
5. The variability range of values of the intervals I_n with amplitudes $A_n>0$ was limited from 0.1 to 6 ms. The greater lengths are represented as silence, plosion distances, etc. This corresponds to the number of samples from 1 to 60 for $f_{pr}=10$ kHz.

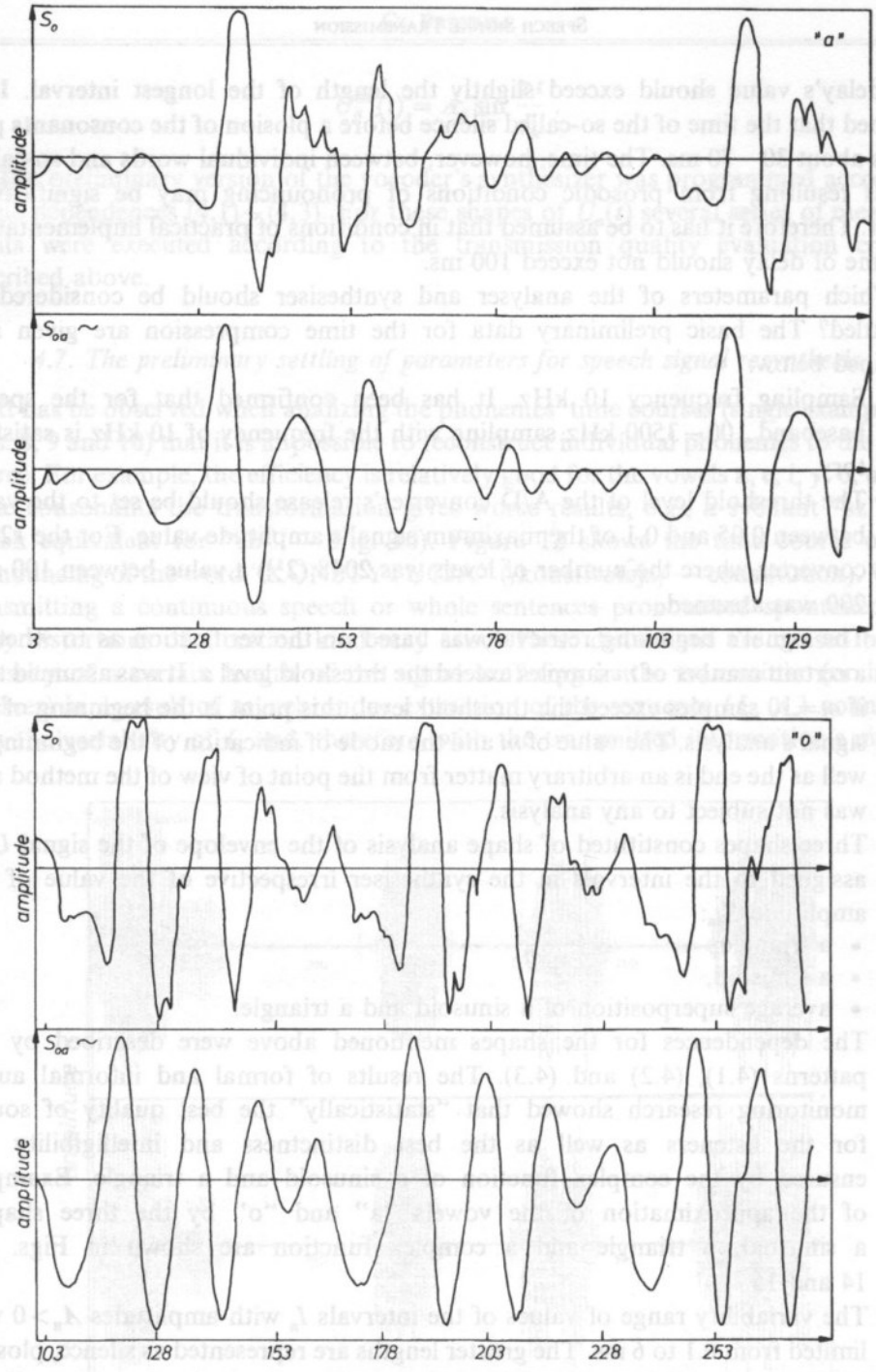


Fig. 13. The time segments of the "a" and "o" vowels approximated by a sinusoid function (the assumption (4.3)).

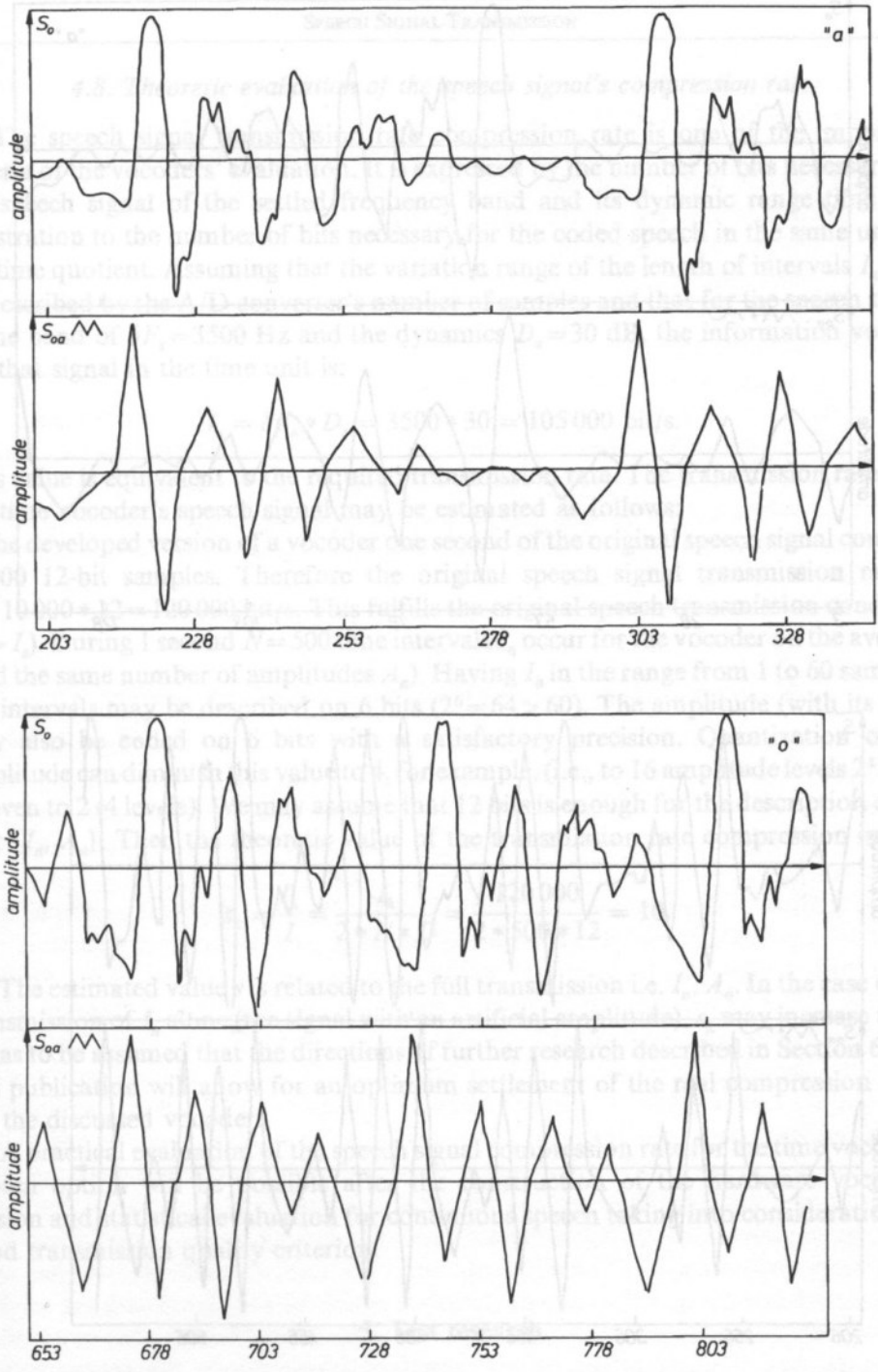


Fig. 14. The time segments of the "a" and "o" vowels approximated by a triangular function (the assumption (4.1)).

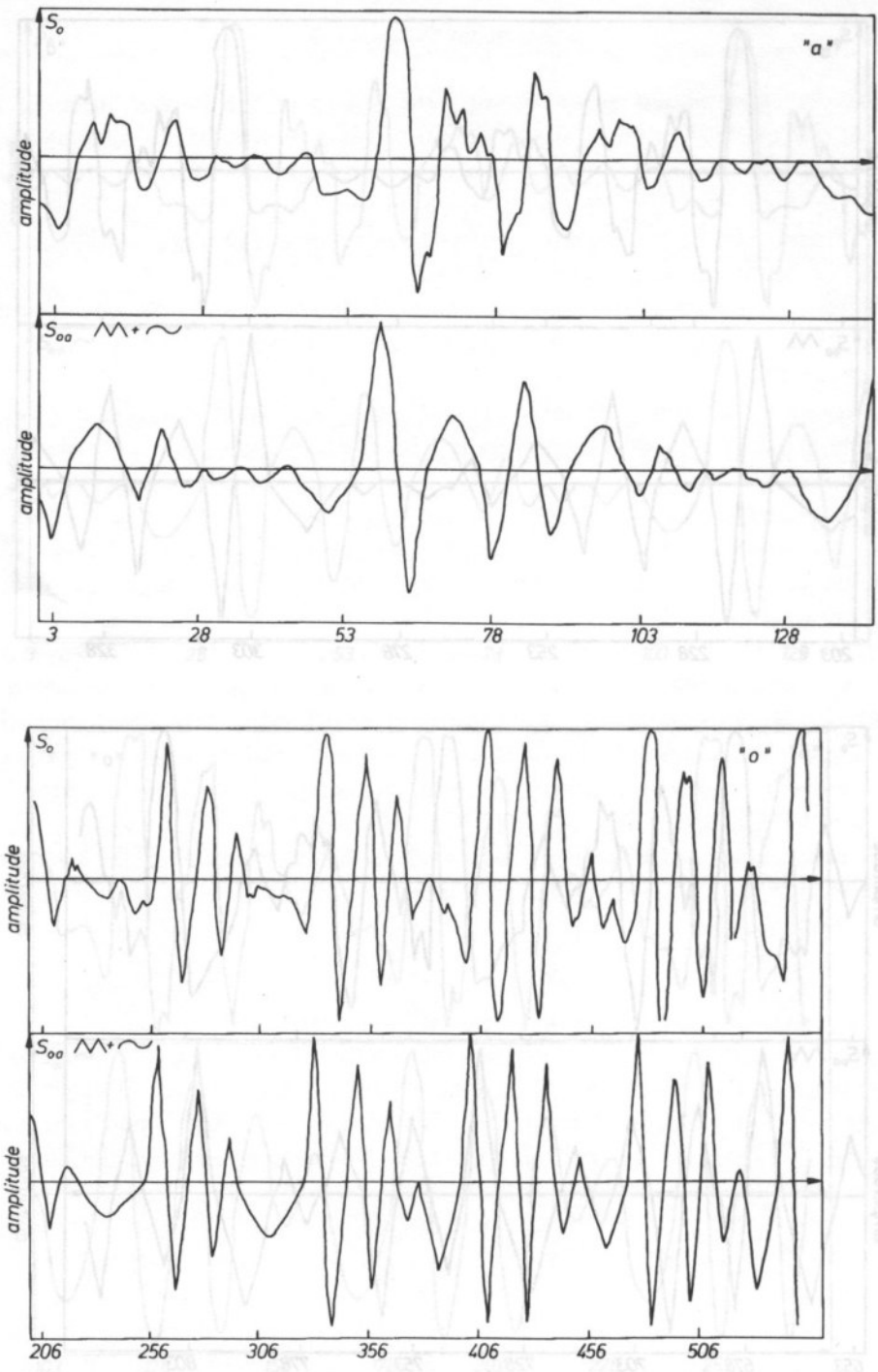


Fig. 15. The time segments of the "a" and "o" vowels approximated by a complex function (the assumption (4.2)).

4.8. Theoretic evaluation of the speech signal's compression rate

The speech signal transmission rate compression rate is one of the important criteria of the vocoders' evaluation. It is expressed by the number of bits necessary for the speech signal of the settled frequency band and its dynamic range time unit registration to the number of bits necessary for the coded speech in the same unit of the time quotient. Assuming that the variation range of the length of intervals I_n may be described by the A/D converter's number of samples and that for the speech signal in the band of $\delta F_s = 3500$ Hz and the dynamics $D_s = 30$ dB, the information volume for that signal in the time unit is:

$$I_s = \delta F_s * D_s = 3500 * 30 = 105\,000 \text{ bit/s}.$$

This value is equivalent to the required transmission rate. The transmission rate I_n of the time vocoder's speech signal may be estimated as follows:
In the developed version of a vocoder one second of the original speech signal contains 10 000 12-bit samples. Therefore the original speech signal transmission rate is $I_k = 10\,000 * 12 = 120\,000$ bit/s. This fulfills the original speech transmission condition ($I_k > I_s$). During 1 second $N = 500$ time intervals I_n occur for the vocoder on the average (and the same number of amplitudes A_n). Having I_n in the range from 1 to 60 samples, the intervals may be described on 6 bits ($2^6 = 64 > 60$). The amplitude (with its sign) may also be coded on 6 bits with a satisfactory precision. Quantization of the amplitude can diminish this value to 4, for example, (i.e., to 16 amplitude levels $2^4 = 16$) or even to 2 (4 levels). We may assume that 12 bits is enough for the description of the pair $\{I_n, A_n\}$. Then the theoretic value of the transmission rate compression is:

$$s_k = \frac{I_k}{I} = \frac{I_k}{2 * N * D} = \frac{120\,000}{2 * 500 * 12} = 10.$$

The estimated value s is related to the full transmission i.e. I_n, A_n. In the case of the transmission of I_n alone (the signal with an artificial amplitude), s_k may increase to 20. It has to be assumed that the directions of further research described in Section 6.2. of this publication will allow for an optimum settlement of the real compression value for the discussed vocoder.

A practical evaluation of the speech signal compression rate for the time vocoder's chosen option will be possible after the construction of the hardware vocoder's version and statistical evaluation for continuous speech taking into consideration the good transmission quality criterion.

5. Test research

5.1. Purpose and methodology of the research

The diversity of transmission quality estimation methods and techniques used for digital systems for coding and the transmission of information is a great obstacle

when comparing directly the quality of such systems. It has been decided to consider the following methods as the estimation criteria for the time vocoder operating according to the newly-introduced method of transmission rate compression:

A. Subjective measurements:
   - average logatom distinctness — SWL,
   - average word intelligibility — SZW,

B. Objective measurements:
   - mean-square difference,
   - cross corelation,
   - sonographic analysis.

The essential purpose of the research mentioned above was the following:

- the verification of correctness of the preliminary choice of the time parameters' $\{A_n, I_n\}$ range of variability. (see Section 4.7 of this publication).
- the statistical objectivized transmission quality estimation expressed in the subjective quality scale (SSJ), logatom distinctness (SWL) or word intelligibility (SZW) as well as in objective measures of the original signal's $S_0 - \{U_o(t)\}$ conformity with the reconstructed signal $S_{oa} - \{U_{oa}(t)\}$.

5.2. Measurements of the time vocoder's distinctness and intelligibility

The correct sound material assortment for subjective (acoustic) estimation tests is a problem of particular significance. The assumption that the test material represents a sound production which is most characteristic of the predicted applications of the discussed teletransmission system should be the basic rule of the test material's assortment [23].

The method of logatom and word lists phonetically and structurally compensated is one of the recommended test types. Phonemes and succession structures are determined in agreement with the proportions in natural speech [23]. Another question in the correct process of subject research is the choice and training of the audio monitoring staff. According to ISO recommendations [25], a person's hearing is correct if its hearing threshold value does not exceed 10 dB for any test frequency less than 4000 Hz and 15 dB within the band from 4000 Hz to 6000 Hz. It is recommended that the audio monitoring group be large enough so that the average results do not change much along with the increase of the group's population. It has been stated that such a minimal number of the group's population is 5 persons of each sex. The mode of the measurements' process preceded by an adequate training was consistent with ISO recommendations [23]. The calculation and the statistical inference modes were also executed according to the rules of statistical methods. The signal with the original amplitude was (S_{oa}) subject to the subjective research. The results (in %) of obtained distinctness and intelligibility along with standard deviations for the approximation using the complex function are shown in Tables 2 and 3.

The results of intelligibility and distinctness tests indicate a good quality of transmission of the speech signal with the original amplitude. This means that the

Table 2. Preliminary distinctness test (SWL) results

The listener	S_{oa} [%]
1	68
2	62
3	70
4	63
5	62
W_L [%]	65
δ_{SWL} [%]	7.48

Table 3. Preliminary results of the intellibility test SZW and the listener's subjective evaluation SOS scaled from 1 to 5

The listener	S_{oa}	
	[%]	SOS (1–5)
1	88	4
2	90	4
3	89	4
4	93	4.5
5	87	4
SZW [%]	89.4	4.2
δ_{SZW} [%]	4.2	0.89

amplitude information transmission is necessary for a good transmission quality. It is known from the speech signal's analysis [4] that this may be slow-changing information which allows further compression and significant improvement of the transmission's quality, close to S_{oa}. The obtained results, the best ones for the complex function (a triangle and a sinusoid) approximating the amplitude, make it possible to obtain the transmission's quality estimated by the listeners (SOS) and given by the results of the SLW and SWZ measurements as good. The preliminary subjective measurements were carried out on a representative sound material which, however, was limited by time consumption. Therefore the obtained results should be considered as preliminary ones indicating the range for the future implementation of the method.

5.3. *Objective measurements*

The tests of objectivized comparisons of the original signal S_o with the reconstructed signals S_{oa} with the original amplitude were carried out utilizing the following:

a) the mean-square distance,

b) the standardized cross-corelation the cross corelation coefficient

c) the sonographic spectrum evaluation.

a) The mean-square difference. The results of R_{sk} for comparisons of the determined speech signal's segments are shown in the Table 4. The mean-square difference is a measure of similarity between the sampled signal courses the original S_o

Table 4. The mean-square difference results for the speech signals' segments (phonemes) of 25 ms length

No.	phonemes	R_{sk} (S_o, S_{oa})
1	/ a /	12277
2	/ e /	32204
3	/ m /	31267
4	/ z /	14232
5	/ ʃ /	62944

and the reconstructed S_{oa}. Owing to the different characteristics of the speech sound courses, the comparative analysis should be limited to the same phonemes. In Table 4 different values of R_{sk} for S_{oa} show different "compatibility" of the signals with the original amplitude that is sent along with the original signals S_o. Not should be made attention of the dispersion of proportions among the values R_{sk} S_o, S_{oa}. This results directly from time structure differences of the phonemes.

b) The cross-correlation. In Table 5 the values of the cross-correlation factor r_{uv} are shown for the chosen segments of the quasi-stationary signals $t=25$ ms. As in the case of the mean-square

Table 5. The cross-correlation factor values

No.	phonemes	$r_{uv} - (S_o - S_{oa})$
1	/ a /	0.91
2	/ e /	0.85
3	/ m /	0.57
4	/ z /	0.88
5	/ ʃ /	0.18
average		0.88

difference, a significantly big dispersion of the cross-correlation factor values was observed for different letters. For the most part the best cross-correlation have the wovels and the sonant consonants. The worst was obtained for sibilants.

c) Sonographic analysis. An example of a speech signal 3-dimensional representation is shown in Fig. 16 (/riba/ — fish) as a spectrogram of the word "RYBA". The

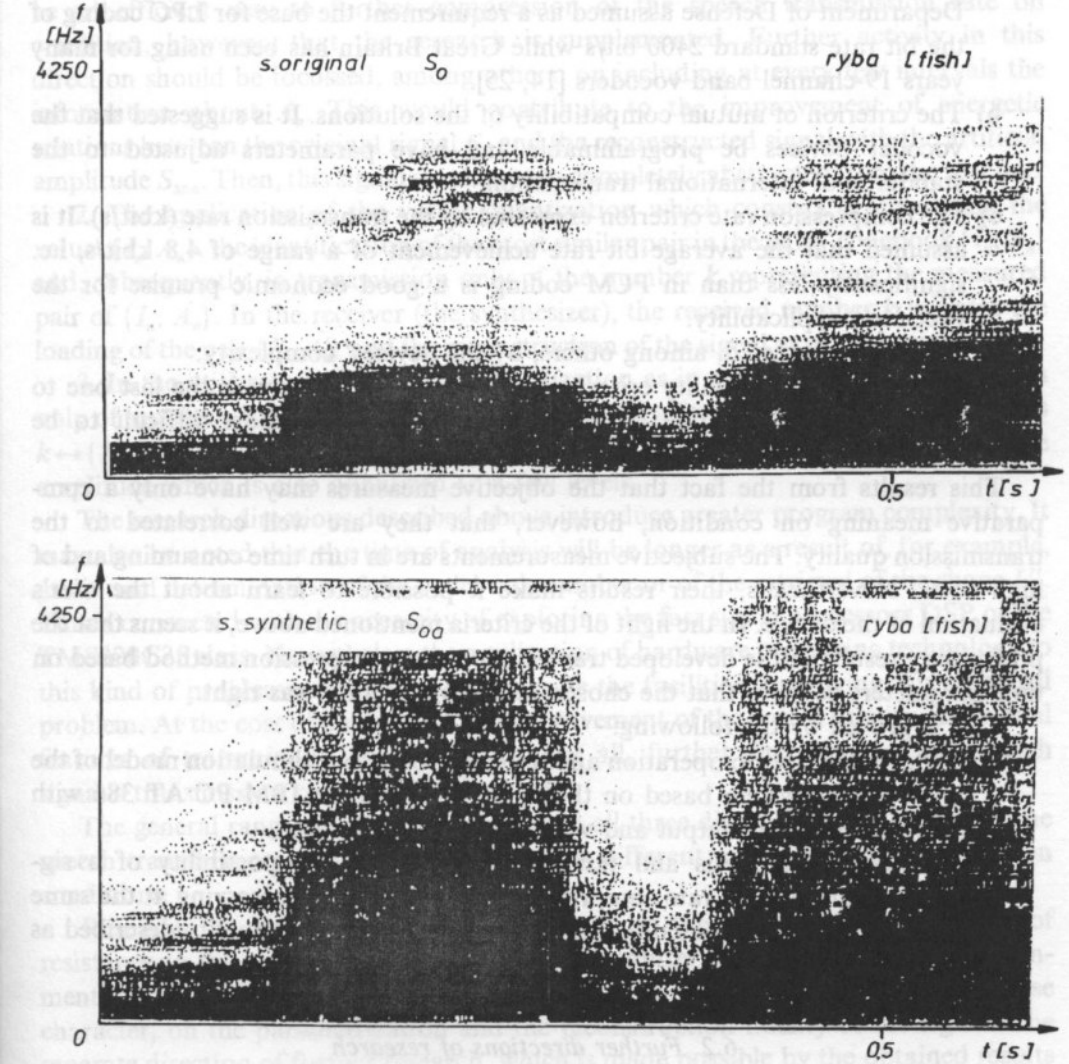


Fig. 16. The spectrogrammes of the pronounciation of "Ryba" (/riba/ — fish) from the top S_o, S_{oa}.

sonographic pictures of the signals confirm the pattern similarity of the reconstructed speech S_{oa} (the bottom spectrogram) with S_o (the upper spectrogram).

6. Summary and conclusions

6.1. Summary

The application of specific parametric coding solutions in vocoders is conditioned by some criteria. The following may be considered as the main ones:

a) The criterion resulting from the specific demands. For example, the US Department of Defense assumed as a requirement the base for LPC coding of the bit rate standard 2400 bit/s while Great Britain has been using for many years 19-channel band vocoders [14, 25].

b) The criterion of mutual compatibility of the solutions. It is suggested that the vocoder devices be programmable and have parameters adjusted to the standards of international transmission.

c) The compression rate criterion expressed by the transmission rate (kbit/s). It is assumed that the average bit rate achievement of a range of 4.8 kbit/s, i.e. significantly less than in PCM coding is a good economic premise for the vocoder's applicability.

d) The costs expressed, among others, by the devices' complexity.

e) The transmission quality criterion. Although this criterion is the last one to be mentioned, it is one of the most important and most difficult to be verified.

This results from the fact that the objective measures may have only a comparative meaning on condition, however, that they are well correlated to the transmission quality. The subjective measurements are in turn time consuming and of high costs. Nevertheless, their results make it possible to learn about the user's evaluation of the device. In the light of the criteria mentioned above, it seems that the preliminary results of the developed transmission rate compression method based on time parameters confirm that the chosen research direction was right.

This is proved by the following:

- the developed, set to operation and tested introductory simulation model of the time vocoder system based on the universal computer IBM PC AT 386 with acoustic input and output and with the proper software,
- the vocoder's analysis and test results indicating the possibility of a significant transmission rate compression, up to 10 times, preserving at the same time the transmission quality falling into the range of the quality described as 4 (good).

6.2. Further directions of research

The developed transmission rate compression method based on the extraction of the time parameters I_n and A_n does not fall back into its basic structure described in this publication, what is an obvious advantage of the method. The performed research and analyses proved that modifications of this method enabling, first of all, further speech signal transmission rate minimalization are possible. Three approaches

may be distinguished, which were mentioned indirectly in the previous sections of this publication:

1. A transmission only of the lengths of the intervals I_n (a signal with an artificial amplitude S_{sza}).

The preliminary analysis of the transmission only of I_n, which was taken into consideration in the "WOKCZAS" program, makes it possible to claim that this may be an efficient way to further compression of the speech transmission rate on condition, however, that the research is supplemented. Further activity in this direction should be focussed, among others, on including at every few intervals the information about A_n. This would contribute to the improvement of energetic relations between the original signal S_o and the reconstructed signal with the artificial amplitude S_{sza}. Then, this signal would not be completely abstracted from the real one.

2. The application of the vector quantization which consists in measuring the values $\{I_n, A_n\}$, the identification of the most similar pair in the pairs' codebook $\{\hat{I}_n, \hat{A}_n\}$ and, subsequently, in transmission only of the number k representing the measured pair of $\{I_n, A_n\}$. In the receiver (the synthesizer), the received number k initiates the loading of the pair $\{I'_n, A'_n\}$ and the reconstruction of the signal as in the basic method.

3. In the utilization of the vector quantization as in point 2: — the transmission only of numbers k of the intervals I_n codebook and a few estimated shapes K_n^i that is $k \leftrightarrow \{I'_n, K_n^i\}$, as well as the information transmission at every m interval about the amplitude which is also quantized to a few levels.

The research directions described above introduce greater program complexity. It may also be noted that the time of analysis will be longer as a result of, for example, additional determination of the codebook number or of the retrieval of the shape U'_n. This is connected with the necessity of exploring the fast signal processors DSP of the TMS320C25 class. Nevertheless the application of hardware processing technology to this kind of problems nowadays may provide the facilities for solving this technical problem. At the cost of complexity the improvement of the sound quality, individual features of transmission fidelity and, above all, further decreasing of the speech signal's transmission rate may be expected.

The general range of research, assigned to all three directions, should enclose the speech transmission quality evaluation under different auditorial and transmission conditions.

Research, in particular, should be focussed on the examination of the rate of resistance to the signal's distortions in the transmitter's and the receiver's environments and on the methods of diminishing the distortions' influence, mainly of a noise character, on the parametrization and the reconstruction fidelity of the signal. The separate direction of further research, which is made possible by the obtained results and the developed "tools" for research is the design of a time vocoder's hardware implementation.

This project was developed on the basis of The Research Project realized on request of the Ministry of National Education, CPBR, GRT-2.

References

[1] D.R. Allen, W.J. Strong and E.P. Palmer, *Experiments on the intelligibility of low frequency speech codes*, JASA **70**, 5, 1248–1255 (1981).

[2] А.Д. Архопова, М.А. Сапожков, *О качестве вокодерной речи*, Акустический Журнал, **16**, 3, 345–353 (1970).

[3] T.P. Barnwell, *Objective measures for speech quality testing*, J. Acoust. Soc. A., **66**, 6, 1658–1663 (1979).

[4] Cz. Basztura, *Sources, signals and acoustical images — processing, analysis and recognition* [in Polish], WKiŁ, Warszawa 1988.

[5] R.E. Crochiere et al., *A study of objective measures for speech waveform coders*, Murray Hill, New Jersey.

[6] W.R. Daumer, *Subjective evaluation of several efficient speech coders*, IEEE Trans. Communic. **COM-30**, No 4 April 1982, 665–662 (1982).

[7] W.R. Daumer and J.R. Cavanaugh, *A subjective comparison of selected digital coders for speech*, Bell Syst. Techn. J. **57**, 9, 3109–3165 (1978).

[8] J.A. Feldman, E.M. Hofstetter, M.L. Mlpass, *A compact, flexible LPC vocoder based on a commercial signal processing microcomputer*, IEEE Trans. on Acoustics, Speech and Signal Processing, vol. **ASSP-31**, 1, 252–257 (1983).

[9] B. Fette, D. Harrison, D. Olson and S.P. Allen, *A family of special purpose microprogrammable digital signal processor IC's in an LPC vocoder system*, IEEE Trans. on Acoustics, Speech and Signal Processing, vol **ASSP-31**, 1, 273–281 (1983).

[10] R. Gubrynowicz, *Application of zero-crossing method in speech signal analysis and automatic word recognition*, (in Polish) IFTR PAS Report 27 (1974).

[11] D.J. Goodman and R.D. Nash, *Subjective quality of the same speech transmission conditions in seven different countries*, IEEE Trans. Commun., **COM-30**, 4, 642–654 (1982).

[12] D.J. Goodman, J.S. Goodman and M. Chen, *Intelligibility and ratings of digitally coded spedch*, IEEE, vol. **ASSP-28**, 5, 403–408 (1978).

[13] A.H. Gray and D. Markel, *Distance measure for speech processing*, IEEE Trans. Acoust. Speech and Sign. Proc., **ASSP-24**, 380–391 (1976).

[14] N. Kitawski, M. Honda and K. Itoh, *Speech-quality assessment method for speech-coding system*, IEEE Communications Magazine, **10**, 26–33 (1984).

[15] B.F. Logan, *Signals designed for recovery after clipping — localization of infinite products*, ATT Bell Laboratories Technical Journal, **63**, 2, 261–285 (1984).

[16] P. Mermelstein, *Evaluation of segmental SNR measure as an indicator of the quality of ADPCM coded speech*, JASA **66**, 6, 1664–1667 (1979).

[17] В.Ф. Михаилов, Л.В. Златоустова, *Измерение параметров речи*, Радио и Связ, Москва 1987.

[18] M. Nakatsui, *Subjective SNR measure for quality assessment of speech coders a cross language study*, MAT., ICA **12 A1–1**, Toronto 1986.

[19] М.В. Назаров, Ю.Н. Прохоров, *Методы цифровой обработки и передачи речевых сигналов*, Радио и связ, Москва 1985.

[20] A.V. Oppenheim [Ed], *Application of digital signal processing*, MIT, Cambridge, Mass., 02139, Prentice-Hall Inc., USA 1978.

[21] N. Osaka and K. Kakehi, *Objective model for evaluating telephone transmission performance*, Review of the Electrical Communications Laboratories, **34**, 4 (1986).

[22] Н.Т. Петрович, М.В. Каблакова, Н.И. Козленко, *Передача сигналов методами КИМ–ОФТ*, Радио и связ, Москва 1985.

[23] *Draft Proposal* ISO/DP 4870, *Recommended methods for intelligibility tests.*

[24] K. Reeder, W.J. Strong and E.P. Palmer, *Preliminary study of a low frequency formant based speech code for the several hearing impaired*, JASA, **61**, 5, 1379 (1977).

[25] М. Сапожков, М. Михайлов, *Вокодерная связ*, Радио и связ, Москва 1983.

[26] M.E. SMITH, K.E. ROBINSON and W.J. STRONG, *Intelligibility and quality of linear predictor and eigen parameter coded speech*, IEEE, vol. **ASSP-29**, 3, 391–395 (1981).
[27] W. SOBCZAK [Ed.], *Problems of teleinformatics*, WKiŁ, Warszawa, 1984.
[28] S. WANTANABE and K. INOMOTO, *On a loudness function of artificial speech*, Acustica, **44**, 4 (1980).
[29] W. WOLFF, *Untersuchungen zur subjectiven Qualitat vom Spektrum des uberlagerten Rauchnes*, Frequenz, **34**, 6–9 (1980).
[30] А.Г. Зюко и другие, *Помехоустойчивость и ефективность систем передачи информации*, Радио и связ, Москва 1985.

ARCHIVES OF ACOUSTICS
19, 1, 37–58 (1994)

AM AND FM DIFFERENCE LIMENS*

A. SĘK and E. OZIMEK
Institute of Acoustics, Adam Mickiewicz University
(60-769 Poznań, ul. Matejki 48/49)

This paper reports two basic experiments aimed at determining difference limens (DLs) for amplitude and frequency modulation as a function of the modulation indices (m, β), modulation frequency and sound pressure level of the modulated signals. Two types of modulator were used. The first one was the sinusoidal signal. In this case the simplest (sine by sine) amplitude and frequency modulation was considered. In the second case the modulator had a constant frequency but its amplitude was randomly changed with mean 0 and standard deviation σ. In a two-alternative forced choice task just noticeable differences of modulation intensity were measured. It was found that AM and FM difference limens increase with the increase of modulation indices of the reference signals. Additionally, it was shown that the difference limens were almost independent of the modulation frequency. They were also independent of the type of modulator used in the investigations.

1. Introduction

Investigations into the perception of amplitude and frequency modulated signals have been extensively described [1–11, 13–25]. One group of papers deals with the thresholds used to detect AM and FM whereas the other group deals with investigations into the intensity of loudness or pitch fluctuation and roughness of the modulated signals.

Modulation thresholds were investigated by ZWICKER [25] who determined AM and FM thresholds with reference to the amplitude and frequency of the carrier signal and to the modulation frequency. This helped to determine three perception ranges of the modulated signals: follow-up, roughness and sidebands separation.

Modulated signals generate different auditory sensations in these areas, what affects the detection threshold of AM and FM. In the case of small modulation frequencies ($f_{mod} < 20$ Hz — follow-up range), changes in the signal amplitude or frequency bring about changes in signal loudness or pitch. When $f_{mod} \in (20 - \text{CMF})$ Hz (CMF — critical modulation frequency [20, 21, 25]), the sensation evoked by the modulated signal is called roughness [11, 22, 23]. For modulation frequencies $f_{mod} > \text{CMF}$, what corresponds to the transition of sidebands of modulated signals

* This research was supported by the grants 20071 1901 and 2 0910 01 from the Committee of Scientific Researches (KBN)

spectra beyond the range of one critical band, the auditory sensation is similar to that evoked by a multi-tone. It is worth stressing here that in this situation thresholds for detecting amplitude and frequency modulation expressed as modulation indices (i.e., m and β) reach the same value.

In investigations into supra-threshold changes of amplitude and frequency modulated signals, the main focus has so far been on the evaluation of the intensity of loudness and pitch fluctuation, and the evaluation of the intensity of the roughness of AM and FM signals [4–6, 8, 11, 17–19, 22–24]. Pitch or loudness fluctuation defines changes of these parameters which occur less than 20 times per second (i.e. $f_{mod} \leqslant 20$ Hz). Roughness occurs when modulation frequency is greater than 20 Hz and at the same time does not exceed the critical modulation frequency (CMF) [20, 21, 25] which depends on the carrier frequency [21].

TERHARDT [22, 23] described the dependence of the intensity of loudness fluctuation on the carrier frequency, modulation frequency and sound pressure level. He found out that the intensity of loudness fluctuation and the intensity of roughness are proportional to the square of the AM index. Likewise, GUIARO and GARAVILLA [8] discovered that the roughness of the AM signal is nonlinearly correlated with modulation depth. SCHOENE [19], on the other hand, showed that the intensity of loudness fluctuation is independent of the sound pressure level and modulation frequency in the range up to $f_{mod} \leqslant 10$ Hz. FASTL [4] found out that the intensity of tonal signal fluctuation and that of a noise band is proportional to the logarithm of the AM index.

The determination of AM difference limens has been discussed less extensively. An AM difference limen (AM DLs) denotes a just noticeable difference between two AM signals which differ only with respect to the amplitude modulation index.

One of the first papers dealing with this problem is the one by SCHOENE [19], which pertains to the dependence of AM difference limens on the modulation index of the reference signal. It was shown that the difference limens are linearly correlated with the AM index of the reference signal. Moreover, it was also shown that AM difference limens are independent of modulation frequency in the range up to $f_{mod} \leqslant 10$ Hz. FLEISCHER [5] has shown a linear dependence of AM difference limens on the AM index. It is worth pointing out that both SCHOENE'S [19] and FLEISCHER'S [5] papers are concerned with AM in the case when both carrier and modulating signals were sinusoids.

OZIMEK and SĘK [17] showed AM difference limens of an octave noise band as a function of the selected physical parameters of the carrier signal and modulating signals. They found out that the dependence of the difference limens on the AM index of the reference signal can be characterized by a local maximum [17].

WAKEFIELD and VIEMEISTER'S paper [24] is one of the recent papers which discusses AM difference limens of a broadband noise. These authors showed that the difference limens of AM do not depend on the sound pressure level and modulation frequency. It is worth stressing that the results obtained by FLEISCHER [5], SSCHOENE [19] and OZIMEK and SĘK [17] are fully compatible, what was also observed by WAKEFIELD and VIEMEISTER [24].

The roughness and intensity of pitch fluctuation of frequency modulated signals have been discussed by TERHARDT [22, 23]. In addition to investigations into the roughness of AM signals, he also studied the problem of roughness of a FM signal. Terhardt suggested that, in accordance with Zwicker–Maiwald's model [25, 15], the general principles of the dependence of roughness and intensity of pitch or loudness fluctuation on the parameters of modulated signals should be similar for both types of modulation (i.e. AM and FM). This hypothesis was confirmed by experiments by TERHARDT [8].

KEMP [11] discussed the perception of roughness in FM signals. He evaluated the roughness of a tonal signal frequency modulated by a sinusoid with respect to the physical parameters of both the carrier and modulation signals. Kemp found out that the magnitude of the roughness depends on modulation frequency, carrier frequency and the sound pressure level. Maximal relative roughness occurs for the frequency modulation 40–80 Hz. Relative roughness is nonlinearly correlated with signal deviation. A similar relation observed was between the amplitude modulation index and roughness in the case of AM signals [19].

FM difference limens were discussed by OZIMEK and SĘK [18]. FM DLs denote a just noticeable difference between two FM signals which differ only with respect to frequency deviation. The authors measured FM difference limens of a tonal signal frequency modulated with another tonal signal with respect to modulation frequency, carrier frequency and the sound pressure level. Futhermore, they also discussed the dependence of the DLs on reference signal deviation. It was observed that FM difference limens were approximately independent of the carrier and modulation frequency. They were also independent of the sound pressure level. However, an increase in reference signal deviation caused an approximately nonlinear increase of the difference limens' value. Ozimek and Sęk also measured DLs using random modulators with uniform and Gaussian probability distributions. Similar difference limens were obtained for both types of modulator.

As follows from the above presented overview, most of the papers discussing investigations into the perception of AM and FM signals with supra-threshold modulation intensity concerned the evaluation of the roughness and intensity of loudness and pitch fluctuation. Some other papers were devoted to AM difference limens and only a few discussed the problem of difference limens of frequency modulated signals. It should be stressed that most of the papers concerned with AM and FM detection or discrimination were based on periodic sinusoidal changes of amplitude and frequency. However, different carrier signals, including a white noise [19, 24] or its different bands [17], were used.

In environmental signals, periodic amplitude and frequency changes occur very seldom and are usually of very short duration. Signals such as music, speech or traffic noise are characterized by quasi-random changes in both amplitude and frequency domains. Hence it seems interesting to determine the difference limens of random changes of some physical parameters obtained by amplitude and frequency modulation.

However, random changes of amplitude and frequency can be of two types. In the case, for example amplitude modulation, a random modulation signal causes irregular changes of both modulation depth and modulation frequency. This simultaneous changes of modulation depth and modulation frequency is significant for the evaluation of a signal and can affect the threshold. Therefore a sinusoidal signal whose amplitude was randomly changed was adopted as a random modulator, similarly to SĘK [21]. In the AM case this modulator caused random changes of the AM index whereas in the FM case the modulator caused random change of frequency deviation. A modulator of this type, generated by a computer, was a set of successive equal periods of function sin (ωt) with an amplitude changing randomly from the period to period.

2. Aim of investigations

The main purpose of the investigations was to determine AM and FM difference limens for a sinusoidal carrier signal with respect to modulation frequency, modulation indices, the type of modulator and sound pressure level.

A just noticeable difference between two amplitude or frequency modulated signals which differed with respect to the value of a respective modulation index were determined. One of them, called the reference signal, was characterized by a constant value of the AM index m_r (in AM case) or deviation d_r (in FM case). The second one, called the test signal was characterized by lower values of modulation indices than the reference signal, i.e., $m_t < m_r$ and $d_t < d_r$.

Measurements of AM and FM difference limens were carried out for the modulation frequency $f_{mod} = 2$, 8, 32 and 128 Hz. These values correspond to the perception ranges of the modulated signals mentioned above, i.e., 2 and 8 Hz correspond to the follow-up range, 32 Hz corresponds to the roughness range and 128 Hz corresponds to the sidebands separation range.

Four values of the AM index of the reference signal (m_r) and four values of the deviation of the reference signal (d_r) were used. A sinusoidal carrier frequency $f_c = 1000$ Hz was used. The sound pressure level was $L_c = 35$, 45, 55, 65, 75 and 85 dB SPL. Two types of modulator were used. One of them was a sinusoid and the other a sinusoid with an amplitude randomly changing from period to period. In the case of AM and FM with random modulators it appeared necessary to analyze the temporal and spectral structure of the modulated signals and to define a direct measure of modulation intensity. The problems are discussed in Section 3 of this paper, in the case of both AM and FM.

3. Temporal and spectral structure of modulated signals

Analytical considerations aimed at determining the temporal and spectral structure of modulated signals, where simple tones are both modulation and carrier signals are reduced to simple trigonometric transformations. Let us assume that the carrier signal has the form:

$$a(t) = A_0 \cos(\omega_0 t), \tag{1}$$

and the modulation signal has the form:

$$b(t) = B \sin(\omega_m t). \tag{2}$$

In the case of amplitude modulation, the temporal form of a modulated signal can be described as follows:

$$a_{AM}(t) = A_0(1 + m \sin \omega_m t) \cos \omega_0 t \tag{3}$$

or:

$$a_{AM}(t) = \frac{A_0 m}{2} \cos(\omega_0 - \omega_m) t + A_0 \cos(\omega_0 t) + \frac{A_0 m}{2} \cos(\omega_0 + \omega_m) t, \tag{4}$$

where

$$m = \frac{kB}{A_0}, \tag{5}$$

denotes the coefficient of amplitude modulation depth and k is a constant.

In the case of frequency modulation the temporal form of a modulated signal can be expressed as follows:

$$a_{FM}(t) = A_0 \cos(\omega_0 t - \beta \cos(\omega_m t)), \tag{6}$$

where

$$\beta = \frac{\Delta\omega}{\omega_m} = \frac{k_1 B}{\omega_m} \tag{7}$$

is a frequency modulation index, $\Delta\omega$ — frequency deviation and k_1 a constant.

The FM signal can be expressed as follows [12]:

$$a_{FM}(t) = A_0 \sum_{-\infty}^{\infty} J_n(\beta) \cos(\omega_0 + n\omega_m) t, \tag{8}$$

where $J_n(\beta)$ is a Bessel function of the first kind, n-th order.

The spectra of AM and FM signals consist of discrete components among which the central one is the component which correspond to the carrier signal and all the others are products of modulation. In the case of AM, the spectrum, in addition to the central component, includes two sidebands with equal amplitudes, situated at a distance of $\pm\omega_m$ from the carrier. In the case of FM, in addition to the component which corresponds to the carrier signal, the spectrum consists of a theoretically unlimited number of sidebands. The sidebands are symmetrical with respect to the central component and they are $\pm\omega_m$ separated in the frequency domain. Odd sidebands with frequencies lower than the carrier are phase shifted by π with respect to sidebands with frequencies higher than the carrier frequency. This causes the so-called monaural phase effect (MPE) [20, 21].

The determination of the spectrum of a simple tone amplitude or frequency modulated by a random signal is a more complex issue. If a random modulation signal is a realization of a stationary ergodic process with a normal probability distribution, an autocorrelation function [12] can be determined for a modulated

signal. Using the Wiener-Chińczyn theorem, we can determine the power spectral density function of the modulated signal. However, this method when used to determine the spectrum of a modulated signal is troublesome since it does not provide a direct measure of modulation intensity. A clear definition of the intensity of modulation for both AM and FM is necessary for measurement purposes. Therefore, in order to determine the spectrum of a sinusoid amplitude modulated by a random signal, earlier consideration on modulation of the tone-tone type were taken into account. However the random modulator is a sinusoidal signal with a randomly changing amplitude. Thus the temporal and spectral structure of the modulated signal does not change significantly as compared with the tone-tone modulation, provided that the amplitude of the modulation signal B in the expression (2) is a randomly changing parameter. Consequently, the values of amplitudes of the sidebands of the AM or FM signals' spectrum change randomly in accordance with amplitude variations of the modulation signal.

In the simplest case of the tone-tone modulation, the amplitude modulation index m (5) or the frequency modulation index β (7) is the measure of intensity modulation. Modulation indices defined in this way cannot be used directly in the case of a modulation by a random signal because such a signal is characterized by an indefinite value of amplitude. Therefore, in the case of AM, the root-mean-square AM index m_{RMS}, which expresses the ratio of RMS values of the modulator and carrier signals, was adopted as the measure of modulation intensity [12]:

$$m_{\mathrm{RMS}} = k\sqrt{\frac{\sigma_{\mathrm{B}}^2}{A_0^2}}. \tag{9}$$

In the case of FM, $\Delta\omega_{\mathrm{RMS}}$ which expresses the root-mean-square value of frequency deviation, was adopted as the measure of modulation intensity [12]:

$$\Delta\omega_{\mathrm{RMS}} = k_1\sigma_{\mathrm{B}}, \tag{10}$$

where σ_{B} denotes an RMS value of a random modulator.

AM difference limens were expressed, as in WAKEFIELD and VIEMEISTER [24] by means of the following expression:

$$\Delta M = 10\,\mathrm{Log}\,(m_r^2 - m_t^2)\ [\mathrm{dB}], \tag{11}$$

where m_r and m_t denote the root-mean-square AM indices for the reference and test signal respectively. The AM index of the reference signal was determined in the same way [5, 24]:

$$M_r = 10\,\mathrm{Log}(m_r^2)\ [\mathrm{dB}], \tag{12}$$

In the FM case difference limens were expressed as differences between the root-mean-square deviations (10) of the reference signal d_r and test signal d_t:

$$\Delta F = d_r - d_t\ [\mathrm{Hz}]. \tag{13}$$

4. Method

The two-alternative forced choice method (2AFC) with Levitt's adaptive procedure [13] was used. The subject listened monaurally to a pair of modulated signals in a random order. One of them was a reference signal with a constant supra-threshold modulation intensity (m_r or d_r) and the other one was a test signal with changing modulation intensity (always smaller than the reference signal, i.e., $m_t < m_r$, $d_t < d_r$). The subject's task was to indicate the signal in which the intensity of amplitude changes (in the case of AM) or intensity of frequency changes (in the case of FM) was greater. The difference in the modulation intensity between the reference signal and the test signal (i.e., $m_r - m_t$ or $d_r - d_t$) was increased by a factor of 1.5 after two successive correct answers or decreased after one incorrect answer. This procedure tracks the point on the psychometric function corresponding to 71% correct. A single measurement of the threshold was completed when 12 reversals were obtained. The threshold value from such a measurement was calculated as an arithmetic mean value of the modulation index at the last eight turnpoints. The results presented in this paper are arithmetic means of ten independent measurements. Three subjects with audiologically normal hearing took part in the investigations.

The signals were numerically generated via a 12-bit D/A converter. Each signal had a 1500 ms duration including raise and decay times of 100 ms each. The inter stimulus interval was 400 ms. The subjects were tested in a sound isolated booth and they used an answering keyboard to give their answers.

5. Results and their analysis

5.1. AM *difference limens*

The dependence of AM difference limens on modulation frequency was the basic one to be determined in this part of the study. The investigations consisted in determining the just noticeable difference between the reference signal, characterized by a constant value of the AM index $m_r = 25$, 50, 75 and 95%, and the test signal in which the AM index m_t was changed in the range from 0 to m_r (i.e. $m_t \leqslant m_r$). The results of these investigations, for the sound pressure level 75 dB SPL, carrier frequency 1000 Hz and for a sinusoidal modulator signal were shown in Fig. 1 for the subjects EO and AS respectively. The AM index of the reference is the parameter of the data.

As can be seen, AM difference limens are very similar for both subjects and they are approximately independent of modulation frequency.

Similar investigations were also carried out for a random modulator, i.e., a sinusoidal signal with an amplitude randomly changing from period to period. In these investigations the same carrier signal was used, i.e., $f_c = 1000$ Hz and $L_c = 75$ dB.

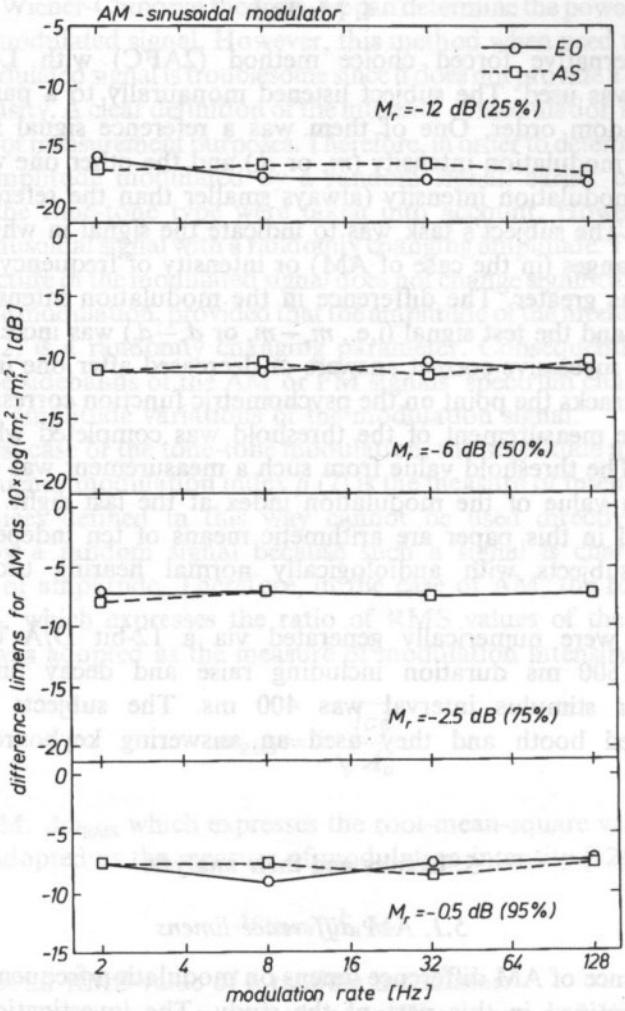


Fig. 1. Dependence of AM difference limens on modulation frequency for a sinusoidal modulator. The AM index is the parameter of the data.

The results were shown in Fig. 2 for the subjects EO and AS. For both subjects AM difference limens for a random modulator are similar and independent of the modulation frequency.

The data obtained for both types of modulator were subjected to analysis of variance ANOVA with the following factors: type of modulator, modulation index of reference signal, subject and frequency modulation. The effect of the modulation index was highly significant [$F(3,55) = 622$, $p < 0$k.0001]. The effect of the subject was not significant [$F(1,55) = 0.938$, $p = 0{,}34$]. Also the effects of the modulating frequency and type of modulator were not significant [$F(3,55) = 0.494$, $p = 0.69$] and [$F(1,55) = 2.47$, $p = 0.67$] respectively.

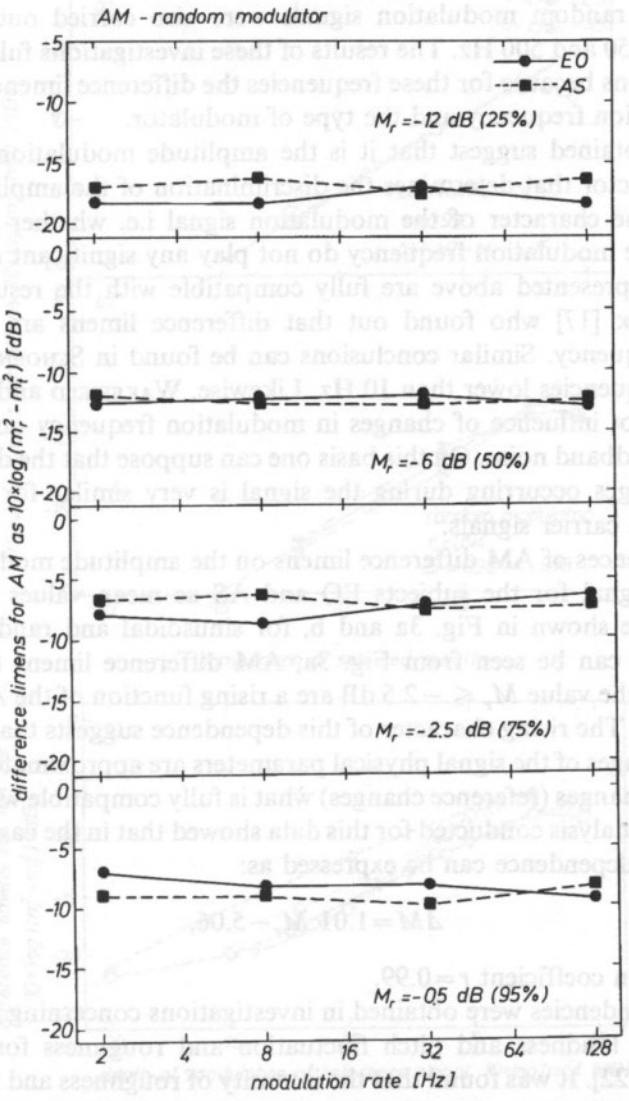


Fig. 2. Dependence of AM difference limens on modulation frequency for a random modulator. The AM index is the parameter of the data.

Generally, it can be stated that the difference limens obtained for sinusoidal and random modulators are very close to each other. This means that discrimination of the amplitude changes does not depend on the type of modulator, i.e., whether it is sinusoidal or random. This helps generalize conclusions drawn from experimental data on the roughness and intensity of amplitude fluctuation for sinusoidal changes and apply them to random changes of amplitude. This is similar to SĘK's [21] conclusion which states that the threshold for detecting amplitude modulation did not depend on the type of modulator. Measurements of AM difference limens for the

sinusoidal and random modulation signals were also carried out for the carrier frequencies of 250 and 500 Hz. The results of these investigations fully confirmed the above conclusions because for these frequencies the difference limens did not depend on the modulation frequency and the type of modulator.

The data obtained suggest that it is the amplitude modulation index which is primarily the factor that determines the discrimination of the amplitude changes in AM signals. The character of the modulation signal i.e. whether it is random or periodic and the modulation frequency do not play any significant role in this case.

The results presented above are fully compatible with the results obtained by OZIMEK and SĘK [17] who found out that difference limens are independent of modulation frequency. Similar conclusions can be found in SCHOENE [19] who used modulation frequencies lower than 10 Hz. Likewise, WAKEFIELD and VIEMEISTER [24] noted the lack ot influence of changes in modulation frequency on AM difference limens of a broadband noise. On this basis one can suppose that the discrimination of amplitude changes occurring during the signal is very similar for many different modulation and carrier signals.

The dependences of AM difference limens on the amplitude modulation index of the reference signal for the subjects EO and AS as mean values for modulation frequencies were shown in Fig. 3a and b, for sinusoidal and random modulators respectively. As can be seen from Fig. 3a, AM difference limens for a sinusoidal modulator and the value $M_r \leqslant -2.5$ dB are a rising function of the AM index of the reference signal. The rising character of this dependence suggests that just noticeable increases in changes of the signal physical parameters are approximately proportional to the existing changes (reference changes) what is fully compatible with Weber's law. The regression analysis conducted for this data showed that in the case of a sinusoidal modulator this dependence can be expressed as:

$$\Delta M = 1.01\ M_r - 5.06, \tag{14}$$

at the correlation coefficient $r = 0.99$.

Similar dependencies were obtained in investigations concerning the valuation of the intensity of loudness and pitch fluctuation and roughness for both types of modulation [11, 22]. It was found that the intensity of roughness and that of loudness and pitch fluctuation are rising (nonlinearly) functions of respective modulation indices.

The dependence of AM difference limens on the AM index of the referrence signal for a random modulator was shown in Fig. 3b. Also in this an case increase in the AM index signal caused an increase in difference limens. This dependence can be approximately expressed as:

$$M = 0.99\ M_r - 5.73. \tag{15}$$

However, for this modulator there were clear differences between the subjects with respect to the difference limens. These differences affected the value of the correlation coefficient which in this case was equal to $r = 0.94$.

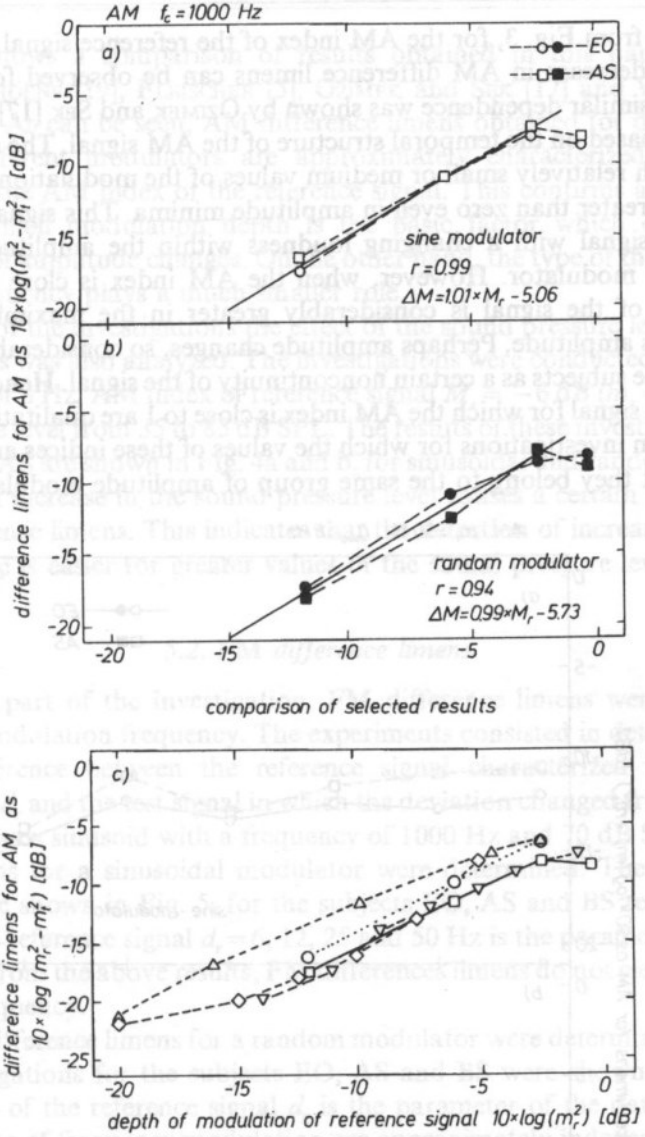


Fig. 3. Dependence of AM difference limens on the AM index of the reference signal for a sinusoidal (a) and random (b) modulator. Solid lines denote the best linear aporoximation, c — comparison of averaged results with literature data.

—□— this paper r = 0.95 $\Delta M = 0.83 * M_r - 3.41$

r = 0.99 $\Delta M = 1.0 * M_r - 5.84$ ---▽--- FLEISCHER [5]

···○··· OZIMEK & SĘK [17] r = 0.98 $\Delta M = 0.99 * M_r - 5.88$

r = 0.99 $\Delta M = 1.04 * M_r - 4.6$ ---◇--- WAKEFIELD & VIEMEISTER [24]

···△··· SCHOENE [19] r = 0.95 $\Delta M = 0.9 * M_r - 5.70$

As follows from Fig. 3, for the AM index of the reference signal M_r greater than −2.5 dB, the decrease in AM difference limens can be observed for both types of modulator. A similar dependence was shown by OZIMEK and SĘK [17]. They also tried to interpret it based on the temporal structure of the AM signal. The amplitude of the AM signal with relatively small or medium values of the modulation depth is always substantially greater than zero even in amplitude minima. This signal is perceived as a continuous signal with a changing loudness within the amplitude maxima and minima of the modulator. However, when the AM index is close to 1, the sound pressure level of the signal is considerably greater in the maximum than in the minimum of its amplitude. Perhaps amplitude changes, so considerably different, are perceived by the subjects as a certain noncontinuity of the signal. Hence investigations of a modulated signal for which the AM index is close to 1 are qualitatively different in perception from investigations for which the values of these indices are small, in spite of the fact that they belong to the same group of amplitude modulated signals.

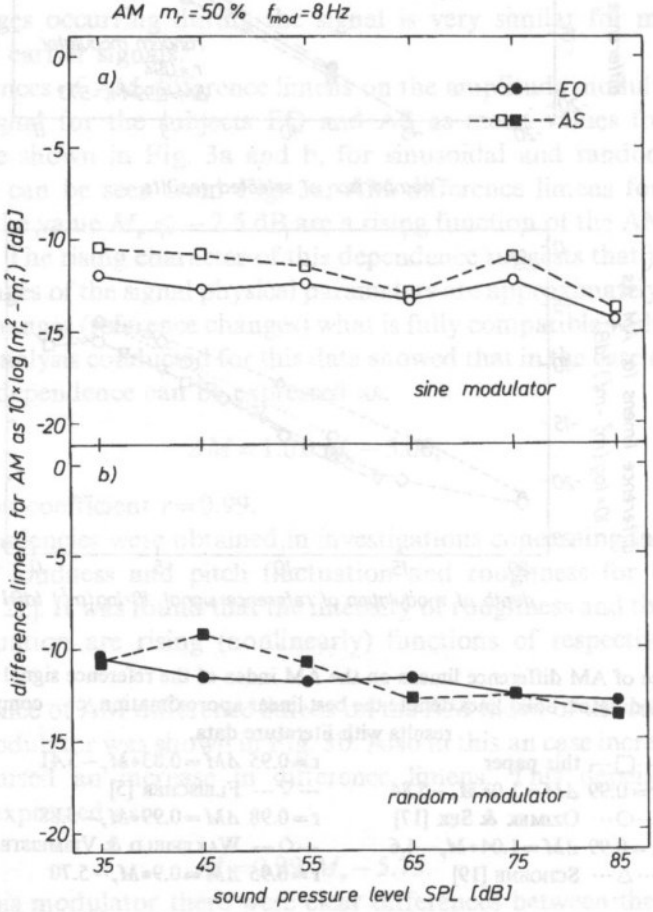


Fig. 4. Dependence of AM difference limens on the sound pressure level for a sinusoidal (a) and a random (b) modulator.

Figure 3c shows a comparison of results obtained in this paper with those obtained by SCHOENE [19], FLEISCHER [5], OZIMEK and SĘK [17] and WAKEFIELD and VIEMEISTER [24]. As can be seen, AM difference limens obtained for different carrier signals and different modulators are approximately characterized by the same dependence on the AM index of the reference signal. This confirms an earlier thesis according to which modulation depth is the basic factor which determines the discrimination of amplitude changes. On the other hand, the type of the modulator or modulation frequency plays a much smaller role.

In this part of the investigations the effect of the sound pressure level on the AM difference limens was also analyzed. The investigations were conducted for a modulation frequency of 8 Hz, AM index of reference signal $M_r = -6$ dB ($m_r = 50\%$) and for a sound pressure level from 35 to 85 dB SPL. The results of these investigations for the subject EO and AS are shown in Fig. 4a and b, for sinusoidal and random modulators respectively. An increase in the sound pressure level causes a certain decrease in the values of difference limens. This indicates that the detection of increasing changes in signal amplitude is easier for greater values of the sound pressure level.

5.2. FM *difference limens*

In the next part of the investigation, FM difference limens were measured as a function of modulation frequency. The experiments consisted in determining a just noticeable difference between the reference signal characterized by a constant deviation value d_r and the test signal in which the deviation changed from 0 to d_r. The carrier signal was a sinusoid with a frequency of 1000 Hz and 70 dB SPL. First, FM difference limens for a sinusoidal modulator were determined. The results of this experiment were shown in Fig. 5, for the subjects EO, AS and BS respectively. The deviation of the reference signal $d_r = 6$, 12, 25 and 50 Hz is the parameter of the data. As can be seen from the above results, FM differences limens do not clearly depend on modulation frequency.

Next, FM difference limens for a random modulator were determined. The results of these investigations for the subjects EO, AS and BS were shown in Fig. 6. The deviation value of the reference signal d_r is the parameter of the data. In this case, difference limens of frequency modulation are approximately independent of modulation frequency. The data obtained for a random and periodic modulation signal were subjected to analysis of variance ANOVA with the following factors: type of modulator, modulation frequency, subject and deviation of the reference signal.

The effect of deviation of the reference signal was very significant [$F(3,86) = 164$, $p < 0.0001$]. The effect of modulation frequency was not significant [$F(3,86) = 1.88$, $p = 0.14$]. This makes possible the confirmation of an earlier idea that FM difference limens do not depend on modulating frequency in a wide range ($2 \div 128$) Hz. The interval covers the three perception ranges of modulated signals i.e., the following-up range, roughness range and the sidebands separation range. FM modulated signals in these ranges generated fairly different sensations. For this reason the thresholds for

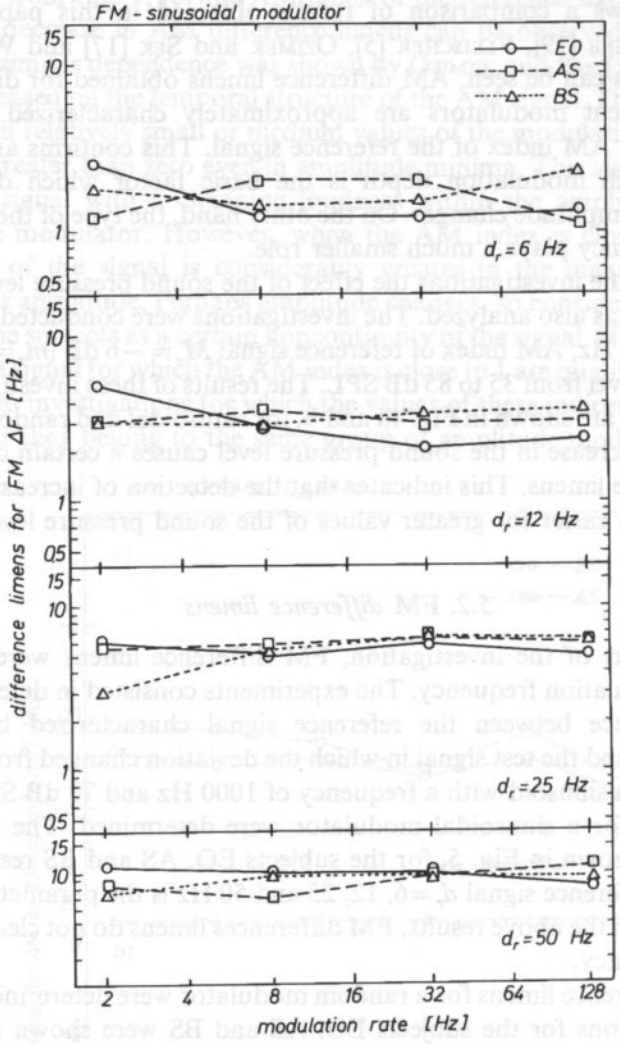


Fig. 5. Dependence of FM difference limens on modulation frequency for a sinusoidal modulator. Deviation of the reference signal is the parameter of the data.

detecting frequency modulation in these ranges are different. On the other hand, as follows from the data, FM difference limens do not depend unequivocally on modulation frequency.

The effect of the type of modulator was not significant [$F(1,86)=0.539$, $p=0.47$]. This suggests that difference limens for both a sinusoidal and random modulator reach approximately equal values. From Figs. 5 and 6 it can be seen that, for the deviation value of the reference signal $d_r=6$ Hz difference limens have values in the range of $2\div2.5$ Hz. A similar situation takes place also for the reference signal deviation values $d_r=12$ and 25 Hz. DLs of FM have in this case values of about $3\div5$

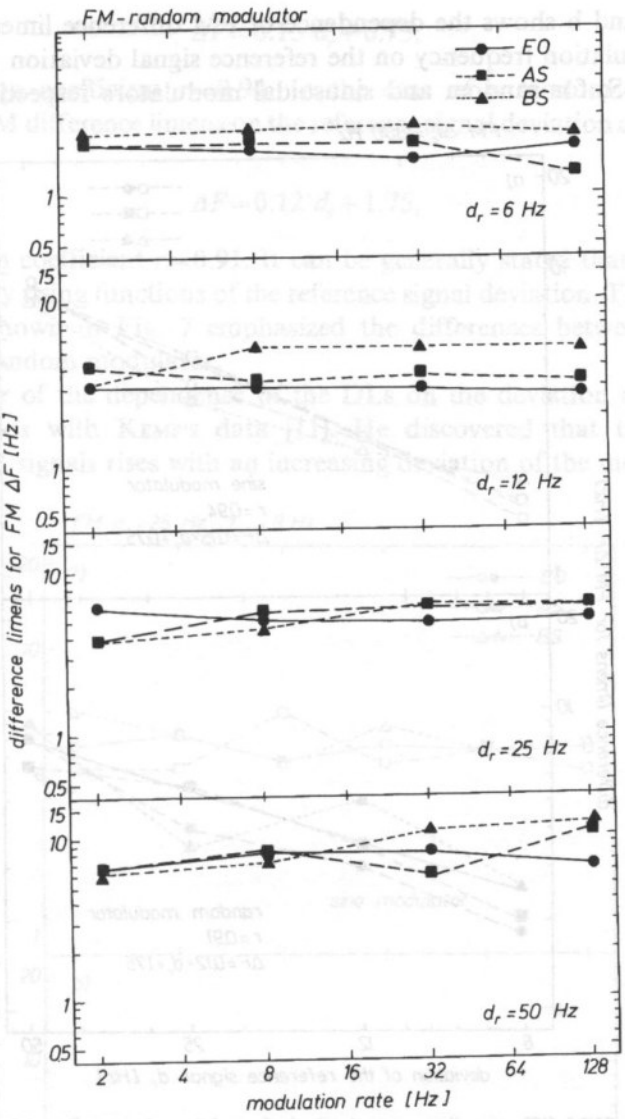


Fig. 6. Dependence of FM difference limens on modulation frequency for a random modulator. Deviation of the reference signal is the parameter of the data.

Hz for both types of modulator. In the case of the greatest value of the reference signal deviation $d_r = 50$ Hz, difference limens reach values in the range of $7 \div 9$ Hz. Thus it can be stated that the character of frequency changes (i.e. whether they are sinusoidal or random) does not significantly affect the discrimination of these changes.

The effect of the subject was not significant [$F(2,86) = 1.17$, $p = 0.321$]. As can be seen from Figs. 5 and 6, the difference limens of FM are scattered more than DLs of AM, particularly in the random modulator case.

Figure 7a and b shows the dependence of FM difference limens averaged with respect to modulation frequency on the reference signal deviation d_r for the subject EO, AS and BS, for random and sinusoidal modulators respectively. The results

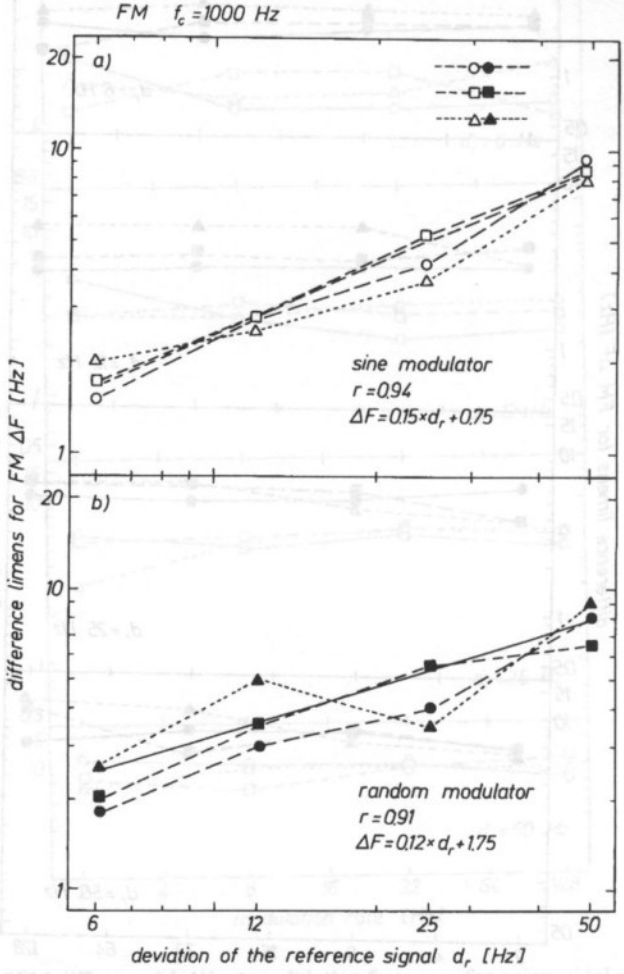


Fig. 7. Dependence of FM difference limens on the deviation of the reference signal for a sinusoidal (a) and a random (b) modulator. A solid line denotes the best linear approximation of the data in accordance with the equations next to the figures.

obtained are compatible for both types of modulator. They show that FM difference limens depend on the deviation of the reference signal. An increase in the reference signal deviation from 6 to 50 Hz causes an increase in difference limens from around 1.5÷2.5 Hz to around 7÷9 Hz. This is a linear dependence because the same increases in the reference signal deviation results similar increases in difference limens. The data were subjected to a regression analysis. In the case of the sinusoidal modulator, the dependence can be expressed as:

$$\Delta F = 0.15\, d_r + 0.75, \tag{16}$$

at the correlation coefficient $r = 0.94$. In the case of a random modulator, the dependence of FM difference limens on the reference signal deviation can be expressed as:

$$\Delta F = 0.12\, d_r + 1.75, \tag{17}$$

at the correlation coefficient $r = 0.91$. It can be generally stated that FM difference limens are linearly rising functions of the reference signal deviation. The way in which the data were shown in Fig. 7 emphasized the differences between the subjects observed for a random modulator.

The character of the dependence of the DLs on the deviation of the reference signal corresponds with KEMP'S data [11]. He discovered that the intensity of roughness of FM signals rises with an increasing deviation of the modulated signal.

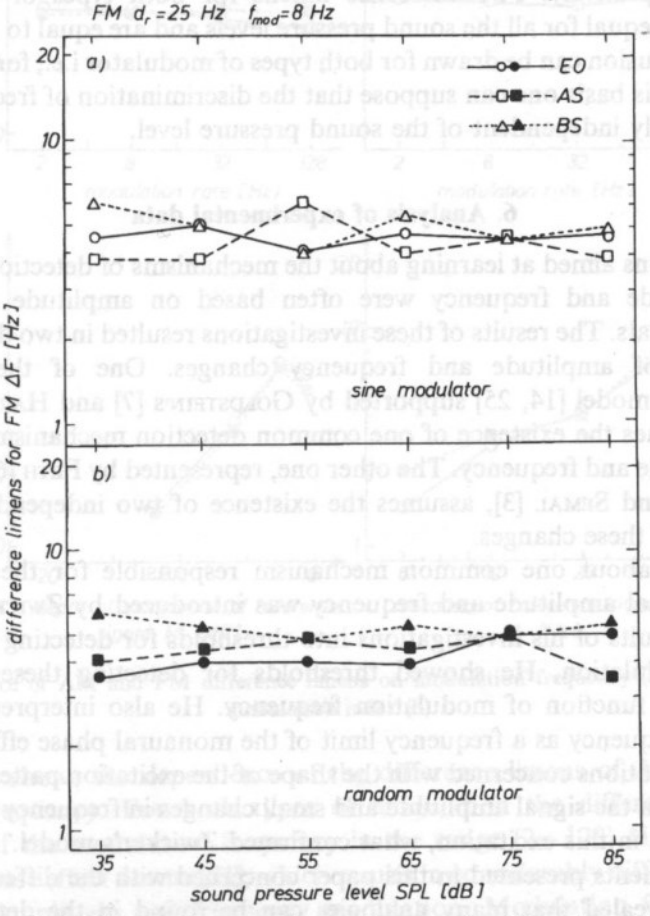


Fig. 8. Dependence of FM difference limens on the sound pressure level for a sinusoidal (a) and a random (b) modulator.

A similar dependence i.e., intensity of roughness as a function of deviation was also observed by TERHARDT [22] with respect to amplitude and frequency modulation. These data are also compatible with the results obtained by OZIMEK and SĘK [18] for a sinusoidal modulator and for a noise with uniform probability distribution.

It should be stressed that the dependence of FM difference limens is analogous to AM difference limens. This can indicates the existence of a very similar discrimination mechanisms of signal amplitude and frequency changes. A similar conclusion was drawn by TERHARDT [22] on the basis of the dependence of the roughness intensity of AM and FM signals on respective modulation indices.

At the last stage of the investigations a dependence of FM difference limens on the sound pressure level was determined. The investigations were conducted for a carrier signal with a frequency of 1000 Hz, reference signal deviation of 25 Hz and modulation frequency of 8 Hz. Then results of these investigations for the subjects EO, AS and BS were shown in Fig. 8a and b, for a sinusoidal and a random modulator respectively. FM difference limens for both types of modulator are approximately equal for all the sound pressure levels and are equal to about 4 – 5 Hz. A similar conclusion can be drawn for both types of modulator i.e., for sinusoidal and random. On this basis one can suppose that the discrimination of frequency changes is approximately independent of the sound pressure level.

6. Analysis of experimental data

Investigations aimed at learning about the mechanisms of detection of changes in signal amplitude and frequency were often based on amplitude and frequency modulated signals. The results of these investigations resulted in two different models of detection of amplitude and frequency changes. One of them, the Zwicker – Maiwald model [14, 25] supported by GOLDSTEIN'S [7] and HARTMANN'S [9, 10] findings, assumes the existence of one common detection mechanism of changes in signal amplitude and frequency. The other one, represented by FETH [6], CONINX [1, 2] and DEMANY and SEMAL [3], assumes the existence of two independent perception mechanisms of these changes.

The claim about one common mechanism responsible for the perception of changes in signal amplitude and frequency was introduced by ZWICKER [25] on the basis of the results of his investigations into thresholds for detecting amplitude and frequency modulation. He showed thresholds for detecting these two types of modulation as function of modulation frequency. He also interpreted the critical modulation frequency as a frequency limit of the monaural phase effect. MAIWALD'S [14, 15] investigations concerned with the shape of the excitation pattern showed that small changes in the signal amplitude and small changes in frequency can cause very similar changes in this excitation, what confirmed Zwicker's model.

The experiments presented in this paper concerned with the difference limens of modulation revealed that many analogies can be found in the detection of supra-threshold amplitude and frequency changes of amplitude or frequency modulated

signals. These analogies suggest that the discrimination of pitch and loudness changes, which was evoked by independent physical parameters of the signal, are governed by similar principles.

Figure 9 shows these analogies schematically. On the lefthand side of the figure the difference limens of amplitude modulation are presented. On the righthand side of the figure the difference limens of frequency modulation are shown. The data in these figures are mean values for the subjects who took part in the experiments.

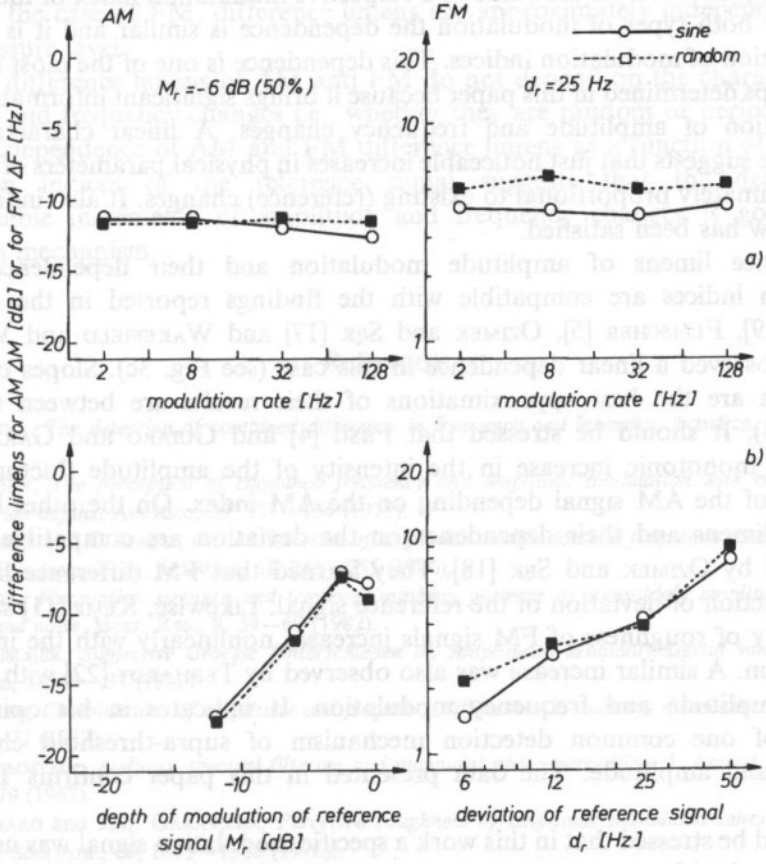


Fig. 9. Dependence of AM and FM difference limens on modulation frequency (a) and a respective modulation index (b).

Figure 9 a shows the dependence of the difference limens of AM and FM on modulation frequency. For both types of modulation the difference limens are independent of the modulation frequency in the range (2÷128) Hz. In this range a monaural phase effect occurs [21], which results in considerably different thresholds for detecting amplitude and frequency modulation. Modulation frequency is the parameter of the modulated signal which determines the rate of changes in the

modulated quantity (i.e., how many times within a time unit the maximal and minimal value of the modulated quantity appears. Since the difference limens are independent of the modulation frequency, one can suppose that the rate of amplitude and frequency changes in AM and FM, respectively, does not play a significant role in the discrimination of these changes. Besides, the discrimination mechanism of amplitude changes is very similar to the discrimination mechanism of frequency changes.

Figure 9 b shows the dependence of the difference limens of amplitude and frequency modulation on the value of a respective modulation index of the reference signal. For both types of modulation the dependence is similar and it is a linearly rising function of modulation indices. This dependence is one of the most important relationships determined in this paper because it brings significant information on the discrimination of amplitude and frequency changes. A linear character of this dependence suggests that just noticeable increases in physical parameters of the signal are approximately proportional to existing (reference) changes. It also indicates that Weber's law has been satisfied.

Difference limens of amplitude modulation and their dependence on the modulation indices are compatible with the findings reported in the literature. SCHOENE [19], FLEISCHER [5], OZIMEK and SĘK [17] and WAKEFIELD and VIEMEISTER [24] also observed a linear dependence in this case (see Fig. 3c). Slopes of straight lines which are the best approximations of their results are between the range $(0.83 \div 1.04)$. It should be stressed that Fastl [4] and GUIARO and GARAVILLA [8] reported a monotonic increase in the intensity of the amplitude fluctuation and roughness of the AM signal depending on the AM index. On the other hand FM difference limens and their dependence on the deviation are compatible with the data found by OZIMEK and SĘK [18]. They learned that FM difference limens are a rising function of deviation of the reference signal. Likewise, KEMP [11] stated that the intensity of roughness of FM signals increases nonlinearly with the increase of the deviation. A similar increase was also observed by TERHARDT [22] with reference to both amplitude and frequency modulation. It indicates in his opinion, the existence of one common detection mechanism of supra-threshold changes of frequency and amplitude. The data presented in this paper confirms Terhardt's conclusion.

It should be stressed that in this work a specific modulation signal was used which was characterized by an amplitude changing randomly from period to period. Hence in an amplitude modulated signal only random changes in modulation depth were observed, whereas in a frequency modulated signal only random changes of deviation were observed. Modulation frequency was constant. As follows from the investigations, the type of modulator (i.e. whether it was random or sinusoidal) did not affect FM and AM difference limens. Values similar to limens obtained for sinusoidal modulation signals were assumed. A similar conclusion can be found in SĘK's [21] paper where thresholds to detect AM and FM for the same types of modulator were determined.

7. Conclusions

The presented results lead to the following conclusions:

1. AM and FM difference limens are rising functions of respective modulation indices. This is compatible with Weber's law and means that just-noticeable changes of modulation depth or frequency deviation are approximately proportional to the existing changes in the reference.

2. AM and FM differenc limens do not depend on modulation frequency.

3. An increase in the sound pressure level causes a slight decrease in AM difference limens. In the case of FM, difference limens are approximately independent of the sound pressure level.

4. The difference limens of AM and FM do not depend on the character of the amplitude and frequency changes i.e., whether they are random or periodic.

5. The dependence of AM and FM difference limens as a function of respective modulation indices of the reference signal suggests that the detection of just-noticeable increments of amplitude and frequency changes is governed by a common mechanism.

References

[1] F. Coninx, *The detection of combined difference in frequency and intensity*, Acustica, **39**, 138–150 (1977).

[2] F. Coninx, *The perception of combined frequency and amplitude modulation with clearly audible modulation depths*, Acustica, **39**, 151–154 (1977).

[3] L. Demany and C. Semal, *On the detection of amplitude modulation and frequency modulation at low modulation frequencies*, Acustica, **61**, 243–255 (1986).

[4] H. Fastl, *Fluctuation strength and temporal masking patterns of sinusoidaly amplitude-modulated broadband noise*, Hear. Res., **8**, 59–69 (1982).

[5] H. Fleischer, *Subjective Groesse Unterschieden im Amplituden-Modulationsgrad von Sinustoenen*, Acustica, **34**, 31–37 (1980).

[6] L.L. Feth, *Combinations of amplitude and frequency differences in auditory discrimination*, Acustica, **26**, 66–77 (1972).

[7] L.J. Goldstein, *Auditory spectral filtering and monaural phase perception*, J. Acoust. Sc. Am., **41**, 458–479 (1967).

[8] M. Guiaro and J.M. Garavilla, *Perceived roughness of amplitude modulated tones and noise*, J. Acoust. Soc. Am., **60**, 1335–1338 (1976).

[9] W.M. Hartmann, *Detection of amplitude modulation*, J. Acoust. Soc., Am., Suppl., **1**, 65, 59 (1978).

[10] W.M. Hartmann, *Detection of frequency modulation at large modulation frequencies*, J. Acoust. Soc. Am., Suppl. 1, 63, p. 66 (1978).

[11] S. Kemp, *Roughness of frequency modulated tones*, Acustica, **50**, 126–137 (1982).

[12] L. Knoch and T. Ekiert, *Modulation and detection* (in Polish), WKŁ, Warszawa 1979.

[13] H. Levitt, *Transformed up-down methods in psychoacoustics*, J. Acoust.Soc.Am., **49**, 467–477 (1971).

[14] D. Maiwald, *Beiehungen zwischen Schallspectrum Mithorschwelle und der Erregung des Gehoers*, Acustica, **18**, 69–80 (1967).

[15] D. Maiwald, *Die Berechnung von Modulationsschwellen mit Hilfe eines Funktionsschemas*, Acustica, 18, 193–207 (1967).

[16] E. OZIMEK and A. SĘK, *Perception of irregular frequency changes in sinusoidal signal*, Acustica, **66**, 146–152 (1988).
[17] E. OZIMEK and A. SĘK, *AM difference limens for noise bands*, Acustica, **66**, 153–166 (1988).
[18] E. OZIMEK and A. SĘK, *Difference limens for FM signals*, Acustica, **71**, 210–217 (1990).
[19] P. SCHOENE, *Messungen zur Schwankungsstoarke von amplitude modulirtensinustonen*, Acustica, **41**, 252–257 (1979).
[20] E. SCHORER, *Critical modulation frequency based on detection of AM and FM tones*, J. Acoust. Soc. Am., **79**, 1054–1057 (1986).
[21] A. SĘK, *Modulation thresholds and critical modulation frequency based on random amplitude and frequency changes*, J. Acoust. Soc. Japan (E) 15, 2, 67–75 (1994).
[22] E. TERHARDT, *On the perception of periodic sound fluctuations roughness*, Acustica, **30**, 201–213 (1974).
[23] E. TERHARDT, *Uber akustische Rauhigkeit und Schwankungsstoarke*, Acustica, **20**, 213–224 (1968).
[24] G.H. WAKEFIELD and N.F. VIEMEISTER, *Discrimination of modulation depth of sinusoidal amplitude modulation SAM noise*, JASA **88**, 3, 1367–1373 (1990).
[25] E. ZWICKER, *Die Grenzen der Hoerbarkreit der Amplitudenmodulation und Frequensmodulation eines Tones*, Acustica, 125–133 (1952).

ARCHIVES OF ACOUSTICS
19, 1, 59–74 (1994)

CRITICAL MODULATION FREQUENCY AND CRITICAL BAND BASED ON RANDOM AMPLITUDE AND FREQUENCY CHANGES

A. SĘK

Institute of Acoustics
Adam Mickiewicz University, Poznań
(60-769 Poznań, Matejki 48/49)

The presented paper deals with determination of the critical modulation frequency (CMF) and critical bandwidth in the case of random amplitude and frequency changes in a sinusoidal signal. The critical modulation frequency is the smallest value of modulation frequency for which thresholds for detecting amplitude and frequency modulation, expressed in appropriate modulation indexes (i.e. m and β), reach the same values. Random amplitude and frequency changes of the simple tone were produced in the amplitude and frequency modulation process by random modulating signals.

The results of the investigations enable to state that the critical modulation frequency is an increasing function of the carrier frequency. It was also shown that psychoacoustically measurable quantities such as detection thresholds, the critical modulation frequency and quantities connected with it (i.e. critical bandwidth, a range of occurrence of a monaural phase effect), do not depend on a modulating signal character (i.e. whether this signal is periodic or random).

1. Introduction

The sensation of hearing caused by a multi-tone all components of which lie within one critical band considerably depends on mutual phase shifts between components of the multi-tone, MATHES and MILLER [9], FLEISHER [2, 3], BUNNEN et al. [1], GOLDSTEIN [6], etc. If one of the components lies outside the specified critical band its initial phase does not significantly affect the sensation caused by the multi-tone. On the other hand, if all components of the multi-tone lie in the range of one critical band, the change of the initial phase of one of them causes a considerable change of the sensation. In the case when the multi-tone is listened to monaurally the mentioned effect is called the monaural phase effect (MPE). So far investigations devoted to the above problem with reference to a three-tone have had an abundant literature and the most of papers deal with the monaural phase effect for sinusoidal signals of the same sound pressure level (see f. ex. MATHES and MILLER [9], FLEISHER [2, 3], BUNNEN et al. [1].)

Researches on the amplitude and frequency modulation (AM and FM) can be an alternative method of investigation of the effect (ZWICKER [24], GOLDSTEIN [6], SCHORER [21]). Spectral structures of AM and FM signals are almost identical when

the frequency modulation index (β) does not exceed the value of 0.3 ($\beta < 0.3$) which takes place on the FM detection threshold. The only difference between AM and FM spectrum is the phase shift of the lower sideband. This difference causes that AM and FM thresholds expressed in appropriate modulation indexes have decidedly different values for low modulatoin frequencies. They asymptotically approach each other with increasing of the modulation frequency. The exceeding of the certain characteristic value of a modulation frequency, so called the critical modulation frequency (CMF), causes that AM and FM thresholds reach identical values. ZWICKER [24] has shown that the CMF is equal to a critical band half — width connected with the carrier frequency. As long as the frequency band of the signal spectrum is no greater than the critical bandwidth (or as long as modulation frequency is less than the critical value) the sidebands' initial phase plays an important role in evaluation of modulated signals CMF. In other words, the CMF is a threshold for detecting phase changes in the modulation frequency domain, ZWICKER [24].

Literature devoted to the CMF has been very scarce so far and most of investigations have been carried out for periodic amplitude and frequency changes. One of the first papers connected with this problem was that by ZWICKER [24] in which the author showed that the critical modulation frequency is equal to the critical band halfwidth. Zwicker's results have been confirmed by GOLDSTEIN [6] who carried out measurements of AM and QFM (quasifrequency modulation) thresholds. One of the recent papers dealing with this problem is that by SCHORER [21]. In this paper the dependence of the critical modulation frequency on the carrier frequency and on the sound pressure level of the signal was discussed. SCHORER [21] found that CMF slightly depends on the sound pressure level and achieves the local minimum for $L = 50$ dB. He also stated that establishing of the width of the critical band on the basis of CMF is a very good method especially for the lowest frequencies of the carrier signal (i.e. no greater than 1–2 kHz). SCHORER [21] has given several reasons for this:

(i) The smallest bandwidth stimulus. The whole stimulus is situated in a very restricted region (in the frequency domain) connected with one critical band and does not excite any other bands;

(ii) Results of measurements should be affected less by variation in the absolute threshold with frequency because stimulus is as narrow as possible. It is difficult to avoid it using a method based on masking;

(iii) Detection of AM and FM takes place on threshold. Thus, off-frequency listening (PATTERSON and NIMMO-SMITH [19]; O'LOUGHLIN and MOORE [16]), even if it takes place, is the smallest possible;

(iv) There are no combination tones (MOORE [10]). Thus the method is very useful for wide range of sound pressure levels of the stimuli.

Previous experiments devoted to CMF were carried out for periodic amplitude and frequency changes. In real signals periodic amplitude and frequency changes occur very seldom and they have a very short duration. Signals such as music, speech, traffic noise are characterized by approximately random changes both in amplitude

and frequency domains. The random character of these changes determines the sensation evoked by an acoustic signal and, as a consequence, influences the detection of the amplitude and frequency changes.

This is the main reason that this paper is devoted to determination of the CMF based on thresholds for detecting random amplitude and frequency changes.

Random changes of these physical parameters were generated in the AM and FM processes using a special class of irregular (random) modulating signals (described wider in Section 2). The establishing of the critical modulation frequency enabled to determine the range in which the monaural phase effect occurs and to determine critical bandwidths for random amplitude and frequency changes.

The spectral and temporal structure of time-varying signals is very important for their perception. Therefore in section 3 of this paper the detailed discussion of a sinusoidal signals modulated by random signals is given.

2. Aim of investigation

The main purpose of the investigations was to determine critical modulation frequency (CMF) and critical bandwidth (CB) based on the thresholds for detecting random amplitude and frequency changes. The author aimed to determine CMF dependence on the carrier frequency and to establish whether CMF depends on a type of the modulation signal. Random changes of amplitude and frequency were obtained in AM and FM processes using the random modulating signals. Furthermore, AM and FM thresholds for a sinusoidal modulating signal (periodic changes of amplitude and frequency) were obtained in order to compare threshold values for both types of the modulating signal.

A simple tone of frequency $f_c = 0.25$, 0.5, 1, 2 and 4 kHz with the sound pressure level of 70 dB SPL was carrier signal. Selection of the modulating signal was an important element in the investigation because according to the assumption it should be a random one. A sinusoidal signal with amplitude randomly changing from period to period in the range $(0 - A_{max})$ was used. Such a signal is characterized by the constant frequency, uniform probability distribution of its amplitude and gaussian distribution of its instantaneous values. Examples of the signals used in the investigation with frequency 4 Hz and corresponding to them distributions of instantaneous values are shown in Fig. 1. For comparison sinusoidal signal and its instantaneous values distribution is also shown in this figure.

A certain factor connected with the dynamics range of amplitude changes, understood as a ratio of its maximum and minimum amplitude is of special importance among parameters describing the signal. In the investigations carried out the value of this coefficient was constant and equal to 34 dB (the quotient of maximum and minimum amplitude value was equal to 50).

For all modulation frequencies (4, 8, 16, 32, 64, 128 Hz) five different time courses of random modulating signal (called R_1, R_2, R_3, R_4, R_5) were used. It was done in

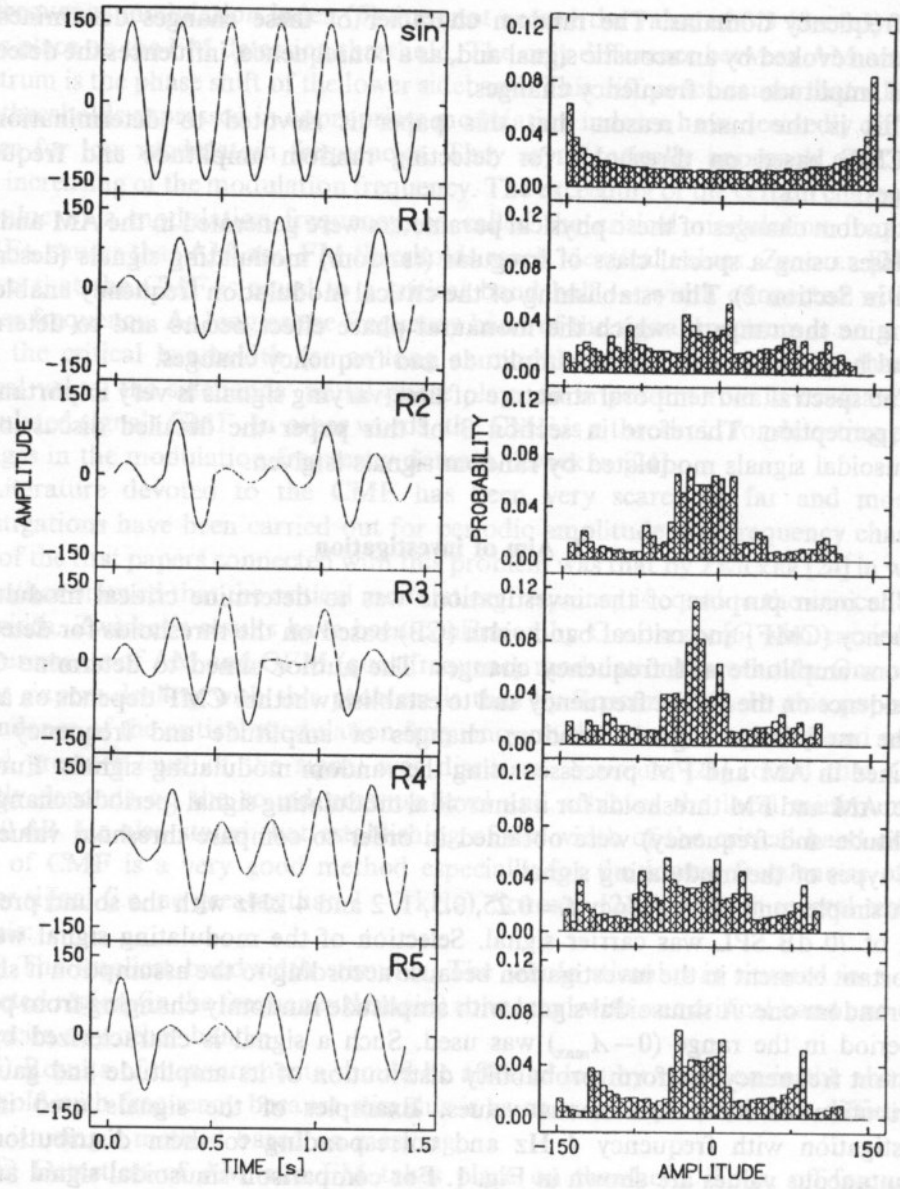


Fig. 1. Temporal structure and distribution of instantaneous values of modulating signals of frequency 4 Hz.

order to establish whether the temporal order of successive maxima and minima of amplitude or frequency changes (in the case of AM and FM, respectively) affects the observed threshold values and consequently whether this order affects the critical modulation frequency. For each value of the modulation frequency modulation signal was calculated separately.

3. Temporal and spectral structure of modulated signals

Analytical considerations aimed to determine the temporal and spectral structure of modulated signals, in the case when simple tones are both the modulating and carrier signal, are reduced to simple trigonometric transformations. Assuming that the carrier signal is of the form:

$$a(t)=A_0\cos(\omega_0 t), \tag{1}$$

and the modulating signal of the form:

$$b(t)=B\sin(\omega_m t), \tag{2}$$

then in the case of amplitude modulation:

$$a_{\mathrm{AM}}(t)=A_0(1+m\sin\omega_m t)\cos\omega_0 t. \tag{3}$$

The spectrum of this signal is of the form:

$$a_{\mathrm{AM}}(t)=\frac{A_0 m}{2}\cos(\omega_0-\omega_m)t+A_0\cos(\omega_0 t)+\frac{A_0 m}{2}\cos(\omega_0+\omega_m)t, \tag{4}$$

where

$$m=\frac{kB}{A_0} \tag{5}$$

is the amplitude modulation index and k is an equipment constant.

In the case of frequency modulation the temporal form of the modulated signal can be expressed as follows:

$$a_{\mathrm{FM}}(t)=A_0\cos(\omega_0 t-\beta\cos(\omega_m t)), \tag{6}$$

where

$$\beta=\frac{\Delta\omega}{\omega_m}=\frac{k_1 B}{\omega_m} \tag{7}$$

is the frequency modulation index, $\Delta\omega$ is the frequency deviation and k_1 is an equipment constant. Assuming, that $\beta\cong 1$, which is true on FM threshold, the spectrum of FM signal could be expresed as follows:

$$a_{\mathrm{FM}}(\mathrm{kt})=-\frac{A_0\beta}{2}\cos(\omega_0-\omega_m)t+A_0\cos(\omega_0 t)+\frac{A_0\beta}{2}\cos(\omega_0+\omega_m)t. \tag{8}$$

Spectra of AM and FM signals are very similar and they are consisted of discrete components, among which the central place is occupied by the component which represents the carrier signal and the remaining components are products of modulation. In the AM case, in addition to the central component, the spectrum consists of

two sidebands with equal positive amplitude value, lying at a distance of $\pm\omega_m$ from the carrier frequency. In the FM case (when $\beta \ll 1$), apart from the component which represents the carrier signal, the spectrum consists of two sidebands with equal amplitudes situated symmetrically with respect to central component at separations of $\pm\omega_m$ in the frequency domain. The sideband with frequency less than the carrier is π phase shifted with respect to the other components. This is the basic difference between AM and FM signals spectrum producing the monaural phase effect which is under investigation.

Spectra determination of a simple tone amplitude or frequency modulated by a random signal, is a problem more complex than was in the case with a tone modulated by another one. If a random modulating signal is a realization of a stationary ergodic process with a normal probability distribution, an autocorrelation function can be determined for a modulated signal, (Knoch and Ekiert (1979)). Using the Wiener-Chińczyn theorem we can determine spectral density of the modulated signal power. This method for determining the spectrum of the modulated signal is troublesome since it dos not permit to obtain a direct measure of modulation intensity. Therefore, in order to determine the spectrum of a simple tone amplitude or frequency modulated by a random signal earlier considerations on modulation of the tone-tone type can be used. Assuming that the modulating signal is a sinusoidal signal with randomly varying amplitude, the amplitude and spectral structure of the modulated signal does not undergo significant changes compared with the case of the tone-tone modulation, with the restriction on that that the amplitude of the modulating signal B from expression (2) is a randomly varying quantity. This affects the values of amplitudes of sidebands of the spectrum of the modulated signal, which are subjected to random changes in accordance with the change of the amplitude of the modulating signal. Nevertheless, the spectral structure of the tonal signal modulated by a sinusoidal signal or a random signal is very similar in these cases.

In the simplest case of the tone-tone modulation, the amplitude modulation index m expressed by equation (5) or the frequency modulation index expressed by equation (7) are the measures of modulation intensity. Quantities defined in this way cannot be directly used in the case of modulation by a random signal since the amplitude of a random signal is unspecified. Therefore, in the AM case, quantity m_{RMS}, expressing the ratio of effective values of modulating and carrier signals (Knoch and Ekiert, 1979), was used as the measure of modulation intensity:

$$m_{\mathrm{RMS}} = K\sqrt{\frac{\sigma^2}{A_o^2}}. \tag{9}$$

In the FM case, on the other hand, quantity β_{RMS} expressing effective index of frequency modulation was used as the measure of modulation intensity:

$$\beta_{\mathrm{RMS}} = k_1\sigma/\omega_m, \tag{10}$$

where σ is the effective value of a random signal.

4. Method

The two-alternative forced choice method (2AFC) with Levitt's adaptive procedure, Levitt [8] was used. The subject listened monaurally to a pair of signals in a random order. One of them was a reference signal — a simple tone without any changes and the second one a test tone i.e. it a sinusoidal — signal. The subject's task was to indicate the modulated signal. Modulation intensity (i.e. m in AM signals or β in FM signals) was increased 1.5 times each time after an incorrect answer or decreased 1.5 times after two successive correct answers. No feedback was used. This procedure tracks the point on the psychometric function corresponding to 71% correct (Levitt, 1971). A single measurement was finished after obtaining 12 turnpoints. The threshold value from a single measurement was calculated as the arithmetic mean value of the modulation index at the last eight turnpoints. The results presented in this paper are mean values of ten separate measurements. Four subjects with an audiologically normal hearing took part in the investigations.

The signals were generated by a computer through a 12-bit digital-to-analog converter. The duration of each signal was 1500 ms, including rise and decay times equal to 100 ms each. A time interval between signals in a pair was 400 ms. During the listening sessions subjects were in an acoustically isolated chamber and they gave answers using the special keyboard.

5. Results of experimental investigations and their analysis

5.1. Thresholds of random amplitude and frequency changes

In the first part of investigations the detection thresholds of AM and FM for a sinusoidal modulating signal were determined. The solid lines in Fig. 2 show the results of experiments for four subjects AS, JT, WR and RP respectively. The investigations were carried out for a sinusoidal carrier signal with frequency $f_c = 1000$ Hz and a sound pressure level $L_c = 70$ dB.

As is shown in Fig. 2 the increase in a modulation frequency between 4 and 64 Hz causes insignificant increase in the thresholds of the amplitude modulation. For a modulation frequency higher than 64 Hz a decrease in the threshold values can be observed. This is associated with crossing of the critical modulation frequency. It is proper to add that it is a very important moment in which the transfer of spectral sidebands beyond the critical band associated with the carrier frequency takes place. Thresholds for detecting AM obtained in this paper are in the good agreement with those obtained by ZWICKER [24] and GOLDSTEIN [6].

Analogically, the thresholds for detecting FM for a sinusoidal modulating signal were measured. The physical parameters of the carrier signal were analogous to those of AM case. The solid lines in the right column of Fig. 2 show the results of these investigations.

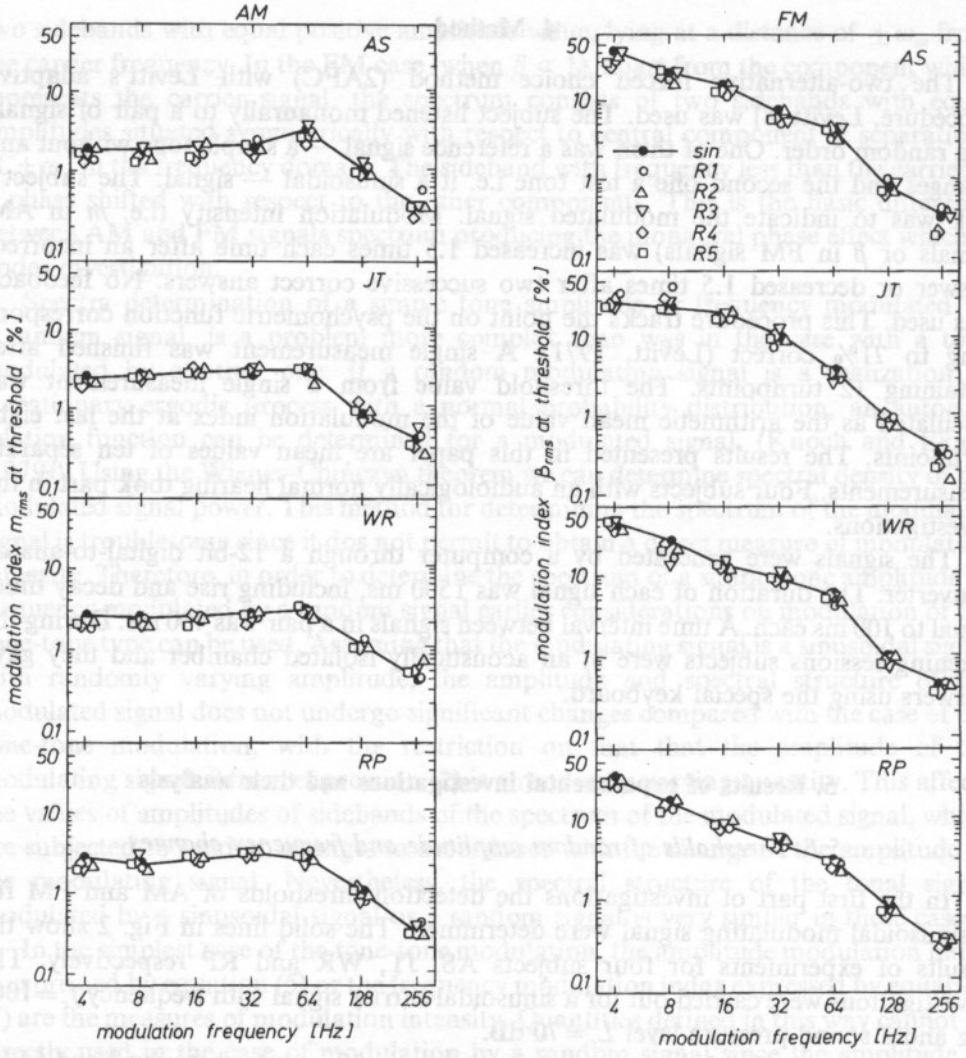


Fig. 2. Thresholds for detecting amplitude and frequency modulation for subjects AS, JT, WR and RP for the carrier frequency of $f_c = 1000$ Hz. The number of the random modulating signal is a parameter of the data.

In the case of FM the increase in modulation frequency causes a regular decrease in the threshold values. The expression of thresholds in the values of the frequency modulation index does not show the existence of any characteristic point connected with the critical modulation frequency. However, this way of data presentation confirms the existence of the nonaural phase effect (it is the range of the modulation frequency where AM and FM thresholds expressed in appropriate indexes differ from each other). When the FM thresholds are expressed as a just noticeably deviation, their dependence on the modulation frequency is completely different (OZIMEK and SĘK [17]), and is similar to those of AM thresholds versus the modulation frequency.

The local maximum which can be seen for the effect described above is connected with the critical modulation frequency (CMF).

The results described in this paper are in agreement for all listeners participating in the experiments. They are also in agreement with results obtained by ZWICKER [24] and GOLDSTEIN [6].

In the main part of the investigation the thresholds for detecting random amplitude and frequency changes of a simple tone were measured. The investigations were carried out for five different realizations of a random modulating signal (see Fig. 1) and their results are shown in Fig. 2. The thresholds for detecting random amplitude changes and random frequency changes versus the modulation frequency are shown in the left and right column of Fig. 2 respectively. The number of the random modulating signal is a parameter of the data. According to the discussion enclosed in Section 3, AM and FM thresholds are expressed in the root mean square values of appropriate modulation indexes i.e. β_{RMS} and m_{RMS}.

The data in Fig. 2 reveal that the thresholds for detecting random amplitude and frequency changes are similar to analogous thresholds for detecting periodic changes of physical parameters of the signal. Furthermore, AM and FM thresholds obtained for all modulating signals used are similar. This means that the number of the random modulating signal does not affect significantly the threshold values. This statement was confirmed by a result of a analysis of variance, that was conducted to this data. There was no significant difference between thresholds for detecting either AM or FM that was obtained for different modulation random signals ($R_1..R_5$). It is worth to point out that this agreement suggests that the thresholds are independent of the temporal structure of the modulated signals. In other words, the temporal order of successively occurring changes of amplitude or frequency does not play an important role in the processes of changes detection. The appropriate modulation index i.e. the amplitude modulation index m in the AM case or the frequency modulation index β in the FM case is the factor which determines detection of any changes.

In the second part of the analysis of variance a comparison between the thresholds obtained for periodic and random changes was made. Taking into account the agreement of the thresholds for detecting either AM or FM for all random modulating signals, the mean values of these thresholds were calculated across the modulators $R_1...R_5$. These mean values were compared with analogous thresholds of the periodic changes. Based on results of this analysis it can be stated that there is no significant difference between thresholds for detecting periodic and random amplitude and frequency changes.

It is worth to point out that the agreement of the threshold values of random and periodic changes was stated for both amplitude and frequency changes. It allows to assume that the thresholds for detecting amplitude and frequency changes produced in the appropriate modulation process (i.e. AM or FM) are independent of the character of the modulating signal. The thresholds expressed in a root mean square values of the appropriate modulation indexes are in agreement for these modulating signals independently of the type of modulating signal.

5.2. The critical modulation frequency

The thresholds for detecting AM and FM were a starting point for calculating CMF. The critical modulation frequency is a characteristic value of the modulation frequency for which the detection thresholds of AM and FM expressed in appropriate modulation indexes (i.e. m and β) reach the same values. In other words, the ratio β/m reach value equal to 1 in a point called CMF.

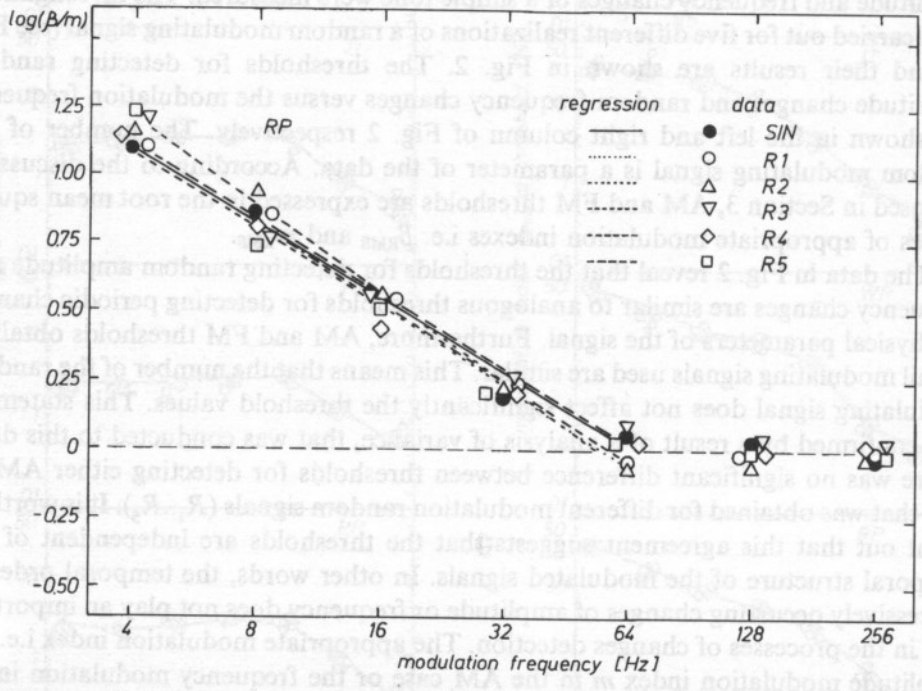


Fig. 3. An example of the curve illustrating the dependence of the β/m quotient on the modulation frequency for the carrier frequency of $f_c = 1000$ Hz. The number of the random modulating signal is a parameter of the data.

In order to obtain the value of CMF for particular carrier frequencies and modulating signals, the thresholds for detecting AM and FM were expressed as β/m quotients versus the modulation frequency. Figure 3 shows an example of the curve obtained for subject RP, for the carrier frequency of $f_c = 1$ kHz. The quotients expressed in logarithmic units are linearly correlated with the modulation frequency, but only in the case when $\beta/m \leqslant 1$. The data satisfying this condition were subjected to a linear regression analysis. As a result of this analysis two coefficients describing the straight line, which was the best approximation of experimental results, were obtained. The values of these coefficients were used for calculation of a value of CMF.

Having the evident form of the function which was the best approximation of the experimental data the values of CMF for both sinusoidal and random modulating

signals were calculated. The CMF values obtained for the same carrier frequency differed from each other for different random modulating signals. This was the reason for which these values of CMF were subjected to an analysis of variance. The main aim of this analysis was to check up whether the CMF values obtained for five different modulating signals do not differ significantly. As the result of this analysis it can be stated that the random modulating signal (from the set of $R_1 ... R_5$) did not significantly influence obtained values of CMF. On this basis one can be stated that the temporal structure of the modulating signal amplitude changes, which determines the temporal order of both amplitude and frequency changes (in the case of AM or FM respectively) does not significantly affect the CMF value. This fact allowed to

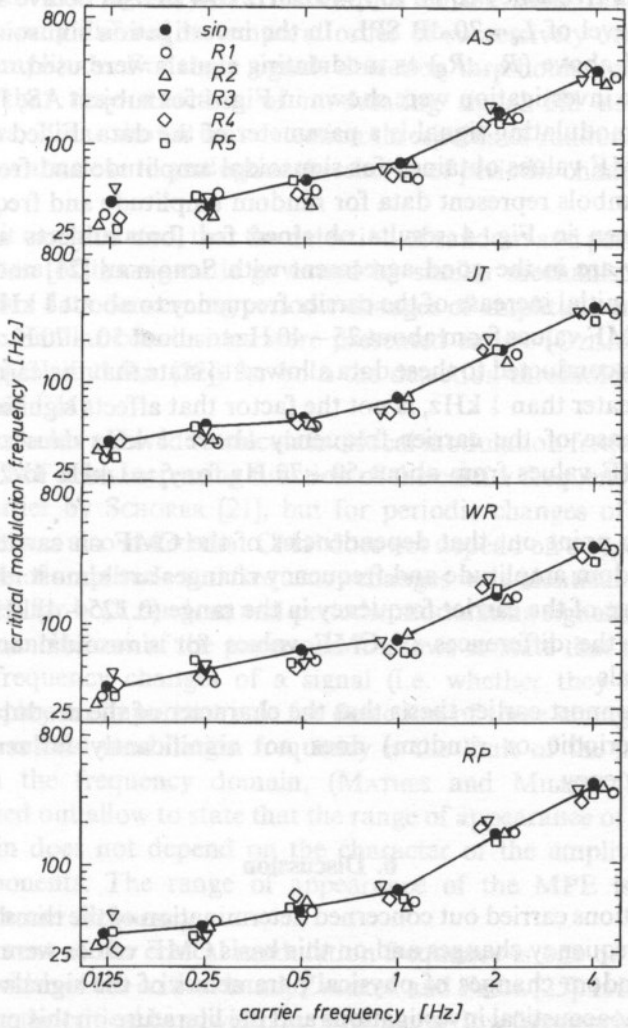


Fig. 4. The dependence of the critical modulation frequency on the carrier frequency for subjects AS, JT, WR, RP respectively. The number of the random modulating signal is a parameter of the data.

average the CMF values obtained for random modulating signals $R_1 ... R_5$ and to make a comparison between the mean value and the CMF value for the sinusoidal modulating signal. As a result of this analysis it can be stated that CMF obtained for periodic and random modulating signals do not differ significantly.

This conclusion enables to state that the character of amplitude and frequency changes of the sinusoidal signal (i.e. whether the changes are periodic or random) does not significantly affect the CMF.

One of the most important relation that was attempted to determine in the investigation was the dependence of the critical modulation frequency on the carrier frequency. Experiments concerned with determination of this relation were carried out for the carrier frequency equal to $f_c = 125$ Hz to 4 kHz in octave steps and for the sound pressure level of $L_c = 70$ dB SPL. In the investigation sinusoidal and random signals described above ($R_1 ... R_5$) as modulating signals were used.

Results of the investigation were shown in Fig. 4 for subject AS, JT, WR and RP. The type of the modulating signal is a parameter of the data. Filled circles and solid line represent CMF values obtained for sinusoidal amplitude and frequency changes whereas open symbols represent data for random amplitude and frequency changes.

As can be seen in Fig. 4 results obtained for four subjects are in the good agreement. They are in the good agreement with SCHORER'S [21] and ZWICKER'S [24] results, too. The initial increase of the carrier frequency to about 1 kHz causes a slight increase of the CMF values from about 25 – 40 Hz to about 50 – 70 Hz. The analysis of variance that was conducted to these data allowed to state that the carrier frequency, if its value is no greater than 1 kHz, is not the factor that affects significantly the CMF values. An increase of the carrier frequency above 1 kHz causes a considerable increase, the CMF values from about 50 – 70 Hz for $f_c = 1$ kHz to 250 – 400 Hz for $f_c = 4$ kHz.

It is worth to point out that dependencies of the CMF on carrier frequency for periodic and random amplitude and frequency changes are almost identical. Thus, it seems, that change of the carrier frequency in the range (0.125 – 4) kHz did not affect considerably on the differences of CMF values for sinusoidal and five random modulation signals.

These facts support earlier thesis that the character of the modulating signal (i.e. whether it is periodic or random) does not significantly influence the critical modulating frequency.

6. Discussion

The investigations carried out concerned determination of the thresholds of random amplitude and frequency changes and on this basis CMF values were calculated. The perception of random changes of physical parameters of the signals is rather a new problem in psychoacoustical investigations and the literature on this problem is scarce. Random amplitude and frequency changes are dominating in the real signals. Periodic changes of these parameters appear very seldom and their duration is very short.

Results of the experimental investigations performed in the paper were obtained for the selected class of the random modulating signals. They were characterized by constant frequency and randomly changing amplitude from period to period. Thus, in the modulated signals the modulation frequency was constant but quantities connected with intensity of modulation were randomly changed (i.e. m in the AM case and β in the FM case).

Detection thresholds of AM and FM as a function of basic parameters of carrier and modulating signals were determined in the first part of the paper. The results of investigation allowed to state that detection thresholds for both random and periodic amplitude and frequency changes are very similar. Five different random modulating signals were used. The signals were realizations of the same random (gaussian) process and they only differed in temporal order of successively occurring maxima and minima of amplitude. For these signals detection thresholds were almost equal. Thus, it seems that time structure of modulating signal did not influence the threshold. Besides, it was shown that detection thresholds of random amplitude and frequency changes are close to analogous thresholds of periodic changes of amplitude and frequency.

On this basis it can be stated, that the detection of random and periodic changes of physical parameters of the signal is governed by similar mechanisms. It allows to generalize literature data concerning periodic changes of amplitude and frequency on random changes. Similar conclusions were presented earlier (OZIMEK and SĘK [18], SĘK and OZIMEK [23], and SĘK [22]) for both the detection thresholds and difference limens of AM and FM.

Obtained thresholds allowed to calculate critical modulation frequency (CMF). It was stated that CMF is an increasing function of the carrier frequency. Such relation was observed earlier by SCHORER [21], but for periodic changes of amplitude and frequency only. It was also stated that CMF does not depend on a type of modulating signal. Thresholds of amplitude and frequency changes were identical to one standard deviation accuracy for both random and periodic modulating signals. It is one of the most important conclusions of the paper and it allows to state that the character of amplitude and frequency changes of a signal (i.e. whether they are periodic or random) do not play an important role for detection of these changes.

However the critical modulation frequency is the limit of the monaural phase effect (MPE) in the frequency domain, (MATHES and MILLER [9]). Results of experiments carried out allow to state that the range of appearance of the MPE in the frequency domain does not depend on the character of the amplitude changes of multi-tone components. The range of appearance of the MPE is an increasing function of the carrier frequency.

The measurement of the critical modulation frequency is one of the methods of determining the width of the critical band, ZWICKER and FASTL [25]. It is a very suitable tool for estimating the critical bandwidth in particular at low frequencies not exceeding 2 kHz SCHORER [21], ZWICKER and FASTL [25]. Thus, on the basis of the CMF values, the widths of the critical band for all carrier frequencies were calculated. Figure 5

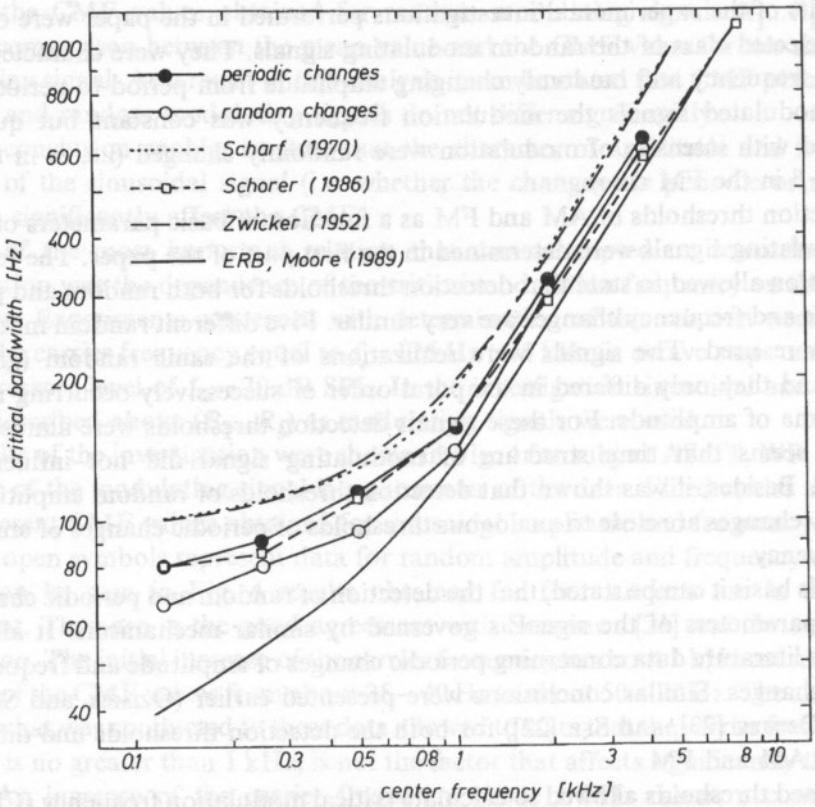


Fig. 5. The comparison between critical bandwidth obtained on the basis of the CMF values calculated for periodic and random amplitude and frequency changes and data presented in literature.

summarizes results for the periodic modulating signal (averaged for four subjects) and for random modulating signals (averaged for signals $R_1 ... R_5$ and for subjects). Results obtained by ZWICKER [24], SCHARF [20] and SCHORER [21] are also shown in Fig. 5.

This comparison of selected literature data with results included in the paper allows to state that the critical bandwidths obtained in the experiments are fully compatible with SCHORER'S [21] results, because in both paper the same method was used. As can be seen results of critical bandwidths presented are slightly smaller than results of ZWICKER and FASTL [25] and SCHARF [20] especially for carrier frequencies smaller than 1 kHz. Results obtained by ZWICKER and FASTL [25] and SCHARF [20] are mean values calculated on the basis of several methods of determining the critical bandwidth.

The critical bandwidth is a very important element of the several models of the perception of acoustic signals, (MAIWALD [15], GOLDSTEIN [6], FLORENTINE and BUUS [4]). It is usually identified as the width of the auditory filter. An alternative attempt to those models is the MOORE and GLASBERG [11] model. An essential element of this

model is auditory filter whose equivalent rectangular bandwidth (ERB) can be calculated based on the formula described by MOORE and GLASBERG [11]. In Fig. 5 the ERB values were also shown as a function of their center frequency. As can be seen for the lowest frequency the ERB is almost by a factor of 2 smaller than the critical bandwidth calculated based on the values of the critical modulation frequency obtained in the paper.

It is worth to point out that using random modulating signal did not influence considerably the critical bandwidth. Similar statement can be found in SĘK and OZIMEK's [23] and SĘK's [22] papers concerned with thresholds for detecting AM and difference limens of modulation. The base measurable quantities such as detection thresholds of amplitude changes (SĘK [22]), detection thresholds of frequency changes (OZIMEK and SĘK [17]), difference limens of AM and FM (SĘK and OZIMEK [23]) do not depend on the character of the modulating signal (i.e. whether it is random or periodic). The above conclusions enable to confirm earlier thesis that the kind of the signal physical parameter changes does not significantly influence the perception of acoustic signals. Thresholds, difference limens, critical modulation frequency and quantities connected with CMF do not depend on the character of amplitude and frequency changes.

7. Conclusions

The results of the investigations enable the formulation of the following conclusions:

1. Threshold for detecting random amplitude and frequency changes are very similar to the analogous thresholds of periodic changes.

2. The critical modulation frequency (CMF) is an increasing function of the carrier frequency.

3. Psychoacoustically measurable quantities such as the detection thresholds, difference limens of modulation, the critical modualtion frequency and quantities connected with it (i.e. the critical bandwidth, the range of occurrence of the monaural phase effect) do not depend on the character of the modulating signal i.e. whether it is periodic or random.

4. The detection of both random and periodic changes of the signal amplitude and frequency is governed by a very similar mechanisms.

The work was supported by State Committee for Scientific Research, project number 2 0910 91 01.

References

[1] T.J.F. BUNNEN, J.M. FESTEN, F.A. BILSEN, and G. VAN DEN BRINK, *Phase effects in three component signal*, J. Acoust. Soc. Am., **55**, 297–303 (1974).

[2] H. FLEISCHER, *Gerade wahrnehmbare Phasenanderungen bei 3-Ton Komplexen*, Acustica, **32**, 44–55 (1975).

[3] H. FLEISCHER, *Über die Wahrenehmbarkeit von Phasenanderungen*, Acustica, **35**, 202–209 (1976).
[4] M. FLORENTINE and S. BUUS, *An excitation-pattern model for intensity discrimination*, J. Acoust. Soc. Am., **70**, 1646–1654 (1981).
[5] B.R. GLASBERG and B.C.J. MOORE, *Derivation of auditory filter shapes from notched-noise data*, Hear. Res., **47**, 103–138 (1990).
[6] L.J. GOLDSTEIN, *Auditory spectral filtering and monaural phase perception*, J. Acoust. Soc. Am., **41**, 458–479 (1967).
[7] W.M. HARTMANN, G.H. HNATH, *Detection of mixed modulation*, Acustica, **50**, 297–312 (1982).
[8] H. LEVITT, *Transformed up-down methods in psychoacoustics*, J. Acoust. Soc. Am., **49**, 467–477 (1971).
[9] R.C. MATHES and R.L. MILLER, *Phase effects in monaural perception*, J. Acoust. Soc. Am., **19**, 780–797 (1949).
[10] B.C.J. MOORE, *An introduction to psychology of hearing*, Academic Press, London 1989.
[11] B.C.J. MOORE and B.R. GLASBERG, *Suggested formulae for calcualting auditory-filter bandwidths and excitation patterns*, J. Acoust. Soc. Am., **74**, 750–753 (1983).
[12] B.C.J. MOORE and B.R. GLASBERG, *Formulae describing frequency selectivity as a function of frequency and level and their use in calculating excitation patterns*, Hear. Res., **28**, 209–225 (1987).
[13] B.C.J. MOORE and B.R. GLASBERG, *Difference limens for phase in normal and hearing-impaired subjects*, J. Acoust. Soc. Am., **86**, 1351–1365 (1989).
[14] B.C.J. MOORE, A. SĘK, *Detection of combined frequency and amplitude modulation*, J. Acoust. Soc. Am., **92**, 3119–3131 (1992).
[15] D. MAIWALD, *Ein Functionsschemades Gehors zur Beschreibung der Erkennbarkeit kleiner Frequenz und Amplitudenanderungen*, Acustica, **18**, 81–92 (1967).
[16] B.J. O'LOUGHLIN and B.C.J. MOORE, *Off-frequency listening: effects on psychoacoustical tuning curves obtained in simultaneous and forward masking*, J. Acoust. Soc. Am., **69**, 1119–1125 (1981).
[17] E. OZIMEK and A. SĘK, *Perception of irregular frequency changes in sinusoidal signal*, Acustica, **66**, 146–152 (1988).
[18] E. OZIMEK and A. SĘK, *Difference limens for FM signals*, Acustica, **71**, 210–217 (1990).
[19] R.D. PATTERSON and I. NIMMO-SMITH, *Off-frequency listening and auditory filter asymmetry*, J. Acoust. Soc. Am., **67**, 229–245 (1980).
[20] B. SCHARF, *Critical bands*, in: *Foundations of modern auditory*, vol. 1 [Ed.] J.V. Tobias Academic Press, New York 1970.
[21] E. SCHORER, *Critical modulation frequency based on detection of AM and FM tones*, J. Acoust. Soc. Am., **79**, 1054–1057 (1986).
[22] A. SĘK, *Perception of irregular amplitude changes of sinusoidal signals*, Acustica, **77**, 262–269 (1993).
[23] A. SĘK and E. OZIMEK, *AM and FM difference limens*, Arch. Acoustics, **19**, 1 (1994).
[24] E. ZWICKER, *Die Grenzen der Horbarkeit der Amplitudenmodulation und Frequensmodulation eines Tones*, Acustica, **2**, 125–133 (1952).
[25] E. ZWICKER and H. FASTL, *Psychoacoustics. Facts and models*, Springer Verlag, Berlin 1990.
[26] E. ZWICKER and A.A. JAROSZEWSKI, *Inverse frequency dependence of simultaneous tone-on-tone masking patterns at low levels*, J. Acoust. Soc. Am., **71**, 1508–1512 (1982).

ARCHIVES OF ACOUSTICS
19, 1, 75—83 (1994)

PROBLEM OF EXCESS ATTENUATION IN ACOUSTIC MEASUREMENTS OF GAS BUBBLE CONCENTRATION IN THE SEA

J. SZCZUCKA
Institute of Oceanology,
Polish Academy of Sciences
(81-712 Sopot, ul. Powstańców Warszawy 55)

This study is aimed at describing the possible errors in the acoustie estimation of gas buble concentration in the sea made by the assumption of the first order scattering approximation, i.e. by neglecting the interaction between individual bubbles enclosed in the water column. An experimental method of counting the bubbles is sketched, results are presented and two different ways of taking into account the excess attenuation in bubbly medium are used for correcting the bubble number, the results being compared. The problem of bubble swarms in the vicinity of the transmitter is also considered.

1. Introduction

Hydroacoustic methods based on the phenomenon of backscattering of the acoustic energy in the sea water are widely applied in the oceanic investigations, especially in estimating plankton and fish biomass. It is usually assumed that the number of scattering objects is small enough to enable the neglecting of the interaction between scatterers, what is connected with the neglecting of the excess attenuation and multiple scattering. In the majority of cases such assumption is valid, but this problem should be considered in each individual case. Usually the determination of gas bubble populations in the sea water is also based on the first order scattering approximation and this simplification can produce some significant errors.

2. Sound backscattering as a method of counting bubbles

One of the methods of determination of gas bubble density in the sea water is based on the interrelation between the measurable volume backscattering strength and the number of resonating bubbles [1], [3]. Sound pulse coming upon the single bubble makes it radiate the secondary spherical wave which returns to the receiver as the backscattered sound. Because bubbles are randomly distributed in the water column, scattering is assumed to be incoherent — the total intensity of backscattered sound is a sum of intensities originating from the individual bubbles. If the number of bubbles in the medium is small (i.e. the average distance between them is large in comparison with their size), it can be assumed that backscattered wave hits

immediately the receiver without being scattered and attenuated by other bubbles. Then the sonar equation for volume backscattering has a well-known simple form [1]:

$$I_{bs} = \frac{I_0}{(z-z_0)^2} S_{bs} \frac{c\tau\psi_D}{2}, \tag{1}$$

where I_{bs} — intensity of the backscattered signal (echo intensity), I_0 — intensity of the emitted signal on the acoustic axis at a distance 1 m from the source, $c\tau$ — spacial length of a pulse, ψ_D — integrated beam width factor, z_0 — depth of transducer, z — actually penetrated depth, S_{bs} — volume backscattering coefficient.

This formula is traditionally used for the determination of the number n of scattering objects, provided that the backscattering cross section σ_{bs} of the individual scatterer is known, $S_{bs}=n\sigma_{bs}$.

If, on the other hand, the number of bubbles is so large that multiple scattering can not be neglected, the sonar equation must include the attenuation factor:

$$I_{bs} = \frac{I_0}{(z-z_0)^2} S_{bs} \frac{c\tau\psi_D}{2} \exp\left(-2\int_{z_0}^{z} S_e(z)\,dz\right), \tag{2}$$

where S_e is the volume extinction (scattering + absorption) coefficient.

In bubbly medium the backscattering coefficient S_{bs} and the extinction coefficient S_e are expressed in the following way [1]:

$$S_{bs} = \frac{\pi}{2} \frac{n(a_R, z)\, a_R^3(z)}{\delta_R}, \tag{3}$$

$$S_e = \frac{2\pi^2 n(a_R, z)\, a_R^3(z)}{\delta_{rR}} = \frac{2\pi^2 n(a_R, z)\, a_R^2(z)}{k}, \tag{4}$$

where a_R — resonant bubble radius (for air bubbles in water a_R [μm]$=3280(1+0.1z)^{1/2}/f$[kHz]), f — incident sound frequency, δ_R — resonant damping constant of individual bubble, $\delta_{rR}=ka_R$ — resonant damping constant due to reradiation only, k — wave number, $n(a_R, z)$ — number of resonant bubbles of radius between a_R and a_R+1μm in 1 m^3 of water at the depth z.

It can be seen (Eq. 1) that the measurement of the echo intensity allows to determine the backscattering coefficient S_{bs} as a function of depth. Subsequently, knowing the values of S_{bs}, we can deduce the values of $n(a_R,z)$ calculating previously values of $a_R(z)$ and $\delta_R(z)$ (Eq. 3). Damping constant δ can be easily evaluated; it depends on the frequency of incident sound, bubble radius, depth and some physico-chemical gas and water constants [1].

3. Results of experiment

Our measurements of gas bubble concentration were carried out in the subsurface layer of the coastal zone of the Baltic Sea [3]. By sounding at frequencies 80, 63, 50

and 40 kHz we could infer vertical profiles of the number of bubbles with radii 40–100 μm. Evaluation of gas bubble density was based on the equations (1) and (3) according to the assumption of single scattering. The data recorded in the shallowest layer $[z_0, z_1]$ were useless because of ringing in the trasmitting-receiving transducer; the data from the greater depths ($z > z_N$) were also useless, because the noise level was reached (at different depth z_N for different frequencies), so only the midrange data from the interval $[z_1, z_N]$ were considered. Fig. 1 shows the obtained number of bubbles resonating at

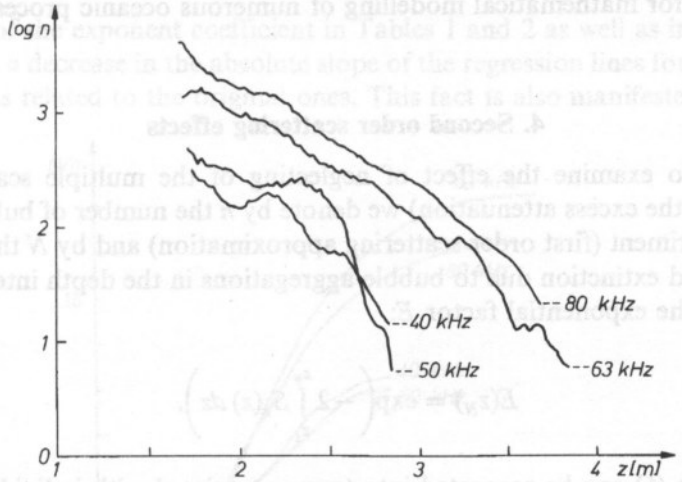


Fig. 1. Vertical profiles of the concentration of gas bubbles resonating at various frequencies used in experiment

different frequencies as a function of depth. It displays the increase of bubble density when the bubble size decreases (bubble radius is inversely proportional to the incident sound frequency). Taking values $n(z)$ for the given frequency and chosen depth levels one can find the dependence of bubble number on the depth of their occurrence $n(z) = A\exp(Bz)$. The coefficients A and B depend on the productivity of bubble source and on the intensity of turbulent mixing in the subsurface water layer connected with the dynamical meteorological conditions. By the linear regression method we have found the following values of the coefficients A and B.

Table 1

f [kHz]	A	B
40	$2.4 \cdot 10^4$	−2.40
50	$1.2 \cdot 10^5$	−2.95
63	$2.0 \cdot 10^5$	−2.57
80	$2.1 \cdot 10^5$	−2.34

These regression lines are depicted as curves 1 in Fig. 5.

Choosing the values of $n(f,z)$ for particular frequencies and selected depth levels, it is possible to determine a functional dependence of the number of bubbles both on the depth and on the bubble size. Approximation by a linear two-variables regression method gave the following relationship:

$$n(a,z) = 2.35 \cdot 10^{11} a^{-3.67} \exp(-2.28z), \tag{5}$$

where the radius a was expressed in micrometres, the depth z in metres and $n(a,z)$ was the number of bubbles of radii $[a,a+1\ \mu m]$ in 1 m^3 of water. Forumula of such type is indispensable for mathematical modelling of numerous oceanic processes involving gas bubbles.

4. Second order scattering effects

In order to examine the effect of neglecting of the multiple scattering (and, consequently, the excess attenuation) we denote by n the number of bubbles inferred from the experiment (first order scattering approximation) and by N the **real** bubble number. Sound extinction due to bubble aggregations in the depth interval $[z_1,z_N]$ is described by the exponential factor E:

$$E(z_N) = \exp\left(-2\int_{z_1}^{z_N} S_e(z)\,dz\right). \tag{6}$$

The integral in (6) can be separated into terms associated with individual sublayers determined by the sampling frequency f_s ($f_s = 45$ kHz, $\Delta t = f_s^{-1}$, $\Delta z = c\Delta t/2 = 1.62$ cm). If we assume that S_e is constant in each individual sublayer, the integration can be replaced by the summation. By inserting the expression (4) for S_e we obtain:

$$\int_{z_1}^{z_N} S_e(z)\,dz = \int_{z_1}^{z_2} S_e(z)\,dz + \ldots + \int_{z_{N-1}}^{z_N} S_e(z)\,dz = \frac{2\pi^2}{k}\Delta z \sum_{j=1}^{N-1} N(z_j)\,a_R^2(z_j). \tag{7}$$

At the beginning we know nothing about the attenuation and we infer the number of bubbles $n(z_1)$ using the first order scattering approximation; then we use $N(z_1) = n(z_1)$ to calculate the attenuation factor $E(z_1)$ in the first sublayer, which influences the determination of the bubble number in the second one. In the next stage we correct the measured concentration $n(z_2)$ by the attenuation factor from the preceding minilayer according to the general formula

$$N(z_i) = E^{-1}(z_{i-1})\,n(z_i), \tag{8}$$

and we calculate the attenuation in the second sublayer, i.e. the correction factor for the next minilayer. Such adaptive processing is conducted for each next sublayer up to the depth z_N. Corrected bubble concentrations can be described by the curves $N(z) = A_1 \exp(B_1 z)$ with the following coefficients:

Table 2

f [kHz]	A_1	B_1
40	$1.5 \cdot 10^4$	−2.13
50	$8.4 \cdot 10^4$	−2.71
63	$1.6 \cdot 10^5$	−2.34
80	$1.7 \cdot 10^5$	−2.13

The formula analogous to (5) takes the form

$$N(a,z)=3.26 \cdot 10^{11} a^{-3.82} \exp(-2.03z) \tag{9}$$

Comparison of the exponent coefficient in Tables 1 and 2 as well as in formulae (5) and (9) shows a decrease in the absolute slope of the regression lines for the corrected concentrations related to the original ones. This fact is also manifested in Fig. 5.

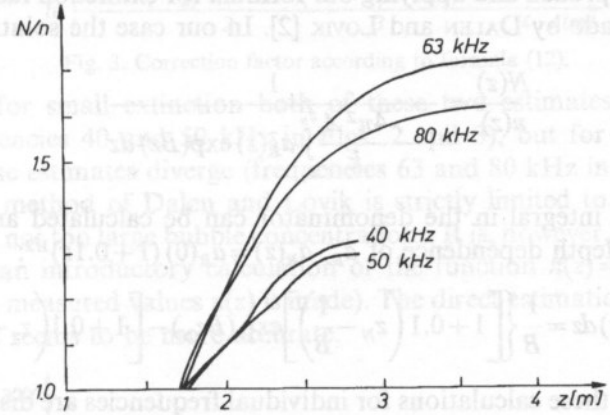


Fig. 2. Correction factor for the bubble number obtained directly from experiment.

The dependence of the correction factor E^{-1} on depth is depicted in Fig. 2. Its value for $f=63$ kHz is higher than for $f=80$ kHz, and for $f=40$ kHz it is higher than for $f=50$ kHz. It is connected with fact that the attenuation coefficient S_e depends both on the bubble density, which rises with increasing frequency, and on the bubble size, which falls (Eq. 4). Bubble concentrations for the frequency pairs mentioned above differ less than the third power of the respective bubble radius and this is the reason for the observed effect.

Another approach to the problem of neglecting of the multiple scattering is presented by DALEN and LOVIK [2]. They simplify this problem by assuming $S_e=N\sigma_e$, what gives the attenuation factor in the form

$$\exp\left(-2\int_0^z N\sigma_e\,dz\right),$$

where σ_e — extinction cross-section of the individual resonant bubble. Formula (8) leads to the following equation:

$$N(z)=n(z)\exp\left(2\int_0^z N\sigma_e\,dz\right). \tag{10}$$

Differentiating (10) with respect to z and multiplying both sides of this expression by n leads to the Bernoulli equation. Its general solution shows that the real number of bubbles N depends on the functional form of n deduced from the mesaurement. Applying the formula $n(z)=A\exp(Bz)$ and assuming that σ_e is constant provides the particular solution:

$$\frac{N(z)}{n(z)}=\frac{1}{1-2\sigma_e A\int_0^z \exp(Bz)\,dz}=\frac{1}{1-\dfrac{2\sigma_e A}{B}[\exp(Bz)-1]}. \tag{11}$$

Following this approach and applying our formula for extinction factor (6) we avoid simplifications made by DALEN and LOVIK [2]. In our case the solution is:

$$\frac{N(z)}{n(z)}=\frac{1}{1-\dfrac{4\pi^2 A}{k}\int_{z_1}^{z_N} a_R^2(z)\exp(Bz)\,dz}. \tag{12}$$

The value of the integral in the denominator can be calculated analytically using additionally the depth dependence of a_R, $a_R(z)=a_R(0)\,(1+0.1z)^{1/2}$:

$$\int_{z_1}^{z_N}(1+0.1z)\exp(Bz)\,dz=\frac{1}{B}\left\{\left[1+0.1\left(z_N-\frac{1}{B}\right)\right]\exp(Bz_N)-\left[1+0.1\left(z_1-\frac{1}{B}\right)\right]\exp(Bz_1)\right\}.$$

The results of these calculations for individual frequencies are displayed in Fig. 3.

Because of the exponential decrease of the bubble number with increasing depth, the ratio N/n reaches the terminal value at some depth dependent on the intensity of turbulent mixing (expressed by the coefficient B). For higher frequencies used in experiment the z_N value is large enough to make this phenomenon easily noticeable (Fig. 3). A linear regression curves $N(z)=A_2\exp(B_2 z)$ computed for values obtained from last estimation have the following coefficients A_2 and B_2 coefficients A_2 and B_2:

Table 3

f [kHz]	A_2	B_2
40	$1.6\cdot 10^4$	−2.14
50	$9.0\cdot 10^4$	−2.74
63	$1.7\cdot 10^5$	−2.34
80	$1.7\cdot 10^5$	−2.13

A two-variables regression gives

$$N(a,z)=3\cdot 10^{11}a^{-3.80}\exp(-2.03z). \tag{13}$$

Since the function (12) has the form $1/(1-x)$, its physical sense is limited to the region $x<1$. It is noteworthy that for $x \ll 1$ the function $\exp(x)$, i.e. N/n in formula (8), and the function $1/(1-x)$, i.e. N/n in formula (12), have the same first order Taylor expansion:

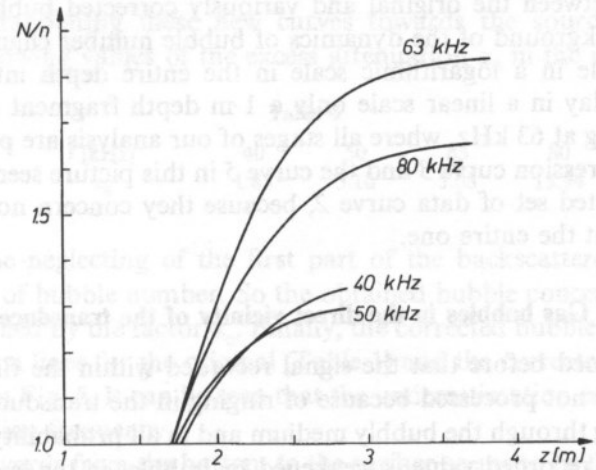


Fig. 3. Correction factor according to formula (12).

$1+x$. In fact, for small extinction both of these two estimates are very similar (compare frequencies 40 and 50 kHz in Figs. 2 and 3), but for greater values of attenuation these estimates diverge (frequencies 63 and 80 kHz in Figs. 2 and 3). It means that the method of Dalen and Lovik is strictly limited to the case of small attenuation, i.e. not too large bubble concentrations. It is, however, very simple in use (provided that an introductory calculation of the function $n(z)=A\exp(Bz)$ on the basis of directly measured values $n(z)$ is made). The direct estimation method is more laborious but it seems to be more accurate.

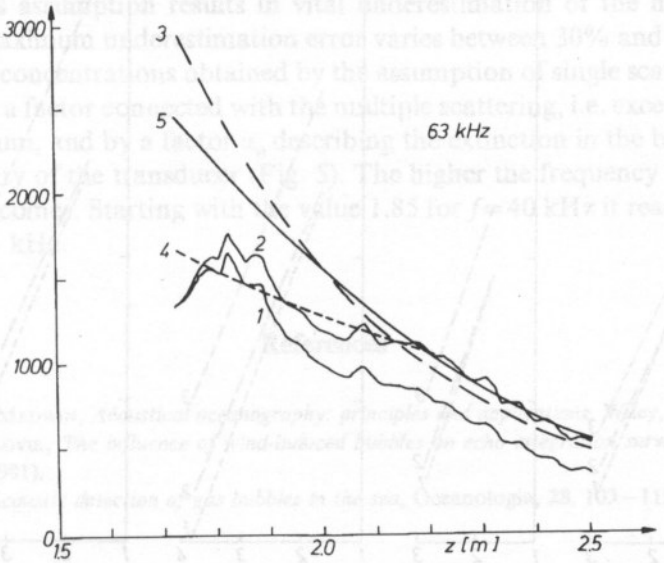


Fig. 4. Number of bubbles resonating at frequency 63 kHz in a chosen small interval of depth, (*1*) — obtained from experiment, (*2*) — after direct correction (Eqs. 6–8) and its (*3*) — total regression and (*4*) — local regression, (*5*) — corrected according to Eq. (12).

Differences between the original and variously corrected bubble concentration curves on the background of the dynamics of bubble number changeability are too slight to be visible in a logarithmic scale in the entire depth interval. So, as an example, we display in a linear scale only a 1 m depth fragment of the density of bubbles resonating at 63 kHz, where all stages of our analysis are presented (Fig. 4). Both the total regression curve *3* and the curve *5* in this picture seem not to be fitted well to the corrected set of data curve *2*, because they concern not only this small depth interval, but the entire one.

5. Gas bubbles in the direct vicinity of the transducer

It was mentioned before that the signal recorded within the first few hundreds microseconds was not processed because of ringing in the transducer. Nevertheless, this signal is going through the bubbly medium and in all probability it is attenuated, so in fact the first recorded value is weakened by bubbles in the layer nearest to the transmitter. Magnitude of this extinction is unknown but it should be significant because of bubble abundance in this area. It can be evaluated by extrapolation of the function $N(z)=A_1 \exp(B_1 z)$ up to the source depth and calculation of the sound attenuation in the interval $[z_0, z_1]$. We can see from Fig. 4, however, that total regression line in the shallowest layer runs remarkably higher than the experimental curve, so an additional regression computation for the first 0.5 m is needed. In fact such local regression gives more realistic values of bubble concentration in this region

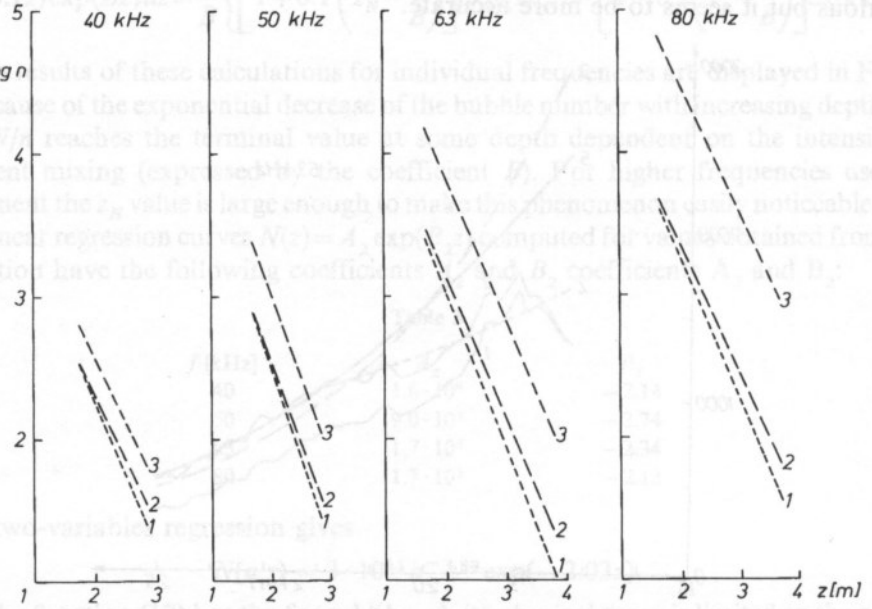


Fig. 5. Regression lines for concentrations of bubbles resonating at frequencies used in experiment: (*1*) obtained directly from the first-order scattering experiment, (*2*) corrected according to Eqs. (6–(8) and (*3*) finally corrected by a factor α_0 describing the extinction in the direct vicinity of the transducer.

(Fig. 4). By extrapolating these new curves towards the source depth we have obtained the following values of the excess attenuation α_0 in the interval $[z_0, z_1]$:

Table 4

f [kHz]	40	50	63	80
α_0	1.85	3.10	5.08	13.94

It means that the neglecting of the first part of the backscattered signal leads to underestimation of bubble number. So the obtained bubble concentration functions should be multiplied by the factor α_0. Finally, the corrected bubble densities together with the regression lines for the original (Table 1) and the corrected (Table 2) sets of data are shown in Fig. 5. It can be seen that the underestimation error increases with increasing sounding frequency.

Sounding upwards from the bottom to the surface seems to be the best solution to this problem. It allows us to avoid the problem of rather dense layer of bubbles occuring in the direct vicinity of the transmitter that is not considered because of the technical break caused by switching the working mode from transmitting to receiving.

6. Summary

To sum up, we have obtained some experimental results concerning gas bubble populations in the coastal zone of the Baltic Sea. The problem of using the first order approximation (neglecting of the multiple scattering) has been considered, and it was shown that this assumption results in vital underestimation of the number of gas bubbles. The maximum underestimation error varies between 30% and 73% (Fig. 2).

Gas bubble concentrations obtained by the assumption of single scattering should be corrected by a factor connected with the multiple scattering, i.e. excess attenuation in bubbly medium, and by a factor α_0 describing the extinction in the bubble layer in the direct vicinity of the transducer (Fig. 5). The higher the frequency is, the greater the factor α_0 becomes. Starting with the value 1.85 for $f=40$ kHz it reaches the value 13.94 for $f=80$ kHz.

References

[1] C.S. CLAY, H. MEDWIN, *Acoustical oceanography: principles and applications*, Wiley, New York 1977.

[2] J. DALEN, A. LOVIK, *The influence of wind-induced bubbles on echo integration surveys*, JASA, **69**, 6, 1653–1659, (1981).

[3] J. SZCZUCKA, *Acoustic detection of gas bubbles in the sea*, Oceanologia, **28**, 103–113 (1989).

ARCHIVES OF ACOUSTICS
19, 1, 85–108 (1994)

ACOUSTIC SIGNATURES OF ORGANIC FILMS FLOATING ON THE SEA SURFACE

S. J. POGORZELSKI
Environmental Acoustics Laboratory
University of Gdańsk

A spectrum of low-frequency amplitude fluctuations of the ultrasonic signal specularly scattered from a wind-driven water surface covered with monomolecular crude oil origin film of well-defined oceanographically relevant viscoelastic properties, was examined in open-sea conditions. The depression of the spectral energy density of wind-created waves by surface films is inferred from the spectra ratio of the acoustic signal fluctuations spectra with/without films, and compared to that predicted by the Marangoni damping theory. Theoretical computations showed that the film filling factor determined for natural sea surface films in coastal waters of the Baltic Sea plays a principal role in determination of the film rheological properties recovered from the acoustic surface probing, whereas a variation of the growth rate of wind waves affected by the film presence is of secondary importance. It has been demonstrated that the relative spectra can be analyzed to characterize the viscoelastic properties of surface films, which for a gasoline film-coated surface turned out to be comparable to these of natural slicks or weathered crude oil spills.

1. Introduction

Organic sea surface films of biogenic origin are particulary predominant in coastal zones [1]. They modify the physics and chemistry of the sea surface and influence remotely sensed optical and microwave imagery [2, 3]. In addition to these natural films or slicks, we find pollutant organic slicks from petroleum spills or municipal effluents [4]. The numerous effects of surface films on air-sea interfacial parameters and exchange processes have been reviewed by GARRETT [3]. One of the most visible and sensible film-induced effects is the attenuation of capillary and short gravity waves. The wave-damping effect [5, 6] which results in a strong resonance-type surface wave damping in the short gravity-capilary region [7] as shown in Section 2. A complete treatment of this problem also involves the physical and chemical properties of the film itself [8]. The result is a resonance-type behaviour of the relative damping coefficient $Y(f)=\alpha_c/\alpha_O$ as a function of surface wave frequency. Here α_c denotes the viscous damping coefficient in the presence of an elastic film, and α_O denotes that of a clean water surface. [9]. The damping ratio can be also extracted from the wind-driven waves spectra (see Sec. 3) measured in the presence $S_c(f)$ and in the absence $S_O(f)$ of a films as a spectral ratio: $Y(f)=S_O(f)/S_c(f)$, as shown in Ref.

[10, 11]. It has to be pointed out that this relation is valid only in the extreme case of the aerodynamically smooth air flow over the wavy surface [12] and if the surface is uniformmly covered with the film [4]. In general, three mechanisms may contribute to surface wave energy dissipation in the film presence: a direct damping influence (viscous damping), modification of wind-wave coupling [13] and modification of wave-wave interactions [14]. The spectra ratio which describes in a quantitative way the film effect is postulated in a form containing terms responsible for surface waves damping and their growth (see Section 2). Moreover, the spectra ratio of wind waves with/without organic films may be expressed by the corresponding ratio of low-frequency amplitude fluctuations spectra of the ultrasonic signal specularly scattered from a wavy surface, as shown in Section 3. Thus these relative spectra can be analyzed to characterize the viscoelastic properties of the spread film using a recently developed Marangoni damping theory. A variation of the growth rate due to the film presence can be evaluated on the basis of the aerodynamic parameters of the air-sea interaction process (see Section 4). These parameters were determined by numerous scientists [15—17].

Natural surface film studies were performed in shallow coastal waters of the Baltic Sea using a novel sampling device [18] for determination of principal structural film parameters (elasticity modulus, film filling factor) affecting to a great extent (compare Section 4), acoustic returns of signals scattered from film-coated surface. Rheological surface properties of an artificial crude oil product were recovered from acoustic surface scattering measurements (see Section 5) by means of the Marangoni damping theory and the best-fit procedure to the data. The parameters characterizing the model slick and its damping signatures are compared to the wave-damping abilities of several artificial organic films of different physicochemical structure. The present study was mainly motivated by the interest in the acoustic signatures of a wave damping effect caused by organic films. One of the aims is also to present a novel acoustic system for continuous measurement of wave-modulating effect from beneath the surface. The presented results are encouraging, however a great deal of supplementary current, meteorological observations and theoretical evaluations will be required to provide the basis for a more detailed interpretation of the acoustic signals in terms of environmental factors in order to develop remote sensing techniques suitable for studies of oil slick behaviour.

2. Surface film damping effect and wind velocity field above slicks

It has already been postulated in previous publications [19, 20] that at least three mechanisms may contribute to surface wave energy dissipation in the presence of a surface film: a direct damping influence (viscous damping), modification of wind-wave coupling [13], and modulation of wave-wave interactions [14, 21, 22]. For small slicks (of the order of tens of meters) and/or light to moderate winds, the situation is not so complicated [9]. In this case one can consider that the influence of the film on short-wave ripples is determined mainly by the anomalous flattening of

free linear surface waves in the presence of SA (surface-active) film [23] which exceeds considerably a conventional viscous attenuation and becomes especially apparent in the centimetr-wavelength range. The first experiments with such slicks verified this assumption [24] at least, for a gentle breeze. The main physical mechanism responsible for losses, in the absence of surface films, is viscous dissipation in the bulk water. For small amplitude water waves, the viscous energy damping coefficient is gives by the familiar expression [23]:

$$\alpha_0 = 4\nu k^2, \tag{1}$$

where ν is the water kinematic viscosity and k is the wave number related to the wave frequency by the Kelvin dispersion equation. In the approximation considered it may be written for a spectral energy density of wind waves affected by such losses S_{ds} [25]:

$$S_{ds} = \alpha_0 S, \tag{2}$$

where S is the hypothetic steady state spectrum in the absence of any losses. When the viscoelastic surface films are presented, expression (2) must be multiplied by a corrective factor — the damping ratio $Y(f)$ [5, 26].

$$Y(f) = \frac{1 - 2\tau + 2\tau^2 - X + Z(X+\tau)}{1 - 2\tau + 2\tau^2 - 2X + 2X^2}, \tag{3}$$

where

$$t = (w_d - 2\omega)^{1/2}, \quad X = \frac{E_0 k^2}{(2\nu\omega^3)^{1/2}}, \quad Z = \frac{E_0 k}{4\nu\gamma\omega}, \tag{4}$$

are dimensionless quantities and $f = \omega/2\pi = (Tk^3/\rho + gk)^{1/2}/2\pi$ is the Kelvin dispersion law. In (3) and (4) T, ρ, and g are the surface tension, density of water, and acceleration due to gravity, respectively. The constant characteristic rheological parameters of the film used are: the dilational elasticity modulus $E_0 = -dT/d(\ln\Gamma)$, where Γ is the surface concentration of a film, forming material. Gibbs excess [27] and the characteristic frequency w_d, which for spread (insoluble) films depends upon the structural relaxation of SA molecules within the monolayer. The film and its viscoelastic properties are completely characterized by E_0 and w_d parameters which are assumed to be slowly varying functions of frequency. Their evaluation changes according to whether the film is soluble or insoluble [2]. Graphically, for each choice of E_0 and w_d parameters, the relative damping $Y(f)$ plots versus frequency as a cuspidate curve. The frequency where the peak occurs is essentially controlled by the value of E_0, while the peak height Y_{max} (damping intensity) is determined by w_d. The apparatus and method for determining these parameters as well as their physical meaning and the values taken have been already described and discussed in detail elsewhere [28]. In brief, a measurement of the pattern $Y(f)$ may lead to characterizing the nature of the surface film and its wave-damping ability as shown in Ref. [9, 29]. Thus the evaluation of the effect of surfactants upon $S(f)$ requires a comparison of the spectrum in the absence $S_0(f)$ and in the presence of SA substances $S_c(f)$ all other

parameters remaining the same. The low amplitude theory mentioned above suggests the adoption of the provisional formula [2, 10, 11, 30]:

$$Y(f) = S_O(f)/S_C(f). \tag{5}$$

It has to be pointed out that the ratio of the spectral energy densities (non-slick/slickarea) defined by (5) is assumed to be equal to the relative damping $Y(f)$ only in the limiting case of a smooth air flow arround the sea surface i.e.. Reynold's number $R_e = hu_*/v_a \leqslant 1$ where h is the effective roughness height, u_* is the friction velocity, and v_a is the kinematic viscosity of air [12, 21]. In general, the relative damping should not be directly interpreted as corresponding to the depression in the spectral energy density S_o/S_c of surface waves as follows from Eqs. (3), (4) and (5). It would be so if the film were uniformly distributed over the surface. However, during the period of the experiment, the film may be partially dispersed by an air flow, waves and tidal currents, so that the surface under study is only in part covered with the viscoelastic film. In this case one shall introduce a fractional filling factor [4, 10, 11, 31] i.e. F — the ratio of the area covered with a film with respect to the total area considered and write for the effective damping ratio $Y_{ef}(f)$ the following expression:

$$Y_{ef}(f) = \frac{1}{1 - F + (F/Y(f))}, \tag{6}$$

where $Y(f)$ is given by Eq. (3). FISCELLA *et al.* [11] have shown that when the radar cell is not uniformly covered by the film, the depression of the spectral energy will be significantly smaller than that expressed by (3) even for a surface coverage of 95% ($F = 0.95$).

Now, let us consider in detail how the wind wave spectrum is affected by the Marangoni damping and the aerodynamical roughness change of the sea surface, both caused by the film presence. Generally, evolution of the wind wave spectrum is governed by energy balance equation [32, 33]:

$$\frac{d}{dt} S(k)/\omega(k) = Q_s - Q_{dis} + Q_{nl}. \tag{7}$$

One should determine how the film in the equation for the energy spectrum $S(k)$ modifies the components Q_s, Q_{dis}, Q_{nl} which describe excitation, dissipation and nonlinear interactions of the spectral excitation, dissipation and nonlinear interactions of the spectral components.

The wind wave spectrum is determined by the combined effect of a number of factors. The first factor is the transformation of the wind velocity field over a slick and, correspondingly, the variation of the regime of wind wave excitation Q_s. Another factor is the anomalous flattening of free linear waves in the presence of a SA film [23] which exceeds considerably a conventional viscous attenuation (Q_{dis}). The resonance excitation by the atmospheric pressure pulsations, PHILIP'S mechanism [34], and MILES

instability [35] are assumed to be the mechanisms of ripple excitation, and the viscous damping taking into account the effect of a SA film, and the nonlinear limitation described phenomenologically (e.g. due to the second harmonic generation in waves) are assumed to be the mechanisms of limitation. The corresponding terms in the right-hand side of Eq. (7) have the form [24]:

$$\omega Q = \frac{k^2}{4\rho^2\omega^2} \Pi_a + \beta(V) S(k) - \alpha S(k) - \delta S^2(k), \tag{8}$$

where Π_a is the pulsation spectrum of the atmospheric pressure, V is the wind speed, β is the Mile's increment, α is the wave decrement as a function of a wave number k, δ is the coefficient of nonlinear spectrum limitation. In the case when the interaction between different spectrum components is neglected, it is assumed that there is a balance between the excitation of ripples and their limitation for each wave number.

When the breeze is gentle (wind speed $V \lesssim 2$—5 m/s), the increment $\beta(V)$ is less than α, and the spectrum level $S(k)$ is low, one can neglect the last component in Eq. (8) finding the stationary value $S(k)$ from the balance condition of the other three terms.

The author characterizes quantitatively the degree of the spectra variability due the film presence by a contrast value $S_0(f)/S_c(f)$. Then, supposing the spectrum of the air pressure pulsation above the slick and nonslick region to be unchangeable, we have [24]:

$$S_0(f)/S_c(f) = \frac{\alpha_0 Y(f) - \beta(V_c)}{\alpha_0 - \beta(V_0)}. \tag{9}$$

Subscripts c and 0 denote that the corresponding values belong to the film-covered and clean water areas, respectively. Since in the centimetre-wavelength range, the decrement $\alpha_c = \alpha_0 Y(f)$ exceeds the value α_0 (by an order of magnitude if the film elasticity is rather large), when calculation the contrast one can neglect the wind variation above the slick $(V_c - V_0)/V_0 \simeq 10$—30% [13] and assume at low to moderate winds $\beta(V_c) \simeq \beta(V_0)$. Under these assumptions Eq. (9) reduces to the previous formula Eq. (5). At wind velocities $V < 2$ m/s, the values of β in Eq. (9) can be neglected; the contrast in this case is easily found from the ratio of decrements. When calculating the contrast, the author used the empirical approximation for the increment β given by PLANT [36] which results from several growth rate experiments showing a quadratic relation to the friction velocity of the wind u_*:

$$\beta = \frac{(0.04 \pm 0.02) u_*^2 \omega \cos\theta}{c^2}, \tag{10}$$

for wind-induced growth rate over the frequency range $g/2\pi V_{10}$ to 20 Hz, where c is the wave phase speed and θ is the angle between wind and wave directions.

The friction velocity of the air flow was determined from vertical wind profiles measured above the water surface at a height of z as shown in Ref. [17]. The wind

profiles over water film-covered surface differ from those of clean water, although both profiles folow a logarithmic distribution. The friction velocity of the wind $u_* = (\tau_s/\rho_a)^{1/2}$ where τ_s is the wind shear stress and ρ_a is the density of air, and the roughness parameter z_0 of the water surface can be derived from the wind profiles near the surface by applying the following dependence:

$$V(z) = (u_*/K) \ln (z/z_o); \tag{11}$$

where K is the Kármán constant = 0.4. The equivalent wind speed at a standard height $z = 10$ m was extrapolated from Eq. (11). The aerodynamic drag coefficient c_d is defined as follows:

$$c_d = \tau_s/\rho_a V_{10}^2 = (u_*/V_{10})^2. \tag{12}$$

The presented principle may be developed into an analytical method by applying an acoustic remote sensor able to deduce spectra ratios (contrast) from acoustic scattering measurements within a polluted and a non-polluted sea area, as postulated in the next section.

3. Characterization of wind-driven surface using specularly scattered sound

The scattering coefficient will be defined here as the ratio of the received intensity, when the acoustic wave is reflected by the surface under study, to the received intensity when the wave is reflected in the specular direction by a plane surface. This definition of the scattering coeffcient is equivalent to the scattering coefficient defined by BECKMANN and SPIZZICHINO [37], for electromagnetic waves. For a detailed discussion about the scattering models applicable to microwave and acoustic scattering at rough surface, the reader is referred to the review articles by OGILVY [38—40].

If the surface may be considered to be perfectly reflecting, as in an underwater scattering from a water/air interface [41], it is now possible to simplify the definition. In this case the intensity measured on the projector axis, at a distance equal to the total path-length used in the scattering measurements, may be substituted into the definition of the scattering coefficient for the speculary reflected intensity from a plane surface. Since the scattered waves are essentially planar over the aperture of the received, the intensity I_s is determined from acoustic pressure measurements $p_s (I_s \sim p^2_s)$. Assuming that the output voltage U of the hydrophone is a linear function of the acoustic pressure $(U \sim p_s)$, the average value of the measured scattering coefficient and/or relative intensity $\delta_m(\theta_i, \theta_r)$ is given by WELTON et al. [42].

$$\delta_m(\theta_i, \theta_r) = \sum_{j=1}^{N} [U_j(\theta_i; \theta_r)]^2 / N U^2_{\text{axis}}, \tag{13}$$

where θ_i, θ_r the angles of incidence and reflection, $U_j(\theta_i, \theta_r)$ is a single, independent sample of the voltage amplitude for a given projector and receiver orientation, N is the number of independent samples taken at a given orientation, and U_{axis} is the projector on axis voltage over the same path-length used to measure $U_j(\theta_i, \theta_r)$.

The scattered field registered in a specular direction from an individual surface may be regarded as composed of coherent and diffuse fields [43]. The smooth surface scatters coherently into the specular direction, as expected, with a little scattered energy away from close to specular.

The effect arises because of constructive interference between all the scattered wavelets from all parts of the surface, in the specular direction. An increase in roughness leads to energy being redistributed out of the coherent component into the more widely spread diffuse field. As a consequence, the main lobe of the scattered energy becomes weaker and the presence of diffuse field leads to significant off-specular amplitudes. In the extreme case, for a very rought surface, there is now only slight variation in signal amplitude with angle of incidence, as the scattered field is totally diffuse and closely isotropic [39]. Coherent and incoherent contributions to the total scattered field depend on the "roughness parameter" for the surface [44], which for specular scatter ($\theta_i = \theta_r$) is defined by Beckmann's g parameter as follows:

$$\sqrt{g} = 4\pi (h/\lambda) \cos\theta_i, \tag{14}$$

where h/λ is the root-mean-square wave height/acoustic wavelength ratio.

The signals scattered from a very rough surface (i.e., $g \gg 10$) are incoherent and, as shown in Ref. [40, 43], there is no coherently scattered field.

The limiting, specularly scattered, relative intensity for a very rough surface (so-called "high-frequency" scattering) is given by the formula: [44, 45]

$$\delta_{hf}(\theta_i) = \frac{\pi A}{32 \langle s^2 \rangle R_1^2 \cos\theta_i} \qquad \text{for } g \geqslant 10, \tag{15}$$

where R_1 is the distance between the transducer and the surface, A is the insonified surface area, $\pi = 3.14...$, $\langle s^2 \rangle$ is the mean- square slope of the surface. The mean square slope of the wind-excited water surface can be derived from the omnidirectional spectral energy density $S(f)$ of the surface waves as follows [21]:

$$\langle s^2 \rangle = \int k^2 S(f)\, \mathrm{d}f. \tag{16}$$

As it can be from Eq. (15), the relative specular scatter for a rough surface in indepent of the driving frequency and is a function only of the geometry of the experiment and the root mean-square slope of the ruffled surface [42].

Recent measurements of high-frequency slope spectra using a wave-following laser surface slope meter have been reported by LUBARD *et al.*, [46] and PLANT [36]. Their results show that such spectra follow an $1/f$ law to a good approximation at a wind speed of 7.5 m/s. If we assume it to hold for a variety of oceanic conditions, then we may approximate the upwind/downwind slope spectrum $S_s(f)$ as follows

$$\begin{aligned} S_s(f) &= 0.75 k^2 S(f), & 0 < f < 1.5 f_m, \\ S_s(f) &= \alpha'/f, & 1.5 f_m < f, \end{aligned} \tag{17}$$

and $$\alpha' \leqslant \frac{0.05}{\ln(13.3/f_m)},$$

where $S(f)$ is any standard spectrum applicable near the dominant wave frequency f_m. Having considered Eqs. (13), (15), (16) and (17) one can note that $U^2(f) \sim 1/S_s(f) = 1/k^2 S(f)$. Thus it is possible to obtain from acoustic scattering measurements performed for a clean and film-covered wavy water surface, assuming the fixed experiment geometry, the following expressions for the spectral energy depression of water waves by surface films:

$$\delta_c / \delta_0 = [U_c(f) / U_0(f)]^2 = S_0(f) / S_c(f), \tag{18}$$

where a surface waves spectrum component of frequency f is taken into account, $U_0(f)$ and $U_c(f)$ are the spectra of the low-frequency amplitude voltage fluctuactions of the ultrasonic signal scattered from a wind-excited surface of clean water and covered with a viscoelastic film, respectively. δ_c, δ_0 are scattering coefficients of acoustic waves derived from scattering measurements for both kinds of the surface mentioned.

4. Sea surface coverage and aerodynamic parameters of wind-surface interaction

To provide data for the proper interpretation of results recovered from remote surface probing, natural surface film studies have been successfully carried out in shallow coastal areas of the Baltic and Mediterranean Seas in 1990 and 1991 [47].

A novel sampling device was constructed by the author [18], which leads to collection of undisturbed film-coated water together with an adjacent subphase layer of a few centimeters in thickness. The film sampler is a submersible rectangular double-walled vessel which "cuts out" a sea area region measuring 45 cm x 35 cm and 8 cm in thickness. Coupled with a torsion wire balance attached to a filter paper Wilhelmy plate, the sampler represents a modified Langmuir trough apparatus to perform force-area isotherm measurements. Preliminary studies performed in shallow offshore waters of the Baltic and Mediterranean Seas showed a good reproductibility of the force-area isotherms, although the film properties (elasticity modulus, film filling factor, isotherm reversibility) are subjected to a large seasonal and areal variability, as widely discussed elsewhere [47]. The filling factor is derived from pressure-area isotherms as shown in Ref [18].

Figure 1 presents F as a function of wind speed V_{10} for measurements carried out in two significantly different sites of the Baltic Sea: A) off Orłowo (Poland) in June, August 1990 and B) off Oksywie (Poland) in April, May 1991. We were largely concerned with a broad category of condensed films in offshore waters. The elasticity modulus averaged over more than 50 samples collected was equal to $\bar{E}_0 = 24.9 \pm 19.9$ (Orłowo) and 21.0 ± 17.4 mN/M (Oksywie). Even though there was no pronounced biological activity in these water, moderate to strong film patches were present at most times. The areal extent and their homogeneity was determined by the velocity and direction of ambient winds. In a highly contaminated area in Orłowo where a high level of anthropogenic pollution is found, large surface regions are uniformly and completely film-covered ($F = 1$) in the wind speed range 0.5—2.5 m/s. For

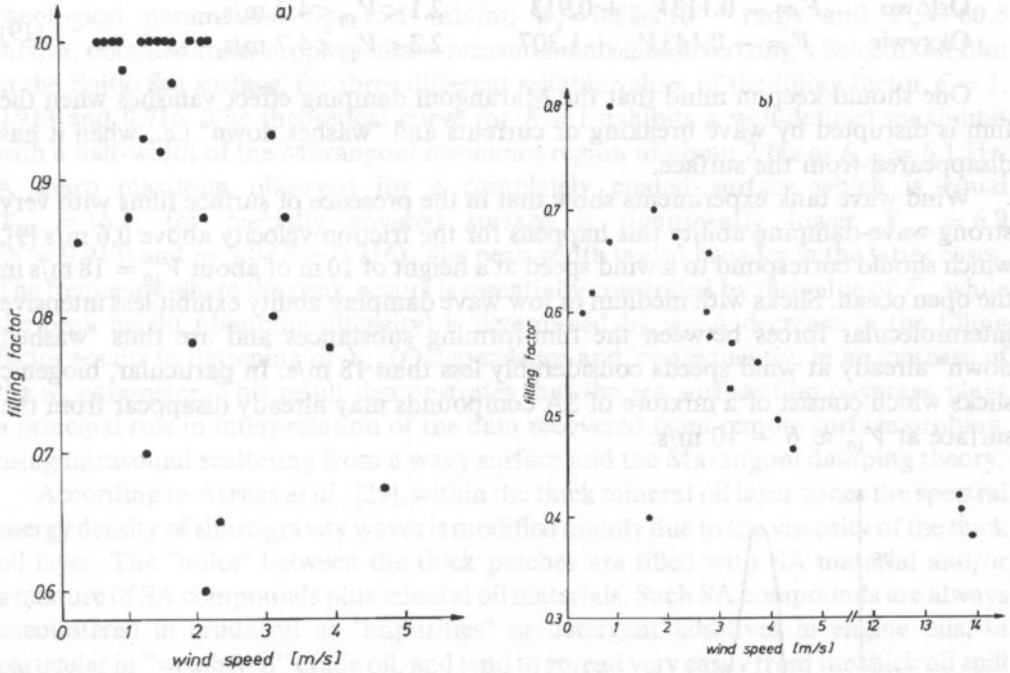


Fig. 1. Film filling factor F as a function of wind speed V_{10} for natural surface film studies performed in shallow coastal waters of the Baltic Sea A) of Orłowo (Poland) in June, August 1990, and B) of Oksywie (Poland) in April, May 1991.

$V_{10}>2.5$ m/s one can observe a rapid decrease of F with increasing wind speed. The mean filling factor registered in the summer time in Orłowo station is equal to 0.913 ± 0.094. The Oksywie site was studied in the spring time after strong wind events. Oksywie is a place distant from human settlements with their municipal effluents, but exhibit the similar $F(V_{10})$ trend, although, the values taken by F in the comparable wind speed range are significantly lower. The latter F values are expected to be found at higher sea states after strong wind events and/or compaction of film patches by surface and tidal currents. The mean filling factor registered in Oksywie and tidal currents. The mean filling factor registered in Oksywie station was equal to 0.716 ± 0.253. All these waters are close to the harbour and are contaminated to a great extent with highly surface-active substances from municipal effluents and oil spills. It is belived that artificial crude oil spills follow the similar $F(V_{10})$ pattern as the naturally-formed slikcs do. Now it is evident that the sea surface can be considered to be uniformly film-covered, as assumed by numerous authors, [6, 29, 52] in lowest wind speed ranges to $V_{10}<2.5$ m/s; for higher winds the filling factor is wind speed-dependent.

It has been found that F appers to be a linear function, as a first approach, of the wind speed (compare Fig. 1. A and B), and can be represented as a least-squares-fit to the data in the following form:

$$\begin{array}{lll} \text{Orłowo} & F = -0.113V_{10} + 0.913 & 2.1 < V_{10} < 4.2 \text{ m/s}, \\ \text{Oksywie} & F = -0.143V_{10} + 1.307 & 2.3 < V_{10} < 4.7 \text{ m/s}, \end{array} \tag{19}$$

One should keep in mind that the Marangoni damping effect vanishes when the film is disrupted by wave breaking or currents and "washes down" i.e., when it has disappeared from the surface.

Wind wave tank experiments show that in the presence of surface films with very strong wave-damping ability this happens for the friction velocity above 0.6 m/s [9], which should correspond to a wind speed at a height of 10 m of about $V_{10} = 18$ m/s in the open ocean. Slicks with medium or low wave damping ability exhibit less intensive intermolecular forces between the film-forming substances and are thus "washed down" already at wind speeds considerably less than 18 m/s. In particular, biogenic slicks which consist of a mixture of SA compounds may already disappear from the surface at $V_{10} \approx 6 - 10$ m/s.

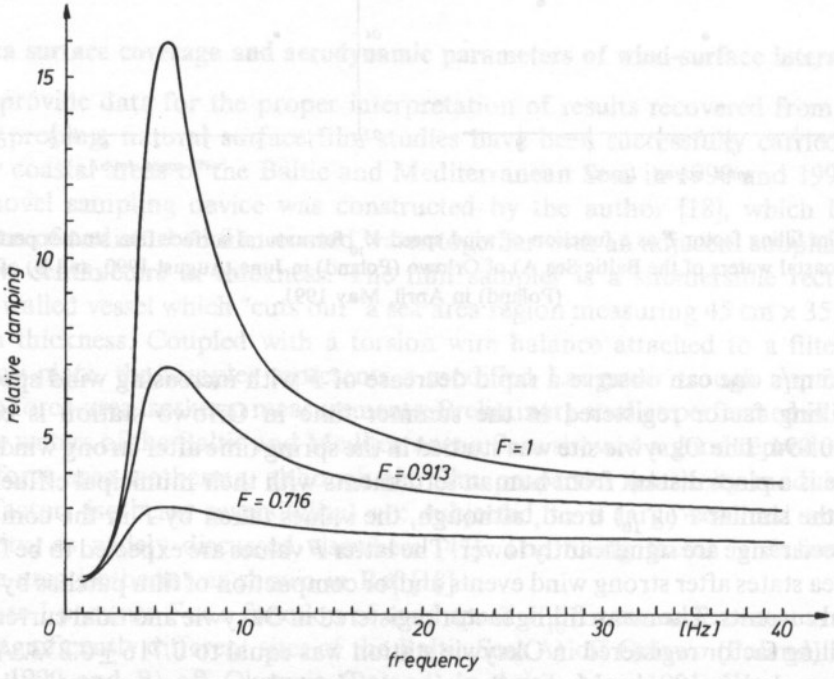


Fig. 2. Effective relative damping Y_{ef} as a function of water wave frequency predicted by the Marangoni damping theory, in which the film parameters ($T_w = 60$ mN/m; $E_o = 15.1$ mN/m and $w_d = 0.87 \times 10^{-3}$ rad/s) originate from supplementary Langmuir through measurements on a Selectol oil film spread over the Baltic Sea surface. The curves correspond to different values of the film filing factor $F = 1$; 0.913, and 0.716, respectively.

Figure 2. illustrates the sea surface coverage effect on the relative damping $Y_{ef}(f)$. Values of Y_{ef} dependence as a function of frequency are calculated by means of the Marangoni damping theory (see Eqs (3), (4) and (6)), by assuming the following film

rheological parameters: $E_0 = 15.1$ mN/m, $w_d = 0.87 \times 10^{-3}$ rad/s and $T_w = 60.5$ mN/m, obtained from supplementary measurements characterizing a Selectol oil film at the Baltic Sea surface, for three different reliable values of the filling factor $F = 1$, 0.913 and 0.716. The theoretical curve for $F = 1$ exhibits a well-defined maximum with a half-width of the Marangoni resonance region of about 2 Hz at $f_{max} = 6.1$ Hz. A sharp maximun observed for a completely coated surface which is equal $Y_{max} = 16.1$, for partially covered surface is significantly lower $Y_{max} = 6.9$ ($F = 0.913$) and $= 3.0$ ($F = 0.716$). The peak width is also broader in the latter cases. The frequency where the peak occurs is essentially controlled by the value of E_0, while the peak height (damping intensity) is determined by w_d. A decrease in the filling factor results in flattening of $Y_{ef}(f)$ dependence and, consequently, in an increase of the w_d parameter. This result demonstrates that the sea surface film coverage plays a principal role in interpretation of the data recovered from remote surface probing, using ultrasound scattering from a wavy surface and the Marangoni damping theory.

According to ALPERS *et al.*, [29], within the thick mineral oil layer zones the spectral energy density of short-gravity waves is modified mainly due to the viscosity of the thick oil layer. The "holes" between the thick patches are filled with SA material and/or a mixture of SA compounds plus mineral oil materials. Such SA compounds are always encountered in crude oil as "impurities" or detergent additives in engine oils, in particular in "weathered" crude oil, and tend to spread very easily from the thick oil spill centers over the surrounding sea surface. Crude oil spill drifting on the sea surface do not only consist of pure hydrocarbon fractions, but they contain considerable amounts of SA compounds which are being formed by photo-oxidation processes and bacterial decomposition [48]. Within the sea surface area covered with these surface-active compounds, surface tension gradients and thus Marangoni wave damping can be induced on an undulating water surface. As a consequence, two zones exhibiting different mechanisms of wave attenuation can be specified within a mineral oil spill.

At present there is some evidence testifying thas SA films transform a wind field above the sea surface. In experiments with artifical slicks of large dimensions [13] carried out at wind velocities 7.5 and 4.5 m/s, a decrease in the roughness parameter z_0 by two to seven times, respectively, and an increase in the mean wind velocity V_{10} by 15 and 30%, were registered. Similar effects were observed in the laboratory wind wave tank [17], where the same variations of V_{10} and z_0 and a small decrease by several per cent in the friction velocity in the presence of a SA film were found.

Table 1 collects the exemplary aerodynamic parameters of the wind-surface interaction process derived from the vertical wind profile, reported by numerous scientists and measured in the experiment reported here, for a film-coated (V_c, u_{*c}, z_{0c}, c_{dc} and β_c), and clean water surface (V_0, u_{*0}, z_{00}, c_{d0} and β_0), computed for film-forming substances of differentiated chemical structure according to Eqs. (10—12). For the covered surfaces one can observe an increase in V_{10} by 5—20%, a decrease in u_* by 7—9% and in z_0 by 7 to 34 times, as well as a decrease in c_d by 23 to 70% if compared to the clean water surface reference case. Values of increment β computed from the empirical equation (10) are two times higher than these

Table 1. Aerodynamic parameters of the wind-sea interaction process, for film-covered and film-free surfaces

Film-forming substance	Wind speed V_{10} (m/s)		Friction velocity u_* (cm/s)		Roughness parameter z_o (10^{-5} m)		Drag coefficient c_d (10^{-4})		Growth rate variation $\beta_o-\beta_c/\beta_c$ (%)	Remarks
	V_o	V_c	u_{*o}	u_{*c}	z_{oo}	z_{oc}	c_{do}	c_{dc}		
Oleyl alcohol	4.1	4.9	15.38	14.34	26.63	1.29	14.0	8.7	15.0	Ref. [15]
Methyl oleate	5.5	6.5	18.88	17.33	7.99	0.32	11.6	7.2	18.6	
Oleyl alcohol	5.3	6.3	18.07	16.64	8.58	0.25	11.7	6.9	17.9	
Sodium lauryl sulfate $C_{12}H_{25}SO_4Na$ ($c=2.6\times10^{-2}$ % w.)	10.0	11.8	34.18	31.71	8.27	0.34	11.6	7.2	16.1	Ref. [17]
Vegetable oil	7.5	9.2	23.34	21.47	2.61	0.34	9.7	5.4	18.1	Ref. [16]
Oleic acid	7.5	8.2	25.00	23.03	3.68	0.14	11.1	7.8	17.8	Ref. [24]
Vegetable oil	6.7	7.3	20.17	18.57	1.63	0.16	9.0	6.6	17.9	
Gasoline 94	2.3	2.7	8.05	7.35	1.08	0.33	12.3	7.2	19.9	this experiment

theoretically-predicted by MILES [17]. A variation of the growth rate $(\beta_0 - \beta_c)/\beta_c$ due to the film presence appears to be of the order of 15–19.9% seems to support the assumption that the surface film effect on the growth rate of wind waves is of secondary importance in considerations concerning suppression of wind waves. The observed changes of parameters in the aerodynamic intereaction phemomenon seem to be dependent on the chemical nature of the film-composing substance in a complicated way.

Recently, the wind velocity field above slicks has been briefly examined in the paper by ERMAKOV *et al.*, [24]. Their approximate estimates show that the corresponding increase of the wind above the slick, referred by numerous authors, can be deduced from the square root of a mean spectral energy depression value in a centimetre wavelength range of surface waves $(\overline{S_0/S_c})_{cm}$ as follows:

$$V \simeq \frac{u_*}{K} \ln (S_o/S_c)^{1/2}_{cm}; \tag{20}$$

here they assumed that $u_* = 0.05\ V_{10}$ as in Ref. [21].

5. Preliminary at-sea experiment. Results and discussion

Laboratory surface film scattering measurements showed that the relative damping dependences $Y(f)$ differ significantly from the theoretically predicted ones [49], in which the surface film parameters originated from the supplementary measurements. Several possible reasons for the discrepancy between the theory and experiment performed under small wind wave tank conditions encouraged the author to carry out the acoustic scattering investigation in open-sea conditions with an artificial crude oil origin film.

An experimental evidence of a resonance-like behaviour of the wave damping ratio in the short-gravity-wave region by monomolecular surface slikcs was first given by CINI *et al.* [50], and confirmed in wind wave tank by HÜHNERFUSS *et al.* [7], and in field measurements using a wave staff and an optical spectrum analyzer [24], as well as a radar backscattering system [10, 11, 30, 51]. Recently it has been found by SINGH *et al.* [52] that the depression of the spectral energy density of surface waves by a mineral oil spill cleary exhibits the typical Marangoni-type behaviour. The reader should note that the substance under study in this report is of similar origin. In stands for a commercially available crude oil product. At first glance, their observation of a comparable strong depression both in the presence of a monomolecular slick and in the presence of an oil spill, is surprising because the bulk of a mineral oil consists of exclusively hydrophobic alkyl and aryl compounds that are not able to give rise to wave-induced surface tensionn gradients. The data of SINGH *et al*, can be explained in the light of the Marangoni theory if SA substances were present is the crude oil spill.

The acoustic system in a form of the free-drifting lightweight buoylike equipment, as shown in Fig. 3., has been already used for remote sensing and monitoring of the sea oil polluted area [53]. It has been found that the system allows for the detection of

the oil spill edge passage observed as a rapid in the time record of scattered signals. In addition, all systematic changes of a wavy surface undulation caused by the presence of oil substances express themselves in the corresponding, regular changes of the scattered signal statistics [54]. Simultaneous analyses of all the statistical distribution parameters could be a starting point for determining the fraction weight of the given substance, its layer thickness, and finally the form of the oil pollutant (monolayer, thick layer or individual dispersed spots [53]).

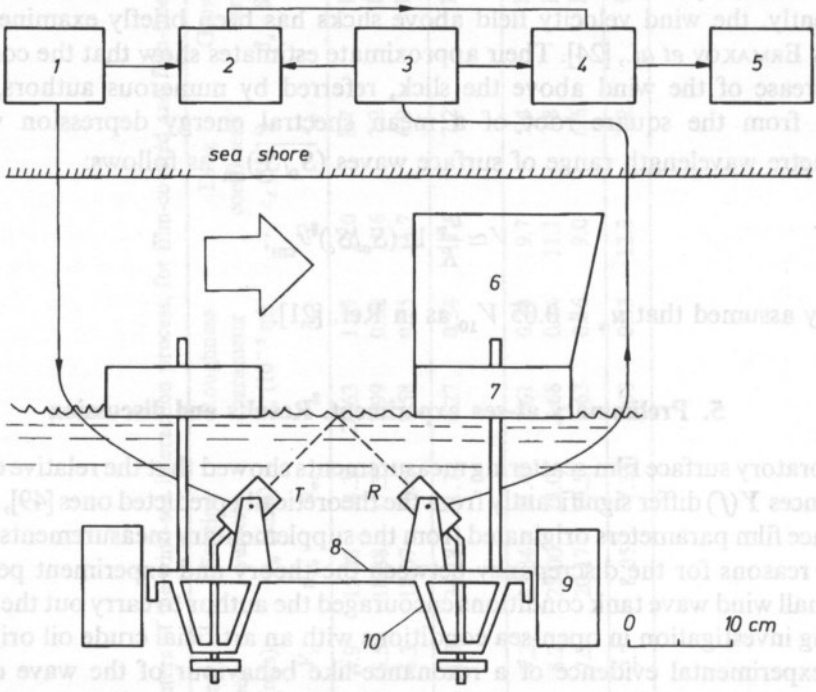


Fig. 3. A schematic diagram of the apparatus and associated electronic of the free-drifting buoylike acoustic system: *1* — ultrasonic transmitter 10 MHz; *2* — oscilloscope; *3* — ultrasonic receiver; *4* — time gating system + signal envelope and peak value detectors; *5* — spectrum analyzer; *6* — orienting wings, *7* — supporting floats; *8* — H-shaped support; *9* — balancing floats; *10* — devices for fixing and orienting transducers; *T* — transmitting and *R* — receiving ultrasonic transducers.

A block diagram of the measuring setup is presented in Fig. 3. The ultrasonic quartz transducers, both transmitting (T) and receiving (R), are placed on the H-shaped support (*8*) of the free-drifting buoy at a depth of 10 cm beneath the water level. Two float wings (*6*) enable self-orientation of the buoy according to the actual wind direction what significantly reduces the influence of surface waves reflected from the supporting floats (*7*) on the measurement. The buoy of a weight of 5 kg is also equipped with balancing floats (*9*) and deviced for fixing and orienting transducers (*10*).

The electronic part of the arrangement is located on the sea shore and connected to the buoy using cables. An ultrasonic transmitter (*1*) operating under pulse regime with a 3 kHz repetition frequency produces series of pulses lasting a few microseconds

and filled with a sine wave of 10 MHz frequency. An acoustic projector has a 3° half-power bandwith at a carrier frequency of 10 MHz. The incident anagle of ultrasonic beam is equal to 45°. The scattered signal is registered in the specular direction. After amplification by an ultrasonic receiver (*3*), for visual inspection of the signal an oscilloscope (*2*) is used. A time gate of the electronic circuit (*4*) enables the envelope and peak value detections of the signal part corresponding only to surface scattering. Then the frequency analysis of the signal amplitude fluctuations is performed by means of an analog tunable band-pass filter (*5*) having a width of 23% (Type 1621 Brüel and Kjaer) in the frequency band 2—40 Hz. The signal was registered for about 3 min in one registration. The spectra were measured 7—10 times and an average spectrum was adopted in further considerations. The standard deviation from the mean in the collected set of data was ranging from 8 to 12% of the registered signal value. A period of the signal examination was limited to about 21—30 min in order to perform the measurement under substantially the same weather conditions and slick state which appeared to evolve in time. Four supporting floats are distans from each other by 25 cm, thus the sea under study surrounded by them has a rectangular shape and is 25 cm wide and 50 cm long. In the center of this "channel" the acoustic beam scattering takes place.

It should be pointed out that the occurrence of floats can affect conditions of surface wave generation, result in not uniformly film-covered surfaces, change the wind stress and produce an additional turbulent flow over the surface screened by the floats. It has been demonstrated in preliminary studies, using fine talc powder deposited at the surface within this area, that the film homogeneity is not affected by the float presence. It is believed that the self-orientation of the buoy, according to the actual wind direction, which leaves a 25 cm — wide sea area unaffected by the floats, provides very similar conditions of surface generation for both film-covered and free sea surfaces. In addition, subsequent discussion concerns the relative scattered signal spectrum referred to the clean water case, where the only surface film effect is an important factor.

The study of an artificial oil slick influence on the amplitude fluctuations spectra of the ultrasonic signal scattered at a wavy sea surface was performed in October, 1989 in the Baltic Sea from aboard the platform built on piles about 200 m off-shore Gdynia in a depth of 15 m.

A quantity of oil placed upon the water surface will spread out by surface tension forces if the spreading coefficient $S_{o/w}$ is positive. This is the net surface tension available to drive the spreading [27]:

$$S_{o/w} = T_w - T_o - T_{o/w}, \tag{21}$$

where T_w is the surface tension of water, T_o is the surface tension of oil, $T_{o/w}$ is the oil-water interfacial tension.

Its positive value suggest the ability of oil substance to form on the water surface a coherent bulk film with thicknesses ranging from monomolecular to one resulting from the amount of the liquid and the surface area available. Many oils, including havier hydrocarbons, have negative spreading coefficients and will not spread on water, and appear to form lenses surrounded by monolayers.

Gasoline 94 was used as an artificial slick-forming material in field measurements. In order to characterize the substance to be deployed, the following additional measurements of structural parameters were performed: density, viscosity, surface tension and interfacial tension in contact with sea water collected from the measuring area. Gasoline 94 stands for a light crude oil derivative ($\rho = 760$ kg/m^3, $\eta = 0.68$ mPa·s) and turned out to have a positive spreading coefficient value against sea water $S_{o/w} = +6$ ($T_o = 20.3$ and $T_{o/w} = 26.3$ mN/m, $T_w = 52.6$ mN/m), and spontaneously formed an uniform slick. An organic substance was deployed from hexan solution at the sea surface, resulting in a slick estimated to be 20 m in radius. A slick dimension, measured from the edge of the spot exposed to the wind up to the measuring point, was of the order of 7–10 m. At the lowest sea states, where the sea surfaces was naturally unruffled, the application of the substance could not be visually determined at all. From a distance the film's presence can only be visually inferred from its surface smoothing effect. Its persistance was quite variable being a function of current, wind, and wave action as well as the accuracy of the initial slick deposition. Slick duration varied between 40 to 100 min., decreasing (as expected) with increasing sea state. The measurements of a mean wind velocity V_5 at a height of 5 m were performed by means of a standard cup anemometer about every 5 minutes. Air and sea water temperatures encountered at the measuring site were 283 and 286 K, respectively. This report deals with the study of small film slicks when the sea is calm at the wind velocity $V_{10} = 2 \pm 0.5$ m/s.

Figure 4 presents the spectral energy depression S_o/S_c, i.e. spectra ratio as a function of water wave frequency in the frequency range 0–40 Hz, for a Gasoline 94 film (circles – experimental points). The dependence was derived from the

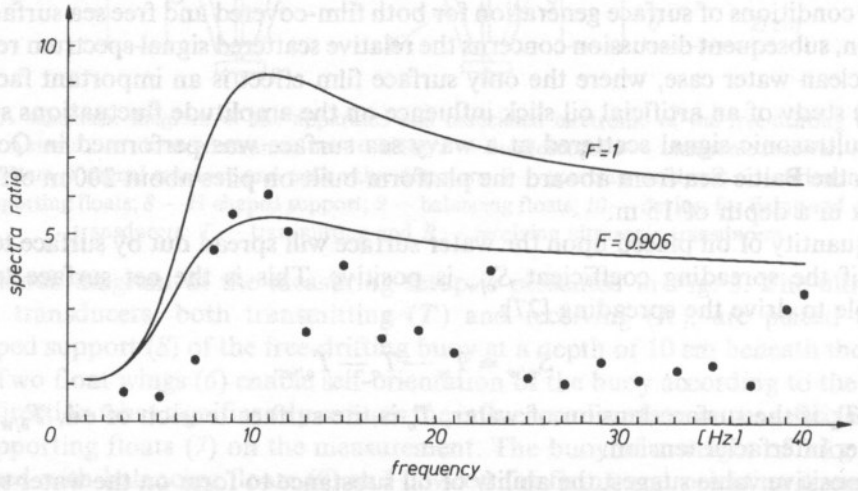


Fig. 4. Spectral energy depression S_o/S_c — spectra ratio of wind-driven waves for the Baltic Sea surface covered with a Gasoline 94 film as a function of water wave frequency as measured by the acoustic scattering system at low wind speeds ($V_{10} = 2.3$ m/s). The lower solid line is the best-fit to the data with the film filling factor $F = 0.906$. The upper line is plotted using the same best-fit-method film parameters (E_o and w_d) but assuming uniform film covering ($F = 1$).

amplitude fluctuations spectra of the ultrasonic scattered signal according to Eq. (18) at a wind velocity of 2.3 m/s. The shape of the spectra ratio clearly exhibits a Marangoni damping phenomenon. However, the experimental points are significantly scattered, especially in the high-frequency region. Thus the intensity and frequency of the spectra ratio peak in Fig. 4 may be interpreted as a maximum of the relative damping Y_{max} and the frequency of the Marangoni damping resonance f_{max}. The coordinates $f_{max}=11$ Hz and $Y_{max}=6.2$, referred to the formulas Eqs. (3), (4) and (5) lead to the rheological film parameters equal to $E_o=5.6$ mN/m and $w_d=3.5$ rad/s. It should be noted that this approximation is based on the assumption that the surface is uniformly film covered ($F=1$). The first solid line in Fig. 4 for $F=0.906$ is the best-fit procedure plot to the experimental points from the frequency range 3–20 Hz surrounding the peak. The viscoelastic film parameters corresponding to this 3-parameter best-fit procedure are equal to $E_o=6.1$ mN/m and $w_d=0.0577$ rad/s. The upper curve is presented for comparison, in which the theoretical equation Y_{ef} describing damping ratio was supplemented with the same parameters but for uniform surface covering $F=1$. The predicted damping intensity is almost two times higher than in the case of the best-fit curve. One can notice that taking into account the film homogeneity correction, leads to obtain apparently different viscoelastic parameters (especially evident for w_d) recovered from the acoustic scattering data. It turned out that the theory-predicted values exceeded those experimentally derived by several per cent in the high-frequency range $f>20$ Hz. This apparent difference could result from the fact that, in this frequency range of surface waves, the registered signal intensity is not simply related to the wind wave spectra (see Eq. (17). The last-mentioned film parameters seem to resemble better the surface properties of crude oil origin films, what has been already shown in laboratory measurements (compare Table 1 in Ref. [28]). The Gasoline 94 film properties were also "in situ" determined using the novel film sampler, for comparison, applied in the same measuring field. The obtained values are in moderate agreement with the data reported above, and are as follows: $E_o=8.3\pm0.3$ mN/m, $w_d=0.0032$ rad/s and $F=0.952$.

Having experimentally measured the relative damping relations, one is now in the position to characterize monomolecular organic films and their smoothing abilities in reference to the results reported by numerous scientists. Table 2 collects the peak frequencies f_{max} and damping intensities Y_{max} of the damping curve derived from laboratory and field experiments carried out with artificial and naturally-formed films of different chemical structure by means of a variety of surface wave field sensors. With regard to monomolecular oleyl alcohol slick OLA, a prononunced wave-damping ability with a maximum at the wave frequency $f_{max}=4.8$ Hz has been observed [55, 56] which implies a relatively large dilational modulus of $E_o=22.5$ mN/m. For comparison, CALLAGHAN *et al.* [57] reported a dilational modulus of 1.65 for North Sea crude oil (NSC) and of 1.76 mN/m for Middle East crude oil (MEC).

Based upon the Marangoni theory, a wave-damping maximum at $f=11.7$ and 11.5 Hz for NSC and MEC crude oil spills can be calculated, what is usually encountered in the presence of weakly damping surface films [48]. Furthermore, the damping intensity

Y_{max} computed by the same theory (for details see Ref. [6]) for NSC $Y_{max}=1.6$ and for MEC $Y_{max}=1.6$ is much lower than the damping intensity of an oleyl alcohol or palmitic acid methyl ester (PME) (see Table 2), which are 27.9 and 40.1, respectively.

According to Hühnerfuss classification [48], the maximum at about 8 Hz are very close to the peak frequency of a natural surface active (NS) compound and triolein (TOLG) "fish oil", which is secreted by plankton and fish. He reported a peak frequency of 6.85 Hz for this natural substance, which implies that it is a material of medium wave-damping ability (see also Murban crude oil, NS and WOS in Table 2).

Table 2. Peak frequencies f_{max} and wave-damping intensities Y_{max} of the spectral energy depression S_o/S_c or relative damping Y dependences, reported from laboratory and open-sea experiments with a variety or organic surface films of different physicochemical structure and origin

Film-forming substance	Film elasticity E_o (mN/m)	Maximum damping frequency f_{max} (Hz)	Damping intensity Y_{max}	Remarks
Hexadecanoic acid methyl ester (PME)	46	3.5	40.1	calculated from the Marangoni theory with E_o and w_d obtained in wave tank measurements
Oleic acid (OLS)	14	6.1	19.5	
Oleyl alcohol (OLA)	22.5	4.8	27.9	Ref. [9, 48]
Hexadecyltrimethylammonium bromide (CEM 3AB)	25.5	4.6	26.3	
Triolein (TOLG) (fish oil)	11.5	6.9	18.3	
Oleic acid methyl ester (OLME)	10.0	8.0	2.3	
North Sea crude oil (NSC)	1.65	11.7	1.6	calculated on the basis of the rheological film properties
Middle East crude oil (MEC)	1.76	11.5	1.6	
"Weathered" crude oil spill (WOS)	10.0	7.8	17.2	Ref. [57]
Olein (technical grade oleic acid)	30.0	4.6	15.8	from the damping of wind waves in open sea experiments
Vegetable oil	10.0	10.0	3.9	Ref. [24]
Natural slick (NS)	6	8.1	3.2	
Murban crude oil	11	8	7.6	Ref. [52]
Natural slick (NS) in polluted waters	25.0	3.8	27.6	from the damping of wind waves
	18.0	4.2	14.3	
	9.0	6.5	5.8	Ref. [4]
Natural slick (NS) in Ligurian Gulf	10.9	7.7	11	from the damping of wind waves
Cetyl alcohol	–	4.0	31.6	Ref. [2]
Methyl alcohol	–	4.3	17.8	
Triton X 100 (concentration 10^{-6} mol)	–	7.1	15.8	

Surface active substances of strong damping ability exhibit maximum wave damping at frequencies between 3.5 and 5 Hz (compare PME, OLA and olein), while weakly

damping substances have this maximum at frequencies between 8 and 10 Hz, or sometimes even above 10 Hz (NSC, MEC, vegetable oil and WOS). According to the above discussion, we have got the surface film of weak wave-damping ability comparable to that of natural slicks of biogenic origin and/or weathered crude oil spills. This result also suggests that caution has to be applied when interpreting acoustic scattering data measured at lower wind speeds i.e., under meteorological conditions which are known to be favourable for the formation or natural organic surface films.

Let us compare the presented results with the values of surface rheological parameters, reported by several authors in laboratory and open-sea experiments with slicks of both artificial and natural origin, derived from the spectral energy depression of wind waves using different surface wave field sensors. The rheological parameters pertaining to an oleic alcohol film, empirically found by FISCELLA *et al.* [10] in wave tank measurements, are $E_o = 9.7$ mN/m and $w_d = 9.65$ rad/s. The work by SINGH *et al.* [52] has presented, for the first time, data which straightforwardly indicate the Marangoni damping behaviour of a crude oil spill. The observed viscoelastic characteristics of the oil employed in the COAATF off Halifax on September 16–17, 1983 are $E_o = 8.0 \pm 0.2$ mN/m and $w_d = 12.0 \pm 0.3$ rad/s, as evaluated by FISCELLA *et al.* [31] by means of the recently developed Marangoni damping theory. Natural sea surface films have been detected and characterized by measurements of short-gravity spectra of wind waves during three experimental periods, one in the Sicilian Channel, and two in the Gulf of Maine [4]. Wind wave spectra with/without organic films were measured "in situ" with a microwave probe. For each of the data set presented in Fig. 5 of LOMBARDINI *et al.* work [4], a theoretical curve is drawn according to Eqs. (3), (4), (5) and (6) with a proper choice of the parameters E_o, w_d and F. For the three cases, the parameters obtained by this best fit method are $E_o = 25.0$; 18.0; 9.0 mN/m, $w_d = 11.0$; 13.0; 6.0 rad/s and $F = 0.996$; 0.880; 0.982. Finally, for coastal waters in the Ligurian Sea, the spectral energy depression curve obtained by photographic techniques by CINI *et al.* [2] allows one to evaluate the rheological parameters to be $E_o = 10.9$ mN/m and $w_d = 0.18$ rad/s. For more experimental data the reader is referred to the paper [2], where the damping ratio $Y(f)$ dependences, registered in the presence of four insoluble films of SA compounds, are depicted in Fig. 4.

The data collected above clearly show that the time of the relaxation process $t_r = 1/2\, w_d$ involved in the surface wave damping effect by monomolecularviscoelastic films is of the order of 0.04 to 2.7 seconds. One should remember that the relaxation time, depending on the type of relaxation process, for soluble films upon diffusional relaxation, and for insoluble ones upon the structural relaxation between intermolecular forces [5], can vary from 10^{-3} s to several minutes [58]. The characteristic time of the relaxation process in the model slick t_r is 8.67 s. For comparison, water waves with a wavelength of 1 cm are damped within the characteristic time $(\alpha_0^{-1}$ (Eq. (1)) equal to 0.4 s.

In the last decade monomolecular surface films, mostly 9-octadecen-1-ol, cis isomer (oley alcohol, an 18 carbon, monounsaturated fatty alcohol) have been

applied as an oceanographic tool for studying the air-sea interaction process at the ocean surface [6, 19]. This material is chosen because it resembles natural slicks in the physico-chemical behaviour [9, 48]. It has been found that long-time scale relaxation times of several minutes measured by the author for crude oil derivatives films [28] were also encountered for an oleyl alcohol film [59].

One can conclude that the Gasoline 94 film spread over the Baltic Sea surface stands for a slick of oceanographically relevant viscoelastic properties. In addition, the experimental spectral energy depression data points could be fitted into the theoretical relation (3) and, after combining Eqs. (3), (5) and (6), the unknown filling factor F can be evaluated, what is of great importance, since it is very difficult to measure this quantity directly which can significantly influence the surface scatter returns [4, 10, 11, 30].

It has been found that the spectra ratio was observed first to increase as the slick drifts over the insonsified area, and then decrease, returning to its preslicked values. It seems that the data collected by this system can be further interpreted to deduce the film weathering and concentration.

Recently, the wind velocity field above slicks has been briefly examined in the paper by ERMAKOV *et al.* [24] (see Sec. 3). Their appproximate estimates show that a corresponding increase of the wind above the slick can be given by the root of a mean spectral energy depression value in a centimetre-wavelength range of surface waves $(S_0/S_c)_{cm}$, what follows from Eq. (20).

Assuming the acoustically determined S_o/S_c values of the order of $2-3$, one can obtain $\Delta V = 0.17$ to 0.30 m/s which corresponds to the wind velocity variation above the studied slick $\Delta V/V \sim 10.4-13.1\%$. This result agrees well with the direct anemometer data of the same order of magnitude (see Table 1). Thus the remote estimating the wind field near the surface covered with a slick, which is of meteorolgical relevance, may be derived from acoustic scattering measurements.

It may be helpful to add that the applied acoustic system consisting of two transducers based on a forward specular scattering geometry has got two features which are of significant interest in future at-sea experiments. The wave staffs measure the surface waves spectra omnidirectionally i.e., integrating over 360° and so-called "spectra of encounter" are derived [19] whereas an acoustic scatterometer senses unidirectionally in reference, to the pointing direction of the acoustic beam. Therefore, a directional pattern of the scattered signal fluctuations spectra with respect to the actual wind direction can be evaluated. The proposed corrections of the "spectra of encounter" in order to get "real spectra" take into account deviations due to the orbital velocity of gravity waves, wave-induced Stokes drift and tidal currents [21]. The spectra of specularly scattered acoustic signals do not have to be corrected [41, 43] and can be directly compared to those obtained at different wind speeds and environmental conditions or theoretically predicted.

These results are encouranging, however a great deal of additional experimental and theoretical evaluation will be required to determine the full capability of the acoustic system for remote sensing of marine surface films under open-sea conditions. A critical element in determining the usefulness of such a system is the quality and thoroughness of the environmental measurements taken simultaneously with the acoustic data.

6. Concluding remarks

Natural surface film studies performed in coastal waters of the Baltic Sea using a novel device for "in situ" sampling and force-area isoterm measurements showed that the film filling factor, affecting the acoustic scattering returns to a great extent, reaches the mean values 0.913 and 0.716 for two different sampling sites studied. In addition, F is a wind dependent quantity. For low wind speeds $V_{10}<2-3$ m/s the sea surface can be assumed to be uniformly and completely film-covered, $F=1$.

Data reported by numerous authors on the vertical wind profile above wavy sea water allowed one to determine the aerodynamic parameters of the air-sea interaction process. For the film-covered surfaces one can notice an increase in V_{10} by 5–20%, a decrease in u_* by 7–9% and in z_0 by 7 to 34 times, as well as a decrease in the drag coefficient c_d by 23 to 70% if compared to the clean surface case. A variation of the growth rate of wind-driven waves $(\beta_0-\beta_c)/\beta_c$ due to the film presence appears to be of the order of 15.0–19.9%, and is of secondary importance in suppression of wind waves. An increase of the wind above the model slick (0.17–0.30 m/s) deduced from the spectra ratio in the centimeters wave range is in agreement with the direct anemometer observation. The measured coordinates $f_{max}=11$ Hz and $Y_{max}=6.2$ referred to the theoretical formulas lead to the rheological Gasoline 94 film surface properties equal to $E_0=6.01$ mN/m and $w_d=0.0577$ rad/s with the filling factor $F=0.906$. The film filling factor plays a principal role in a proper determination of the film properties recovered from the acoustic surface probing. According to the discussion by Hühnerfuss, we have the surface film of weak wave-damping ability comparable to that of natural slicks of biogenic origin and/or weathered crude oil spills.

Acknowledgments

I wish to thank Prof. R. Cini and Prof. G. Loglio with their co-workers (University of Florence, Italy) and Prof. H. Hühnerfuss (University of Hamburg, Germany) for fruitful discussions on the Marangoni damping and structure of surface films. I am also grateful for. An Individual Mobility Grant (contract No. IMG-PLT-0143-90) provided by the European Community under the Tempus program. Valuable comments of the Reviewer are also acknowledged.

References

[1] W.R. BARGER, W.H. DANIEL, and W.D. GARRETT , *Surface chemical properties of banded sea slicks*, Deep-Sea Res. **21**, 83–89 (1974).

[2] R. CINI, P.P. LOMBARDINI, and H. HÜHNERFUSS, *Remote sensing of marine slicks utilizing their influence on wave spectra*, Int. J. Remote Sensing **4**, 101–110 (1983).

[3] W.D. GARRETT, *The physicochemical effects of organic films at the sea surface and their role in the interpretation of remotely sensed imagery*, ONRL Workshop Proc. Rep. C-11-86 (U.S. Office of Naval Res., London, 1986).

[4] P.P. LOMBARDINI, B. FISCELLA, P. TRIVERO, C. CAPPA, and W.D. GARRETT, *Modulation of spectra of short gravity waves by sea surface films: Slick detection and characterization with a microwave probe*, J. Atmos. Ocean Tech. **6**, 882–890 (1989).

[5] P. CINI, P.P. LOMBARDINI, C. MANFREDI, and E. CINI, *Ripple damping due to monomolecular films*, J. Colloid Interface Sci. **119**, 74–80 (1987).

[6] H. HÜHNERFUSS, W. WALTER, P.A. LANGE and W. ALPERS, *Attenuation of wind waves by monomolecular sea slicks and the Marangoni effect*, J. Geophys. Res., **92**, 3961–3963 (1987).

[7] H. HÜHNERFUSS, W. ALPERS, P.A. LANGE and W. WALTER, *Attenuation of wind waves by artificial surface films of different chemical structure*, J. Geophys. Res. Lett. **8**, 1184–1186 (1981).

[8] F.L. HERR and J. WILLIAMS, *Role of surfactant films on interfacial properties of the sea surface*, ONRL Workshop Proc. Rep. C-11-86, (U.S. Office of Naval Res. London, 1986).

[9] W. ALPERS and H. HÜHNERFUSS, *The damping of ocean waves by surface films: A new look at an old problem*, J. Geophys. Res. **94**, 6251–6265 (1989).

[10] B. FISCELLA, P.P. LOMBARDINI and P. TRIVERO, *Ripple damping on water surface covered by a spreading film: Theory and experiment*, Nuovo Cimento, **8C**, 491–500 (1985).

[11] B. FISCELLA, P.P. LOMBARDINI and P. TRIVERO, P. PAVESE and R. CINI, *Measurements of the damping effect of a spreading film on wind-excited sea ripples using a two-frequency radar*, Nuovo Cimento, **8C**, 175–183 (1985).

[12] H. SCHLICHTING, *Boundary-layer theory*, (Mc Graw–Hill, New York 1968).

[13] W.R. BARGER, W.D. GARRETT, E.L. MOLLO-CHRISTENSEN and K.W. RUGGLES, *Effects of an artificial sea slick upon the atmosphere and the ocean*, J. Appl. Meteorol., **9**, 196–400 (1970).

[14] G.J. KOMEN, *Nonlinear contributions to the frequency spectrum of wind generated water waves*, J. Phys. Oceanogr. **10**, 779–790 (1980).

[15] H. HÜHNERFUSS, W.D. GARRETT and F.E. HOGE, *The discrimination between crude oil spills and monomolecular sea slicks by an airborne lidar*, Int. J. Remote Sensing **7**, 137–150 (1986).

[16] S.A. ERMAKOV, A.M. ZUJKOVA and S.G. SALASHIN, (in Russian) *Transformation of wind-wave spectra in film slicks*, Izv. AN USSR Fiz. Atmos. Okeana **23**, 707–715 (1987).

[17] H. MITSUYASU and T. HONDA, *Wind-induced growth of water waves*, J. Fluid Mech. **123**, 425–442, (1982).

[18] S.J. POGORZELSKI, *Isotherms of natural sea surface films: A novel device for sampling and properties studies*, Rev. Sci. Instrum. **63**, 3773–3776 (1992).

[19] H. HÜHNERFUSS, W. ALPERS, A. CROSS, W.D. GARRETT, W.C. KELLER, P.A. LANGE, W.J. PLANT, F. SCHLUDE and D.L. SCHULER, *The modification of X and L band radar signals by monomolecular sea slicks*, J. Geophys. Res. **88**, 9817–9822 (1983).

[20] J. WU, *Suppression of oceanic ripples by surfactant — spectral effects deduced from sun-glitter, wave staff and microwave measurements*, J. Phys. Oceanogr. **19**, 238–245 (1989).

[21] O.M. PHILLIPS, *The Dynamics of the Upper Ocean*, (Cambridge University Press, New York, 1977).

[22] H.Y. WU, E.Y. HSU and R.L. STREET, *Experimental study of nonlinear wave-wave interaction and white-cap dissipation of wind-generated waves*, Dyn. Atmos. Oceans, **3**, 55–78 (1979).

[23] V.G. LEVICH, *Physicochemical hydrodynamics* (Prentice-Hall, New York, 1962).

[24] S.A. ERMAKOV, A.M. ZUJKOVA, A.R. PANCHENKO, S.G. SALASHIN, T.G. TALIPOVA and V.I. TITOV, *Surface film effect on short wind waves*, Dyn. Atmos. Oceans **10**, 31–50 (1986).

[25] O.M. PHILLIPS, *Spectral and statistical properties of the equilibrium range in wind-generated gravity waves*, J. Fluid Mech. **156**, 505–531 (1985).

[26] R. CINI and P.P. LOMBARDINI, *Damping effect of monolayers on surface wave motion in a liquid*, J. Colloid Interface Sci. **65**, 387–389 (1978).

[27] A.W. ADAMSON, *Physical chemistry of surfaces*, (Wiley, New York, 1982).

[28] S.J. POGORZELSKI, *Monomolecular organic film effect on wind-driven waves deduced form ultrasound scattering*, J. Acoust. Soc. Am. **90**, Pt. 1, 965–972 (1991).

[29] W. ALPERS and H. HÜHNERFUSS, *Radar signature of oil films floating on the sea surface and the Marangoni effect*, J. Geophys. Res. **93**, 3642–3648 (1988).

[30] P. TRIVERO and C. CAPPA, *Sea wave spectra by electromagnetic techniques*, Nuovo Cimento, **10C**, 409–418 (1987).

[31] B. FISCELLA, P.P. LOMBARDINI, P. TRIVERO and C. CAPPA, *Sea return at C and K_u bands*, Nuovo Cimento, **10C**, 381–385 (1987).

[32] K. HASSELMANN, *Weak-interaction theory of ocean waves in: Developments in Fluid Dynamics*, M. Holt, ed. (Academic Press, New York), 2, 117–182.

[33] S.A. ERMAKOV and E.N. PELINOVSKY, *Variation of the spectrum of wind ripple on coastal waters under the action of internal waves*, Dyn. Atmos. Ocean **8**, 95–100 (1984).

[34] O.M. PHILLIPS, *On the generation of waves by turbulent wind*, J. Fluid Mech. **2**, 417–445 (1957).

[35] J.W. MILES, *On the generation of surface waves by shear flows*, J. Fluid Mech., **3**, 185–204 (1957).

[36] W.J. PLANT, *A relationship between wind stress and waves slope*, J. Geophys. Res. **87**, 1961–1967 (1982).

[37] P. BECKMANN and A. SPIZZICHINO, *The scattering of electromagnetic waves from rough surfaces*, (Mc Millan, New York 1963), 22–23.

[38] J.A. OGILVY, *Wave scattering from rough surfaces*, Rep. Prog. Phys. **50**, 1553–1608 (1987).

[39] J.A. OGILVY, *Computer simulation of acoustic wave scattering from rough surfaces*, J. Phys. D: Appl. Phys. **21**, 260–277 (1988).

[40] J.A. OGILVY, *Model for the ultrasonic inspection of rough defects*, Ultrasonics **27**, 69–79 (1989).

[41] C.S. CLAY and H. MEDWIN, *Acoustical oceanography: principles and applications*, (Wiley, New York, 1977).

[42] P.J. WELTON, H.G. FREY and P. MOORE, *Experimental measurements of the scattering of acoustic waves by rough surfaces*, J. Acoust. Soc. Am. **52**, 1553–1563 (1972).

[43] M.L. BOYD, and R.L. DEAVENPORT, *Forward and specular scattering from a rough surface: Theory and experiment*, J. Acoust. Soc. Am. **53**, 791–801 (1972).

[44] H. MEDWIN, *Specular scattering of underwater sound from a wind-driven surface*, J. Acoust. Soc. Am. **41**, 1485–1495 (1967).

[45] I. TOLSTOY and C.S. CLAY, *Ocean acoustics*, (Mc Graw–Hill, New York, 1966).

[46] S.C. LUBARD, J.E. KRIMMEL, L.R. THEBAUD, D.D. EVANS and H. SHEMDIN, *Optical image and laser slope meter intercomparisons of high-frequency waves*, J. Geophys. Res. **85**, 4996–5002 (1980).

[47] S.J. POGORZELSKI, A.M. STORTINI and G. LOGLIO, *Natural surface film studies in shallow coastal waters of the Baltic and Mediterranean Seas*, Cont. Shelf Res. (1993) (in press).

[48] H. HÜHNERFUSS, *The molecular structure of the system water monomolecular surface film and its influence on water wave damping*, Habilitation thesis, University of Hamburg, Federal Republic of Germany (1986).

[49] S.J. POGORZELSKI, *Suppression of wind waves by monomolecular films of crude oil origin deduced from acoustic surface scattering*, Dyn. Atmos. Oceans (1992) (submitted).

[50] R. CINI and P.P. LOMBARDINI, *Experimental evidence of a maximum in the frequency domain of the ratio of ripple attenuation in monolayered water to that in pure water*, J. Colloid Interface Sci. **81**, 125–131 (1981).

[51] H. HÜHNERFUSS, W. ALPERS and F. WITTE, *Layers of different thicknesses in mineral oil spills detected by grey level textures of red aperture radar images*, Int. J. Remote Sensing **10**, 1093–1099 (1989).

[52] K.P. SINGH, A.L. GRAY, R.K. HAWKINS and R.A. O'NEIL, *The influence of surface oil on C- and K_u band ocean backscatter*, IEEE Trans. Geosci. Remote Sens., **GE-24**, 738–744 (1986).

[53] S.J. POGORZELSKI, *The influence of crude oil spills on the sea surface on ultrasound scattering*, Oceanologia **31**, 107–118 (1991).

[54] S.J. POGORZELSKI, *Detection of oil-derivative contamination of water surfaces by statistical analysis of scattered acoustical signals*, J. Acoust. Soc. Am., **85**, 2383–2387 (1989).

[55] H. HÜHNERFUSS, P.A. LANGE and W. WALTER, *Relaxation effects in monolayers and their contribution to water wave damping*, I. *Wave-induced phase shifts*, J. Colloid Interface Sci. **108**, 430–441 (1985).

[56] H. HÜHNERFUSS, P.A. LANGE and W. WALTER, *Relaxation effects in monolayers and their contribution to water wave damping*, II. *The Marangoni phenomenon and gravity wave attenuation*, J. Colloid Interface Sci. **108**, 442–450 (1985).

[57] I.C. CALLAGHAN, C.M. GOULD, R.J. HAMILTON and E.L. NEUSTADTER, *The relationship between the dilational rheology and crude oil foam stability. I. Preliminary studies*, Colloids Surfaces **8**, 17–28 (1983).

[58] E.H. LUCASSEN-REYNDERS, *Effects of surfactnat aggregation in soluble monolayers on dynamic surface properties*, Colloids Surfaces **25**, 231–249 (1987).

[59] H. HÜHNERFUSS and W. WALTER, *The thermal anomaly of relaxation effects in monomolecular surface films*, J. Colloid Interface Sci. **97**, 476–480 (1984).

ARCHIVES OF ACOUSTICS
19, 1, 109–122 (1994)

ELECTROACOUSTIC ANALOGIES APPLIED TO ACOUSTIC OSCILLATOR ANALYSIS

G. BUDZYŃSKI AND M. SANKIEWICZ

Department of Sound Engineering
Gdansk Technical University
(80-952 Gdańsk)

Analogies were, so far, rarely applied to acoustic oscillator analysis, however, their use turns out to be advantageous. It allows, due to comparison of analogous electric and acoustic oscillators operation, for better understanding of the process of oscillation maintenance, as well as, the interpretation of the circuits behaviour. Particular conclusions are presented relative to examples of a bowed string, and of a labial pipe oscillators.

1. Introduction

It is almost incredible that one of the most ancient acoustic devices invented by our prehistoric ancestors to produce musical sound, the flue pipe, furnishes, so far, an unsolved problem for scientists attempting to explain fully the involved mechanism of oscillation maintenance. A more recent example of a similar problem is delivered by the mechanism of bowed string oscillations. Meanwhile, there is a large family of electric oscillators, thoroughly investigated, with very well known characteristics. Why the experience gained in the domain of electric oscillators could not be applied to mechanic and acoustic self-oscillating devices?

The most probable negative answer bases on formal constraints of electroacoustical analogies, which are limited to linear elements only, while self-oscillating circuits contain, as a rule, nonlinear elements. Nevertheless, as shown beneath, the analogies can be enlarged, at least qualitatively, to the comparative analyses of oscillator circuits of the mechanical, acoustical and electrical nature.

First of all, a brief review of electroacoustical analogies is necessary, including some comments and enlargements of their usual applications.

2. Electroacoustic analogies

Analogies between mechanic, acoustic and electric quantities were studied already in XIX century. Many authors paid particular attention to that matter and numerous textbooks contain presentations of analogies, laid down as appropriate tables of

corresponding quantities and equivalent circuits [5], [7], [12], [13]. Although the theory of analogies originated by Lord Kelvin and by Firestone remains a valid and exhaustive basis of those publications a care should be kept in their use, because of some intricacy in particular presentations, and even few mistakes cointained therein.

A most complete and inspiring presentation of that matter has been given by MALECKI [8], where he even enlarged the well known concepts on the domain of field quantities. Quoting here all his consideration would consume too much place, thus only an abridged information on analogies, indispensable for this article, is given beneath.

Formal analogies, i.e. those based on similarity of equations describing electrical, mechanical and acoustical phenomena, are usually denoted as a table of corresponding quantities. For the case of the motional (corrected, or Firestone's) analogies, in contrast to the dynamical (classical, or Kelvin's) ones, the following quantities can be listed, as quoted in the Table I.

Table I.
Motional analogies

Electrical —		Mechanical —		Acoustical Quantities	
voltage	[V]	velocity	[m/s]	volume-velocity	[m^3/s]
current	[A]	force	[N]	pressure	[N/m^2]
charge	[C]	force impulse	[Ns]	pressure impulse	[Ns/m^2]
flux	[Wb]	displacement	[m]	volume displacement	[m^3]
inductance	[H]	compliance	[s^2/kg]	compliance	[m^4s^2/kg]
capacitance	[F]	mass	[kg]	inertance	[kg/m^4]
resistance	[Ω]	conductance	[s/kg]	conductance	[m^4s/kg]
conductance	[S]	resistance	[kg/s]	resistance	[kg/m^4s]

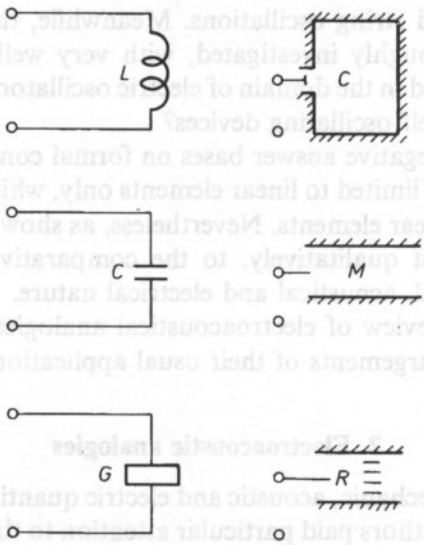


Fig. 1. Examples of motional analogy between two-poles.

Only the motional analogies, in contrast to dynamical ones, are considered here, because they keep unaffected topological features of the corresponding quantities. Moreover, they introduce impedances instead of admittances, and inversely. These properties are both favourable for a more clear presentation of the concept of analysis. Besides, keeping to one family of analogies reduces a probability of possible mistakes, numerous in practice, as mentioned above, even in textbooks.

Employing the correspondent quantities of the table given above, various equivalent circuits of the acoustical two-poles may be described as analogues to the electrical two-poles, e.g. those shown in Fig. 1, under assumption of linearity of all circuit elements.

3. Two-pole oscillator circuits

Further analogies may be considered, concerning oscillator circuits, composed of two-poles: the first one having non-linear maintenance characteristics, and the second one built of linear elements with resonant properties. Here, however, due to non-linearity of the maintaining two-poles the full analogy is possible only then, when shapes and scales of their non-linear characteristics are exactly analogous. Figure 2 shows the two basic types of equivalent circuits of electrical two-pole oscillators: the parallel,

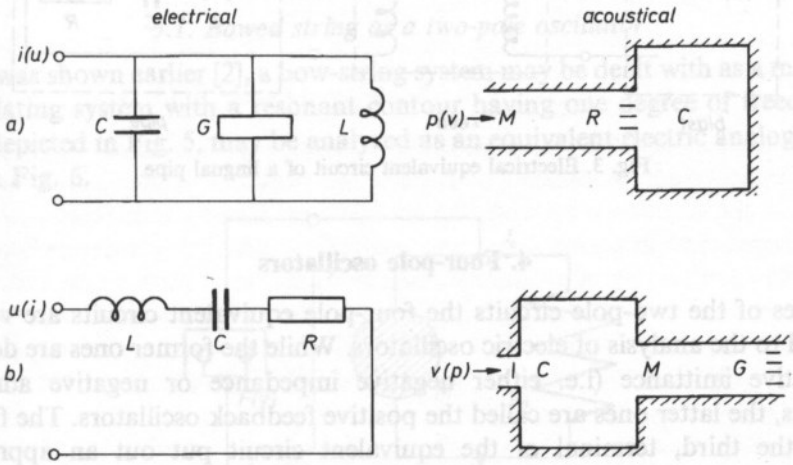


Fig. 2. Two basic types of two-pole oscillators
a) parallel controlled, b) controlled in series.

voltage-controlled oscillator, and the series, current controlled oscillator circuit, as well as their acoustical analogues. Mechanical analogies are omitted here, because their equivalent circuit are identical to acoustical ones, except differences in diagram symbols.

Even when strict analogy between non-linear two-poles is not reached, it may be sufficient to take advantage of the similarity of maintenance and of stability conditions of the oscillator circuit, as well as of its circuit variables.

A particular attention is to be paid to the two acoustical circuits shown in Fig. 2. While the parallel controlled oscillator represents the case of a labial pipe maintained by a jet action (Fig. 2a), which is a main topic of our consideration, the acoustical circuit controlled in series should require a special maintaining two-pole, pressure controlled (Fig. 2b). As such a device is unknown in practice, so the respective two-pole characteristics $v(p)$ is rather hypothetic.

On the other hand, we know that a pipe resonator may be maintained in oscillation by excitation applied to the pipe closed end, which case is just equivalent to a series resonant circuit. It is commonly known that such wind instruments like lingual pipe with a tuned resonator, or like clarinet, are excited from their closed ends, yet by means of a reed. The reed acts as a mechanical lever, which is equivalent to an electrical transformer, transforming a low impedance of the series resonant circuit into a high impedance needed to match a maintaining two-pole, of the same kind as that of the Fig. 2a). Thus, an appropriate equivalent circuit for the tuned lingual pipe is that shown in Fig. 3.

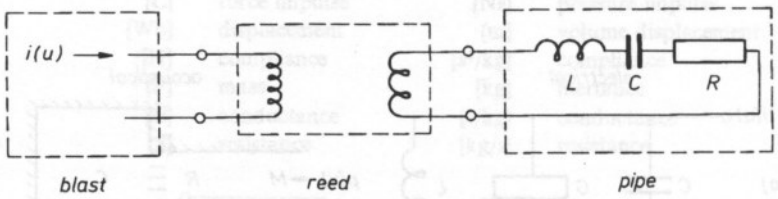


Fig. 3. Electrical equivalent circuit of a lingual pipe.

4. Four-pole oscillators

Besides of the two-pole circuits the four-pole equivalent circuits are very often employed to the analysis of electric oscillators. While the former ones are denoted as the negative imittance (i.e. either negative impedance or negative admittance) oscillators, the latter ones are called the positive feedback oscillators. The fourth, or at least the third, terminal of the equivalent circuit put out an appropriately phase-shifted variable from a divided or tapped resonant contour into the maintaining four-pole (at least three-pole), see Fig. 4.

Although many mechanical oscillators can be represented by four-pole equivalents, no acoustical examples of such circuits are known in practice, because acoustical resonators generally have neither taps nor branches. Therefore, attempts to analyze acoustical oscillators as feedback circuits, as e.g. is practized in edgetone theory, do not seem to be justified.

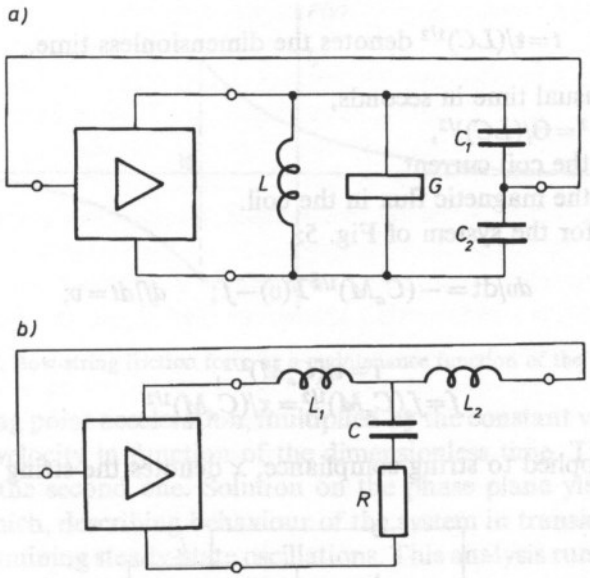


Fig. 4. Four-pole (positive feedback) electric oscillator circuits; a) with parallel, tapped resonator, b) with series, branched resonator.

5. Examples of oscillator analysis

5.1. Bowed string as a two-pole oscillator

As it was shown earlier [2], a bow-string system may be dealt with as a mechanical self-oscillating system with a resonant contour having one degree of freedom. The system, depicted in Fig. 5, may be analyzed as an equivalent electric analogue circuit shown in Fig. 6.

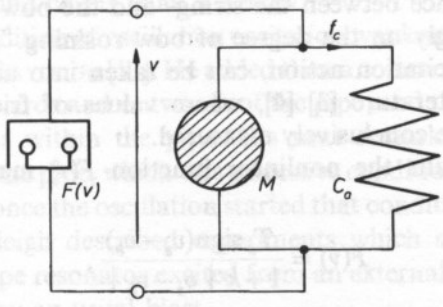


Fig. 5. Mechanical system representing a bow-string oscillator.

For the circuit of Fig. 6:

$$du/dt = -(L/C)^{1/2} I(u) - i; \qquad di/dt = u; \tag{1}$$

where

$t = \mathbf{t}/(LC)^{1/2}$ denotes the dimensionless time,

whilst

$\mathbf{t}$ denotes usual time in seconds,

$i = i_L(LC)^{1/2} = \Theta/(LC)^{1/2}$,

i_L denotes the coil current,

Θ denotes the magnetic flux in the coil.

Analogically, for the system of Fig. 5:

$$dv/\mathrm{d}t = -(C_a M)^{1/2} F(v) - f; \qquad df/dt = v; \tag{2}$$

where

$$t = \mathbf{t}/(C_a M)^{1/2},$$
$$f = f_c(C_a M)^{1/2} = x/(C_a M)^{1/2},$$

f_c denotes force applied to string compliance, x denotes the string displacement.

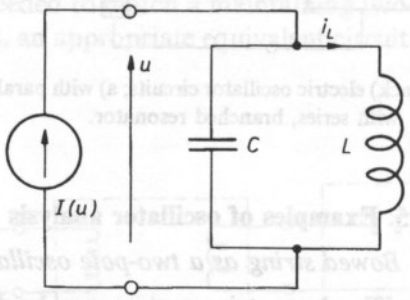


Fig. 6. Electrical circuit equivalent to the system of Fig. 5.

The nonlinear function $F(v)$ from the equation (2) represents the force of friction between the bow and the string. The friction is a function of the relative bow-string velocity, i.e. the difference between the string- and the bow-velocities $(v_s - v_b)$. The function depends strongly on the degree of bow rosining. This dependence, being essential for bowing excitation action, can be taken into account based on investigations reported in literature [3] [4], where values of friction force vs. relative bow-string velocity were conclusively measured.

Based on those results the nonlinear function $F(v)$ may be described by the following expression:

$$F(v) = \frac{T_o \operatorname{sign}(v_s - v_b)}{1 + k \mid v_s - v_b \mid}, \tag{3}$$

where

T_o denotes static friction force at $(v_s - v_b) = 0$

k is a coefficient depending on rosining.

The shape of function $F(v)$ is depicted in Fig. 7.

Substituting expression (3) into (2) gives a set of equations describing oscillations of the string point under the bow, on the phase plane. The variables of the set

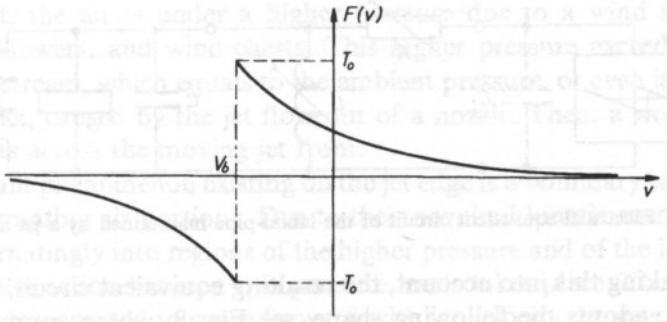


Fig. 7. Bow-string friction force as a maintenance function of the system.

represent the string point acceleration, multiplied by the constant value $(C_a M)^{1/2}$, and the string point velocity in function of the dimensionless time. The first variable is a differential of the second one. Solution on the phase plane yields trajectories of a phase point, which, describing behaviour of the system in transient states, tends to a limit-cycle determining steady-state oscillations. This analysis runs quite similarly to that one of an electric analogous oscillator.

5.2. Labial pipe as a two-pole oscillator

A sounding labial pipe may be represented by the acoustical equivalent circuit shown in Fig. 2 a). Theoretically the pipe resonator should be treated as a system with distributed constants, however, under usually accepted approximation it can be represented by a simple parallel equivalent circuit with lumped constants. When connected to a two-pole, which represents an air-jet maintenance action, the pipe resonator is controlled by the volume-velocity at the open pipe end, i.e. the air column at this end is accelerated under influenced of the pressure, delivered by the air jet, pumped through the wind chests from a bellow.

This topological condition of oscillation excitation was known from long ago. It was already stated by Lord RAYLEIGH [11]. He added then a comment concerning alternative deflections of the jet inwards and outwards of the pipe, and thought this motion to be maintaining oscillations within the pipe. His next remark concerned an accurate adjustment of the jet to the pipe, which was a decisive condition of oscillation onset. He noticed, however, that once the oscillation started that condition became less exacting.

Besides, Lord Rayleigh described experiments which showed that the natural frequency of the flue pipe resonator excited from an external source is lower than the frequency maintained by an usual blast.

Those remarks and, first of all, many experimental observations convince us, that a coupling two-pole is an unavoidable element of the equivalent circuit, which should represent mainly a compliance L', due to an influence of the static pressure exerted by the blast within the labium chamber, and a resistance R, due to losses of an air flow through the flue. Values of those coupling elements depend on operating conditions

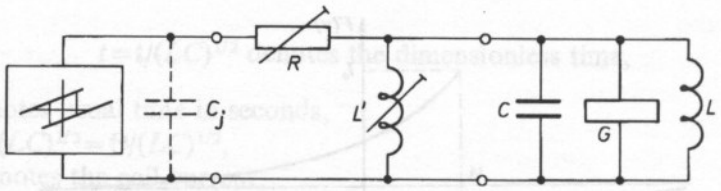


Fig. 8. Electrical equivalent circuit of the labial-pipe maintained by a jet action.

of the blast. Taking this into account, the resulting equivalent circuit, turned into its electrical form, adopts the following shape, see Fig. 8, where, moreover, an extra inertance C_j, marked with broken lines, has been added.

The inertance C_j is an analogue of the, so called, dynatron capacitance, a fictitious element playing an important role in the theory of oscillators. Under its influence, the operating frequency of a negative conductance, voltage controlled oscillator with simple resonant LGC circuit, is always lower than the natural frequency of the resonant circuit [6].

We have measured the operating frequency and the natural frequency for eight labial pipes of various types and pitches, and we have found all operating frequencies lower than natural ones. Thus, this property of acoustical oscillators is again analogous to electrical circuits.

Thanks to above proposed equivalent circuit configuration, the mentioned discrepant observations [11] or other similar measurement results [10] may be interpreted as occuring in circuits having low value of C_j and relatively high value of L'. In such circiuts the operating frequency may be indeed higher than the natural one. Moreover, the dependence of the compliance L' on wind pressure causes a reduction of the resulting circuit compliance with pressure increase, which, in turn, augments the operating frequency. Such mechanism explains why in acoustical oscillator the operating frequency increases with increasing wind pressure.

The jet action is nonlinear when its velocity is above certain limit value of the Reynolds number, which determines the turbulent air flow. The Reynolds number for a jet velocity $v$inside a pipe with a diameter d is:

$$\mathrm{Re} = \rho/\mu$$

where ρ/μ is the kinematic viscosity of the air.
Limit values to be outvalued as a turbulence condition, quoted in the literature, are inconsistent and discrepant. The resulting jet velocities, however, should be in the range not less than about several m/s.

The nonlinearity is an elementary condition for possible oscillation maintenance by a two-pole. It is difficult to describe its further properties, because of the dynamical character of the flow. However, assuming that an observer moves together with the jet, it is possible to create a simple, quasi-static velocity-pressure characteristics of the jet action.

Inside a jet, the air is under a higher pressure due to a wind supply system, composed of blowers, and wind-chests. This higher pressure exceeds the pressure around the jet stream, which equals to the ambient pressure, or even is lower, thanks to sucking effect, caused by the jet flow out of a nozzle. Then, a working pressure difference exists across the moving jet front.

An important phenomenon existing on the jet edge is a boundary layer, composed of vortices of rotating air portions. Due to their acquired kinetic energy of rotation, they enter alternatingly into regions of the higher pressure and of the lower one, then a rapid switching action between two pressure levels takes place. Those phenomena are, of course, transported in space forward with jet velocity, however, they may be treated as independent of their position in space, within a certain time, sufficiently long in comparison to several oscillation periods.

The main simplification of the jet action characteristics depends on an assumption of the plane front edge of the jet. Then only one coordinate is variable, while those representing the second and the third dimension, adopt constant values, thus, although ineffective in circuit description, they permit to keep the proper unit denominations for the corresponding quantities. None the less, the value of an air resistance, keeping its dimension, remains a real quantity.

Basing on those assumptions, a quasi-static characteristics of the jet action may be sketched, see Fig. 9. The two straight sections, marked with a heavy line, and denoted by *H* (for higher) and *L* (for lower pressure) represent the relation between pressure and particle velocity, for the two operational levels of static pressure.

Due to mentioned process of turbulence in the jet boundary layer, some portions of air at higher pressure are injected into the lower pressure region, from where, after a time, some portions are again injected into the higher pressure region. This was schematically depicted in Fig. 9. Such injection-switching may occur at velocities either slightly above the mean jest velocity, taken here as zero value on the abscissae axis, or equal, or slightly below it, so as it is shown on the diagram with broken lines.

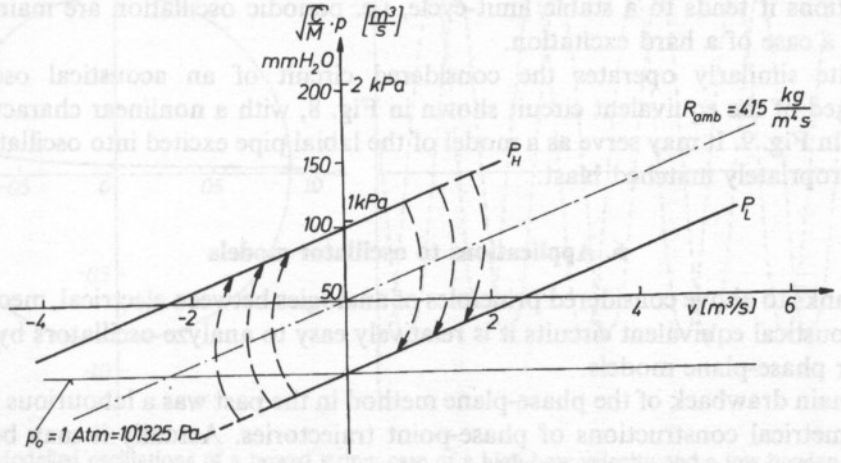


Fig. 9. A quasi-static characteristics of a jet action.

The derived, quasi-static characteristics of the jet action is quite similar to a static current-voltage characteristics of a two-pole, compound of a N-shaped, symmetric, negative conductance two-pole, with a resistance in series. The components and the resultant nonlinear characteristics are shown in Fig. 10. Authors investigated oscillators of this kind several years ago, and found them advantageous in various applications, due, mainly, to their stable operation [1].

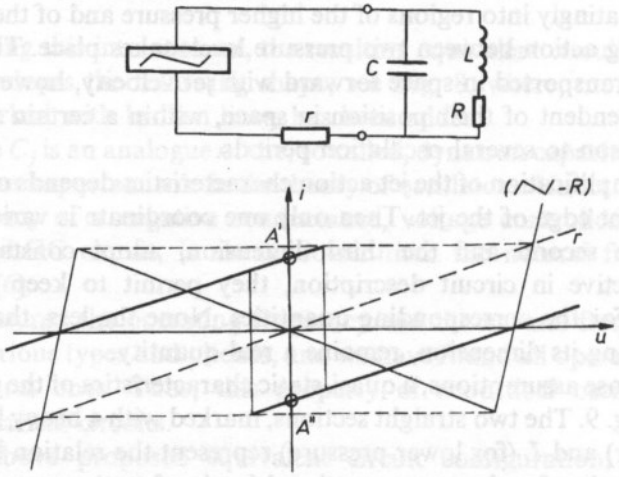


Fig. 10. A negative conductance two-pole characteristics with bistable operating point.

Treating the coordinates of the diagram shown in Fig. 10 as phase-plane coordinates, i.e. assuming they have been appropriately reduced and compensated, according to a parallel connected resonant circuit, it is possible to draw phase point trajectories, which show us the circuit oscillatory behaviour. For small displacements the phase point tends to one of the stable operating points, A' or A''. For larger elongations it tends to a stable limit-cycle, i.e. periodic oscillation are maintained. This is a case of a hard excitation.

Quite similarly operates the considered circuit of an acoustical oscillator, combined of the equivalent circuit shown in Fig. 8, with a nonlinear characteristics shown in Fig. 9. It may serve as a model of the labial pipe excited into oscillations by an appropriately matched blast.

6. Applications to oscillator models

Thanks to above considered principles of analogies between electrical, mechanical and acoustical equivalent circuits it is relatively easy to analyze oscillators by means of their phase-plane models.

A main drawback of the phase-plane method in the past was a labourious process of geometrical constructions of phase-point trajectories. Actually it may be easily replaced by an appropriate computer program.

Beneath, a few selected examples of the phase-plane solutions are presented. Models were executed by means of a PC AT computer with a special program written in the Turbo Pascal language. For calculations the Runge-Kutta method was applied.

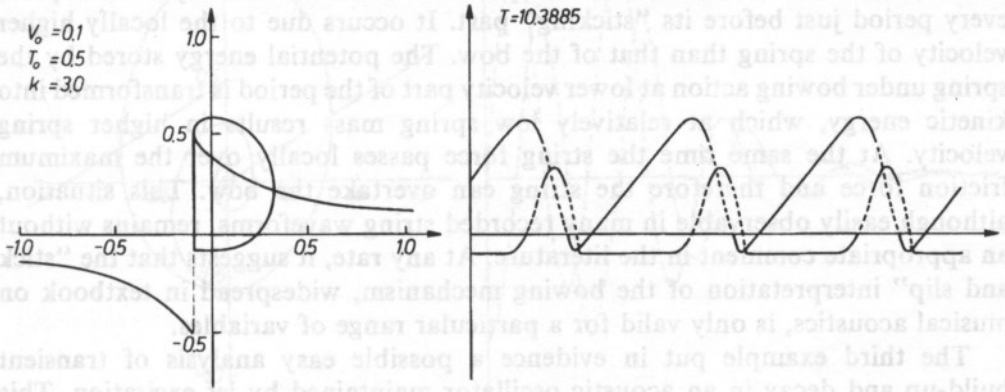


Fig. 11. Modelled oscillations of a bowed string; case of a bow velocity and a high bowing force; — saw-tooth waveforms represent force vs. time, — cut-off sines represent velocity vs. time.

The first example represents the case of string vibrations maintained by bow action at very low bowing velocity, see Fig. 11. Both waveforms, of velocity and of displacements are quite similar to those observed by many authors in such conditions, i.e. for a bow position very near to the bridge. A high value of applied bow force enhances the typical saw-tooth shape of displacement waveform.

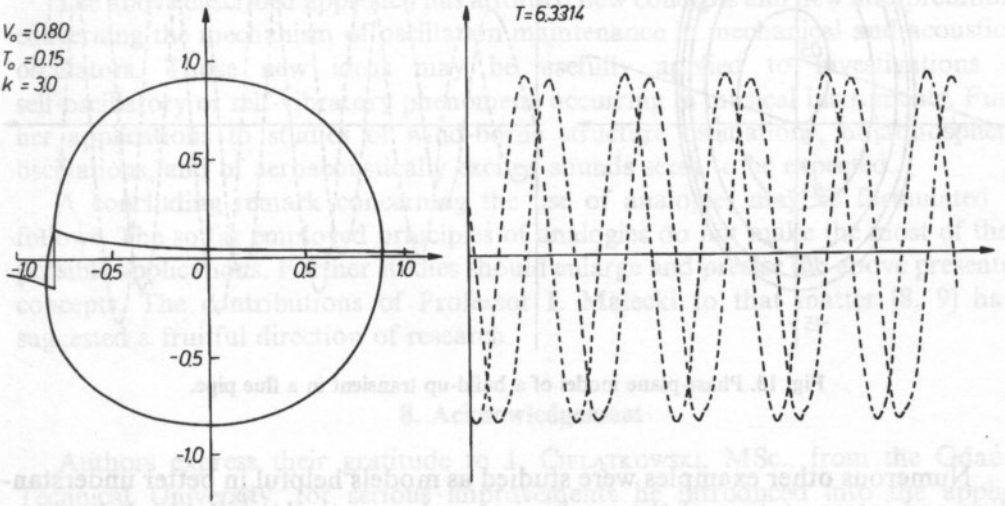


Fig. 12. Modelled oscillations of a bowed string; case of a high bow velocity and a low bowing force; — velocity waveform shows typical distortions caused in intervals when string ahead of bow.

The second example concerns the conditions extremely different from the first one, see Fig. 12. The bow force is smaller at high bowing velocity. The waveforms result of almost sinusoidal shape. The "sticking" parts of period became very short, almost unremarkable. Besides, a peculiar distortion of the cycle takes place every period just before its "sticking" part. It occurs due to the locally higher velocity of the spring than that of the bow. The potential energy stored by the spring under bowing action at lower velocity part of the period is transformed into kinetic energy, which at relatively low spring mass results in higher spring velocity. At the same time the string force passes locally over the maximum friction force and therefore the string can overtake the bow. This situation, although easily observable in many recorded string waveforms, remains without an appropriate comment in the literature. At any rate, it suggests that the "stick and slip" interpretation of the bowing mechanism, widespread in textbook on musical acoustics, is only valid for a particular range of variables.

The third example put in evidence a possible easy analysis of transient build-up and decay in an acoustic oscillator maintained by jet excitation. This analysis, shown in Fig. 13 and 14, is, of course, oversimplified, because it assumes an abrupt start of the jet flow, as well as an abrupt stop of its action, however, it makes evident a distinctly longer build-up — than a decay-duration.

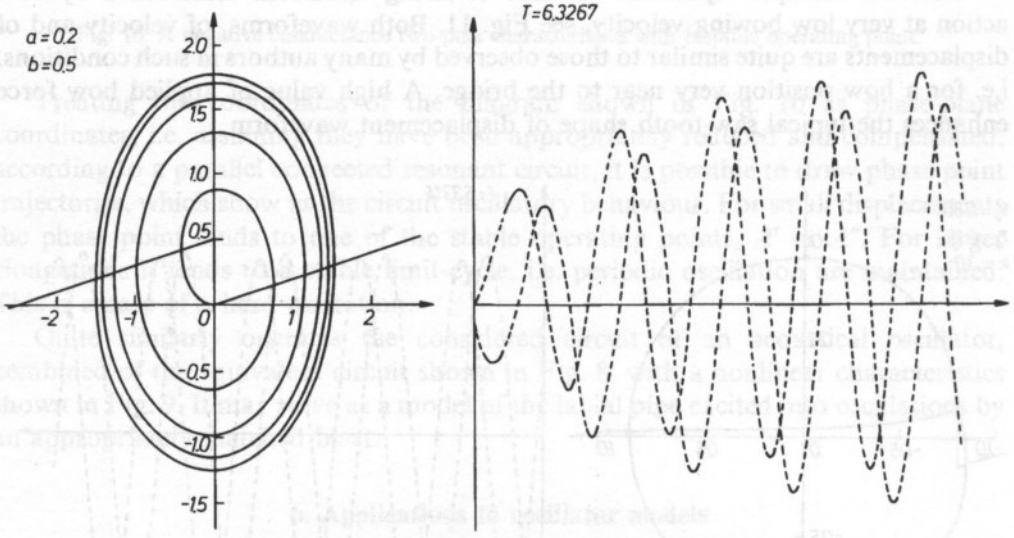


Fig. 13. Phase-plane model of a build-up transient in a flue pipe.

Numerous other examples were studied as models helpful in better understanding and interpretation of the mechanical and acoustical oscillators behaviour, thanks to application of analogies.

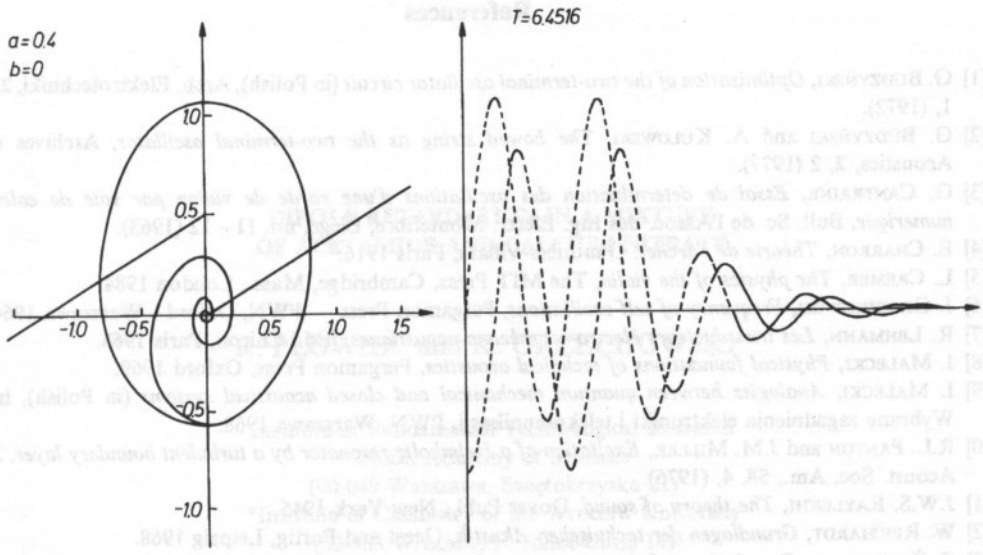


Fig. 14. Phase-plane model of a decay transient in a flue pipe.

7. Conclusions

The presented enlargement of the application of the electroacoustical analogies onto a class of nonlinear two-poles, able to maintain oscillation in resonant circuit, has allowed to analyze easily mechanical and acoustical oscillators. It has been achieved thanks to possibility of comparisons to very well known and easy to investigate properties of electrical oscillators.

The above described approach has afforded new concepts and new interpretations concerning the mechanism of oscillation maintenance in mechanical and acoustical oscillators. Those new ideas may be usefully applied to investigations of self-oscillatory or self-vibratory phenomena occurring in musical instruments. Further applications to studies of wind-borne structure oscillations, of atmospheric oscillations, and of aeroacoustically excited sounds seem to be expected.

A concluding remark concerning the use of analogies may be formulated as follows. The so far employed principles of analogies do not make the most of their possible applications. Further studies should enlarge and precise the above presented concepts. The contributions of Professor I. Malecki to that matter [8, 9] have suggested a fruitful direction of research.

8. Acknowledgement

Authors express their gratitude to J. CIELĄTKOWSKI, MSc., from the Gdańsk Technical University, for serious improvements he introduced into the applied computer programs, which permitted to study efficiently the modelled acoustical oscillators.

References

[1] G. BUDZYŃSKI, *Optimization of the two-terminal oscillator circuit* (in Polish), Arch. Elektrotechniki, **21**, 1, (1972).

[2] G. BUDZYŃSKI and A. KULOWSKI, *The bowed string as the two-terminal oscillator*, Archives of Acoustics, **2**, 2 (1977).

[3] G. CANTRAINE, *Essai de determination des oscillations d'une corde de violon par voie de calcul numerique*, Bull. Sc. de l'Assoc. des Ing. Electr. Montefiore, Liège, no. 11–12 (1963).

[4] E. CHARRON, *Théorie de l'archet*, Gauthiers-Villars, Paris 1916.

[5] L. CREMER, *The physics of the violin*, The MIT Press, Cambridge, Mass., London 1984.

[6] J. GROSZKOWSKI, *Frequency of self-oscillations*, Pergamon Press — PWN, Oxford–Warszawa 1964.

[7] R. LEHMANN, *Les transducteurs électro- et mécano-acoustiques*, [Ed.] Chiron, Paris 1963.

[8] I. MALECKI, *Physical foundations of technical acoustics*, Pergamon Press, Oxford 1969.

[9] I. MALECKI, *Analogies between quantum mechanical and closed acoustical systems* (in Polish), in: Wybrane zagadnienia elektroniki i telekomunikacji, PWN, Warszawa 1968.

[10] R.L. PANTON and J.M. MILLER, *Excitation of a Helmholtz resonator by a turbulent boundary layer*, J. Acoust. Soc. Am., **58**, 4, (1975).

[11] J.W.S. RAYLEIGH, *The theory of sound*, Dover Publ., New York 1945.

[12] W. REICHARDT, *Grundlagen der technischen Akustik*, Geest and Portig, Leipzig 1968.

[13] Z. ŻYSZKOWSKI, *Foundations of electroacoustics*, (in Polish) WNT, Warszawa 1966.

ARCHIVES OF ACOUSTICS
19, 1, 123–130 (1994)

DIPOLE RETARDATION IN A MIXTURE OF ACETAMIDE AND CALCIUM NITRATE

R. PŁOWIEC and K. ORZECHOWSKI*

Institute of Fundamental Technological Research
Polish Academy of Sciences
(00-049 Warszawa, Swiętokrzyska 21)
*Institute of Chemistry of the Wrocław University
(50-383 Wrocław, F. Joliot-Curie 14)

The results of investigations of the dielectric properties of an acetamide and calcium nitrate mixture ($x_{acetam} = 0.7644$) are presented. They are to be complementary to ultrasonic and viscoelastic studies made previously [4]. It has been established that two relaxation processes exist. They are referred to the previously described ultrasonic relaxation spectra and retardation in the viscoelastic spectrum.

1. Introduction

Mixtures of acetamide and inorganic salts containing mono- and bivalent cations tend to create high viscosity supercooled liquids. This makes it possible to investigate structures of these systems in metastable, glassy state.

In previous published articles [1, 2] the acoustical and viscoelastic properties of a mixture of acetamide and sodium thiocyanate were investigated. The results of these investigations lead to the separation of the relaxation process of a structural nature. One supposes that this process is connected with the formation of polymeric chains in acetamide "coupled" by ions of metals. Dielectric relaxation measurements, made for the same mixture, support the existence of polymeric structures in glassy state [3]. Two dielectric relaxation times were found. The low-frequency process with a high dielectric increment was interpreted as a result of the cooperative movement of ions along the polymeric chain. The second one, which occurs in the megacycles/s range, was connected with the relaxation of $Na(CH_3CONH_2)_2^+$ groups in a high-viscosity environment.

Mixtures of acetamide and bivalent ions also form glassy states, but their stability is much lower than similar ones with monovalent ions. It can be expected that, also in this case, polymeric structures connected by metal ions are created. However,

investigations of the acoustical and viscoelastic properties of the mixture of acetamide and calcium nitrate have not given a correct answer as to whether the proposed polymerization process occurs in this mixture. In the used frequency range intensive relaxation of the chemical nature has been observed; however it can also be the relaxation connected with the vicinity of the phase transormation, called eutectic relaxation [4]. It can be assumed that the great electric charge of the metal ion leads to a stiffening of the polymer structure and shifting of the structural relaxation into the lower-frequency range. To check this idea dielectric relaxation investigations have been taken up.

In the absence of an external electric field, liquids which consist of molecules with a constant dipole momentum do not indicate spontaneous polarization. Under the influence of the applied electric field, the tendency of molecule rotation and their arrangement in coincidence with the direction of the electric field appears. This means that the resultant polarization originates and coincides with the direction of the field. The energy of the interaction between the dipole and electric field is many times smaller than thermic movement energy. This fact allows only for small deviation from random orientation of the dipole. Electric permeability is defined as

$$e = \frac{1}{e_0}\frac{P}{E} + 1, \tag{1}$$

where e is the electric permeability, e_0 is electric permeability in the vacuum, P is the polarization and E is electric field intensity. Electric permeability depends on the frequency of the applied field. At lower frequency, when the dipole reorientation follows changes of the electric field, electric permeability achieves the maximum. When the frequency of the electric field increases, than the dipoles do not follow these field changes and, consequently electric permeability becomes the complex value. The dependence of the complex permeability on frequency for the dielectric relaxation process with single relaxation time is given by Debye's equation:

$$e^*(\omega) = e_\infty + \frac{e_{st} - e_\infty}{1 + j\omega t_D}, \tag{2}$$

where e_∞ and e_{st} are the electric permeabilities measured at high and low frequencies, respectively, ω is the angular frequency and t_D is the time of the dielectric relaxation. Separating the complex permeability in the real and imaginary parts gives

$$\begin{aligned} e'(\omega) &= e_\infty + \frac{e_{st} - e_\infty}{1 + \omega^2 t_D^2}, \\ e''(\omega) &= \frac{e_{st} - e_\infty}{1 + \omega^2 t_D^2}\omega t_D, \end{aligned} \tag{3}$$

The frequency dependence of the real and imaginary parts of the permeability is presented in Fig. 1. This dependence is similar to that for a complex compliance

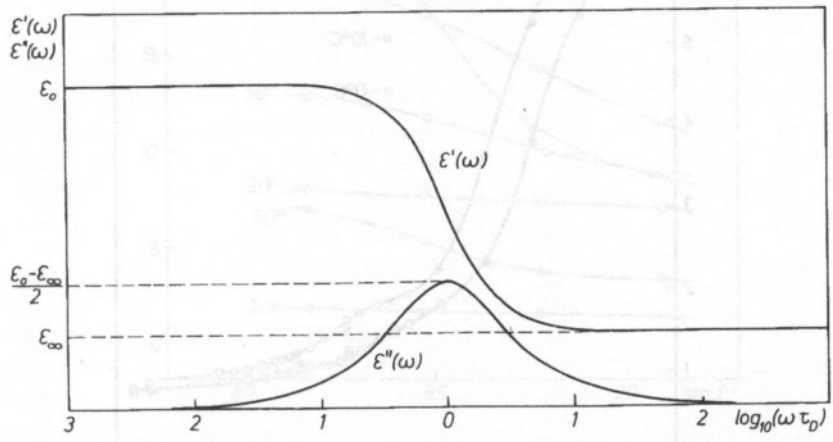


Fig. 1. The dependence of the complex electrical permeability on frequency for the monodispersive relaxation process.

$J^*(\omega)$. It should be underlined that this is only a formal similarity and results from a uniform description of the relaxation phenomena. Nevertheless, the process called dielectric relaxation can, through nomenclature analogy used in mechanical investigations, be called dielectric retardation.

2. Results of measurements

The electric permeability of the mixture of acetamide and calcium nitrate ($x_{\text{acetam}}=0.7644$) was measured in the ranges of temperature from -5 to 20°C and frequency from 0.5 kHz to 300 MHz. In the range of high temperature (20 – 110°C), the measurements were made only for frequencies 150 – 200 MHz. In the frequency ranges 0.5 – 10 kHz and 15 kHz – 4.788 MHz the Semiautomatic Tesla Bridge and Wayn-Kerr Radio Frequency Bridge type B601 were used respectively. In measurements at higher frequencies the wave-guide methods were applied; at 150 and 300 MHz the coaxial concentric line was used, while the waveguide with a rectangular section was applied at 600 and 1200 MHz.

The results of measurements are shown in Tables 1 – 3 and in Figs. 2 – 4. In the low-frequency range, the dielectric losses were calculated from the formula: $e''=1.8\,\delta(f)$ where $\delta(f)$ is the specific conductivity, f is the frequency. In Figs. 2 and 3 the black points marek the results of measurements carrried out in the same system at Prof. R. Petrhrick's lab at the Strathclyde University of Glasgow (Great Britain) [5], which covered the low-frequency range $10^{-2}-10^{5}$ Hz.

Accuracy of the measured permeability and dielectric losses was estimated at 5%.

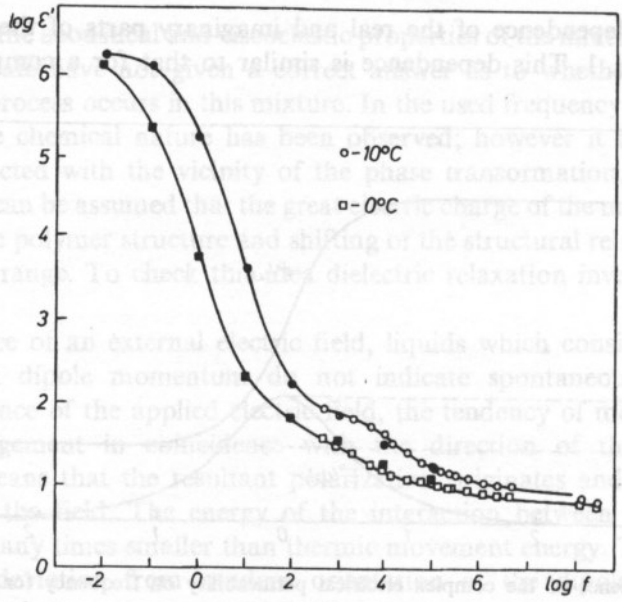


Fig. 2. The dependence of the real part of electrical permeability on frequency at the temperatures 0°C (−□−) and 10°C (−○−). The black points represent the values from Ref. [5].

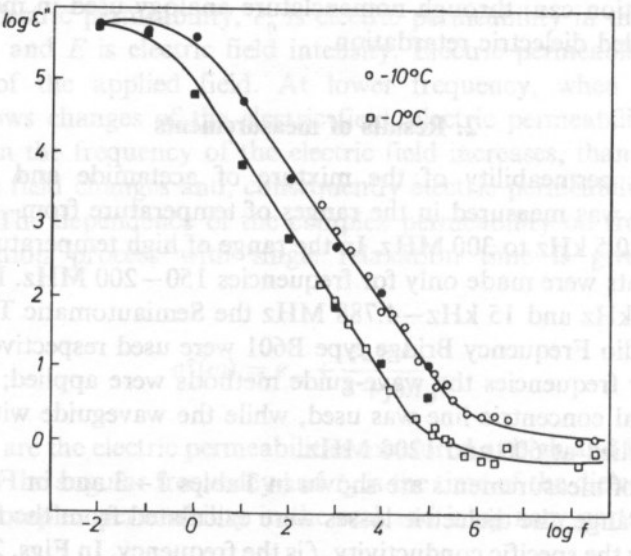


Fig. 3. The dependence of the imaginary part of electrical permeability on the frequency at the temperatures 0°C (−□−) and 10°C (−○−). The black points represent the values from Ref. [5].

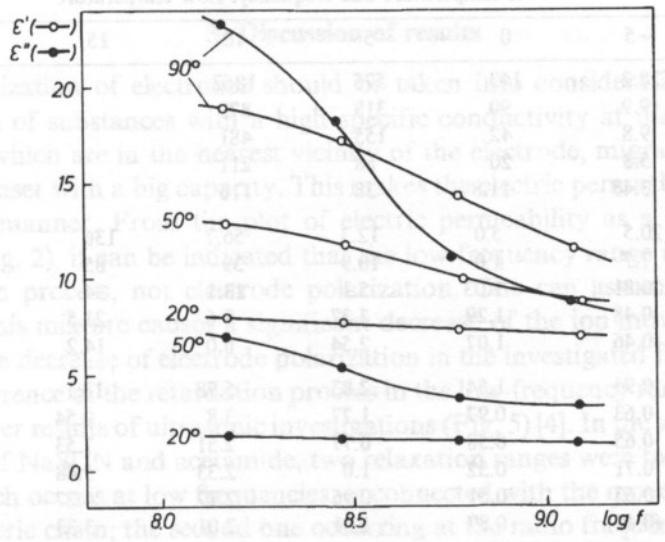


Fig. 4. The dependence of real and imaginary parts of electrical permeability on the frequency at high temperature.

Table 1. The real part of electrical permeability for the acetamide and calcium nitrate mixture as a function of temperature and frequency. Low temperature

f [kHz]	−5	0	5	10	15	20*	°C
0.5	19	37	63	89.5			
1.0	15.5	30	53	75.5			
2.0	13	24	42	66			
5.0	11.5	18	31.5	52.5			
10	10	15	24.5	42			
15	9.5	13	19	33	57	75	
30	9.2	11.5	16	25	41	65	
60	9.2	11.5	14.5	19	29	52	
120	8.5	10	12.5	16.5	26	43	
240	8.5	9.5	11.5	15	22	33	
149	8.5	10	12.2	15.4	22	33.3	
300	8.1	9.4	11.1	13.8	18.6	29.5	
596	7.6	8.8	10.3	12.2	15.4	21.2	
1211	7.4	8.4	9.8	11.1	13.5	17.5	
2382	7.0	8.0	9.7	11.0	12.8	15.5	
4788	6.9	7.7	9.0	10.4	12.0	14.2	
150000	6	6.5	6.9	7.3	7.8	8.2	
300000	5.6	6.1	6.4	6.7	7.1	7.4	

* — extrapolated

Table 2. The imaginary part of electrical permeability for the tested mixture dielectrtic losses as a function of temperature and frequency. Low temperature

f [kHz]	−5	0	5	10	15	20	°C
0.5	28.2	143	575	1862			
1	19.9	90	318	872			
2	9.8	42	155	481			
5	5.3	20	68	211			
10	3.45	11.8	38	116			
15	20.5	5.0	12.3	56.7	136	258	
30	7.7	1.0	10.9	39	85	147	
60	0.81	2.03	5.1	13.1	34	90	
120	0.48	1.29	3.37	8.6	21.5	52.5	
240	0.46	1.07	2.54	6.0	14.2	34	
149	0.94	1.54	2.83	5.78	13.1	33.4	
300	0.63	0.97	1.77	3.8	9.54	28.2	
594	0.65	0.58	0.77	2.51	7.53	17.5	
1211	0.71	0.52	1.0	2.33	5.48	11.4	
2382	0.63	0.51	0.96	2.17	4.27	7.38	
4788	(0.47)	0.87	1.31	2.01	3.33	5.65	
150000	0.34	0.5	0.7	0.96	1.24	1.56	
300000	0.6	0.74	0.94	1.15	1.4	1.64	

* — extrapolated

Table 3. The real and imaginary parts of electrical permeability for the tested mixture as a function of temperature and frequency. High temperature

T/°C	e'	e''	T/°C	e'	e''
	f=150 MHz			f=600 MHz	
101.7	19.6	31.7	108.3	14.9	13.4
91.4	18.8	25.4	98.3	14.5	11.8
80,9	16.8	18.6	86.0	13.4	9.1
70.1	15.8	13.4	74.8	12.2	7.0
59.3	14.3	8.6	63.5	11.0	5.3
49.3	12.2	5.8	53.5	9.9	3.9
39.5	10.6	4.0	43.9	8.8	2.7
30.2	8.9	2.7	31.2	7.7	1.2
22.8	7.9	1.8	22.8	7.3	1.0
8.9	6.6	1.0	12.1	6.3	0.8
	f=300 MHz			f=1200 MHz	
102.2	18.4	22.7	109.5	12.5	11.2
92.2	17.1	18.5	98.5	11.8	10.0
80.6	16.2	13.3	87.0	11.3	7.5
70.1	14.5	10.1	74.5	10.0	5.9
59.6	13.0	7.3	63.0	9.4	4.4
49.8	11.3	5.1	52.7	8.7	3.3
40.0	9.9	3.5	41.7	7.8	2.3
30.7	8.7	2.3	32.9	7.2	1.8
22.8	7.5	1.7	21.4	6.5	1.3
9.8	6.4	0.9	13.8	6.1	1.0

3. Discussion of results

The polarization of electrodes should be taken into consideration in dielectric investigations of substances with a high specific conductivity at the low-frequency range. Ions, which are in the nearest vicinity of the electrode, migrate to it forming a layer condenser with a big capacity. This makes the electric permeability increase in an apparent manner. From the plot of electric permeability as a function of the frequency (Fig. 2), it can be indicated that the low-frequency range is dominated by the relaxation process, not electrode polarization. One can assume that the high viscosity of this mixture causes a significant decrease of the ion movement. In turn, this causes the decrease of electrode polarization in the investigated frequency range.

The occurrence of the retardation process in the low-frequency range corresponds with the former results of ultrasonic investigations (Fig. 5) [4]. In the work concerning the mixture of NaSCN and acetamide, two relaxation ranges were found [1–3]. One of them, which occurs at low frequencies, is connected with the movement of cations in the polymeric chain; the second one occurring at the radio frequencies — with the reorientation $Na(CH_3CONH_2)_2^+$ groups. It seems that the low-frequency retardation, observed in the mixture of $Ca(NO_3)_2$ and acetamide, is caused by the cooperative movement of the solvated cations in the limits of the forming polymeric chains. From the measurements of $e'(f)$ and $e''(f)$, the relaxation time of this process can be estimated as $t_{(0°C)}=8$ s and $t_{(10°C)}=0.8$ s. These values are bigger nearly one order than those received in the mixture of NaSCN and acetamide ($t_{(0°C)}=0.13$ s and $t_{(10°C)}=0.06$ s) [3]. This increase of relaxation times is clear on the basis of the proposed mechanism of relaxation. The relaxation time can be connected with the ionic charge (Ze) and their mobility (K) in the following way [3]:

$$t=Ze\xi^2/nKT, \tag{4}$$

Thus the relaxation time grows when the charge of the ion, which "couples" the polymer chains, also grows. The dielectric increment, which is the difference between permeability measured at low and high frequencies, can be connected with the ion charge [3]:

$$\Delta e=e_{st}-e_{\infty}=\frac{4\pi nN}{6kT}(\xi Ze)^2. \tag{5}$$

Assuming the constancy of the correlation coefficient n, the ratio of the dielectric increments equals the ration of squares of the ion charges:

$$\frac{\Delta e(Na^+)}{\Delta e(Ca^{+2})}=\frac{Z^2(Na^+)}{Z^2(Ca^{+2})}=\frac{1}{4} \tag{6}$$

The estimation above corresponds with the experimental results:

$$\Delta e(Na^+)\simeq 4\cdot 10^6\ [3], \qquad \Delta e(Na^+)/\Delta e(Ca^{+2})=1/4,$$
$$\Delta e(Ca^{+2})\simeq 16\cdot 10^6 \quad \text{(Fig. 2)},$$

It confirms the existence of the proposed model of the cation movement in the polymeric chain as a phenomenon which is responsible for retardation in the low-frequency range.

In the mixture of NaSCN and acetamide, the existence of the additonal relaxation process with a relatively high dielectric increment at the megacycles/s range ($\Delta e = 40$ at $T = 1.6°C$) [3] was established. In the mixture investigated here, the existence of this kind of process has not been found. Only for a temperature higher than 20°C can one observe a significant increase of electric permeability and losses (Fig. 4). It seems that this relaxatoin process may be hidden in the low-frequency one with a high dielectric increment. The shift of this relaxation to the low-frequency range may result from a cross linking of the polymer with the help of great-charge ions.

4. Conclusions

The electric permeability measurement, which were made for the mixture of acetamide and calcium nitrate, enable to support the polymeric structure model proposed for this liquid. Also it was indicated that low-frequency retardation is connected with the cooperation movement of Ca^{+2} ions in the polymeric chain. In the higher-frequency range the high dielectric increment of this process dominates the second relaxation, the existence of which was established by acoustical methods [4].

References

[1] G. BERCHIESI, G. VITALI, P. PASSAMONTI and R. PŁOWIEC, J. Chem. Soc. Faraday Trans. 2, **79**, 1257 (1983).

[2] R. PŁOWIEC, A. AMICO and G. BERCHIESI, J> Chem. Soc., Faraday Trans., **81**, 217 (1985).

[3] A. AMICO, G. BERCHIESI, C. CAMETTI and A. Di BIASIO, J. Chem. Soc., Faraday Trans. 2, **83**, 619 (1987).

[4] G. BERCHIESI, G. VITALI, R. PŁOWIEC and S. BAROCCI, J. Chem., Soc., Faraday Trans. 2, **85**, 635 (1989).

[5] to be published

ARCHIVES OF ACOUSTICS
19, 1, 131–135 (1994)

LITHIUM TETRABORATE SINGLE CRYSTALS AS A SUBSTRATE FOR SAW DEVICES

J. FILIPIAK, A. MAJCHROWSKI AND T. ŁUKASIEWICZ

Institute of Technical Physics Military Technical Academy
Warszawa, Kaliskiego 2

Single crystals of lithium tetraborate ($Li_2B_4O_7$) as a potential substrate for SAW devices have been investigated. Single crystals of lithium tetraborate have been grown according to the Czochralski technique. They combine a high electromechanical coupling coefficient for SAW with a small temperature coefficient of SAW velocity.

1. Introduction

New piezoelectric materials have been investigated to find one which could combine the properties of two materials commonly used in surface acoustic wave (SAW) devices: $LiNbO_3$ with a high electromechanical coupling coefficient and $\alpha-SiO_2$ with zero temperature coefficients of time delay for some cuts. Recently lithium tetraborate $Li_2B_4O_7$ has attracted much attention as a promising SAW material [1]. The material is tetragonal and belongs to the 4 mm point group. It melts congruently at 917°C what allows for its crystals growth from the melt. The crystals of $Li_2B_4O_7$ are colourless, optically uniaxial negative and nonferroelectric. Low density ($\rho=2451$ kg/m^3) is combined with quite high hardness (6 in the Mosh scale) [1].

2. Crystal growth

$Li_2B_4O_7$ single crystals were grown from the melt using the Czochralski method [2]. A platinum crucible was used. The starting material was prepared by melting boric acid H_3BO_3 and lithium carbonate Li_2CO_3 until water and carbon dioxide were completely removed from the melt which became totally transparent. Chemical analysis showed that during the growth process evaporation of the compound enriched in B_2O_3 occurred. To reduce the formation of cores in the central parts of the crystals due to changes in the stoichiometry of the melt, 1 mol% excess of B_2O_3 was used. First, small crystalline samples of $Li_2B_4O_7$ were obtained on a platinum were, then [001] seeds were cut from the samples. The seeds were mounted in a platinum holder to avoid reaction with the alund tube. The resistance furnace giving sharp temperature gradients was used to make $Li_2B_4O_7$ crystallizaton possible. Molten lithium tetraborate has a very high viscosity and shows a tendency to solidify as

a glass in conditions of small temperature gradients. Only high temperature gradients in the vicinity of the surface of the melt ($\cong$100 K/cm) allowed to obtain $Li_2B_4O_7$ single crystals of good quality. The seed was rotated with 5 rpm and was pulled up with a speed of 0.5 mm/h. Transparent single crystals 5 cm long and 1.5 cm in diameter were grown. Only the central parts of the crystals were slightly defected by thin cores. The quality of the crystals was investigated by means of the chemical etching method in a mixture of water and glycerine (1 : 1) at 25°C for 24 hours. The density of dislocations revealed in a (001) plane, using this method, was 6×10^3 cm^{-2}. The method was also used to eliminate crystals partially defected by twinning.

A plate 30 mm long in a [001] direction, 10 mm wide in a (110) plane and 3 mm thick was cut from a $Li_2B_4O_7$ single crystal. To evaluate the usefulness of the material as a substrate for SAW devices, the following parameters were investigated:

1. The velocity of the surface wave — v;
2. Electromechanical coupling coefficient — k^2;
3. Dielectric constant (resultant) — ε;
4. Temperature coefficient of delay $\Delta\tau/\Delta T$;
5. Attenuation of the surface wave in the material;
6. The quantity of bulk waves generated in the material;

To determine the above parameters, two co-working, simple periodic interdigital transducers with 20 pairs of electrodes and with an aperture of 2910 μm were made on the $Li_2B_4O_7$ substrate. The distance between their geometrical centres was 16.81 mm, when the distance between successive electrodes d was 28 μm. The electrodes were made from aliminium and had a 2000 Å thickness.

This kind of transducer generates surface acoustic waves with a fundamental frequency described by the following formula:

$$f = \frac{v}{2d}, \tag{1}$$

and also on its fifth harmonic.

The velocity of the surface wave v was determined from the slope of the phase characteristic of the transducers in function of the frequency ω as 3.5 mm/μs.

Using the measurements of the statical capacity of the transducer in the Wayne–Kerr bridge, the dielectric constant was found as $\varepsilon = 8.5$. It is twice as big as for (Y,X) quartz and five times smaller than for (Y,Z) lithium niobate. A high electromechanical coupling coefficient k^2 connected with a small dielectric constant is useful when high frequencies and long transducers are used, because the resistance effect is minimalized.

The transducers conductance measured by the Wayne–Kerr bridge near the fundamental frequency gave the value of the electromechanical coupling coefficient $k^2 = 1.7\%$. The coefficient is ten times bigger than for (Y,X) quartz and three times smaller than for (Y,Z) lithium niobate.

The simplest way to obtain information about bulk waves generated in the transducer is the measurement of the conductance of the tranducer σ. Measurements in the wide band of frequency are very difficult and time-consuming, hence measurements were performed of the reflection coefficient of the wave from the transducer with respect to 50 Ω. Its value is proportional to $|\sigma+j\omega c|$, so it includes the same information as the conductance σ. The measurements were made using the Wiltron measuring set. The shape of the coefficient versus the frequency curve is shown in Fig. 1. The surface wave is generated at 62.5 and

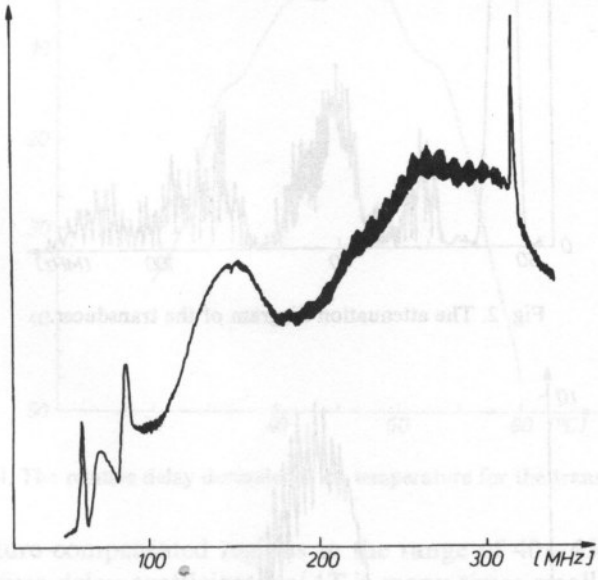


Fig. 1. The reflection coefficient dependence on the frequency for the $Li_2B_4O_7$ transducer.

312.5 MHz. On the contrary, bulk waves fill the area between the fundamental frequency and its fifth harmonic. They are present almost throughout the area and their quantity exceeds the quantity of surface waves. From this point of view lithium tetraborate is not superior to (Y,X) quartz and (Y,Z) lithium niobate. These effects restrict the possibilities of constructing SAW transducers to cases in which their bands are below their structural frequency because the attenuation of transducers above their structural frequency is influenced by such large bulk waves. The influence is shown in Fig. 2. Figure 3 shows a characteristic of the transducer in its working band. The characteristic is independent of bulk waves. Its shape and irregularities (caused by the wave reflections from the transducers) are totally compatible with the theoretical curve. Filter losses are equal to 16 dB on the structural frequency and 36 dB on the fifth harmonic. The quantities are compatible with the theory and give the evidence that the material attenuation of the wave in this range of frequencies is negligible.

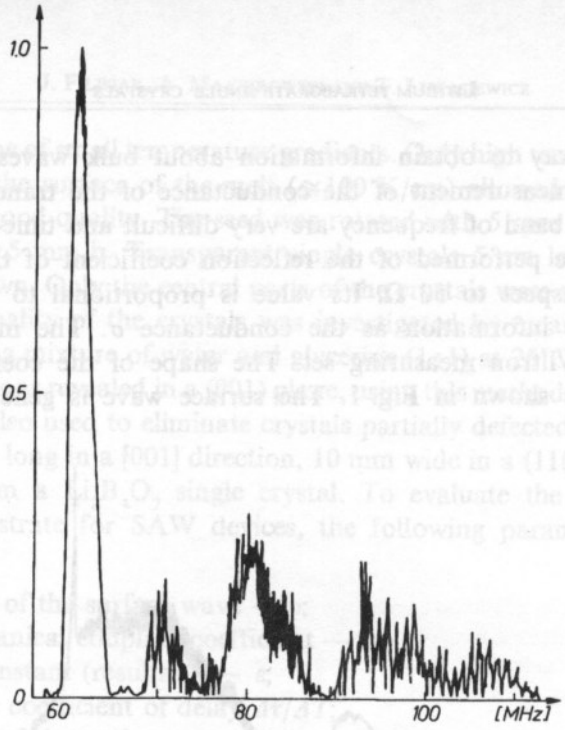


Fig. 2. The attenuation diagram of the transducer.

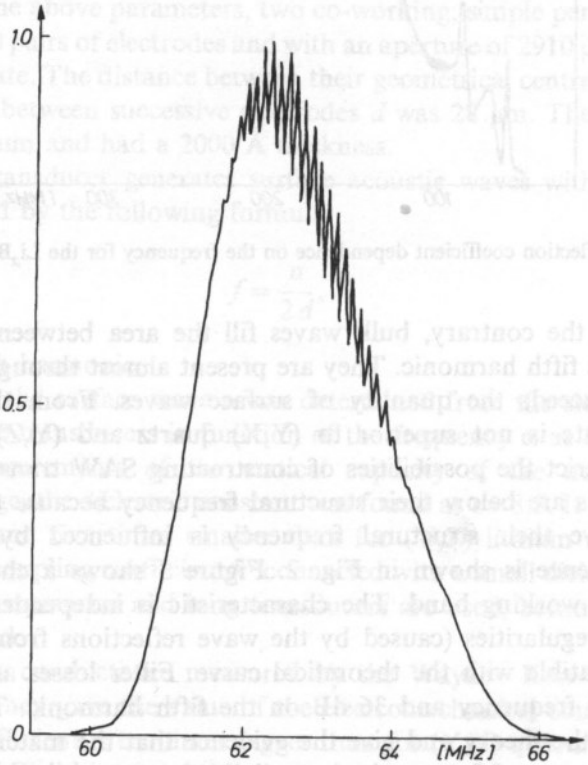


Fig. 3. The transducer characteristics in its working band.

The coefficient of temperature delay $\Delta\tau/\Delta T$ was measured in the generator set in which the SAW unit worked as a delay line in the loop of the amplifier feedback. The generator was set at 62.5 MHz. Measurements of the working frequency of the generator versus temperature in the range of 20–80°C were made. The relationship has a parabolic form (Fig. 4) and the vertex of the parabola is at 50°C. Besides

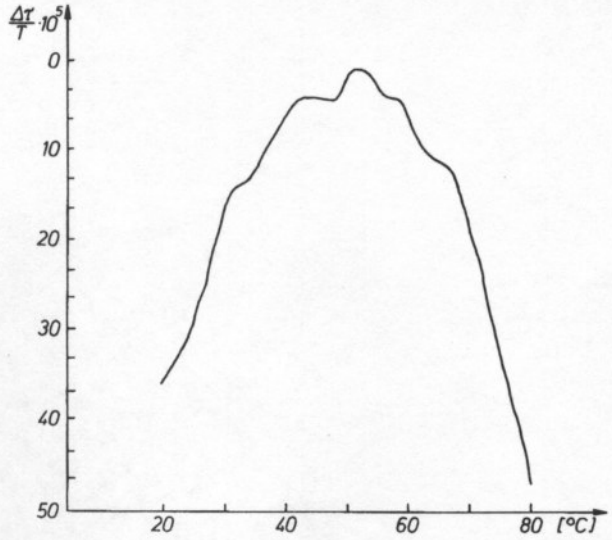


Fig. 4. The relative delay dependence on temperature for the transducer

existing temperature compensated regions in the range of 40–45°C and 50°C, the $Li_2B_4O_7$ temperature delay coefficient $\Delta\tau/\Delta T$ is many times smaller than for (Y,X) quartz ($\Delta\tau/\Delta T$=20 ppm) and (Y,Z) lithium niobate ($\Delta\tau/\Delta T$=90 ppm). From this point of view, the properties of lithium tetraborate distinctly exceed the properties of (Y,X) quartz and (Y,Z) lithium niobate.

References

[1] A. BALLATO, J. KOSIŃSKI and T. ŁUKASZEK, IEEE 1989 Ultrasonics Symposium, p. 441.
[2] T. ŁUKASIEWICZ and A. MAJCHROWSKI, Materials Letters, **11**, 8, 9, 281–283, 1991.