ARCHIVES OF ACOUSTICS
24, 1, 3–24 (1999)

INVESTIGATING POLYNOMIAL APPROXIMATIONS FOR THE SPECTRA OF THE PIPE ORGAN SOUND

A. KACZMAREK, A. CZYŻEWSKI and B. KOSTEK

Sound Engineering Department,
Faculty of Electronics, Telecommunications and Informatics,
Gdańsk University of Technology
(80-952 Gdańsk, Poland)

An attempt was made to find some features describing the shape of the spectra of organ pipe sounds in terms of their subjective discernibility. A multi-channel recording system was employed in order to represent reliable pipe sounds in a realistic acoustic interior. A precise method for the determination of the spectral representation of pipe sounds was introduced. The polynomial approximation of the spectral envelope was found to be an effective tool, allowing the study of differences between sounds produced by organ pipes of various types belonging to some selected instruments. The paired comparison subjective testing procedure was applied in order to assess the similarities between sounds synthesized using polynomial smoothed spectra and the original organ sound patterns. The statistical processing of test results revealed that a direct relation exists between the type of organ pipe and the minimum order of the approximating polynomial that can be used to represent the pipe sound spectrum, as determined by the positive opinions of the experts. The applied pipe organ sound recording and processing methods, subjective testing procedures and experiment results are discussed in the paper.

1. Introduction

Numerous scientists have been researching the acoustic features of organ instruments for many years. The historical instruments, particularly, attract a great deal of attention. By following the publications on the subject over the last decade, one can observe an increasing number of researchers dealing with the subject. A partial reason for this is probably due to the development of measurement technique. On the other hand, the methodology of subjective testing is also observed to have been developing. Without it, any achievement in the field of acoustic instrumentation would be difficult to verify. Hence, the approach to the subject is twofold: objective analysis and auditory assessment. Much research was and still is striving to achieve homogeneity between the two concepts; that is, to come up with such rules for the interpretation of objective analysis results so that they would as much as possible be in accord with the subjective assessment. So far, subjective assessment has been the supreme authority in terms of sound quality evaluation.

Following are listed some facts that might help to give an idea of the scale of problems with regard to organs as related to the above aspects:

— Organs were constructed by organ builders in various epochs, usually in accordance with the contemporary concept determining the style of play on the instrument. However, what is essential here is not just the date of creating the instrument, but also the dates of its renovation. To be considered is the fact that a great many historical organs have been damaged in the course of history, and their later renovation was not always done in accordance with the construction rules applied by the builder of the original version of the instrument.

— Firms, which constructed organs, used various techniques and different technical solutions even though the epoch was the same. Differences include the technologies of pipe manufacture, the rules of scaling and voicing of the pipes, the selection of organ stops, the procedures of interior acoustic adaptation, etc.

— What is also extremely significant is how the room affects the acoustic image of the instrument located in it. An organ is usually fitted in the interior of a church whose characteristic feature is a long reverberation time. The acoustics of the interior affects not only the transient phases of sound, easily observed upon auditory assessment, but it also affects the steady-state phase. It is not as easy with organs as it is with other instruments to record sound for analytical purposes. It has to be taken into account that the sound is recorded in highly reverberant conditions and that it is not always in a uniformly distributed acoustic field. The outcome is a strong dependence of the measurement result on the location of the microphone recording the sound.

There are a great number of organ instruments in existence, a large percentage of which are of high historical value. Currently, most of these instruments have no acoustic documentation containing objective analyses which would allow the assessment of their sound features. This situation also concerns many instruments in Poland, where the authors performed the herein discussed investigations. The results of the research which aimed to identify a solution to the problem of objective classification and assessment of organ stops are presented in this paper.

2. Recording the organ sound

This paper focuses on the analysis of the sounds of organ pipes on the basis of their steady-states. Recording in real interiors is complicated due to a large number of acoustic wave reflections, with the number getting larger as the reverberation time gets longer for the frequency under examination. The method for the recording of consecutive organ pipes of the studied stop must, under these conditions, take into account the element of coincidence resulting from the changing of microphone position when recording individual organ stops.

The interiors of the churches where the investigated organ instruments are located have one more distinctive feature which introduces an additional obstacle to sound recording and analysis, especially when recording quiet sounds and ones containing harmonics of a low level. What is referred to here is background noise coming from external

sources and from the blower that each of the instruments is equipped with. This intruding noise affects the level of the acoustic noise floor which must be considered upon interpreting the results of analysis [29]. Therefore, the value of the signal-to-noise ratio should be taken into account. Consequently, the results of spectral analysis had to be limited to such components of sound that exceeded the level of the noise floor, at least in the adjacent band range. The problem is to interpret the results of spectral analysis for harmonics of low amplitude, close to the value of the noise floor level. The spectra obtained from analysis show a loss of energy in the higher harmonics. Due to this, there is a limitation in terms of the number of spectrum components to be analyzed.

As a consequence of the above considerations, a system was adopted which was based on multi-channel organ sound recording. To reduce the effect of microphone location on the results of the analysis, it was decided to use several different recordings with microphones placed in various points. Increasing the number of recording points brings about much better accuracy in the measurements, and within limits it leads to results which exactly reflect the emitted sound. To obtain the sound spectrum that a pipe actually generates, the energy for each component would need to be summed up from the entire room. Since such an operation is clearly unfeasible and impractical, a four-channel digital recording was employed which is a compromise between the accuracy and the enormous amount of work needed for this analysis.

Recordings were made in the following Gdańsk churches: St. Mary's Basilica; St. Nicholas' Basilica; Holy Trinity Church; The Church of Jesus' Holiest Heart; and The Church of St. Peter and St. Paul. The majority of the recordings were made for the full (chromatic) musical scale, in ascending order, for the following organ stops: Principal 8′, Principal 4′, Octave 8′, Octave 4′, Subbas 16′, Flute 4′, Viola da Gamba 8′, Quintadena 8′, Trumpet 8′, Posaune 16′.

With the use of software written specially for multi-channel digital recording and sound editing on a UNIX workstation, sound files were created and then edited using selected fragments from multi-channel recordings of organ pipes belonging to the same stop. To do this, a UNIX computer was equipped with a 4-channel A/D conversion card which allowed 16-bit recording with a sampling frequency of 22.05 kHz.

3. Representation of the organ sound

A computer program written at the Sound Engineering Department of the Gdańsk University of Technology performed analyses described in this section. The first task of this program was to read sound files made after editing. Next, the program performed calculations and analysis of the steady-state spectrum on the basis of the discrete Fourier transform (DFT). When distortions caused by the background noise coming from external sources and from the blower are present in the registered signal, increasing the time window so that it encompasses several periods of the signal can provide a correction of the results of the analysis [17]. On the other hand, the DFT analysis loses accuracy when performed over several periods. To make possible the precise tuning of the time window to the period of the signal or its multiples, a variable length time window procedure was

used. Data on the particular harmonics and on distortions are extracted using the above calculations. The best estimation of the signal will take place when the exact value of the fundamental frequency is known. However, in the case under discussion only an approximate value is known. This is because we must consider the possibility of organ sound mistuning with comparison to musical pitch based on the equal temperament scale. For this reason, a procedure to determine the fundamental frequency should be applied. The fundamental frequency may be obtained by changing the length of the time window in a way that leads to better tuning to the sound. After such a procedure, the spectrum and the signal-to-noise (S/N [dB]) ratio are computed according to the formula:

$$\frac{\mathrm{S}}{\mathrm{N}} = 10 \cdot \log \frac{E_S}{E_N}, \tag{3.1}$$

where:

$$E_S = \sum_{i=k}^{k \cdot n} E_i \tag{3.2}$$

represents the total energy of the harmonic spectral components, and

$$E_N = \sum_{i=1}^{k-1} \sum_{j=0}^{n} E_{i+j \cdot k} \tag{3.3}$$

represents the total energy of parasite spectral components, E_i – energy of the k-th spectrum harmonic, $E_{i+j \cdot k}$ – energy of the parasite spectral component, k – multiple of signal periods within a chosen window length, n – number of the highest spectral component.

The value of the S/N ratio is then compared to the previously computed one. The result leads to acquiring the highest possible value of the useful signal-to-noise ratio in consecutive steps until reaching the required accuracy. The main tasks performed by the developed algorithm are presented in Fig. 1. Examples of this type of analysis are presented in Fig. 2. The numbered lines in Fig. 2a and b are interpreted by the program as components of the useful signal, the others as noise. This procedure is completed for all sounds of the analyzed organ stop.

Based on the obtained results of the analyses, there are some observations useful for designing further stages of the research:

1. Spectra of the particular pipes comprising one organ stop are not identical, and they show certain discrepancies in terms of the frequency and amplitude of consecutive harmonics [13];

2. A gradual, systematic change of the spectrum is visible as one moves to extreme ends of the musical scale. This tendency is best visible for flue pipes, and is in this case the effect of scaling [13].

A statistical verification of the results was applied to the obtained data. For this purpose, a computer program was developed which performs such computations as:

- average spectrum for each organ stop, where the range of averaging may pertain to a selected part of the musical scale. This was used, for example, in researching the effect of scaling on the sound timbre. The amplitudes of components are also subjected

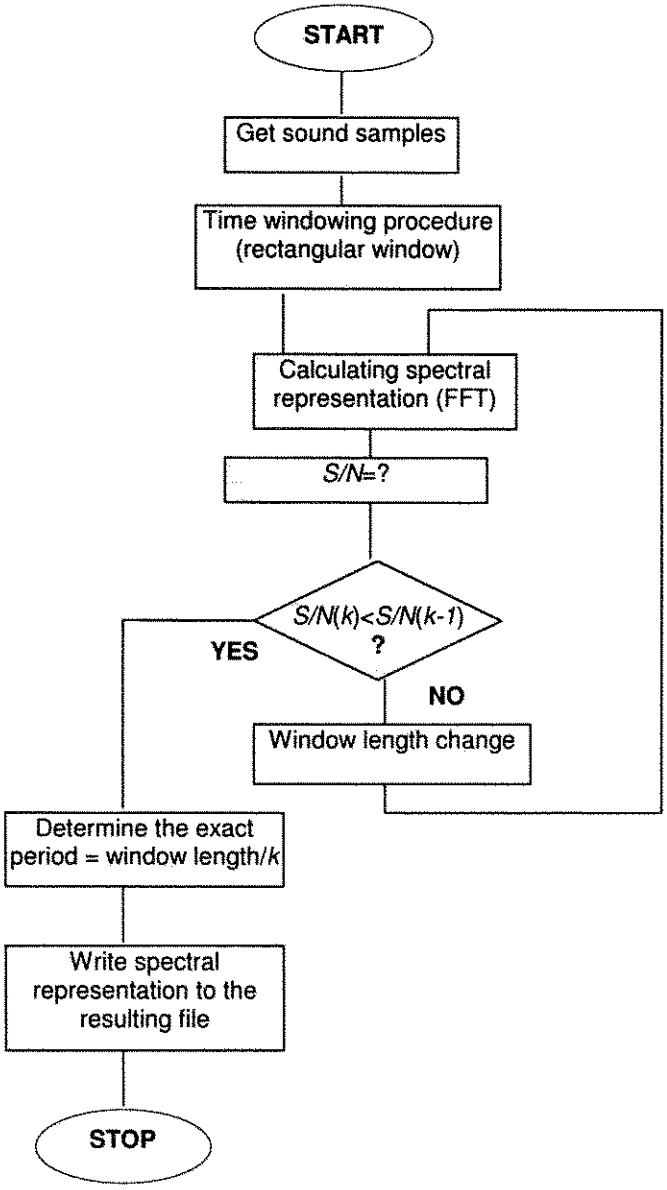


Fig. 1. Algorithm of spectrum determination using the window adjustment procedure.

to averaging computations. However, prior to averaging computations, the spectra are normalized in terms of their energy. The result is that only the relative amplitude levels of the components are considered;

• the variance computed for each harmonic is examined in order to study the dispersion in the shape of the spectrum within the given organ stop;

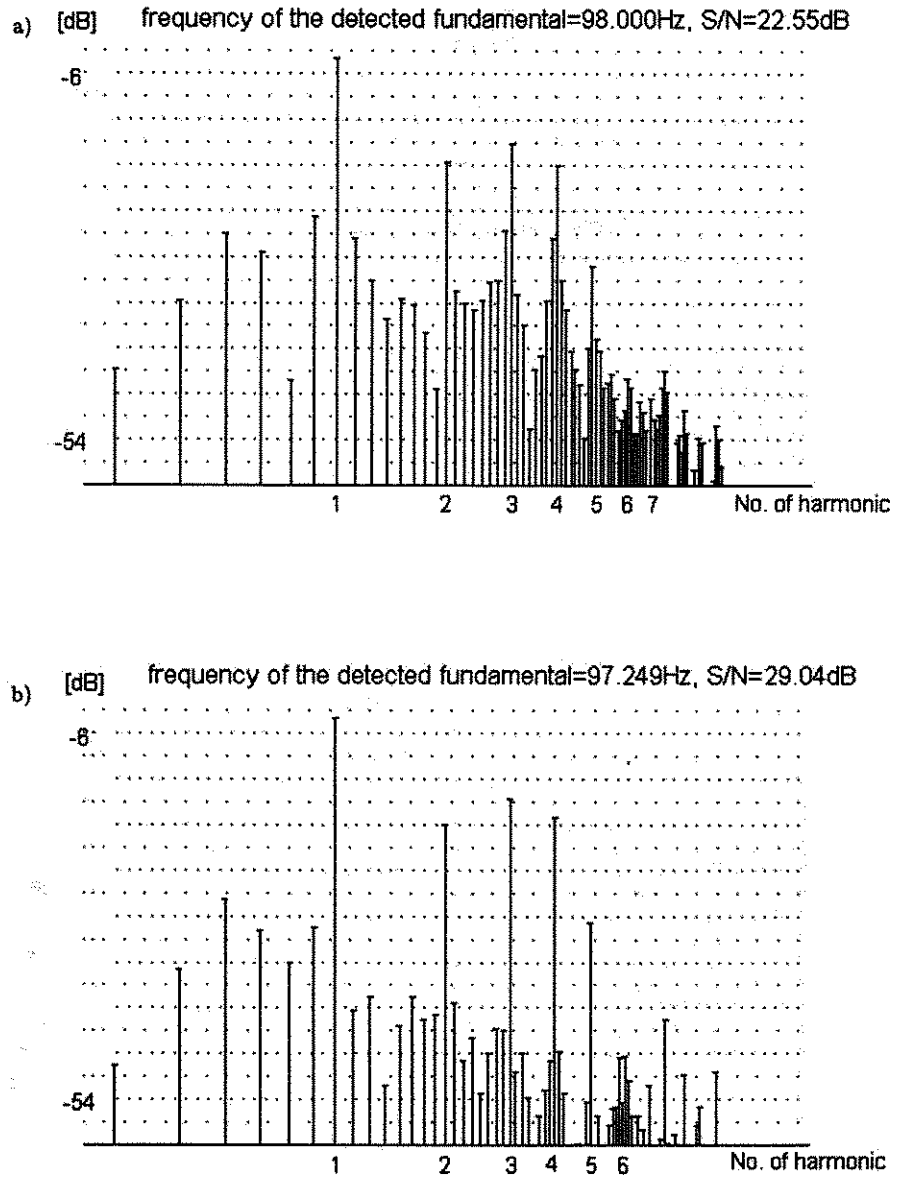


Fig. 2. Spectral analysis of the organ sounds: a) result of spectral analysis of the organ pipe sound, b) result of spectral analysis of the organ pipe sound when the tuned window length procedure was used in order to maximize the S/N ratio.

• the calculation of confidence intervals for each harmonic in order to determine estimation boundaries of the average spectrum, using the formula:

$$\frac{c_n}{2} = \frac{t^{\alpha}_{p-1} \cdot S_n}{\sqrt{p-1}}, \tag{3.4}$$

where: c_n – width of the confidence interval for the amplitude of n-th component of the average spectrum, t^{α}_{p-1} – boundary value of t Student statistics for $p-1$ degree of freedom and of significance level α, S_n – standard deviation for the amplitude of the n-th component computed from measurements according to formula (3.5):

$$S_n = \sqrt{\frac{1}{p} \cdot \sum_{i=1}^{p} \left(A_n(i) - \overline{A}_n\right)^2}, \tag{3.5}$$

where: $\overline{A}_n$ – n-th amplitude value of the n-th component of the average spectrum, n – number of the spectrum component, p – total number of sound patterns, $A_n(i)$ – n-th amplitude value of the harmonic of the i-th sound pattern spectrum after energy normalization.

The computed confidence interval for the n-th component can be written down in the form:

$$\left\langle \overline{A}_n - \frac{t^{\alpha}_{p-1} \cdot S_n}{\sqrt{p-1}},\ \overline{A}_n + \frac{t^{\alpha}_{p-1} \cdot S_n}{\sqrt{p-1}} \right\rangle. \tag{3.6}$$

The experimentally chosen value of α for the confidence intervals was equal to 0.05. Using a graphical presentation of the computed confidence intervals for each harmonic separately, a specific confidence "sleeve" is created for the entire average spectrum.

Based on the results obtained using the above specified methods, it is possible to show that it is justified to apply the multi-channel methodology of sound analysis in an acoustic interior. To do this, it is sufficient to compare the dispersion for the particular harmonics or the widths of respective confidence intervals for the average spectrum computed from a single-channel recording with corresponding values originating from a multi-channel recording. This will be shown in next section.

4. Features of organ sound

Upon reviewing the acoustical literature, it should be noted that apart from such basic terms as loudness or pitch perception, there are several or even a dozen subjective terms for sound timbre, such as brightness, depth, sharpness, warmth, etc., which do not have such simple objective equivalents as these two basic attributes [4, 26, 32]. Some of them, however, are classified as objective through the fact that a majority of researchers agreed on a certain mathematical formula to describe them, e.g. brightness [2, 22]. Other definitions of sound parameters arose on the basis of proposals made by acousticians; they were formulated on the basis of features of musical instrument sounds [1, 3, 5, 10, 18, 32, 34].

In the case of organ sounds, the applied analytical approach can be divided into two types:

- investigating whole instruments to find common features for various schools to which the builders belonged [23, 24, 30];
- investigating particular types of pipes to identify features which are distinctive for organ stops [29, 31].

Earlier research results were the outcome of applying an analog technique where the most frequently used analyzers ensured an octave or a one-third-octave resolution in the field of the spectrum [6]. This helped to define and implement several parametric representations of sound features. In these analyses, a twofold approach can be observed: long-term averaging of sounds of the entire instrument scale or of a selected stop [11, 12, 30, 33]; or an elaborate recording of sounds from single pipes [4, 8, 27, 29]. Results of the octave analysis helped, for example, to isolate certain differences between organ instruments originating from manufacturers who belonged to various schools. The basis for this was a five-element vector of parameters consisting directly of measurement values in five octaves of the acoustic band [23, 24]. Later research showed that based on these results it is possible to reduce the length of the parameter vector while at the same time maintaining classification values. This procedure was based on the matrix of a so-called cosine transformation and eventually helped obtain a three-element vector [30]. Application of the one-third-octave analysis was used to perform feature extraction of sounds from single organ pipes by means of a division of the spectrum into groups of harmonics. This methodology led to presentation of a subjective assessment of sound timbre in terms of polar coordinates [29].

Research on the transient state produced a whole series of parameters such as rising time, releasing, overblowing, etc. [14, 15, 18, 19, 35]. Apart from these, the method of describing sound timbre evolution using a respective grouping of harmonics, called the tristimulus method, appeared [31]. The research on the transient states of musical sounds [8, 9], however, especially on the attack transients, is more closely related to articulation effects in organ playing [14, 15, 16, 18, 20, 28]. Therefore it is not the main subject of this paper. Nonetheless, the authors have participated in these types of research experiments [18, 19, 20].

The idea of describing sound timbre through a set of three parameters has recently inspired research related to sounds coming from recordings of a whole musical group [25]. In this work, however, the main considerations pertain to single organ pipes. In the case of such sounds, there is a well-known method of describing this feature, called spectrum irregularity (French: *irregularité du spectre*) [22].

As will be shown in the next section, builders of organ instruments were most probably guided by the acoustic effect related to the sound spectrum when seen as a whole. For example, the sound timbre can be described through defining the size of decay of the higher harmonics in relation to the lower ones. In this case, a parameter may be proposed: the directional coefficient of a line adjusted for a set of points matching the maxima of the harmonics, expressed in dB/log f. Application of this linear approximation of the spectrum may be the point of departure to a better description of sound timbres of organ pipes. Feature extraction based on groups of harmonics is in a way a reference to such an approach [31]. However, using this extraction as a basis, it is difficult to perform additive synthesis which for the steady-state requires not only the knowledge of what the fundamental frequency is, but also the amplitudes of the particular components, and in some cases the knowledge of what the phase variations are. Therefore, it seems that it will be more beneficial to identify a more accurate description of the envelope of the sound spectrum.

The authors' proposal of such a description is based on an approximation of the spectral envelope shape using polynomials [13, 17]. This approach seems to be correct, especially in applications involving richer spectra such as the sound spectrum of reed pipes. The applied approximation is based on minimizing the mean-square error in the range of the analyzed spectrum, and is computed by using the following proposed relation:

$$E = \sqrt{\sum_{i=1}^{m} (20 \cdot \log_{10} A(i) - W_r(\log_2 i))^2}, \tag{4.1}$$

while:

$$E_r(\log_2 i) = \sum_{j=0}^{r} a_j \cdot (\log_2 i)^j, \tag{4.2}$$

where: E – mean-square error, i – number of the harmonic, $i = 1, 2, ...m$, m – number of the highest harmonic, $A(i)$ – value of the amplitude of i-th component, $W_r(\log_2 i)$ – value of the polynomial for i-th component, a_j – j-th term of the polynomial, j – number of the consecutive term of the polynomial, r – order of the polynomial.

Computations which minimize the error are performed by consecutive substitution of $r = 1, 2, \ldots$ etc. up to the assumed maximum value, successively obtaining the coefficients $a_1, a_2, a_3, \ldots$. Based on formula (4.1), the approximation is performed on the spectrum presented in the log-log scale, which causes the consecutively computed coefficients a_j, where $j = 1, 2, 3, \ldots$ to be expressed respectively in dB/octave, dB/octave^2, dB/octave^3, etc. These coefficients have a clear physical interpretation, e.g. the first defines the decay of higher harmonics in the spectrum, whereas the second indicates a gain or a loss of the middle part of the spectrum in relation to its lower or higher parts.

Moreover, by raising the approximation order, more coefficients are obtained which describe more precisely the spectrum of the real sound. Such examples of an ascending order polynomial approximation of the steady-state spectrum of an organ pipe sound pattern will be shown in the next section.

5. Statistical analysis of the recorded material

To determine whether the computed parameters – coefficients of polynomials approximating the spectrum – can be treated as distinctive features, a statistical tool was used in the study, in the form of Behrens-Fisher's statistics [7, 13]. The basic assumption is that of mean equality in two normally distributed populations. The following values are computed from the formulae:

- mean parameter values:

$$\overline{X} = \frac{1}{n} \cdot \sum_{i=1}^{n} X_i, \qquad \overline{Y} = \frac{1}{m} \cdot \sum_{i=1}^{m} Y_i, \tag{5.1}$$

- and variance estimators of respective random variables:

$$S_1^2 = \frac{1}{n-1} \cdot \sum_{i=1}^{n} \left(X_i - \overline{X}\right)^2, \qquad S_2^2 = \frac{1}{m-1} \cdot \sum_{i=1}^{m} \left(Y_i - \overline{Y}\right)^2, \tag{5.2}$$

where: n, m – cardinality of populations X and Y accordingly, $S_1, S_2 \neq 0$.

On the basis of the values obtained from Eq. (5.1) and (5.2), the statistics are computed from the following relation:

$$V = \frac{\overline{X} - \overline{Y}}{\sqrt{S_1^2/n + S_2^2/m}}, \tag{5.3}$$

which can then be compared to a respective boundary value from statistical tables under the assumed significance level. This requires additional computation of the statistical parameter c from the formula:

$$c = \frac{S_1^2/n}{S_1^2/n + S_2^2/m}, \tag{5.4}$$

which reflects the distribution of statistics V and the determination of the boundary value.

Statistics V, computed for each parameter separately, can also be treated as a measure of the distance between the compared classes in the analyzed space of parameters. It is this interpretation that was also applied in the described experiments. For the assumed and fixed n and m, it is possible to compare the values computed for various parameters related to the investigated populations.

The above described test using Behrens-Fisher's statistics was applied to study the multi-channel methodology, which can also be examined using the confidence interval method described in Sec. 3. The latter method is illustrated in Fig. 3. The figures present the average spectra of selected types of pipes from the organ at St. Mary's Basilica in Gdańsk. For each spectral component, a confidence interval was found at a level of 0.95. The range of sounds was reduced to one octave, in this case the 12 sounds of the chromatic scale from C2 to B2. The spectra in Fig. 3 are presented in the log-log scale. The x-axis is related to normalized frequency in terms of the frequency of the fundamental component. The y-axis represents the normalized energy of the particular harmonics expressed in [dB]. Comparing the width of confidence intervals computed for average spectra of both single and four-channel recording proves the advantages of applying multi-channel recording. In the latter case the confidence interval width is much narrower than in the case of a single channel recording, what signifies a smaller dispersion of computed spectral components (see Fig. 3).

This observation is further confirmed by the results of the statistical analysis (Table 1) and analysis, which is based on the Behrens-Fisher's statistics (Fig. 4). The approximation parameters which were used as initial data in the test were computed on the basis of the same results. In the latter case, a limitation of four stops was introduced. The obtained results show that four-channel recording usually leads to a smaller dispersion and consequently to a better separation between the parameter values in comparison to single-channel recording.

To show the differences in the spectra of sounds from various parts of the musical scale, average spectra were computed and feature extraction was performed for these sounds. This has been illustrated in Fig. 5 in graphical form as a diagram of an approximating polynomial in points matching spectral lines. Examples of the results of the

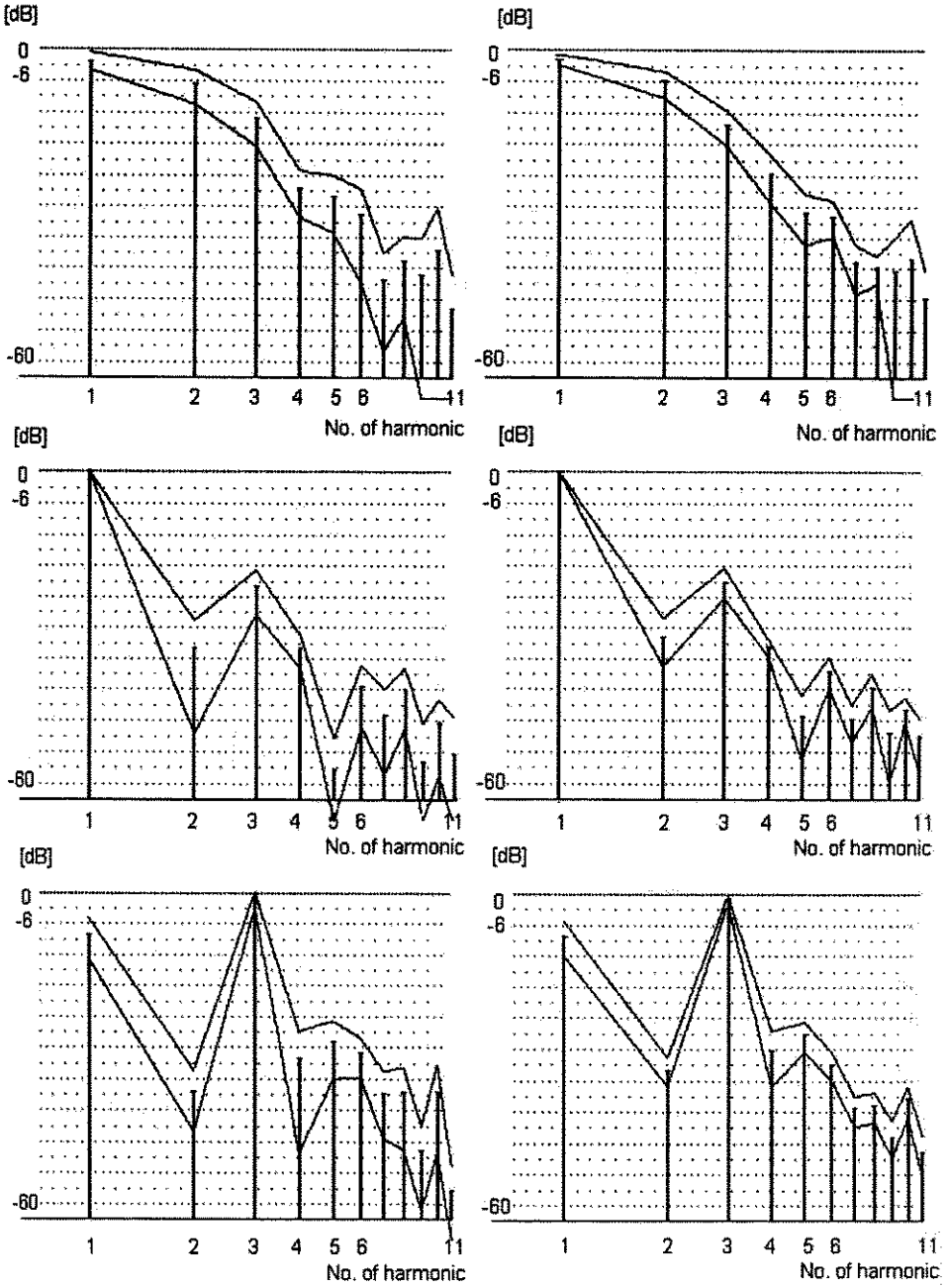


Fig. 3. Comparison between average spectra of sounds from the chromatic scale of the octave from C2 to B2 of three selected flue pipe organ stops at St. Mary's Basilica in Gdańsk. Going from the top, in successive order: Principal 8′; Subbas 16′; and Quintadena 8′. On the left, the result based on a single channel recording; on the right, based on a four-channel recording.

Table 1. Average values and dispersions for parameters obtained on the basis of single- and four-channel recording.

Principal 8′ – single-channel recording					
Parameter No.	1	2	3	4	5
mean value	−17.092	−2.621	2.138	−0.111	1.531
dispersion	3.454	3.458	2.558	2.139	4.322
Principal 8′ – four-channel recording					
Parameter No.	1	2	3	4	5
mean value	−15.542	−2.085	1.805	−0.377	−1.003
dispersion	2.415	2.074	1.007	1.421	1.440

analysis based on the Behrens-Fisher's statistics for the stops originating from different instruments are given in Table 2.

Table 2. Results of the comparisons of sound parameters between the same stops from various organ instruments using Behrens-Fishers' statistics.

Principal 8′: object I – object IV					
Parameter No.	1	2	3	4	5
Statistics V	1.488	−1.385	−1.501	0.655	1.479
c parameter	0.420	0.330	0.135	0.169	0.171
Principal 4′: object IV – object III					
Parameter No.	1	2	3	4	5
Statistics V	**9.590**	−0.581	0.315	−0.236	−2.810
c parameter	0.419	0.166	0.167	0.116	0.269
Posaune 16′: object I – object II					
Parameter No.	1	2	3	4	5
Statistics V	**4.671**	1.622	0.391	0.466	−2.415
c parameter	0.387	0.567	0.622	0.387	0.612
Posaune 16′: object I – object III					
Parameter No.	1	2	3	4	5
Statistics V	2.513	2.687	−0.207	**4.700**	−0.726
c parameter	0.179	0.385	0.827	0.138	0.465
Posaune 16′: object II – object III					
Parameter No.	1	2	3	4	5
Statistics V	−0.634	1.417	−0.800	**4.254**	1.514
c parameter	0.257	0.323	0.744	0.202	0.356
Subbas 16′: object II – object I					
Parameter No.	1	2	3	4	5
Statistics V	−1.702	−1.041	−1.301	1.643	1.174
c parameter	0.527	0.684	0.328	0.355	0.441

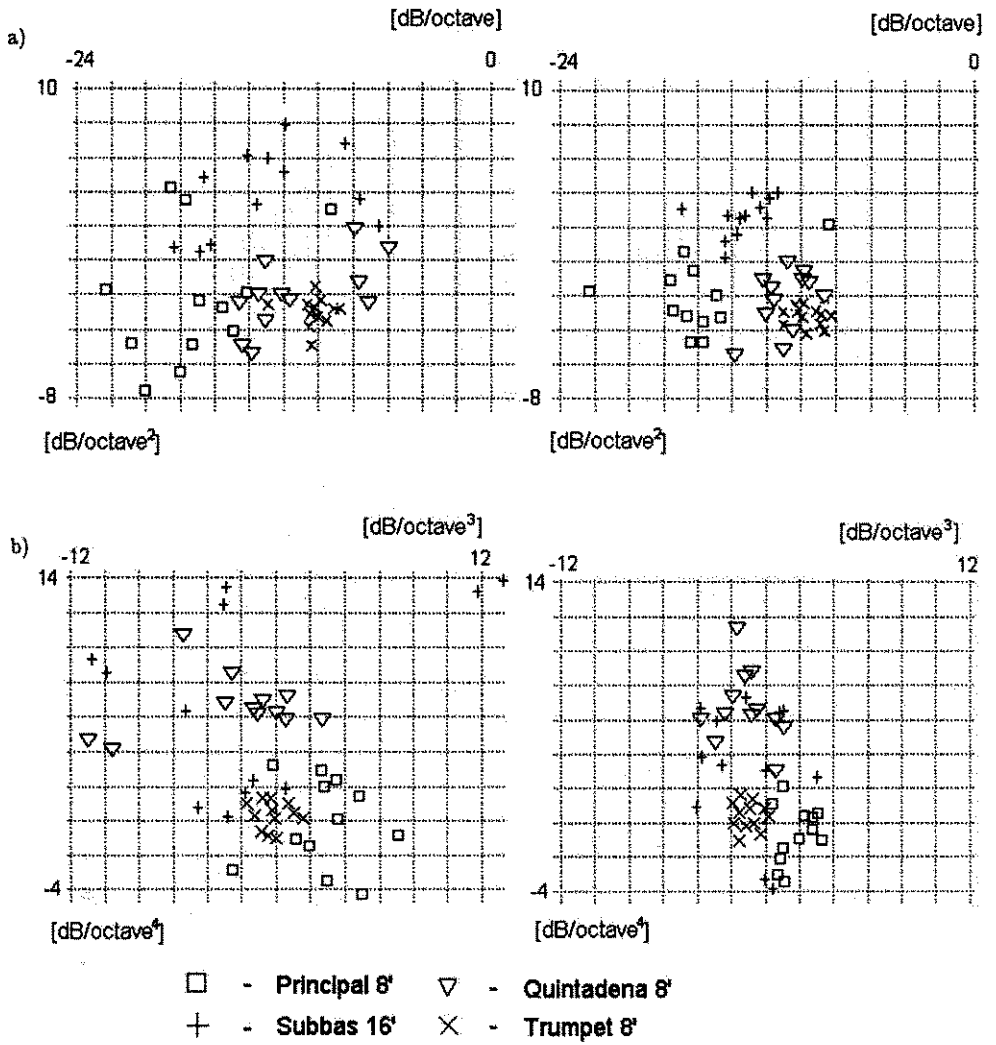


Fig. 4. Graphical presentation enabling the comparison of parameter distribution for the analyzed sounds – on the left, single-channel recording; on the right, four-channel recording: a) compared parameters are coefficients which were previously introduced: namely, linear, expressed in dB/octave; and squared, expressed in dB/octave^2, b) compared parameters are coefficients which were previously introduced: namely, 3-rd order, expressed in dB/octave^3; and 4-th order, expressed in dB/octave^4.

The chromatic scale of the octave from C3 to B3 (12 sounds) was studied for four and eight foot stops, and the chromatic scale of the octave from C2 to B2 (12 sounds) for sixteen foot stops (see Table 2). The compared stops come from organs located in Gdańsk churches, according to the following symbols: object I – St. Mary's Basilica; object II – St. Nicholas' Basilica; object III – The Church of St. Peter and St. Paul; object IV – The Church of Jesus' Holiest Heart. Values for which the value of statistics V exceeded the range from 2.845 to 3.169 were distinguished. The following assumptions

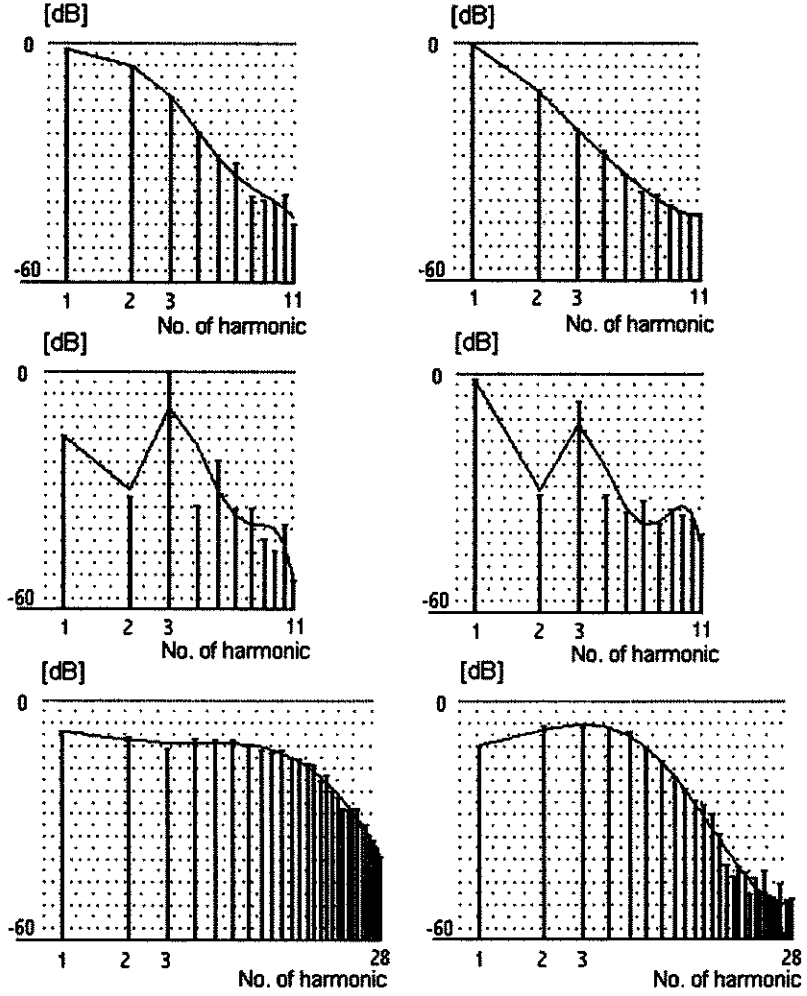


Fig. 5. Comparison between the average spectra of sounds from the chromatic scale of the octave from C2 to B2, and the octave from C5 to B5 of three selected organ stops at St. Nicholas' Basilica in Gdańsk. Going from the top, in successive order: Principal 8′; Quintadena 8′; and Trumpet 8′. On the left, the octave from C2 to B2; on the right, the octave from C5 to B5. Envelopes of the spectra are approximated with a 5-th order polynomial. Axes x and y are scaled like in Fig. 2.

were adopted: the statistical parameter c ranged from $(0, 1)$; the number of freedom degree 10; the significance level was adopted at 0.01.

The applied experiments showed that for the majority of parameters, the assumption of normal distribution of the investigated populations was correct. Some exceptions exist, however, such as: parameter 1, upon comparing Principal 4′ of objects IV with III; parameter 1, upon comparing Posaune 16′ of objects I with II; and parameter 4, upon comparing Posaune 16′ of objects I with III and II with III. This may be the result of differences due to construction details in the compared organ instruments.

6. Subjective verification

The paired comparison, non-parametric test procedure was applied [21] in order to verify whether the applied system of feature extraction provides parameters related to subjective sound perception. In the case of such a test, it was not necessary to determine particular features to be assessed by the experts (they produced overall preference scores).

For this purpose, digitally recorded real signals were used as signal patterns, and signals synthesized on the basis of additive synthesis of approximation parameters were used as test signals. The sounds used for comparison had spectra computed using approximating polynomials of orders within the range from one to five. An elementary test was to provide the real signal as the pattern and two signals synthesized on the basis of various approximation orders. All the signals had duration of 2 seconds, and the intervals between them were 0.5 seconds. The intervals between the elements of the test were 5 seconds. This was also the time designated for a subject response. The test question was, "*Which of the two synthesized signals has a timbre more similar to the model signal pattern?*"

On the basis of the experts' answers, a preferential diagram was obtained which turned out to be a rising curve for all the tests, showing a preference for the higher order approximation. Resulting from the statistical analysis of the results, the significance of this preference was also calculated. This value indirectly provides information on the approximation order that realizes a good reproduction of the investigated sound timbre. On this basis, it was possible to determine the saturation point of the preference curve.

Sounds of three various types of pipes were selected for testing: Principal; Viola da Gamba; and Trumpet. Two musical notes were used, C and B, with corresponding fundamental frequencies at about 130.8 and 246.9 Hz. The applied approximation of the order from 1 to 5 led to producing 10 pairs of sounds for one part of the test. As a result, a single testing session presented 20 pairs for comparison. The tests were carried out for two groups of experts, with a total number of 22 people. Altogether 12 test sessions were carried out.

Figures 6 and 7 illustrate graphically the exemplary spectra selected for subjective tests: real and synthetic ones.

Analysis of the results obtained as the end product of the tests was performed using a program written in the C language on a UNIX computer. Sets of answers from the experts provided direct test results that were then analyzed statistically. The schema of computations is the following:

- Summing up the number of votes given by the particular experts for each of the objects,
- Determining the stability of each experts' choices (parameter z_1),
- Determining the sum of votes given to an object by the successive experts,
- Determining the number of votes given to an object by successive experts to both parts of the test,
- Determining the statistics χ^2 by comparing the results of both parts of the test,

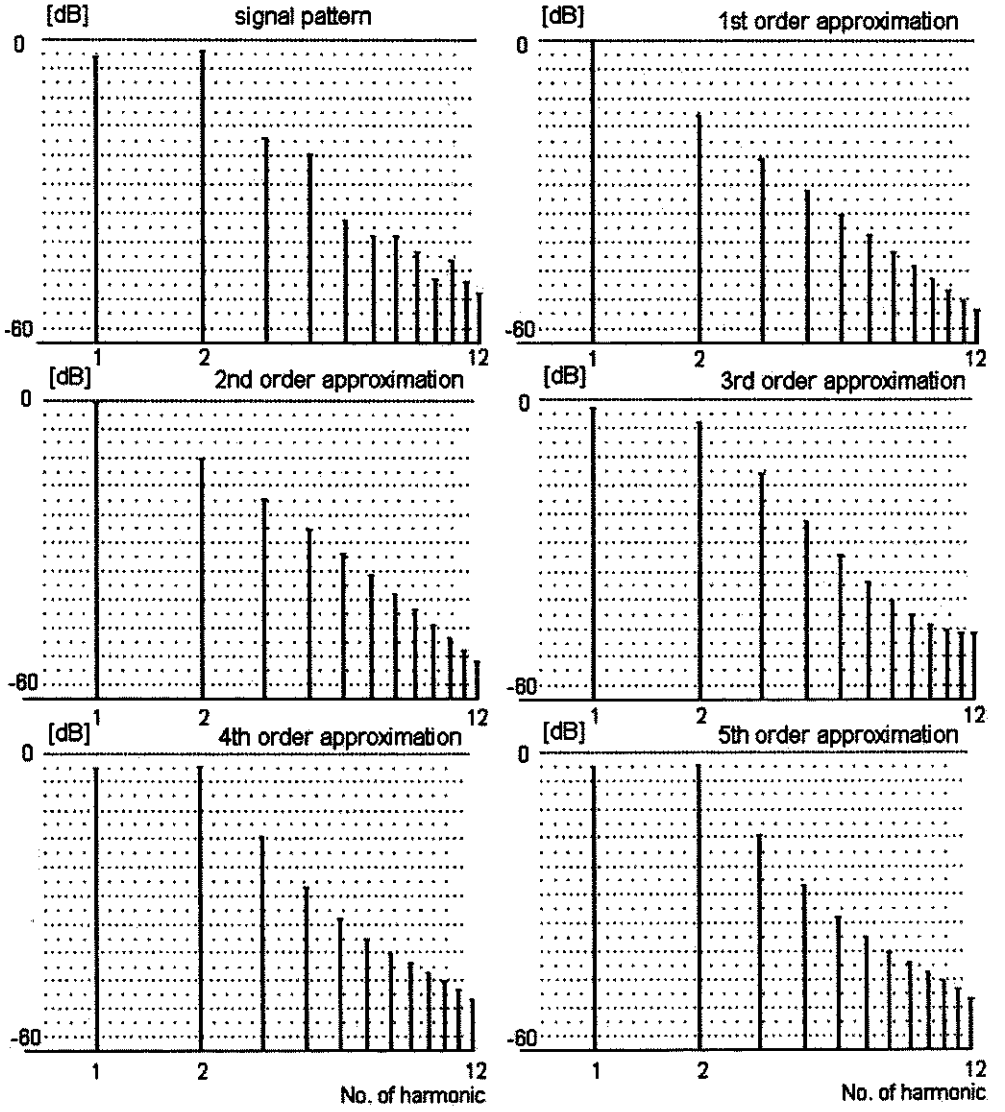


Fig. 6. Spectra of sounds used in subjective test No. 1: the pattern signal (the note C from the Principal 8′ organ stop at St. Mary's Basilica) at the left upper end; and the sounds generated on the basis of the polynomial approximation results, with rising polynomial order.

• Determining the number of experts who interpret a given pair in a different way depending on the part of the test (parameter z_2),

• Examining the significance of the differences between the objects in a pair at the assumed level of significance (parameter z_3),

Since the test was carried out for two groups of experts, the number of results was doubled and comparison between the groups becomes possible through determining the

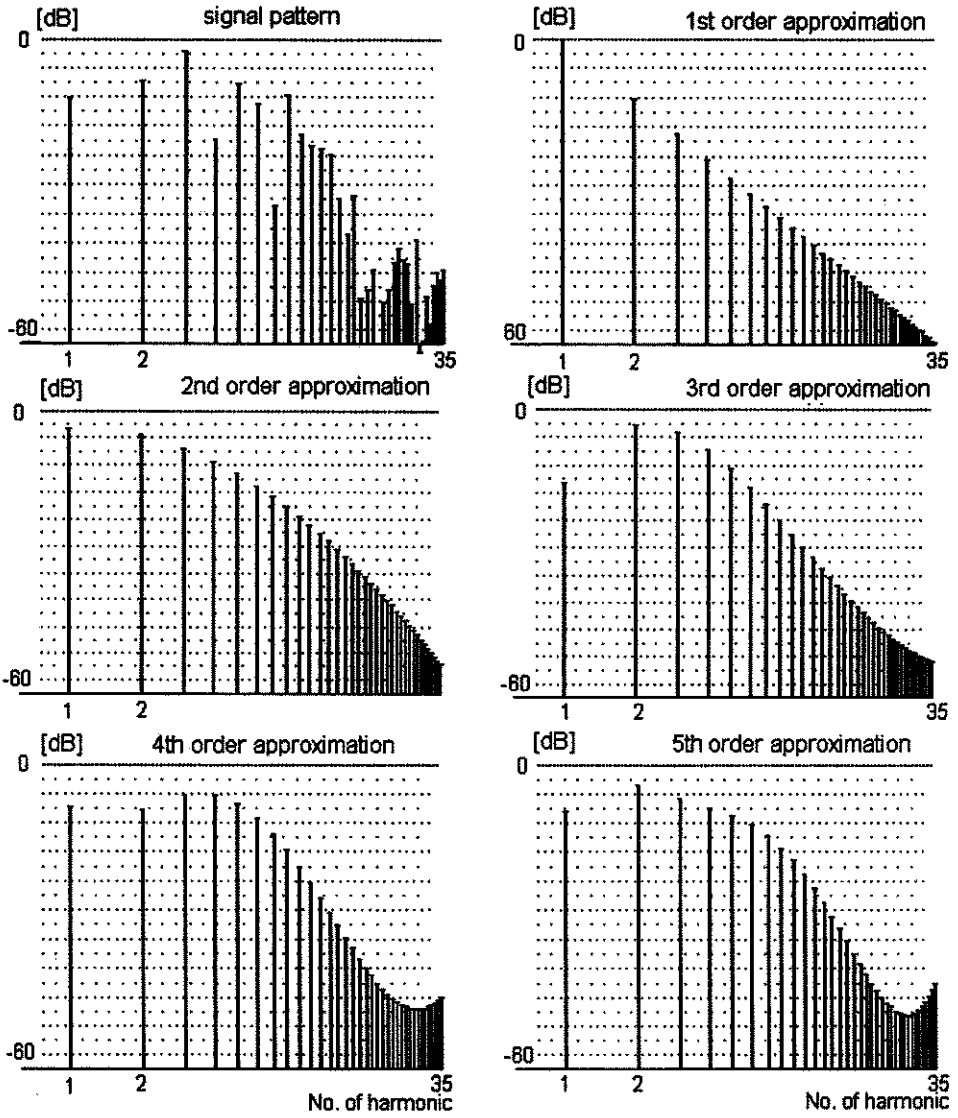


Fig. 7. Spectra of sounds used in subjective test No. 1: the pattern signal (the note B from the Trumpet 8′ organ stop at St. Mary's Basilica) at the left upper end; and the sounds generated on the basis of the polynomial approximation results, with rising polynomial order.

statistics χ^2 by comparing the results obtained from the two groups of experts with each other.

Examples of results from test 1 are given in Table 3. The upper part of the table contains the particular numbers of votes obtained by the objects. A column for parameter z_1 was distinguished – the number of errors made by each expert. It also shows the total sums of votes for the objects, and also for parts of the tests. The lower part of the

Table 3. Results of test No. 1 for group No. 1.

OBJECT	A	B	C	D	E	z_1
Expert 1	3	3	2	8	4	3
Expert 2	5	3	1	6	5	3
Expert 3	1	1	6	6	6	1
Expert 4	4	2	1	8	5	1
Expert 5	3	2	2	6	7	1
Expert 6	1	4	4	6	5	2
Expert 7	3	1	4	8	4	3
Expert 8	4	3	0	7	6	3
Expert 9	1	2	4	5	8	2
Expert 10	3	2	2	6	7	2
Expert 11	1	3	4	5	7	3
Expert 12	1	3	3	6	7	2
sum	30	29	33	77	71	240
part 1	14	15	15	39	37	120
part 2	16	14	18	38	34	120

PAIR	AB	AC	AD	AE	BC	BD	BE	CD	CE	DE
z_2	6	3	0	3	1	1	3	2	5	2
z_3	−	−	+	+	−	+	+	+	+	−

table shows the values of parameters z_2 and z_3 for all pairs of the compared objects. Parameter z_2 gives the total number of errors for a given pair, whereas for parameter z_3 the mark + was given in the case of meeting the criterion of differences significance; that is, exceeding the boundary value of 1.96 by the value of statistics. The results were also depicted in the form of examples of preference diagrams in Fig. 8. Successive objects from A to E are synthesized sounds related to the rising approximation order from 1 to 5.

To verify the assumption with regard to the effect of auditory memory on perception, a comparison of the results from both parts of the test had to be carried out. A similar comparison of the results from the two groups of experts provided the answer to the question of conformity of the observed tendencies. Both of these comparisons were performed using Pearson's test χ^2. This is one of the methods used to study the independence of features of population elements. A comparison of the results of the auditory monitoring tests between the groups of experts, which show conformity of interpretation of the tested sounds, is presented in Table 4. Additionally, a comparison between the results of both parts of the tests shows a slight effect of auditory memory on the ability to differentiate the timbre of the tested sounds (Table 5).

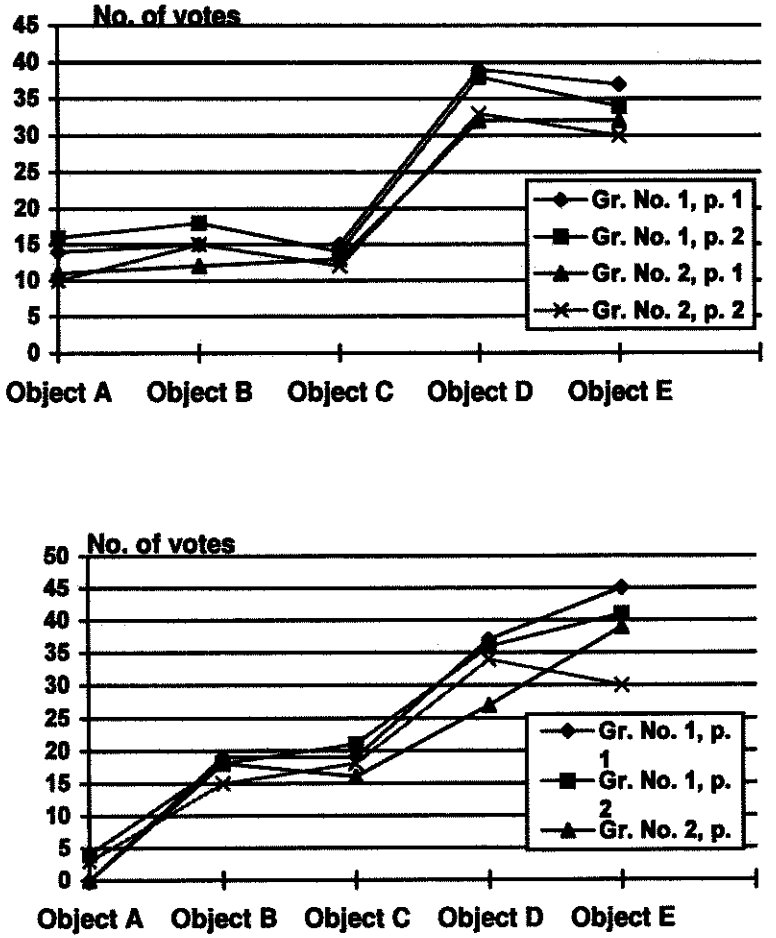


Fig. 8. Examples of preference diagrams for assessment tests: at the top, test No. 1; below, test No. 6.

Table 4. Comparison of the results of the auditory tests between groups of experts.

	Test 1	Test 2	Test 3	Test 4	Test 5	Test 6
Test 1		**12.976**	**14.052**	**26.703**	**28.063**	**15.974**
Test 2	**17.247**		8.742	4.328	**14.062**	7.217
Test 3	**19.417**	**11.660**		**21.058**	**11.613**	**11.197**
Test 4	**31.414**	2.681	**18.962**		**20.601**	**13.928**
Test 5	**39.943**	**18.233**	**13.463**	**17.798**		5.944
Test 6	**23.063**	9.913	**15.034**	**13.336**	**11.096**	

Table 5. Comparison between the results of both parts of the tests.

	Group 1	Group 2
Test 1	0.580	0.501
Test 2	3.114	4.466
Test 3	3.180	1.737
Test 4	8.588	5.022
Test 5	1.366	0.725
Test 6	4.327	5.368

7. Results and conclusions

The presented system of sound feature extraction based on smoothing the spectrum of the steady-state through polynomial approximation allows the comparison of sounds produced by various organ pipes. Its effectiveness was confirmed through statistical analysis of databases obtained as the result of feature extraction with this method, and also, perhaps more importantly, through the results of non-parametric subjective tests. The obtained preference curves are usually of a rising character, showing an increase in the similarity of the sound timbre together with an increase in the order of polynomial approximation. The exceptions to this tendency are local, pertain only to the final part of the diagram and are statistically insignificant. Also observed is the effect of saturation of preference curves, different for various types of pipes. This effect describes when the preference curve reaches a state, after an initial fast growth, in which a further change becomes statistically insignificant. There are various manifestations for various stops, and a statistically significant difference was found between reed and flue stops. These results are mutually confirmed by both groups of experts. The polynomials approximating spectral envelopes enable synthesis of sound that imitates the analyzed real sound. The fact that there is a saturation point of the preference curve may suggest that a significant similarity has been reached between the timbre of sounds synthesized on the basis of the polynomial spectrum representation and the timbre of the original sounds. Consequently, it is possible to reduce the approximation order to the number tested in the study, or to one that is not much greater.

The system of feature extraction developed in this study does not eliminate the possibility of extending the analytical apparatus for the purpose of sound classification. It is also possible to modify this system in order to better approximate the spectrum. Especially interesting is the concept of spectrum division into two parts, even and odd, and then separate approximations of their envelopes. There is a significant number of stopped pipes which as a rule should not generate any even harmonics. In this case, an attempt to approximate the entire spectrum will probably result in a bigger error than after its division into the even and odd parts. Such an approximation would require the application of two polynomials, but it is nonetheless probable that the final vector of parameters would include fewer elements while securing a better accuracy of

approximation. Moreover, a wavelet transform can provide a platform for representing some spectral properties of organ pipe sounds more accurately than a Fourier transform. These are topics of planned future research.

References

[1] J.W. BEAUCHAMP, *Time-variant spectra of violin tones*, J. Acoust. Soc. Am. , **56**, 995–1004 (1974).

[2] J.W. BEAUCHAMP, *UNIX Workstation software for analysis, Graphics, Modification and synthesis of musical sounds*, 94th AES Convention, Preprint 3479, Berlin 1993.

[3] K.W. BERGER, *Some factors in the recognition of timbre*, J. Acoust. Soc. Am., **36**, 1888–1891 (1963).

[4] G. VON BISMARCK, *Timbre of steady sounds: A factorial investigation of its verbal attributes*, Acustica, **30**, 146–159 (1974).

[5] A. DE BRUIJN, *Timbre-classification of complex tones*, Acustica, **40**, 108–114 (1978).

[6] G. BUDZYŃSKI, *Akustische Untersuchungen an Orgeln in Gdansk*, 6 Tagung Akustik in Berlin. Kurzfassungen der Vorträge und Poster, 79–92, Berlin, Gessel 1987.

[7] A. CZYŻEWSKI, A. KACZMAREK, B. KOSTEK, *A method of recognition of parametrized binary representations of audio signals*, 94th AES Convention, Preprint 3555, Berlin 1993, J. Audio Eng. Soc., **41**, 5 (1993).

[8] H. FLETCHER, E.D. BLACKHAM, D.A. CHRISTENSEN, *Quality of organ tones*, J. Acoust. Soc. Am., **35**, 314–325 (1963).

[9] M.D. FREEDMAN, *Analysis of musical instrument tones*, J. Acoust. Soc. Am., **41**, 793–806 (1967).

[10] M. GREY, *Multidimensional perceptual scaling of musical timbres*, J. Acoust. Soc. Am., **61**, 1270–1277 (1977).

[11] E.V. JANSSON, *Long-time-average-spectra applied to analysis of music. Part III: A simple method for surveyable analysis of complex sounds sources by means of a reverberation chamber*, Acustica, **34**, 275–280 (1976).

[12] E.V. JANSSON, J. SUNDBERG, *Long-time-average-spectra applied to analysis of music. Part I: Method and general application*, Acustica, **34**, 15–19 (1975).

[13] A. KACZMAREK, B. KOSTEK, *Brauchbarkeit der Polynomglättung des Klangspektrums zur automatischen Klassifizierung von Orgelpfeifen*, Deutsche Arbeitsgemeinschaft für Akustik, DAGA 93', Fortschritte der Akustik, Teil B, 908–911, Frankfurt am Main 1993.

[14] J.S. KEELER, *Piecewise-periodic analysis sounds and musical transients*, IEEE Trans. on Audio and Electroacoustics, **AU-20**, 5, 338–344 (1972).

[15] J.S. KEELER, *The attack transients of some organ pipes*, IEEE Trans. on Audio and Electroacoustics, **AU-20**, 5, 378–391 (1972).

[16] B. KOSTEK, *Articulation-related features in the pipe organ sound*, Archives of Acoustics, **22**, 2, 219–244 (1997).

[17] B. KOSTEK, A. KACZMAREK, *Rechneranalysator für Schallsignale*, Deutsche Arbeitsgemeinschaft für Akustik, DAGA 94', Fortschritte der Akustik, Teil C, 997–1000, Dresden 1994.

[18] B. KOSTEK, *Application des resaux de neurones pour l'analyse de l'articulation musicale*, Journal de Physique, Coll. C5, Suppl. JP III, I, 597–600 (1994).

[19] B. KOSTEK, A. CZYŻEWSKI, *Investigation of articulation features in organ pipe sound*, Archives of Acoustics, **18**, 2, 3, 417–434 (1993).

[20] B. KOSTEK, *Computer control of the pipe organ and the quality of the instrument*, J. Acoust. Soc. Am., **94**, 4, 2453–2454 (1993).

[21] B. KOSTEK, A. KACZMAREK, *Polynomglättung des Klangspektrums von Orgelpfeifen – subjektiver Test*, Deutsche Arbeitsgemeinschaft für Akustik, DAGA 93', Fortschritte der Akustik, Teil B, 912–915, Frankfurt am Main 1993.

[22] J. KRIMPHOFF, S. MCADAMS, S. WINSBERG, *Caractérisation du timbre des sons complexes. II. Analyses acoustiques et quantification psychophysique*, Journal de Physique IV, Coll. C5, Suppl. JP III, **4**, 625–628 (1994).

[23] W. LOTTERMOSER, *Vergleichende Untersuchungen an Orgeln*, Acustica, **3**, 129 (1953).

[24] W. LOTTERMOSER, J. MEYER, *Orgelakustik in Einzeldarstellungen*, Teil I. Verlag Das Musikinstrument, Frankfurt am Main 1960.

[25] S.N. MALLOCH, A.M. CAMPBELL, *An investigation of musical timbre*, Journal de Physique IV, Coll. C5, Suppl. J. de Physique III, **4**, 589–592 (1994).

[26] A. MELKA, *Methodological approaches to the investigation of musical timbre*, Journal de Physique IV, Coll. C5, Supp. JP III, **4**, 569–576 (1994).

[27] D.M.A. MERCER, *The voicing of organ flue pipes*, J. Acoust. Soc. Am., **23**, 45–54 (1951).

[28] A.W. NOLLE, C.P. BONER, *The initial transients of organ pipes*, J. Acoust. Soc. Am., **13**, 149–154 (1941).

[29] C. PADGHAM, *The scaling of timbre of the pipe organ*, Acustica, **60**, 3, 189–204 (1986).

[30] R. PLOMP, J. DE LAAT, *Comparisons of organ in a spectrum space*, Acustica, **55**, 3, 193–194 (1984).

[31] H.F. POLLARD, E.V. JANSSON, *A Tristimulus method for the specification of musical timbre*, Acustica, **51**, 162–171 (1982).

[32] R.L. PRATT, P.E. DOAK, *A subjective rating scale for timbre*, J. Acoust. Vib., **45**, 317–328 (1976).

[33] J. SUNDBERG, E.V. JANSSON, *Long-time-average-spectra applied to analysis of music. Part II. An analysis of organ stops*, Acustica, **34**, 269–274 (1976).

[34] V. RIOUX, M. YOKOTA, D. VÄSTFJÄLL, M. KLEINER, *Preliminary study of an organ builder's perception of a flue pipe sound*, Proc. International Symposium on Musical Acoustics, Leavenworth, USA, 273–278 (1998).

[35] S. YOSHIKAWA, *Jet and vortex behaviors during the starting transient in organ pipes*, Acta Acustica, **85**, Suppl. 1 (1996).

ARCHIVES OF ACOUSTICS
24, 1, 25–38 (1999)

A PRELIMINARY STUDY ON GREEK ESOPHAGEAL SPEECH AND A METHOD FOR QUALITY AND INTELLIGIBILITY ENHANCEMENT

C. PASTIADIS and G. PAPANIKOLAOU

Laboratory of Electroacoustics,
Aristotle University of Thessaloniki
(GREECE)

The present work is a preliminary study on Greek esophageal speech and is mainly concerned with the investigation of major features such as pitch, formant frequencies, and speech power envelopes. The implementation in esophageal speech of various well-known techniques for normal voice analysis is overviewed. An improved method for resynthesizing voiced sounds (such as vowels or nasal consonants) by convolution of an ARMA estimate of the speaker's vocal tract impulse response and a periodic glottal waveform is proposed as a tool for voice quality enhancement. Fundamental frequency values were confirmed to be close to previous works' findings. F1 and F2 formant alterations due to laryngectomy were not detected compared to normal speech values. However, speech power envelopes tended to be flatter as the speaker's training stage was higher. The proposed method for speech enhancement proved able enough to preserve speaker characteristics and provide cues for higher quality reproduction of vowels as well as nasals.

1. Introduction

Esophageal speech is produced by laryngectomized people who utter by expelling air constricted in their esophagus. The expelled air forces the cricopharyngeal cartilage to oscillate in a manner that imitates the vocal folds' operation. Although proper training may help esophageal speakers utter intelligibly enough, a severe degradation of voice quality after laryngectomy usually occurs. The voice quality of esophageal speech may be described as harsh, rough and low. The amount of aperiodicity is high and the voice is often very noisy.

Esophageal speech features have been investigated in the past. A part of previous works focused on pitch and intensity characteristics and their perceptual aspects [11, 12, 18, 19, 24].

Efforts for esophageal speech quality and intelligibility enhancement have also been reported [1, 2, 9, 16, 17, 23].

The present work is a preliminary study on Greek esophageal speech and is mainly concerned with the investigation of major features such as pitch (or fundamental frequency F0), formant frequencies, and speech power envelopes. The implementation of various normal voice analysis techniques in esophageal speech is overviewed [14]. An

improved method for resynthesizing voiced sounds (such as vowels or nasal consonants) is proposed as a tool for voice quality enhancement [14].

2. Subjects and recordings

Nine male alaryngeal speakers were used as subjects for uttering various CV sequences and full sentences. CV sequences comprised of all Greek vowels and stop consonants.

Speakers were selected among groups of various grades of speech production training and had no treatment prior to voice recording.

Utterances were spoken at normal rate and level and were recorded on a DAT recorder through an electret condenser microphone placed at a distance of approximately 20 cm from the speaker's mouth. A sampling rate of 11 kHz was adopted for further computer processing.

3. Investigation of speech features

Fundamental or pitch frequency, speech power envelope during phonations and formant frequencies are considered as some of the most important features for speech comprehension, training and clinical evaluation of voice. These features were investigated and the methods employed together with analysis results are presented.

3.1. F0 *investigation*

F0 or fundamental frequency was investigated for all nine speakers during vowel phonations in discrete CV contexts and isolated vowel segments through whole sentences.

Estimation of F0 was performed using well known methods for F0 extraction on normal speech, such as the Autocorrelation Method (biased and unbiased), the Center-Clipped Autocorrelation and 3-level Center-Clipped Autocorrelation methods, the Average Magnitude Difference Function method (AMDF), the Cepstrum and a so-called "Hubert– Envelope" Method.

3.1.1. The autocorrelation method. The autocorrelation method [6] identifies F0 or period of voiced speech by finding the lag at which the autocorrelation function of the speech signal is maximized, that is:

$$T_0 = \frac{1}{f_0} = \max_m \{\Phi(m)\} \cdot T_s \,, \tag{3.1}$$

where

$$\Phi(m) = \frac{1}{N} \sum_{n=0}^{N-1-|m|} s(n)s(n+|m|), \qquad m = 0, \ldots, N-1 \tag{3.2}$$

with $s(n)$ = speech signal, N = number of samples, T_s = sampling period.

We may also use the unbiased autocorrelation estimator:

$$\Phi(m) = \frac{1}{N-|m|} \sum_{n=0}^{N-1-|m|} s(n)s(n+|m|), \qquad |m| = 0, ..., N-1. \tag{3.3}$$

Estimation of the first maximum of the autocorrelation function using the biased and unbiased estimators is shown in Figs. 1b, 1c, respectively.

3.1.2. The center-clipped and 3-level center-clipped autocorrelation methods. Further improvement in the estimation of maxima of the autocorrelation function may be achieved using either center-clipping [21] or 3-level center-clipping on the speech signal:

$$s'(n) = \begin{cases} s(n) - C^+ & s(n) > C^+, \\ 0 & C^- \le s(n) \le C^+, \\ s(n) - C^- & s(n) < C^-, \end{cases} \tag{3.4}$$

$s'(n)$ = center-clipped speech signal, or

$$s'(n) = \begin{cases} 1 & s(n) > C^+, \\ 0 & C^- \le s(n) \le C^+, \\ -1 & s(n) < C^-, \end{cases} \tag{3.5}$$

$s'(n)$ = 3-level center-clipped speech signal and C^+, C^- are threshold values.

Estimated maxima of the autocorrelation tunction after center-clipping of the speech signal are shown in Fig. 1d.

3.1.3. The AMDF *(Average Magnitude Difference Function) method.* The AMDF method [7] tries to locate a strong minimum of the AMDF

$$\text{AMDF}(m) = \sum_n |s(n) - s(n+m)|, \tag{3.6}$$

$s(n)$ = speech signal.

For a strictly periodic speech signal the AMDF would take on a value of zero at $m = T_0/T_s$, $T_0 = 1/f_0$.

For quasi-periodic signals a strong minimum usually occurs at the period lag.

Results of this method on an esophageal speech sample are shown in Fig. 1e.

3.1.4. The cepstrum. This well known technique relies on the fact that time of maximum at the high-quefrency region of the cepstrum represents the period of the speech signal [6, 13]. Results of this method employed on esophageal speech are shown in Fig. 1f.

3.1.5. The "Hilbert-Envelope" method. This method [3] performs maximum-likelihood epoch determination as the basis for the estimation of glottal closure (epoch) instants (GCI) in normal voices and implements a Hilbert transformation for the improvement of its performance and reliability. It is also posed that the method may cover

most speech signals (even under noisy conditions). Since it is capable of estimating closure instants, it can sense period-to-period variations or nonstationary period variations within longer frames.

The method is employed on esophageal speech under the assumption that even in this kind of severely damaged vocal function there must exist a moment that a main pulse excites the whole nasopharyngeal system. Thus, the method tries to locate the maximally possible instants of main excitation.

The method initiates with the formation of a so-called Maximum-Likelihood Epoch Determination signal ("MLED-signal") which is proven to be a cross-correlation between the speech signal and the impulse response of the nasopharyngeal system filter due to an epoch, that is

$$\hat{f}(k) = \sum_{n=0}^{N-1} s(n+k)\hat{s}(n), \tag{3.7}$$

where $s(n)$ = speech signal, and

$$\hat{s}(n) = \begin{cases} G & n = 0, \\ -\sum_{i=1}^{p} a_i \hat{s}(n-i) & 0 < n \leq \infty, \\ 0 & \text{otherwise} \end{cases} \tag{3.8}$$

which is virtually the nasopharyngeal filter's impulse response obtained by AR modeling of the speech signal and a_i are the model's coefficients. The use of a high order AR modeling in esophageal speech is motivated by the fact that the glottal function is generally unknown and the nasopharyngeal system's function may include zeroes that should normally be represented by ARMA modeling. A selection of $p \geq 40$ was made in order to compromise between accuracy and computation time, although a value of at least $N/5$ where N = record length is suggested [10].

GCI's are identified as the time indices of maxima of the MLED signal. To reduce ambiguity in selection of maxima the MLED is multiplied by a Hubert–Envelope of itself,

$$\hat{f}(k) \cdot \hat{g}(k),$$

where

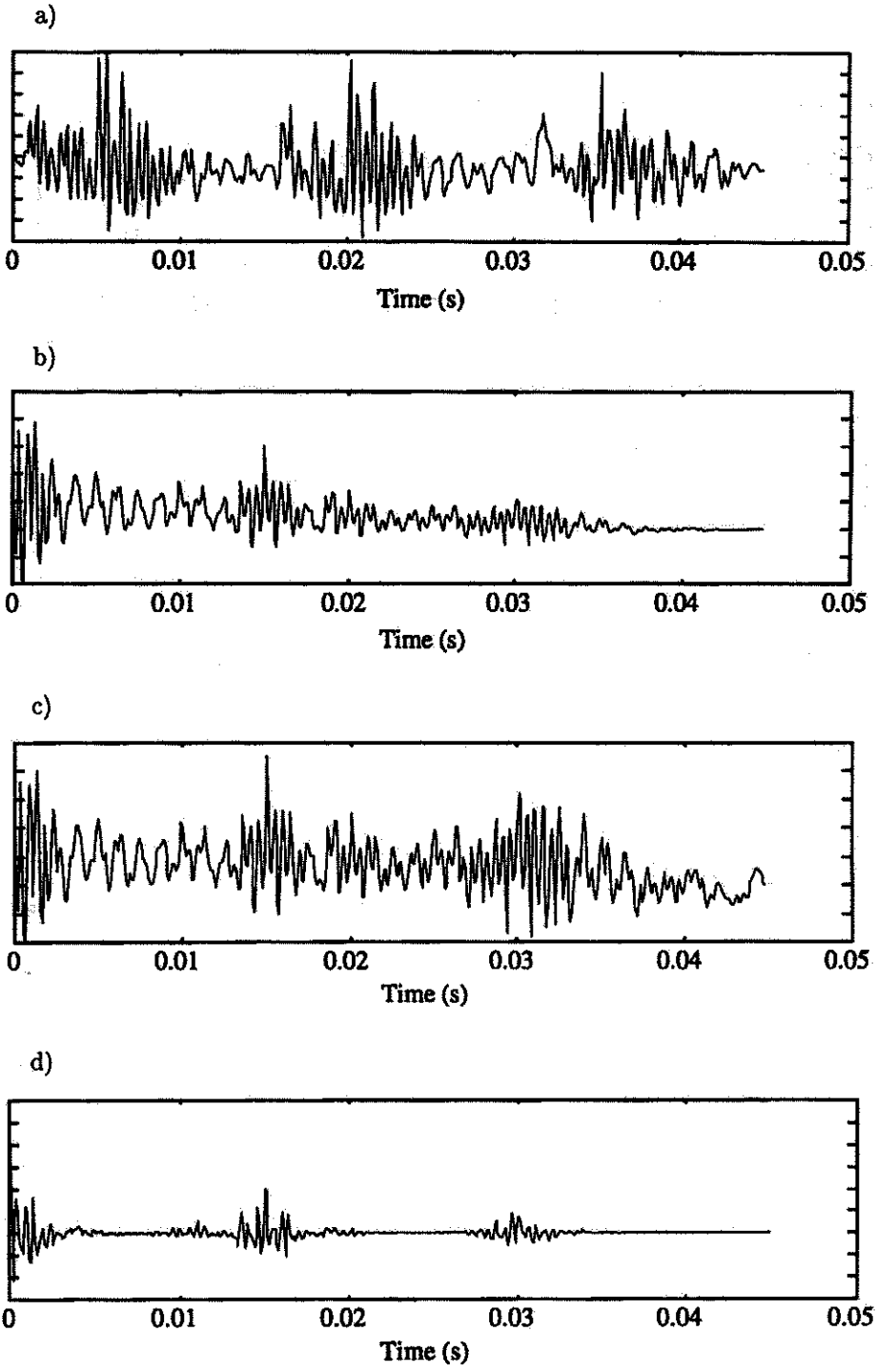
$$\hat{g}(k) = \sqrt{\left[\hat{f}^2(k) + \hat{f}_H^2(k)\right]}, \qquad \text{and} \qquad \hat{f}_H(k) = HT(\hat{f}(k)). \tag{3.9}$$

The Hilbert–Envelope proves able to emphasize the contrast between the main epoch pulse and other possible sub-pulses that indicate sub-optimal excitation instants.

The results of this algorithm are shown in Fig. 1g, where values of the MLED signal lower than a predetermined threshold were set to zero.

As it is observed, optimal instants of excitation are possible to locate in esophageal speech too.

Since this method of F0 estimation relies on the interpretation of fundamental frequency as the inverse of between-excitation period, it can provide information on


a)
0
0.01
0.02
0.03
0.04
0.05
Time (s)
b)
0
0.01
0.02
0.03
0.04
0.05
Time (s)
c)
0
0.01
0.02
0.03
0.04
0.05
Time (s)
d)
0
0.01
0.02
0.03
0.04
0.05
Time (s)

[Fig. 1]

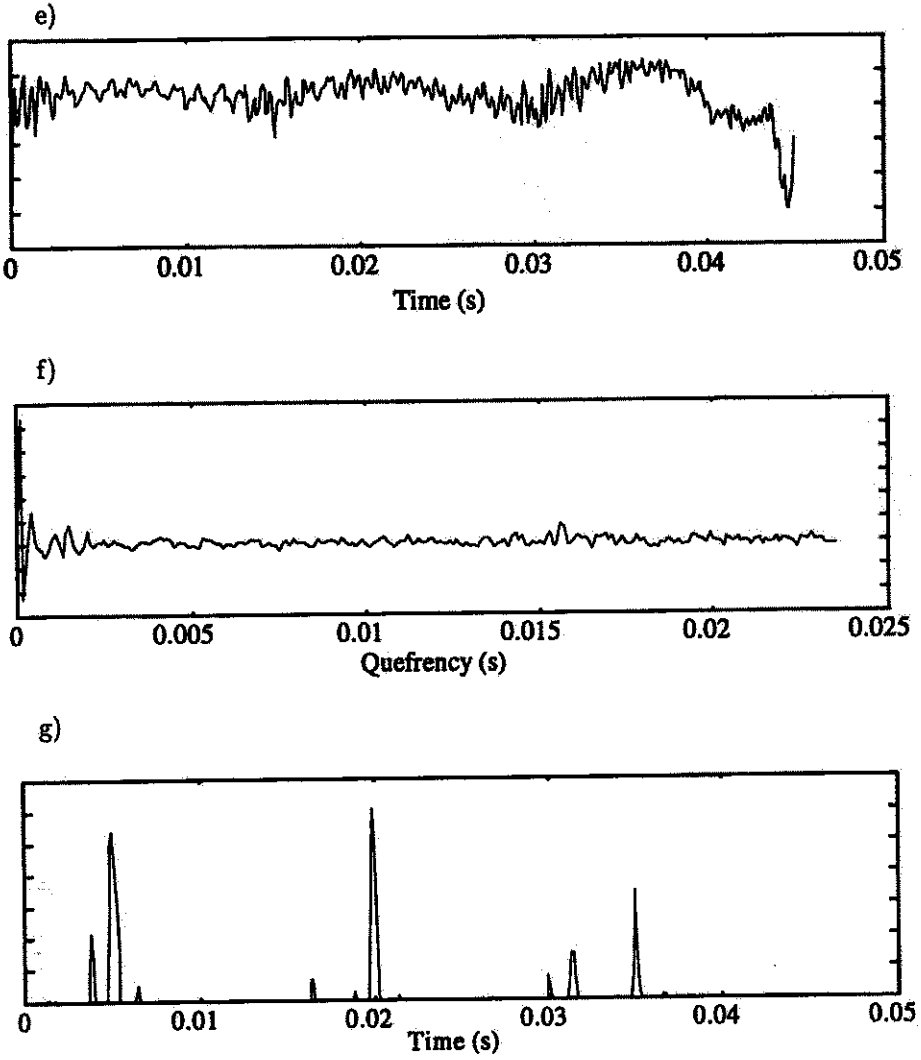


Fig. 1. Esophageal vowel /e/ (a), and fundamental period estimates using: the biased autocorrelation (b), the unbiased autocorrelation (c), the center-clipped autocorrelation (d), the AMDF (e), the Cepstrum (f) and the "Hilbert Envelope" (g) methods.

period-by-period F0, and thus it can be used for cycle-to-cycle estimation of glottal activity fluctuation (such as jitter or shimmer) [22]. The existence of various degrees of ambiguity in selection of GCI's is currently being investigated as a tool for the clinical evaluation of voice quality and/or training procedure's progress.

3.1.6. Results and discussion on F0 *investigation.* Though all the employed extraction methods are capable of estimating rough F0 values, ambiguity is lower when estimates are taken using the Center-Clipped Autocorrelation methods and/or the "Hilbert–

Envelope" method, which seem to provide more consistent F0 estimates to manually extracted ones. However, since the Autocorrelation-based methods are short-term average methods, they are not applicable to cycle-to-cycle variations investigation [22], whereas the "Hilbert–Envelope" method seems more appropriate.

Mean intra-speaker F0 (estimated using the Center-Clipped Autocorrelation method) values and standard deviation together with mean F0 values and standard deviation between all speakers are presented in Table 1.

Table 1. Intra-speaker F0 mean and standard deviation values together with inter-speaker mean and standard deviation values.

Speaker	Mean F0 (Hz)	S.D. (Hz)
1	57.5	9.0
2	64.9	9.8
3	86.4	19.9
4	88.4	19.9
5	73.8	15.1
6	82.2	19.0
7	89.9	18.4
8	43.3	9.0
9	68.2	12.1
inter-speaker Mean value	**72.7**	**14.7**
inter-speaker S.D.	**14.9**	**4.4**

As observed, mean F0 for all speakers is found at 72.7 Hz, which verifies previous works' findings that esophageal speech is about 1 octave lower than normal speech [11, 18, 19, 24].

3.2. Speech power envelope investigation

The slope of speech power envelope during phonations exhibits major perceptual and clinical interest [19], since it may provide information on voice dynamics and the speaker's training progress.

Speech power envelopes of /pa/ utterances from all speakers were obtained using a pitch-period wide integration window, and mean slopes of phonations were computed. Results are presented in Table 2.

Mean value between all speakers was found to be −86 dB/sec, with standard deviation 13.7 dB/sec.

Additional information about the stage of training of each one of the speakers showed a tendency of decrease in slopes (numerically higher slope values) with training past.

Table 2. Phonation slopes for the utterance /pa/.

Speaker	Mean phonation slope (dB/sec)
1	−81.3
2	−78.9
3	−78.8
4	−74.3
5	−92.9
6	−116.1
7	−70
8	−100
9	−81.3
Mean Value	−86
Standard Deviation	13.7

3.3. Vowels' F1, F2 *investigation*

First two formant frequencies (F1, F2) were investigated using LPC, for all 5 Greek vowels α, ε, ι, o, ov. As it is well known, Greek vowels differ from other languages' vowels in that they are not rounded and thus are displaced in the F1/F2 space [8].

Table 3 gives mean F1, F2 values for all speakers together with mean values of normal Greek speech.

Table 3. F1, F2 mean values for all Greek vowels for esophageal and normal speech [7].

	Esophageal speech		Normal speech	
Greek Vowel	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)
α	732	1390	~ 700	~ 1300
ε	521	1750	~ 475	~ 1700
ι	385	1823	~ 300	~ 2000
o	510	992	~ 450	~ 850
ov	420	1095	~ 350	~ 900

A general coincidence between normal and esophageal speech formant frequencies values is observed, which proposes that formant extraction methods for speech recognition may be used in esophageal speech too.

4. A method for esophageal speech quality and intelligibility enhancement

As already stated, esophageal speech is severely degraded. Perceptual judgments of esophageal voice characterize it as harsh, rough and low. Generally, esophageal speakers are able to produce intonational contrasts but listeners do not readily perceive the variation.

Previous works on esophageal voice rehabilitation had followed both surgical and speech signal processing procedures. Among the most well known surgical methods is one that uses a valve prosthesis that permits air to flow from trachea to esophagus [191. In the signal processing domain, recent works report methods of spectral substitution of esophageal speech [1, 2], resynthesizing speech using LPC spectral estimation [16, 17, 23], and elimination of undesirable phenomena (such as injection noise) during voice production [9].

The method proposed in this work uses ARMA estimation on esophageal speech. Since in esophageal speech the excitation signal is generally unknown, an all-pole model alone may not accurately represent the speech production procedure. Moreover, a pure all-pole model of relatively low order may not be sufficient for the analysis of voiced consonantal sounds that include significant zeroes in their spectral envelope such as nasals (/m/, /n/). ARMA modeling allows for the extraction of major spectral envelope profiles by extracting both formants (poles) and anti-formants (zeroes) frequencies and bandwidths. Estimation of parameters of an ARMA model that describes esophageal speech production is performed using a Least Squares Modified Yule–Walker Equations (LSMYWE) approach [10]. Next, voiced speech resynthesizing is performed using convolution of the oronasal tract filter's impulse response (obtained with the use of ARMA modeling and selection of appropriate pole/zeroes pairs) with a waveform that represents normal voices' vocal fold vibratory function [20].

The proposed method's functional diagram is shown in Fig. 2.

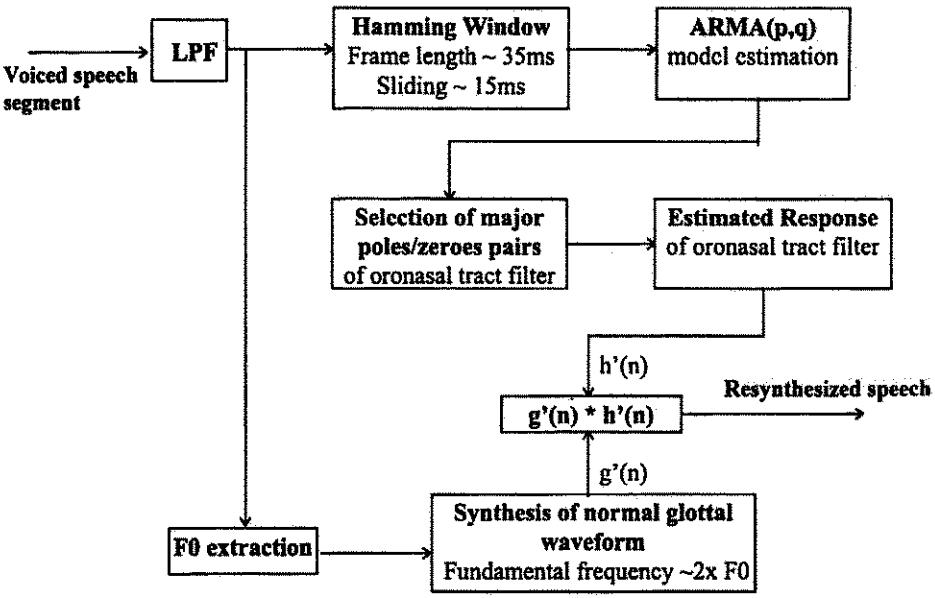


Fig. 2. Block diagram of the proposed method for voice quality enhancement.

As shown, a selection of voiced speech segments precedes the implementation of the method.

The ARMA modeled spectral envelope is of the form

$$S(z) = \frac{\sum_{i=0}^{q} b_i z^{-i}}{1 + \sum_{i=1}^{p} a_i z^{-i}}, \qquad b_0 = 1, \quad (p, q) = \text{ARMA model order.} \tag{4.1}$$

A $p \leq 16$ and $q \leq 10$ pair of values may be selected.

The selection of major poles and zeroes is made under the following assumptions:

a. Oronasal filter's poles are complex conjugate pairs with relatively high frequency-to-bandwidth ratios.

b. Significant zeroes also appear as complex conjugate pairs with relatively high frequency-to-bandwidth ratios, whereas real valued zeroes may reflect radiation and/or possible glottal waveform's spectral characteristics.

More specifically, the estimated filter's impulse response z-transform is:

$$V(z) = A \cdot \frac{\prod_{i=1}^{q'} (1 - z_i z^{-1})(1 - z_i^* z^{-1})}{\prod_{i=1}^{p'} (1 - p_i z^{-1})(1 - p_i^* z^{-1})}, \tag{4.2}$$

where z_i, p_i are selected zeroes and poles from the estimated ARMA model of esophageal speech according to the previous assumptions and $|p_i| < 1$.

Figure 3 shows the results of the LSMYWE analysis on steady-state portions of both the consonantal and vowelic regions of two original esophageal CV speech utterances /na/ and /me/. A 11 kHz sampling rate and a model order of $p = 12$ and $q = 5$ were selected. The analysis of the C part of the utterances seems to confirm previous works' findings on acoustic analysis of nasals. As it can be seen from the power spectra and the pole-zeros chart, a low frequency pole (nasal murmur) appears in the region of about 300 Hz for both /m/ and /n/. Moreover, the presence of side branch resonators (as the oral cavity in the case of nasals) introduces zeros in the spectrum of the uttered phoneme [15, 5]. Although these zeros are not very prominent, they produce a smoother energy distribution over different frequency ranges between /m/ and /n/ [15]. The zero introduced in the case of /n/ lies in the mid-frequency region (over 1 kHz), whereas a broader zero appears in the case of /m/ causing a more even energy distribution over the low-frequency region. In the case of the V utterances (/a/ and /e/), the formantic structure for each phoneme appears clearly, with formant values close to typical ones. The low-frequency nasal murmur is removed. Zeros in the estimated vocal tract response seem to exist close to mid-high formants, whereas the rest of them lie either on the real axis or they are of larger bandwidth; thus, they do not interfere significantly with the all-pole configuration of the vocal tract during vowel production.

A synthetic waveform that represents normal voice's glottal vibration is used for convolution with the previously estimated response. Triangular-like waveforms are preferred to pulse-like ones, since they seem to produce more intelligible synthetic speech

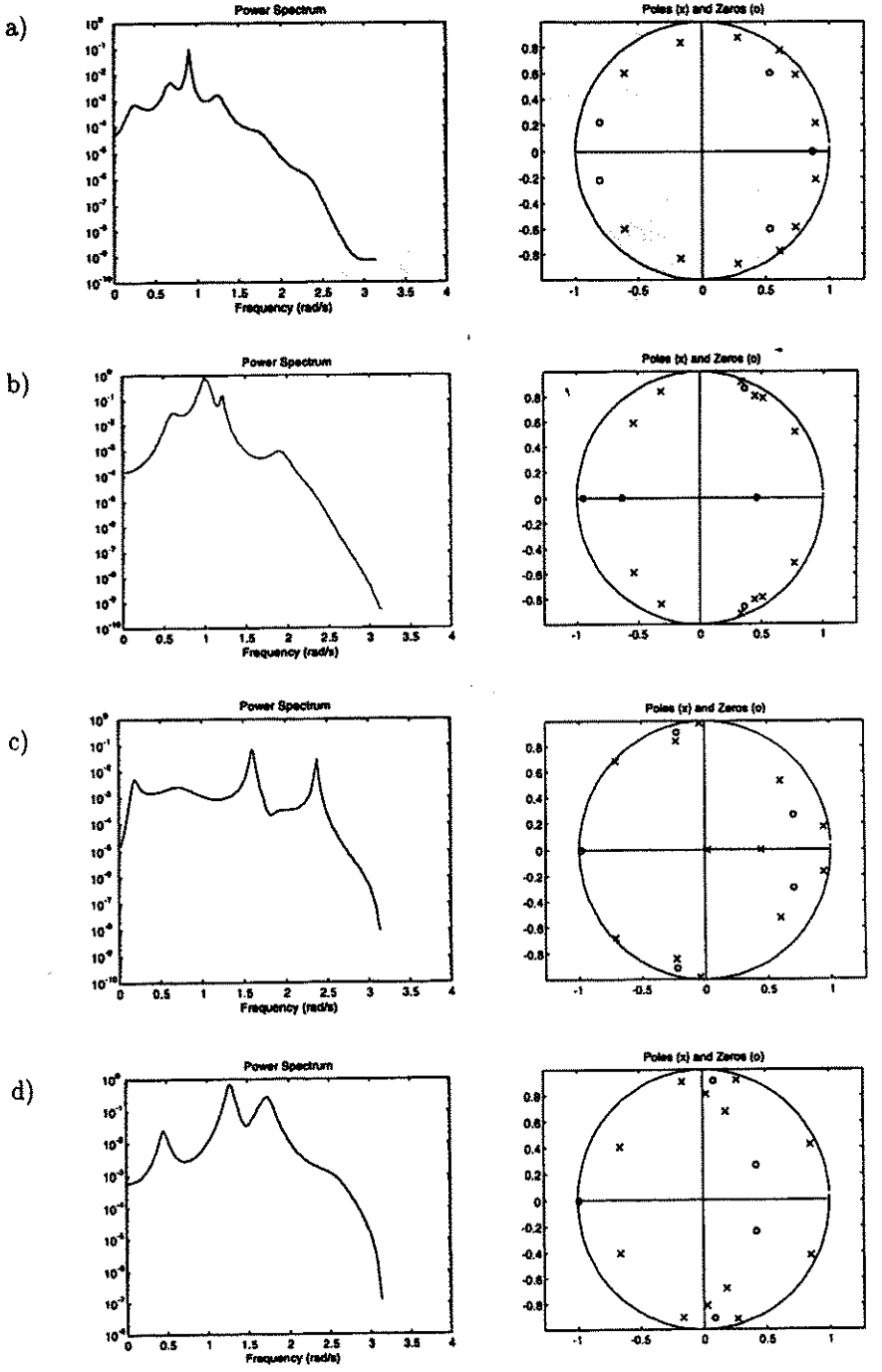


Fig. 3. Vocal tract transfer functions and pole-zero plots for: /n/ in /na/ (a), /a/ in /na/ (b), /m/ in /me/ (c), /e/ in /me/ (d), as obtained by a (12,5)-order ARMA model using LSMYWE.

[4, 20]. Fundamental frequency is set at about double (1 octave higher) than the mean F0 estimated from the esophageal speech signal. The inclusion of jitter and/or shimmer perturbations on the synthetic glottal signal may significantly improve the quality of produced speech.

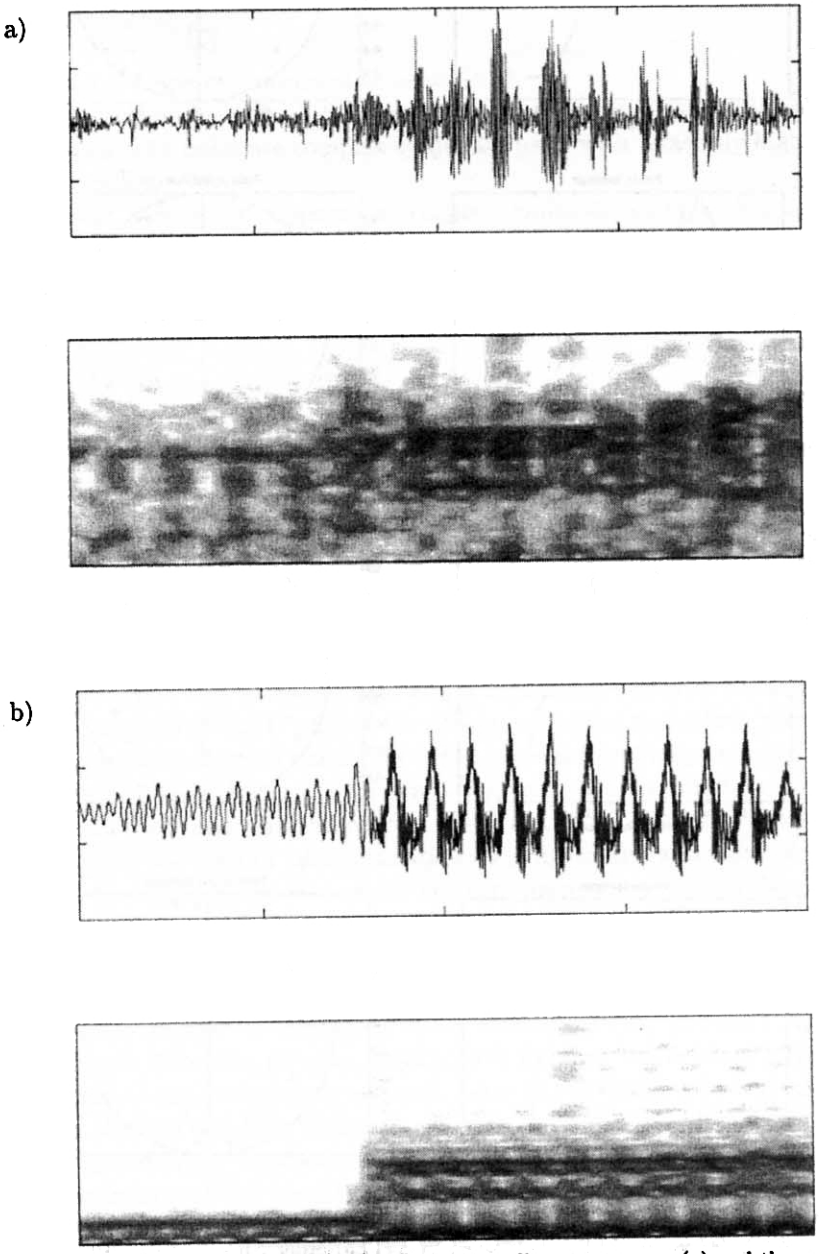


Fig. 4. Esophageal speech utterance /na/ and corresponding spectogram (a) and the resynthesized signal and its corresponding spectogram (b).

The synthesized speech signal is computed in overlapping frames using triangular weighting of the current and next frame's analysis window. Figure 4 displays an original esophageal /na/ utterance and the resynthesized one together with their corresponding spectograms. As it can be observed, the formantic structure is preserved, and a higher and regular F0 pattern appears after the resynthesis procedure.

Early results of subjective tests on the synthetic speech produced with the proposed method show that great improvement in voice quality and spoken utterances' intelligibility is achieved, including voiced consonantal sounds. Also, the resynthesized speech preserves the cues needed for speaker identification. The LSMYWE method shows sufficient ability in tracking both poles and zeroes of the vocal tract acoustic filter. A technique which uses more generalized speech production models such as ARMAX and Box–Jenkins models together with smoothing formant and antiformant trajectories prior to convolution with the synthetic glottal signal is currently investigated.

5. Conclusions and further work

This work focused on the investigation and methods for extraction of major features for Greek esophageal speech.

Fundamental frequencies were found at about 70 Hz confirming the results of previous works arguing that esophageal speech sounds about 1 octave lower than normal speech. The Center-clipped Autocorrelation and the "Hilbert–Envelope" methods were found to be more efficient among various well-known methods of F0 extraction.

Slopes of power envelope vs. time during phonations served as a measure of voice dynamics and training progress. Mean values were estimated at about $-90\,\mathrm{dB/s}$ exhibiting a progressive increase from the less to the more trained speakers.

Formant profiles were studied for all Greek vowels and results showed a general coincidence with values of normal speech and verified the general distinctive features of Greek vowels to other languages' ones. The facts enforce the use of formant extraction method for esophageal speech recognition.

A new method for esophageal speech enhancement which is based on resynthesizing voiced segments using ARMA modeling of vocal tract and a higher pitch glottal waveform signal was also proposed and exhibited encouraging results.

Concurrent work includes study of major esophageal voice features such as jitter, shimmer, S/N ratios, intonational characteristics, e.t.c. under various conditions of the patient's surgical treatment procedure, period of training, e.t.c.

The proposed method for voice quality enhancement is revised under the use of extended ARMAX and Box–Jenkins models and variations of formant trajectories' extraction methods and glottal waveform signals. Accordingly, subjective quality tests are going to take place for the assessment of resynthesized speech. Further, the method is intended to be implemented on a DSP for real-time use.

References

[1] N. Bi, Y. Qi, *Alaryngeal speech enhancement based on spectral substitution*, ASA 127th Meeting M.I.T. 1994 June 6-10.

[2] N. Bi, Y. Qi, *Application of speech conversion to alaryngeal speech enhancement*, IEEE Trans. Speech & Audio Processing, **5**, 2, 97–105 (1997).

[3] Y.M. Cheng, D. O'Shaughnessy, *Automatic and reliable estimation of glottal closure instant and period*, IEEE Trans. AS SP, **37**, 12, 1805–1815, December 1989.

[4] P. Cook, *Identification of Control Parameters in an Articulatory vocal tract model with applications to the synthesis of singing*, Ph.D. Thesis, Stanford Univ., 1991.

[5] J. Flanagan, *Speech analysis, synthesis and perception*, Springer Verlag, 1972.

[6] S. Furui, *Digital speech processing, synthesis and recognition*, Marcel Dekker, Inc., 1989.

[7] S. Furui, M. Sondhi Mohan, *Advances in speech signal processing*, Marcel Dekker, Inc., 1992.

[8] A. Iivonen, *Articulatory vowel gesture presented in a psychoacoustical F1/F2-space*, Studies in Logopedics and Phonetics, Univ. Of Helsinki, **3**, 19–44 (1992).

[9] H. Jarkin, M. Galler, N. Niedzielski, *Enhancement of esophageal speech by injection noise rejection*, Proc. ICASSP'97, Munich, 1997.

[10] S. Kay, *Modern spectral estimation*, Prentice-Hall Signal Processing Series, 1988.

[11] A. Leinonen, *Intonational patterns and voice quality in esophageal speech*, Studies in Logopedics and Phonetics, Univ. of Helsinki, **3**, 151–159 (1992).

[12] K. Mourikis, *Phonetic rehabilitation of alaryngeal people*, 4th European Interuniversity Symposium of "Head and Neck Cancer: Improvements of Logoregional Control", Thessaloniki 1994.

[13] A. Oppenheim, R. Schafer, *Discrete time signal processing*, Prentice-Hall International 1989.

[14] C. Pastiadis, G. Papanikolaou, *A preliminary study on Greek esophageal speech and a method for voice quality enhancement*, AES 102nd Convention Preprint, Munich, March 1997.

[15] Y. Qi, R. Fox, *Analysis of nasal consonants using perceptual linear prediction*, Journal of the Acoustical Society of America, **91**, 3, 1718–1726 (1992).

[16] Y. Qi, *Replacing tracheoesophageal voicing sources using LPC synthesis*, Journal of the Acoustical Society of America, **88**, 3, 1228–1235 (1990).

[17] Y. Qi, B. Weinberg, N. Bi, *Enhancement of female esophageal and tracheoesophageal speech*, Journal of the Acoustical Society of America, **98**, 5, 2461–2465, 1995.

[18] Y. Qi, B. Weinberg, *Characteristics of voicing source waveforms produced by esophageal and tracheoesophageal speakers*, Journal of Speech & Hearing Research, **38**, 536–548 (1973).

[19] J. Robbins, H. Fisher, F. Blom, M. Singer, *A comparative study of normal, esophageal and tracheoesophageal speech production*, Journal of Speech and Hearing Disorders, **49**, 202–210 (1984).

[20] P. Rubin, L. Goldstein, *Articulatory synthesis (ASY)*, Haskins Laboratories.

[21] M. Mohan Sondhi, *New methods of pitch extraction*, IEEE Trans. Audio and Electroacoustics, **16**, 16, 262–266 (1968).

[22] I. Titze, H. Liang, *Comparison of F0 extraction methods for high-precision voice perturbation measurements*, Journal of Speech and Hearing Research, **36**, 1120–1133 (1993).

[23] R. Tull, J. Rutledge, J. Mahler, *Female alaryngeal speech enhancement for improved speaker identification using linear predictive synthesis*, ASA 129th Meeting Washington D.C. 1995 May 30–June 6.

[24] B. Weinberg, Y. Horii, B. Smith, *Long-time spectral and intensity characteristics of esophageal speech*, Journal of the Acoustical Society of America, **67**, 5, 1781–1784 (1980).

ARCHIVES OF ACOUSTICS
24, 1, 39–47 (1999)

AN INFLUENCE OF A MODULATING SIGNAL STARTING PHASE ON THE MODULATION DETECTION

A. P. SĘK and E. B. SKRODZKA

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

This paper is concerned with an influence of a modulator starting phase on psychometric functions of amplitude modulation (AM) and frequency modulation (FM) detection, for low modulation rates, i.e. $f_{mod} = 2$ and 5 Hz. For a sinusoidal carrier at a frequency of 0.25, 1 and 6 kHz it is shown that there is no effect of a sinusoidal modulator starting phase. Data shown in this paper are consistent with predictions of models based on excitation pattern changes and with a concept according to which AM/FM detection at low modulation rates depends mainly on maximum and minimum values of signal amplitude/frequency. Time intervals between beginning of a signal and the moment when the signal reaches maximum/minimum value of amplitude/frequency and a pattern of physical parameter changes as a function of time seem to be less crucial when the changes are continuos in time.

1. Introduction

Measurements of just noticeable amplitude or frequency changes of an acoustic signal are a methods of determination of sensorial limits of the human auditory system. Apart from measurements of difference limens (DL), or just noticeable differences (jnds) of two tones in amplitude or frequency domain, detection thresholds of amplitude modulation (AM) and frequency modulation (FM) are determined as well. Detection of modulation is often considered in a context of excitation pattern changes [10, 27] especially in the situation when amplitude and frequency changes are so slow, that they produce noticeable temporary changes in loudness or pitch [24]. In papers concerning this problem an analysis of the auditory system based on maximum and minimum values of physical parameters of a signal is presented: it is assumed that an acoustic stimulus evokes an excitation pattern in the peripheral auditory system. Extreme levels of the excitation play a crucial role in signal changes detection. The excitation pattern reflects internal activity of the auditory system and it is conceived either as an envelope of the basilar membrane's displacement or as a number of neural spikes per second as a function of their characteristic frequency. It is also assumed that detection of changes in amplitude or frequency occurs when excitation pattern changes at any point (as in Zwicker's model [27]) or when excitation pattern changes of all active auditory filters exceeds certain value (FLORENTINE and BUUS model [2]). The problem of modulation detection is

described in many studies. However, there are not many papers concerning an effect of the starting phase of the modulator on the detection of modulation. Literature related to this problem is very scarce (except papers concerning mixed modulation detection [4, 13]). Such an effect is not expected for a modulation rate, f_{mod}, corresponding to a period shorter than duration of a signal or in a case when the modulation rate is high enough to evoke multitone sensation. Thus, the effect of the starting phase of the modulator should not be observed both in the roughness region [6, 20, 21, 26] where $20 < f_{mod} < \text{CMF}$ (CMF – a critical modulation rate [22]) and in the spectral region ($f_{mod} > \text{CMF}$), where modulation detection is based mainly on the spectral structure of the modulated signal. However, for very low modulation rates, or when the period of the modulating signal is comparable to the duration of the signal, or when just a few periods of the modulator are present in the stimulus, an influence of the starting phase of the modulator on modulation detection thresholds cannot be excluded. For stimuli lasting hundreds of milliseconds such a situation takes place when modulation is perceived on the basis of very slow fluctuations in amplitude or frequency only. In such a case extreme values of amplitude/frequency are heard as following changes in pitch or loudness. Thus, the starting phase of the modulating signal and instantaneous values of amplitude/frequency dependent on a starting phase may significantly influence sensation caused by the stimulus. Since the problem of the influence of the starting phase of the modulator is still ambiguous, in many psychoacoustical experiments concerning modulation detection/discrimination a random value of the modulator starting phase is used to avoid its systematic effect on experimental results [11, 20, 22, 25].

Two different mechanisms are usually considered to explain perception of AM and FM signals at a low modulation rate. One of them is based on changes in excitation level (so-called place mechanism) and the second one, additional in some sense, is based on a phase locking phenomenon [16, 19, 24] (temporal mechanism). This additional mechanism, which probably analyses time intervals between succeeding neural spikes in auditory neurones, operates only for low modulation rates ($f_{mod} < 5$ Hz) and for carriers no higher than 5 kHz [24]. This mechanism may evoke a dual result. First, the starting phase of the FM modulator may influence the modulation detection of signals whose frequency is lower or higher than 5 kHz, i.e. it may play an important role both in an area where phase locking is crucial for FM detection and in an area where phase locking does not work. Second, taking into account the fact that this additional mechanism operates in the frequency domain only, it can be expected that the starting phase of the modulator should have a different effect for FM than for AM.

The main aim of the presented work is to answer the following question: Does the starting phase of the AM or FM modulator influence the modulation detection threshold for very low modulation rates? And: Does this effect depend on type of modulation and the carrier frequency?

2. Psychometric functions for modulation detection

Measurements of psychometric functions for AM and FM detection were carried out in the aim to point out an influence of the starting phase of the AM or FM modulator on

detection thresholds. Such measurements describe probability of detection in amplitude or frequency changes as a function of an appropriate modulation index (m or β). This is more sensitive method of measuring of an influence of any parameter on the detection than direct threshold measurements with an up-down procedure [7]. Hence, the main aim of our investigations was to determine psychometric functions for modulation detection as a function of the modulation index m for AM and modulation index β for FM, for different values of the starting phase of the sinusoidal modulating signal.

2.1. Method

Psychometric functions for AM and FM were determined in separate experiments using a 2 AFC method. Subjects were presented with pairs of signals. One of the signals in a pair was AM or FM modulated while the second one was a pure tone carrier. These two signals were presented in random order. Subjects' task was to point out the modulated signal.

The starting phase of the modulator in each trial was randomly chosen from a discrete set of values i.e.: 0, $\pi/2$, π and $3\pi/2$. Signals with the different starting phase were presented in random order and during each experimental run (60 pairs each) number of presentations of each starting phase was the same. Five different values of modulation indices (m or β) were used and they occurred in random order. The highest modulation index corresponded usually to 85–95% correct while its lowest value corresponded usually to 55–60% correct. Values of modulation indices were measured in preliminary measurements. Each value of the modulation index was presented to the subjects at least 200 times.

A 70-dB pure tone at a frequency of 0.25, 1 and 6 kHz was a carrier signal. Modulation rate, f_{mod}, was equal to 2 or 5 Hz. Duration of the signal (including 20-ms raised-cosine rise/fall ramps) was 1 s. 1-s long signal included at least two periods of the modulator. Time interval between signals in a pair was equal to 300 ms. All signals were digitally generated by means of 16-bit AD converter (Tucker and Davis Technology) at a 50 kHz sampling rate and they were presented monaurally by means of Sennheiser HD 414 headphones in an acoustically isolated room. Three subjects with audiologically normal hearing took part in experiments.

2.2. Results and discussion

Probabilities of correct answers obtained during experiments corresponding to appropriate values of modulation indices m and β were converted into detectability domain d' [3, 7]. Generally, data obtained for each subject could be expressed in a following way:

$$d' = K x^{\gamma}, \tag{1}$$

where x denotes FM or AM modulation index, K and γ are constants having different values for different carrier and subject. For functions being the best approximation of experimental data values of exponents γ are included in a range of (1.83 – 2.12) for AM and in a range of (1.78 – 2.21) for FM. However to carry out a further analysis, the

experimental data were approximated by a square function, i.e. for $\gamma = 2$. For this exponent the smallest average value of the standard deviation between experimental data and their approximations was found. This exponent is also motivated theoretically [5] and such approximation has been usually used for analysis of similar data [12, 15, 19, 23].

Detectability d' for FM and AM signals as a function of the corresponding modulation index i.e. $d'(\beta^2)$ and $d'(1000 \cdot m^2)$ is shown in Fig. 1. Since the data obtained for all subjects taking part in experiments were qualitatively consistent, averaged results for $f_{mod} = 2$ Hz are shown there. AM data are presented in the left column while FM data are shown in the right one. Following rows of the figure illustrate data for 0.25, 1 and 6-kHz carrier, respectively. The starting phase of the modulator (0, $\pi/2$, π and $3\pi/2$) is a parameter of the data.

From data presented in Fig. 1 it can be stated that an increase in d', which is related with an increase of an appropriate modulation index, reflects a monotonic increase in psychometric functions. In the case of the AM signal the pattern of results is approximately the same for all carrier frequencies (0.25, 1 and 6 kHz). Moreover the detectability d' seems to be independent of the starting phase of the modulator. There is no such influence for both subthreshold ($d' \leq 1.16$) and suprathreshold ($d' > 1.16$) values of the AM modulation index. The situation is somehow different in FM case (see the right column of Fig. 1). For near-threshold values of FM modulation index no influence of the starting phase of the modulator on modulation detection is observed. However, for suprathreshold values of FM index some non-regular influence of the starting phase of the modulator is observed. This effect is visible especially for 6-kHz carrier.

Raw experimental data could not be subjected to an analysis of variance (ANOVA) because values of d' for different carriers were obtained for modulation indices coming from different ranges. Therefore, in the first step of the analysis psychometric functions obtained in the experiment were approximated by monotonically increasing functions, passing through the origin, according to Eq. (1). Then, the slope coefficients, K_{AM} and K_{FM}, were subjected to a within-subject analysis variance with the following factors: starting phase of the modulating signal (0, $\pi/2$, π and $3\pi/2$) and carrier frequency (0.25, 1 and 6 kHz).

For amplitude modulation AM, the starting phase of the modulator was not statistically significant [$F(3,6) = 0.23$, $p = 0.875$]. This result confirms that threshold for AM detection does not depend on the starting phase of the modulator. The carrier frequency was not statistically significant, too [$F(2,4) = 2.98$, $p = 0.161$]. This result is consistent with a finding that AM detection threshold is independent of the carrier frequency at low modulation rates [1, 17, 25, 27]. It is worth to add that an interaction between the carrier frequency and the starting phase of the modulator was not statistically significant, [$F(6,12) = 0.11$, $p = 0.994$].

For frequency modulation FM, the starting phase of the modulator was not statistically significant [$F(3,6) = 1.07$, $p = 0.431$]. Thus, similarly to AM case, it allows to state that the starting phase of the modulator is not a crucial parameter for FM detection too. However, as opposite to AM, it was found that the carrier frequency was highly significant [$F(2,4)] = 52.67$, $p = 0.001$]. This result reflects a dependence of the

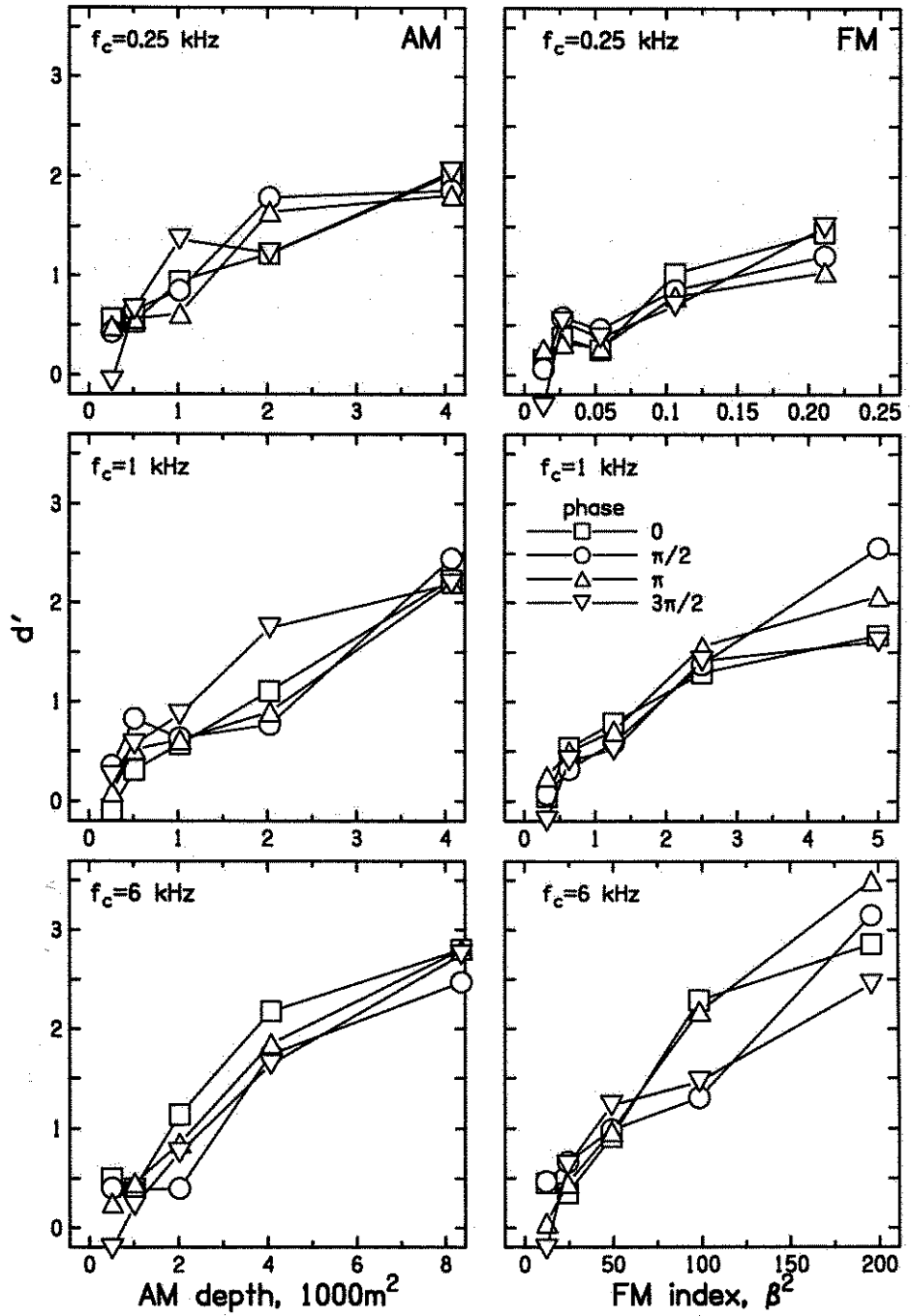


Fig. 1. Psychometric functions for AM (left column) and FM (right column) detection expressed in terms of detectability d' as a function of 1000 m^2 and β^2 for AM and FM respectively. Following rows show the data for a sinusoidal carrier at a frequency of 0.25, 1 and 6 kHz. Data were averaged across three subjects and the parameter is the value of the starting phase of the modulator: 0, $\pi/2$, π or $3\pi/2$.

FM detection thresholds on the carrier frequency when the thresholds are expressed in terms of FM index, β. This result is in a good agreement with literature [10, 27]. An interaction between the starting phase of the modulator and the carrier frequency was not statistically significant [$F(6,12) = 1.24$, $p = 0.351$]. It suggests that a scatter of slope values, observed for different values of the starting phase, suprathreshold values of FM modulation index and a carrier frequency, are not crucial.

Data presented above enable us to state that the starting phase of the modulator is not a crucial parameter that determines the shape of psychometric functions and it does not influence the detection thresholds of AM or FM. It seems that when the subject's task is to detect amplitude/frequency changes in a signal, the starting phase of amplitude or frequency changes is not the most important factor. In such cases the basic detection cue is a difference between maximum and minimum value of the discriminated

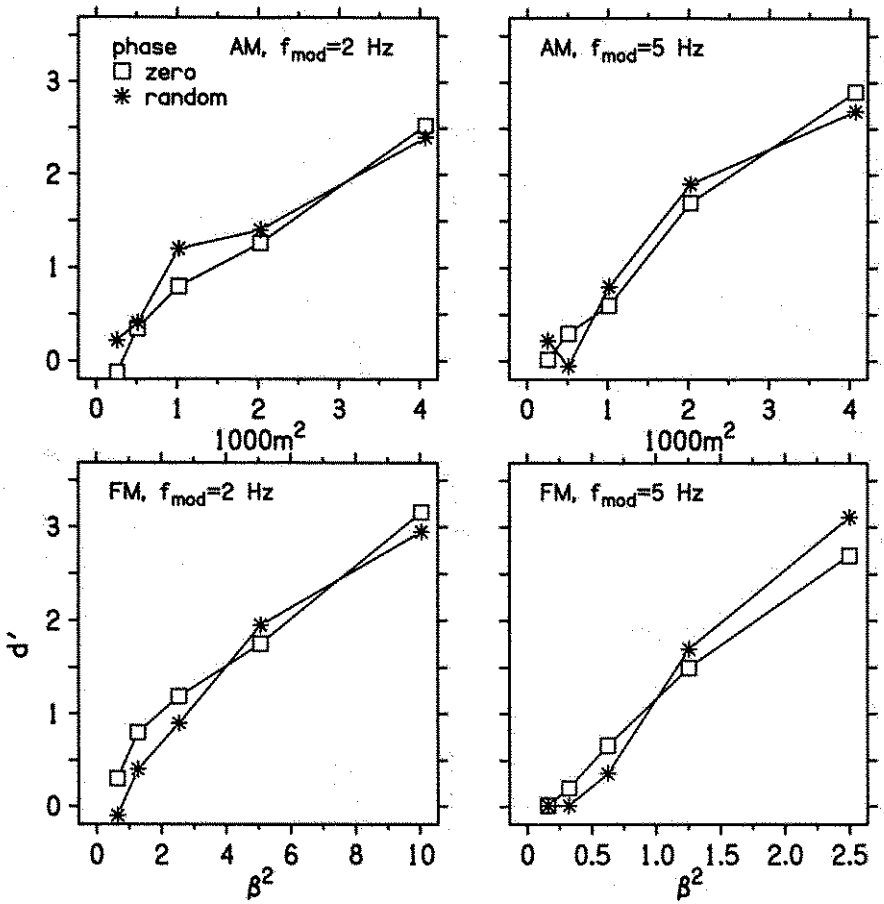


Fig. 2. Psychometric functions expressed in terms of detectability d' as a function of 1000 m^2 for AM and β^2 for FM (lower row) for 1-kHz sinusoidal carrier and for modulation rate of 2 Hz (left column) and 5 Hz (right column). Data were averaged across three subjects and the parameter is the value of the starting phase of the modulator: constant (zero) or randomly chosen from (0–2π) range.

parameter. Lack of the influence of the starting phase of the modulator on psychometric functions is consistent with models based on analysis of extreme values of the excitation pattern (i.e. Zwicker's model [28] or a group of integration models [3]).

Presented above experimental data were obtained for a randomly chosen phase value, taken from four discrete values, i.e. 0, $\pi/2$, π and $3\pi/2$. Although, these starting phases represent the most important starting values and directions of changes in amplitude/frequency which may occur at the beginning of the modulated signal, they are not "random" in a sense of equal probability that phase was taken by chance from the range of $(0-2\pi)$. Therefore an additional experiment (an analogue to described above) was performed. In this experiment psychometric functions for modulation detection, for randomly chosen starting phase from the range of $(0-2\pi)$ and for zero starting phase were compared. The method and signal parameters were similar to those described above. 1-kHz sinusoid was a carrier and modulation rate was equal to 2 and 5 Hz. Results of these additional measurements averaged for three subjects (the same as in previous experiment) are shown in Fig. 2. The upper row shows results obtained for AM and the lower one for FM. In the left column data for 2-Hz modulation rate are shown while in the right column – data for 5-Hz modulation rate are presented. The modulation starting phase is the parameter of the data: it was either random or zero.

From data presented in Fig. 2 it can be seen that psychometric functions expressed in terms of detectability d' are similar for both random and zero starting phase of the modulator. This observation confirms results obtained for the starting phase randomly chosen from four discret values.

3. Summary and conclusions

Experimental data presented in this paper correspond well to detection models based on an analysis of differences in excitation level at extreme values of physical parameters of a signal. It seems that the auditory system analyses short, overlapping time intervals (i.e. by means of the sliding window in time domain [18]) and compares excitation levels evoked by the acoustic signal in these time intervals. However, despite of the duration of the time interval, which is probably related to integration time of the auditory system, information about extreme excitations is stored at higher levels of the auditory system long enough to be further compared with a current excitation coming from the periphery. In the auditory system the time during which the extreme excitation is stored is much longer than integration time, because AM detection thresholds are constant in the wide range of the modulation rate [1]. If this time was short, it would be impossible to detect slow changes in amplitude/frequency of a signal. A sequence of extreme values of physical parameters of the signal and corresponding extreme excitation levels, or the excitation starting value (related to the starting value of the parameter under analysis) does not influence the detection of changes in time. It seems that all models based on the difference in excitation level, corresponding to extreme values of investigated physical parameter and according to which its sequence and the starting phase do not play a crucial role (i.e. FLORENTINE–BUUS model [2] or non-optimal excitation model [19, 24]) describe

the auditory system in a proper way as far as detection of slow amplitude/frequency changes is concerned. If the succession of the temporal order of extreme values of physical parameters in the acoustic signal does not influence the threshold of these changes, then the detection of amplitude/frequency changes should be independent of the time pattern of these changes and should depend on differences in extreme values of the parameter only. Indeed, such relation was observed for linear changes in signal frequency for a wide frequency range [8, 9, 14]: thresholds of such changes were independent of time and direction of changes. The crucial factors were the difference in frequency between the beginning and the end of the signal and its centre frequency.

From the results presented in this paper it can be generally stated that the starting phase of the sinusoidal modulator at a rate of 2 or 5 Hz, superimposed (as AM or FM modulator) on a sinusoidal carrier does not critically influence psychometric functions for AM and FM detection. Both, random and constant starting phases bring about similar results. Even though the auditory system can evaluate changes in physical parameters based on local extrema (in time), the actual course of these changes does not influence the threshold.

However this is true for changes that do not contain any additional detection cues, i.e. the spread of signal spectrum caused by rapid amplitude/frequency changes.

Acknowledgement

This work supported by State Committe for Scientific Research, grant No. 7 TO7B 05216.

References

[1] T. DAU, B. KOLLMEIER and A. KOHLRAUSCH, *Modeling auditory processing of amplitude modulation: I. Detection and masking with narrowband carriers*, J. Acoust. Soc. Am., **102**, 2892–2895 (1997).

[2] M. FLORENTINE and S. BUUS, *An excitation-pattern model for intensity discrimination*, J. Acoust. Soc. Am., **70**, 1646–1654 (1981).

[3] D.M. GREEN and J.A. SWETS, *Signal detection theory and psychophysics*, Krieger, New York 1974.

[4] W.M. HARTMANN and G.M. HNATH, *Detection of mixed modulation*, Acustica, **50**, 297–312 (1982).

[5] W.M. HARTMANN and M.A. KLEIN, *Theory of modulation detection for low modulation frequencies*, J. Acoust. Soc. Am., **67**, 935–946 (1980).

[6] S. KEMP, *Roughness of frequency modulated tones*, Acustica, **50**, 126–133 (1982).

[7] N.A. MACMILLAN and C.D. CREELMAN, *Detection theory: A user's guide*, Cambridge University Press, Cambridge, England 1991.

[8] J.P. MADDEN and K.M. FIRE, *Detection and discrimination of frequency glides as a function of direction, duration, frequency span and center frequency*, J. Acoust. Soc. Am., **102**, 2920–2924 (1997).

[9] J.P. MADDEN and K.M. FIRE, *Detection and discrimination of gliding tones as a function of frequency transition and center frequency*, J. Acoust. Soc. Am., **100**, 3754–3760 (1996).

[10] B.C.J. MOORE, *An introduction to the psychology of hearing*, 4th Ed., Academic Press, London 1997.

[11] B.C.J. MOORE and A. SĘK, *Discrimination of modulation type (AM or FM) with and without background noise*, J. Acoust. Soc. Am., **96**, 726–732 (1994).

[12] B.C.J. MOORE and A. SĘK, *Effects of carrier frequency, modulation rate and modulation waveform on the detection of modulation and the discrimination of modulation type (AM vs FM)*, J. Acoust. Soc. Am., **97**, 2468–2478 (1995).

[13] B.C.J. MOORE and A. SĘK, *Detection of combined frequency and amplitude modulation*, J. Acoust. Soc. Am., **92**, 3119–3131 (1992).

[14] B.C.J. MOORE and A. SĘK, *Discrimination of frequency glides with superimposed random glides in level*, J. Acoust. Soc. Am., **104**, 1, 411–421 (1998).

[15] B.C.J. MOORE and A. SĘK, *Effects of carrier frequency and background noise on the detection of mixed modulation*, J. Acoust. Soc. Am., **96**, 741–751 (1994).

[16] B.C.J. MOORE and A. SĘK, *Effects of carrier frequency, modulation rate and modulation waveform on the detection of modulation and the discrimination of modulation type (AM vs FM)*, J. Acoust. Soc. Am., **97**, 2468–2478 (1995).

[17] E. OZIMEK, J. KONIECZNY, Y. SUZUKI and T. SONE, *Random changes in envelope of AM tones and their detection*, J. Acoust. Soc. Jpn. (E), **19**, 83–91 (1998).

[18] C.J. PLACK and B.C.J. MOORE, *Temporal window shape as a function of frequency and level*, J. Acoust. Soc. Am., **87**, 2178–2187 (1990).

[19] A. SĘK, *Detection and discrimination of modulation at a low modulation rate*, Archives of Acoustics, **23**, 3, 363–378 (1998).

[20] A. SĘK, *Modulation thresholds and critical modulation frequency based on random amplitude and frequency changes*, J. Acoust. Soc. Jpn. (E), **15**, 67–75 (1994).

[21] A. SĘK, *Perception of irregular amplitude changes of sinusoidal signal*, Acustica, **77**, 262–269 (1993).

[22] A. SĘK and B.C.J. MOORE, *The critical modulation frequency and its relationship to auditory filtering at low frequencies*, J. Acoust. Soc. Am., **95**, 2606–2615 (1994).

[23] A. SĘK and B.C.J. MOORE, *Detection of mixed modulation using correlated and uncorrelated noise modulators*, J. Acoust. Soc. Am., **95**, 3511–3518 (1994).

[24] A. SĘK and B.C.J. MOORE, *The detection of modulation at low modulation rates*, [in:] Auditory Perception: Some Principles and Applications, U. JORASZ [Ed.], Adam Mickiewicz University Press, Poznań 1996.

[25] A.P. SĘK and E.B. SKRODZKA, *Krzywe strojenia modulacji*, Audiofonologia, accepted (1997).

[26] E. TERHARDT, *On the perception of periodic sound fluctuations (roughness)*, Acustica, **30**, 201–213 (1974).

[27] E. ZWICKER and H. FASTL, *Psychoacoustics – facts and models*, Springer-Verlag, Berlin 1990.

[28] E. ZWICKER and R. FELDTKELLER, *Das Ohr als Nachtrichtenempfänger*, Hirzel-Verlag, Stuttgart 1967.

ARCHIVES OF ACOUSTICS
24, 1, 49–56 (1999)

SINGLE PHOTON COUNTING IN BRILLOUIN LASER LIGHT SCATTERING EXPERIMENTS

T. BŁACHOWICZ

Institute of Physics
Silesian University of Technology
(44-100 Gliwice, Krzywoustego 2)
e-mail: blachowicz@tytan.matfiz.polsl.gliwice.pl

What follows is a description of the electronics and simple mechanical equipment needed for single photon counting. It is intended as an introduction and guide to assist users in the measurement of Brillouin laser light scattering on acoustic hypersonic waves. Simple method of a numerical filtration of the low-intensity-level optical signal is described in detail.

1. Introduction

Observation of the acoustic wave in the hypersonic range consists in measuring changes of photon energy inelastically scattered in annihilation and creation processes by phonons lying at the beginning of the first Brillouin zone. However the intensity of this optical signal is very low with respect to the intensity of light incident on the investigated sample, so an appropriate registration method must be applied. One of the most sensitive techniques that can perform this task is a single photon counting SPC [1–3], well known from its wide application to different materials such as crystals, thin layers and superlattices structures [4–9].

The main purpose of this paper is to instruct users how to perform such experiments and how to build their own experimental setup, especially for observing acoustic waves in the hypersonic range by the Brillouin laser light scattering.

2. Experimental apparatus

The general scheme of an electronic for the SPC method is presented in Fig. 1. The signal from a fast photomultiplier, for example, a Hamamatsu P-series, is amplified and then formed to a rectangular shape. An amplitude discriminator with upper and lower thresholds is then applied, followed by a counter sensitive for TTL standard signal level connected with a computer. In the case of presented experimental equipment the amplifier, the shaper and the discriminator are stored in a PTI-614 analog-digital module

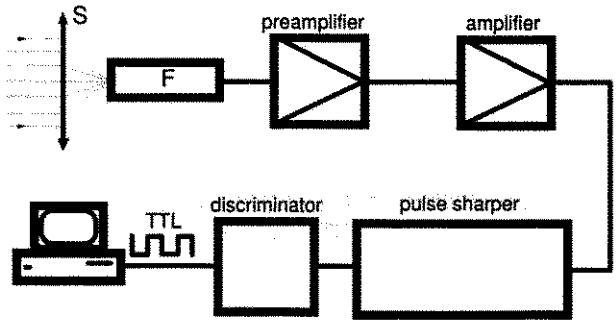


Fig. 1. Block diagram of electronic equipment for the SPC method. S – lens, F – photomultiplier. An optical signal comes through a lens to the photomultiplier. At the output of the electronic equipment the signal is in TTL standard and is registered by the PC computer.

produced by PTI Inc. Table 1 provides the main parameters of this unit. The Hamamatsu R-4220P photomultiplier was placed inside the PTI-614. Table 2 provides its parameters. One of the most important parameters of the PTI-614 module is a pulse pair resolution equal to 250 ns. This means that any photon or a group of photons coming into the photomultiplier in a period of time shorter than 250 ns will not be recognized as a new count. The signal coming to the counter (in our case, a standard counter used for nuclear experiments), possessed a stochastic nature due to the thermal noise of the photomultiplier and the random nature of the low-intensity-level scattered light.

Table 1. Parameters of the 614 PMT analog/digital unit.

Drift of the signal (%/hour)	Maximum count rate (MHz)	Pulse pair resolution (ns)	Raise time of the pulse (ns)	Fall time od the pulse (ns)	Pulse width (ns)
0.03	4	250	20	100	220

Table 2. Parameters of the R4220P photomultiplier.

Photocatode	Amplification	Spectral response (nm)	Raise time of the signal (ns)
Low noise "bialkali" (Na-K-Sb)	$1.2 \cdot 10^7$	185 - 710 (max. 410)	2.2 ns

For these reasons a numerical filtration must be performed, assuming that the noise is conformable to the Poisson distribution (Fig. 2). The signal is also a function of the voltage supply of the photomultiplier (Fig. 3). The main reason for performing filtration is to obtain a smoothed Brillouin spectrum. So a weighted average was used of some neighboring values of counts, creating an "average window". This averaging was achieved as follows:

An error of a single measurement is set equal to

$$\Delta N_i = \sqrt{N_i}. \tag{1}$$

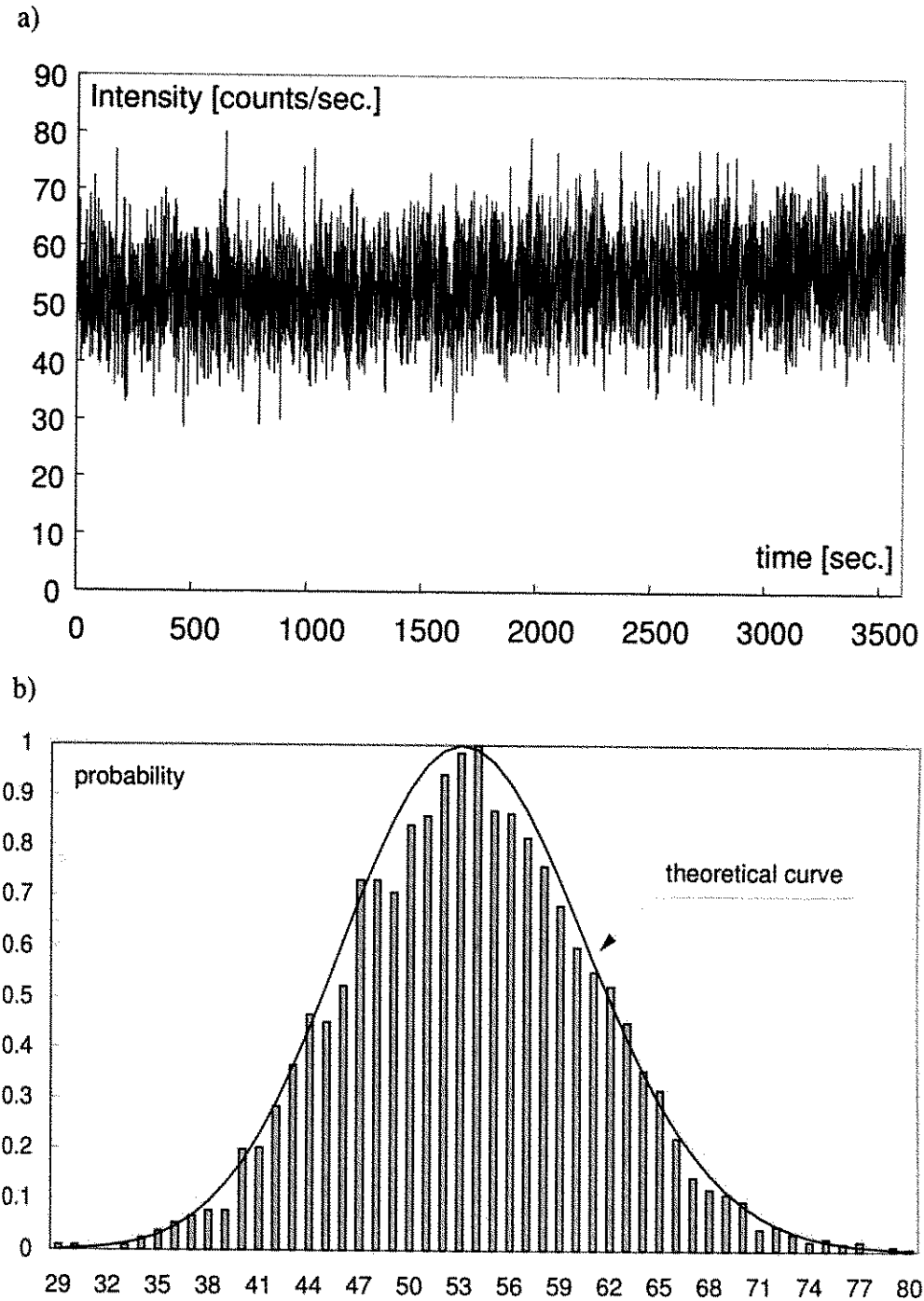


Fig. 2. Noise of the photomultiplier with no optical signal (a) and a histogram of its Poisson distribution (b). Time of measurement was equal to one hour.

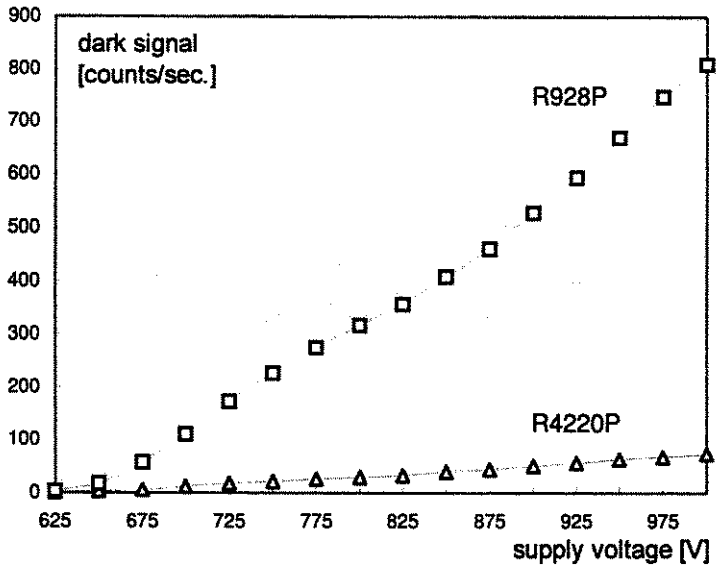


Fig. 3. Comparison of dark signals from two photomultipliers: R928P and R4220P. The second photomultiplier is more useful for low-level optical signal detection.

A relative error of a single measurement is equal to

$$\frac{\Delta N_i}{N_i} = \frac{\sqrt{N_i}}{N_i}. \tag{2}$$

Then, a weighted average value from n single measurements is equal to

$$\overline{N} = \frac{\sum_{i=1}^{i=n} \omega_i N_i}{\sum_{i=1}^{i=n} \omega_i}, \tag{3}$$

where the weight of a single measurement ω_i is equal to the reciprocal of a relative error, of this measurement, raised to the second power. By substituting Eq. (2) into Eq. (3) we obtain

$$\overline{N} = \frac{\sum_{i=1}^{i=n} N_i^2}{\sum_{i=1}^{i=n} N_i}, \tag{4}$$

where $\overline{N}$ is a weighted average value from the "filtration window" of the width equal to n. The above procedure must be performed from the first point of data of initial position to the position equal to $(N_T - n + 1)$, where N_T is the total number of registered points.

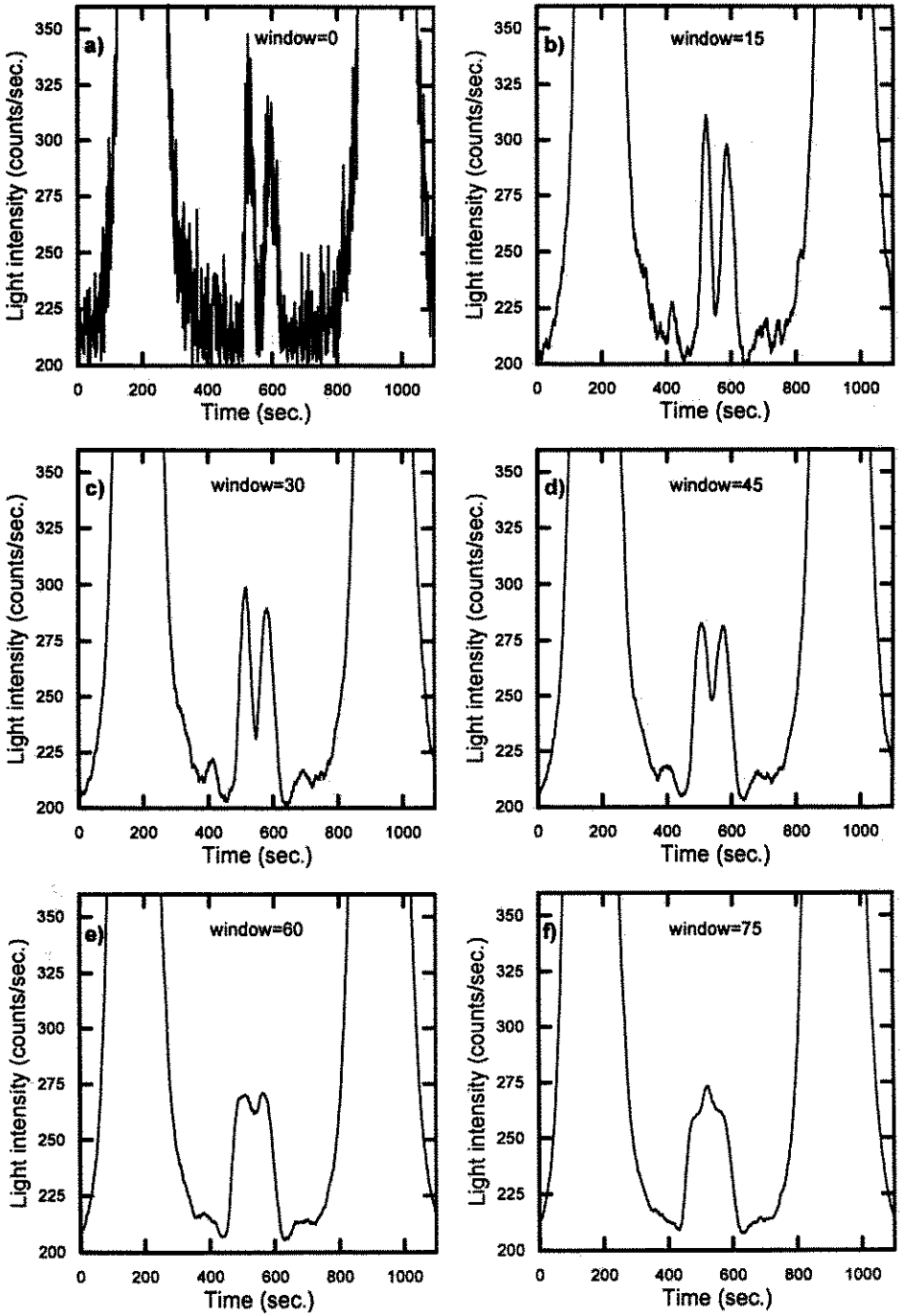


Fig. 4. Examples of the Brillouin spectra from the $LiTaO_3$ crystal after filtration by different widths of the filtration window. Each picture was created from 1100 data points.

A separate problem is the choice of an appropriate width of the "filtration window". If the width is too large, important properties of the spectrum may be lost (Fig. 4). A workable value, discovered in practice, takes the width to be comparable to half the width of a Poisson distribution of dark counts at the same input voltage as for experiments with real scattered light. As we can see from Fig. 4, the best choice of a width of the "filtration window" is the "d" case. A filtration width equal to 75 loses physical information completely.

A problem arises in connecting the PTI-614 module to the computer. This can be solved by the use of any frame grabber or external module. In our case the counter was a relatively old module giving on its output a signal in a BCD standard. An interface was designed which connects the counter with a parallel port of a PC class computer. The interface was controlled by a Pascal language program. The main task of the interface was the successive reading and transmission of numbers from the counter to the computer.

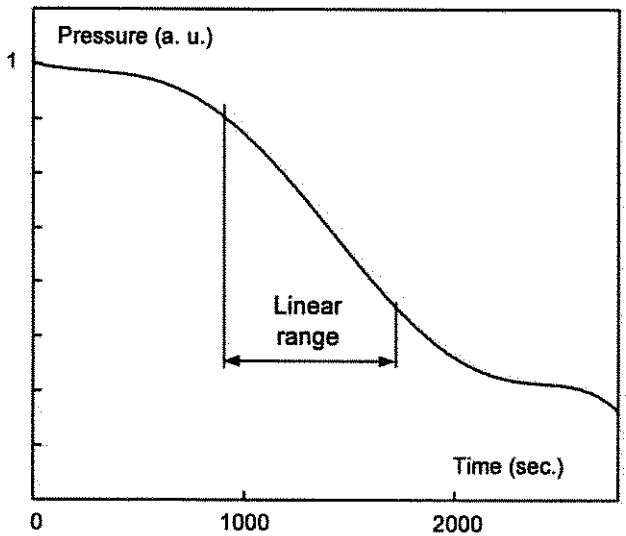


Fig. 5. Dependence of the pressure in interferometer chamber on a time. The middle of the plot provides a linear dependence useful for measurement.

Another problem is the linear scanning of the Brillouin spectrum in time. The total phase difference between the mirrors of the Fabry–Perot interferometer must be scanned or controlled. In most cases, this is done by the piezoelectric method [10]. An older, less precise, but easier method is a pressure scanning [11], where an interferometer is placed in a chamber connected with a pump and with a second controlling chamber useful for valves and a capillary mounting. This allows us to remove a gas from the chambers and then leak it slowly back. Most importantly, this produces a linear range of pressure changes (Fig. 5), useful for our purposes. It means that the Brillouin spectrum can be linear scanned. Obviously, this is a function of the mechanical parameters of the equipment. In the case of our apparatus, the total volume of the chambers was equal to $6 \cdot 10^{-2} m^3$ and the diameter of the capillary was of the order of 10^{-5} m.

3. Examples of a registered Brillouin spectrum

As was mentioned at the beginning, Brillouin light scattering consists in measuring changes of photon frequency inelastically scattered on acoustical phonons lying at the beginning of the first Brillouin zone. At the quantum level annihilation and creation processes are responsible for the typical picture of the Brillouin spectrum where lines of lowered and increased frequency can be visible. Figure 6 provides the experimental spectra of acoustic phonons registered by the use of the SPC in the transparent for laser wavelength $LiTaO_3$ crystal. Lines described as longitudinal (L) and quasi-transverse waves (T_1 and T_2) were spatially separated from high intensity elastically scattered light by the Fabry–Perot interferometer — a standard method in every Brillouin or Raman spectroscopy experiment. Table 3 provides the results of measurements in the [100] crystalographic direction (compare Fig. 6 b). The averaged values are equal to 25.60 ± 0.17 GHz and 20.36 ± 0.15 GHz.

Table 3. Example results of measurements of quasi-transverse waves frequency in the $LiTaO_3$ crystal – [100] direction.

Frequency of the first quasi-transverse wave (GHz)	Frequency of the second quasi-transverse wave (GHz)
25.62	20.40
25.68	20.46
25.44	20.33
25.46	20.41
25.81	20.38

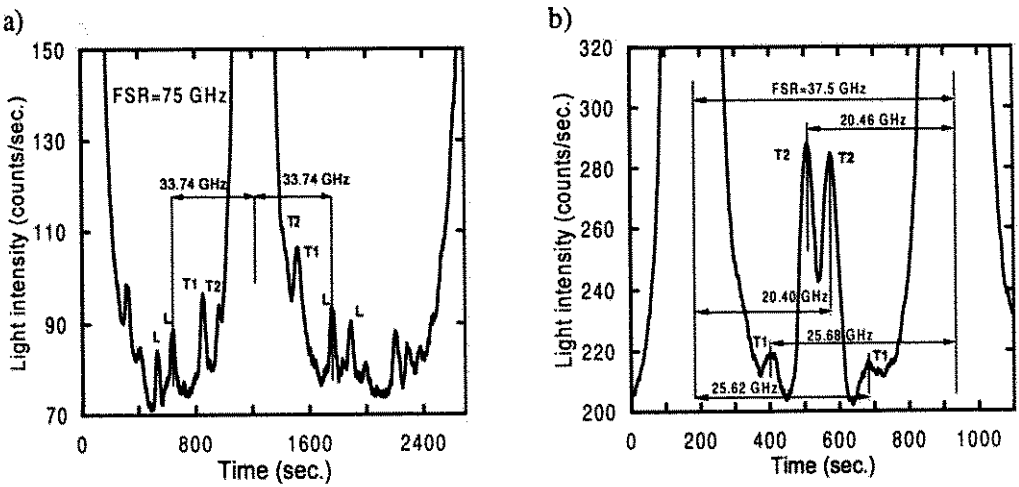


Fig. 6. Two examples of Brillouin spectra from the $LiTaO_3$ crystal for two different full spectral ranges: a) FSR = 75 GHz, the pressure range is linear only around the middle Rayleigh line. b) FSR = 37.5 GHz. Descriptions: T_1, T_2 – quasi-transverse waves, L – quasi-longitudinal wave.

The systematic error in phonon frequency measurement caused by the choice of the interferometer's full spectral range and numerical treatment of data was equal to 0.15 GHz. It is interesting to compare the amplitude of the line from elastic scattering with the amplitude of Brillouin signal with respect to the noise level. The high resolution capacity of the SPC method is evident.

4. Conclusion

Single-photon counting is one of the most sensitive methods for very low-level optical signal detection. What was described, in a simple way, is how to perform such an experiment for the registration of light inelastically scattered on acoustic hypersonic waves. Most of the equipment was designed and made in laboratory. By the appropriate choice of volume of the pressure chamber and diameter of the capillary, it was possible to obtain a linear range of Fabry–Perot scanning, which is crucial in such experiments. Simple numerical filtration was used to smooth the registered spectra. The described use of the SPC method can be applied not only in hypersonic acoustics [12], but also in other field of physics, chemistry and biology.

References

[1] S. CORA *et al.*, Proc. 10th International Symp. Techn. Committee Photon-Detectors. Berlin (West) 1982, p. 14.

[2] W. DEMTRÖDER, *Laser spectroscopy*, Springer-Verlag, Berlin 1981.

[3] *Technical information Cat-No TPMO90001E01, How to perform photon counting using photomultiplier tubes*, Hamamatsu, March 1993.

[4] T. BŁACHOWICZ, Z. KLESZCZEWSKI, *Observation of hypersonic acoustic waves in a* $LiTaO_3$ *crystal*, Archives of Acoustics, **22**, 3, 351–357 (1997).

[5] Y. TAKAGI, R.W. GAMMON, *Brillouin scattering in thin samples: Observation of backscattering components by* 90° *scattering*, J. Appl. Phys., **61**, 5, 2030–2034 (1987).

[6] B. HILLEBRANDS, P. BAUMGART, G. GÜNTHERODT, *In situ Brillouin scattering from surface-anisotropy-dominated Damon–Eshbach modes in ultrathin epitaxial* Fe(110) *layers*, Phys. Rev., **B36**, 4, 2450–2453 (1987).

[7] J.A. BELL, W.R. BENNETT, R. ZANONI, G.I. STEGEMAN, C.M. FALCO, F. NIZZOLI, *Elastic constants of* Mo/Ta *superlattices measured by Brillouin scattering*, Phys. Rev., **B35**, 8, 4127–4130 (1987).

[8] M. HUES, R. BHADRA, M. GRIMSDITCH, E. FULLERTON, I.K. SCHULLER, *Effect of high-energy ion irradiation on the elastic moduli of* Ag/Co *superlattices*, Phys. Rev., **B39**, 17, 12966–12968 (1989).

[9] S. KUMAR, R. BHADRA, A. FARTASH, M. GRIMSDITCH, C. KIM, B. QADRI, A. EDELSTEIN, *Brillouin scattering from ion-beam-sputtered* Cu/Ni *superlattices*, Phys. Rev., **B44**, 11, 5905–5907 (1991).

[10] R. MOCK, B. HILLEBRANDS, R. SANDERCOCK, *Construction and performance of a Brillouin scattering set-up using triple-pass tandem Fabry–Perot interferometer*, J. Phys. E: Sci. Instrum., **20**, 656–659 (1987).

[11] T. BŁACHOWICZ, *Application of Brillouin scattering to the analysis of acoustic properties of the piezoelectric crystals* [in Polish], Ph.D. Thesis, Silesian University of Technology in Gliwice, Poland 1997.

[12] T. BŁACHOWICZ, Z. KLESZCZEWSKI, *Elastic constants of the lithium tantalate crystal in the hypersonic range*, Acoust. Lett., **20**, 10, 221–223 (1997).

ARCHIVES OF ACOUSTICS
24, 1, 57–74 (1999)

LOSSES IN A PIEZOELECTRIC CERAMIC IN HIGH FIELDS

M. SZALEWSKI

Institute of Fundamental Technological Research
Polish Academy of Sciences
(00-049 Warszawa, Świętokrzyska 21)

The paper concerns the problem of measurements of electrical and mechanical losses in a piezoelectric ceramic driven by high electric field and vibrating with high vibration velocity. Deformations and discontinuities of resonance curves appear in this range of the fields. Therefore the application of typical methods of measurements of quality factors is difficult or even impossible. The author proposes to apply the measurements of voltage ratios in a piezoelectric transformer instead of the measurements of resonator quality factors. In the paper the relations connecting the voltage ratios in the piezoelectric transformer with the values of electrical and mechanical losses as well as experimental results are presented. The formulae have been derived using the KLM equivalent circuit for two limits of the transformer load resistance. The obtained results are compared with the results obtained using one of the earlier known measurement method in the range of its applicability.

1. Introduction

Electrical and mechanical losses in a piezoelectric ceramic and methods of their measurements were the subject of many theoretical and experimental works. However the problem of measurements of losses in the range of high electrical and mechanical fields has not been solved. This problem is important because piezoelectric ceramics are widely used in high-power devices, such as piezoelectric actuators, ultrasonic motors, piezoelectric transformers and sending transducers in hydroacoustic devices. In recent years one can observe the development of high-power piezoelectric ceramic devices and the growing number of their applications. Experimental investigations of multicomponent ceramics have proved that the ceramic compositions with excellent electromechanical properties under a relatively low vibration level did not necessarily guarantee good operation under a relatively high vibration level [38]. Therefore investigations of properties of a piezoelectric ceramic in high fields are still the subject of many researches, e.g. [5, 28, 40].

In a high electric field, when a vibration velocity is high, resonance curves deform at first and next discontinuities of resonance curves appear. These effects make difficult or even impossible to determine losses applying classical methods of measurements of a quality factor. We have tried to solve this problem applying measurements of a voltage ratio in a piezoelectric transformer instead of measurements of a resonator quality

factor. Relations between voltage ratios in a piezoelectric transformer and mechanical and electrical losses are presented in Sec. 4. These equations have been derived using the KLM equivalent circuit. Experimental results and their comparison with the results obtained using one of the earlier known methods (described in Sec. 3), in the range of its applicability, are presented in Sec. 5.

2. Origin of losses in a piezoelectric ceramic

As yet there is no complete doubtless description of the mechanism of electrical and mechanical losses in a piezoelectric ceramic. Many theories exist attributing a leading part to different effects. According to [27] mechanical losses are caused by the scattering at grain boundaries, by internal friction in ceramic grains and domains and by micro- or macrocrackings. According to [4] mechanical losses are mainly due to the motional resistivity and electrical losses are mainly due to bulk grain and grain boundary effects, domain walls effects are very important in the frequency range above 10^3 Hz. According to [11, 12] both types of losses are induced by 90° ferroelectric domain walls moving under the influence of electrical fields or mechanical stresses. The physical nature of electrical losses is explained by the damping of moving 90° domain walls. For frequencies up to 10^8 Hz the origin of damping of the wall motion is ascribed to point defects within the 90° domain wall. At higher frequencies damping results from the reflection of thermal-lattice waves impinging on a moving 90° domain wall. The ratio of electrical and mechanical losses depends on material properties like spontaneous polarization, spontaneous striction, elastic compliance, dielectric constant [12]. The analysis of effects induced by the domains motion allows also to explain the strong increase of electrical losses at frequencies above 10^8 Hz and their decrease at very high frequencies [11]. The description of the phenomena accompanying the domains motion in a polycrystalline ceramic is difficult because many factors influence these processes — external electric fields and mechanical stresses, temperature, grain size, grain boundaries, internal stresses, defects, microcracks, ceramic composition, ageing, ferroelectric fatigue, chemical inhomogeneities, porosity. The existing theories both microscopic and phenomenological are too complicated to apply them in practice. Besides as a rule quantities that cannot be measured in real polycrystalline materials exist in these theories [8, 43].

Changes of losses caused by compositional modifications of ceramics confirm that losses are mainly due to the domain walls motion. Electrical and mechanical losses increase in $Pb(Zr,Ti)O_3$ ceramics with donor additives. Donor additives induce the increase of domain walls mobility, then even small electric fields or mechanical stresses can displace domain walls. Acceptor additives reduce domain walls mobility and electrical and mechanical losses in result [19].

Electrical and mechanical losses increase in result of the increase of driving electric field, vibration velocity, temperature. Ageing causes the decrease of losses, also this effect is due to the domain walls motion [26].

3. Measurements of losses in high fields

Various methods of the measurements of losses are known. In many cases only a mechanical quality factor Q_m is measured. One assumes that electrical losses are negligibly small in relation to mechanical losses. This assumption is satisfied when a transducer is excited by low electrical field to resonant vibrations and its acoustical load is small [6, 9]. The mechanical quality factor can be determined from the width of transducer resonance curve on the 3-dB level, from the characteristic frequencies of the transducer admittance circle or from the measurements of the absolute admittance as a function of frequency [27, 30]. Electrical losses predominate when a transducer is driven by high electric field at frequencies much below its resonance [42]. They can be measured using Schering bridge [7]. However for high-coupling piezoelectric ceramics of a low mechanical Q the measured electrical losses contain an excess portion resulting from the mechanical losses. These additional losses are caused by a piezoelectrically induced quasi-static strain due to low-frequency electric field [34]. The authors of [23] have presented a method of the determination of total losses (sum of electrical and mechanical losses) from measurements of a real part of an electric impedance of a transducer. The measurements were done only for low electric fields, the measured losses were mainly mechanical.

The above-mentioned methods can be applied only in the linear range of the transducer work. In high electric fields, when the vibration velocity is high, the resonance curves deform — decrease of the resonance frequency, asymmetry. The stiffness of the ceramic decreases. The ceramic is not obeying Hooke's law, the material may be classified as a soft spring [25, 36, 42]. When the electric field continues to increase the curves become discontinuous, so-called jump phenomenon appears. The jump phenomenon appears at different frequencies for a upgoing or downgoing frequency sweep during measurements (hysteresis effect). This phenomenon was thoroughly investigated experimentally and theoretically for quartz resonators [10], for a piezoelectric ceramic it was described for the first time by K. NEGISHI [29]. This nonlinear behaviour of the piezoelectric ceramic can be described theoretically extending up to the second order the constitutive piezoelectric equations [13]. The discontinuities are also observed during the measurements of transducer admittance circles or absolute admittance as a function of frequency [39]. The determination of the quality factor from the width of the resonance curve on the 3-dB level gives large error in the case of the asymmmetric curves. For the curves with the discontinuities it makes no sense.

The range of the fields applicable during the measurements is limited by irreversible effects inducing durable changes of ceramic properties. Large mechanical stresses, high electric fields and thermal effects connected with them can cause the ferroelectric fatigue, partial depoling of the ceramic, changes in ageing effects [42].

In recent years the authors from Japan published the series of papers (e.g. [14–16, 37–39]) concerning the applications of the quality factor measurements in the resonance and antiresonance to the determination of the electrical and mechanical losses in wide range of input voltages. For example they obtained for length extensional vibrations of

a thin, long and narrow plate poled along the thickness direction [14, 15, 18]:

$$\operatorname{tg}\delta_m = \frac{1}{Q_B}, \tag{3.1}$$

$$\operatorname{tg}\delta_e = \gamma\left(\frac{1}{Q_A} - \frac{\omega_B}{\omega_A}\frac{1}{Q_B}\right), \tag{3.2}$$

$$\gamma = \frac{1-k_{31}^2}{k_{31}^2}\frac{\pi^2}{8}, \tag{3.3}$$

where: Q_A – the quality factor measured at the electrical resonance angular frequency ω_A under mechanically free conditions (or at mechanical resonance angular frequency ω_A giving maximum vibration velocity under electrically short-circuit conditions), Q_B – the quality factor measured at the electrical antiresonance angular frequency ω_B under mechanically free conditions (or at mechanical resonance angular frequency ω_B giving maximum vibration velocity under electrically open-circuit conditions), k_{31} – electromechanical coupling coefficient.

Similar relations were obtained for thickness vibrations of a plate [20].

The authors of the above-mentioned method determined the quality factors using the frequency perturbation method [16, 33]. In this way they could apply their method in the range of the occuring of the resonance curve deformations. However it does not solve the problem of the loss measurements in the range of the occuring of the resonance curves discontinuities. We have tried to solve this problem applying the measurements of voltage ratios in a piezoelectric transformer instead of the measurements of a resonator quality factor.

4. Application of a piezoelectric transformer to loss measurements

Various designs of piezoelectric transformers are known. They have various shapes of individual parts [24]. For the loss measurements we applied a ring-shaped ceramic piezoelectric transformer – Fig. 1. The thickness and width of a ring were small in comparison with its radius. Such a shape has an important advantage - the stress and strain distribution is uniform in the whole ring [3, 42]. The piezoelectric ceramic was poled along the thickness direction. One pair of vacuum evaporated silver electrodes constituted the input of the transformer, the second pair – its output (Fig. 1).

The foundations of the theory of piezoelectric transformers were done by C.A. ROSEN [35]. Applying Mason's equivalent circuit Rosen presented an equivalent circuit of a piezoelectric transformer and he derived the equations relating a voltage amplification and components of the equivalent circuit. Rosen's analysis cannot be applied in the case discussed in this paper. Firstly – Rosen did not consider electrical losses assuming that they were negligibly small. Secondly – he analysed the parallepiped – shaped transformer with two mechanically free ends.

Piezoelectric transformers have been already applied for the measurements of various material constants. In [41] piezoelectric constants and an electromechanical coupling coefficient of ceramics operating in the range of high mechanical strains were measured.

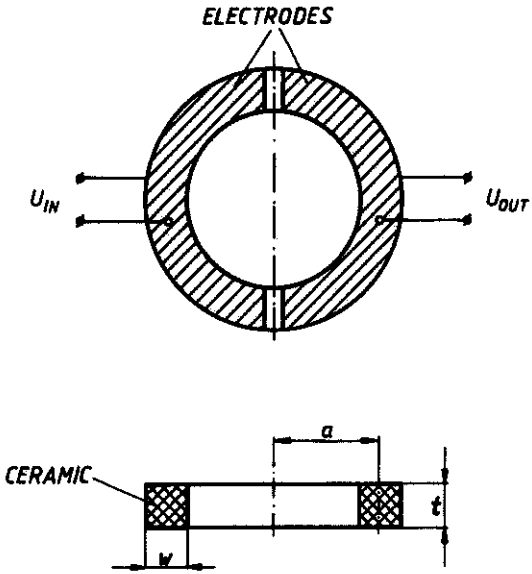


Fig. 1. Ring-shaped piezoelectric transformer.

Rosen's theoretical analysis [35] was applied. In [32] piezoelectric and elastic constants of piezoelectric polymers and composites were measured. In both cases electrical losses were not considered.

The equivalent circuit of the ring-shaped piezoelectric transformer can be derived if each half of the ring is replaced by the KLM equivalent circuit [21, 22] – Fig. 2. We assume that both halves of the ring are identic and we neglect the gaps between the electrodes. In the applied transformers the gap width was equal 1 mm – about 1% of the mean circumference of the ring. In a very strict analysis the gaps between electrodes may be taken into consideration [17] and inserted into the KLM circuit [22].

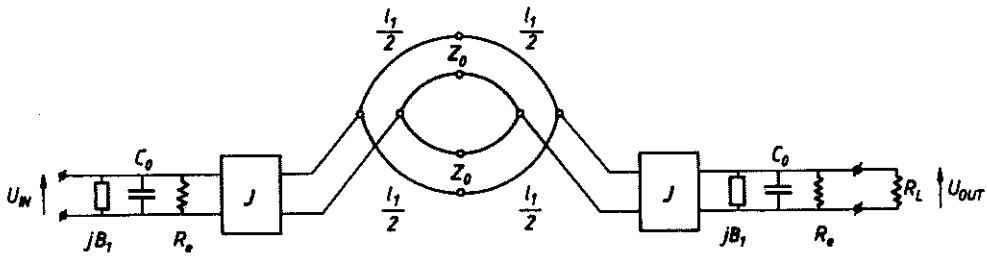


Fig. 2. KLM equivalent circuit of a ring-shaped piezoelectric transformer.

For the ring at the resonance frequency the values of the elements of the circuit presented in Fig. 2 are as follows:

$$l_1 = \frac{\lambda}{2} = \pi a, \tag{4.1}$$

$$B_1 = \frac{1}{Z_0} M^2 \sin\left(\frac{l_1\omega}{v}\right) = 0, \tag{4.2}$$

$$J = |B_2|, \tag{4.3}$$

$$B_2 = -\frac{\phi}{Z_0} \sin\left(\frac{l_1\omega}{2v}\right) = -\frac{\phi}{Z_0}, \tag{4.4}$$

$$\phi = \frac{\pi w d_{31}}{s_{11}^E}, \tag{4.5}$$

$$C_0 = \frac{\pi a w \varepsilon_{33}^T}{t}\left(1 - k_{31}^2\right), \tag{4.6}$$

$$Z_0 = \pi w t \sqrt{\frac{\varrho}{s_{11}^E}} \tag{4.7}$$

a, w, t – ring dimensions (cf. Fig. 1), R_L – load resistance, λ – wavelength, Z_0 – characteristic impedance, v – acoustic wave velocity, ω – angular frequency, J – inverter parameter, d_{31} – piezoelectric constant, s_{11}^E – elastic compliance, ε_{33}^T – permittivity, k_{31} – piezoelectric coupling coefficient, ϱ – density. The definitions of the material constant and the subscripts and superscripts are as in [3], B_1, B_2, J, M are defined as in [21, 22].

Electrical losses are represented by the resistance R_e [3], mechanical losses – by the complex propagation constant of the transmission line (central part of the KLM equivalent circuit – Fig. 2):

$$\gamma = \alpha + j\beta, \qquad \alpha = \frac{\pi}{\lambda Q_m}, \qquad \beta = \frac{2\pi}{\lambda}. \tag{4.8}$$

Certain simplifications are necessary for the derivation of the equations describing the voltage ratio in the piezoelectric transformer. We assumed that $\alpha l_1 = \pi/(2Q_m)$ was small (true for typical values of Q_m of PZT-type ceramics) and that the transformer was driven by the source of ac voltage of constant amplitude. Rosen [35] ascertained that such assumptions were admissible in the analysis of piezoelectric ceramic transformers. We assumed also, as we have already mentioned, that both halves of the transformer were identic and we neglected the gaps between the electrodes. The simplifications allow to determine the influence of physical properties of the ceramic on the operation of the transformer and to express the voltage ratio as a function of the values that can be measured.

Voltage ratios in piezoelectric transformers can be calculated applying relations known from the electrical circuit theory and the theory of transmission lines (see Appendix) to the KLM circuit presented in Fig. 2. Under above-mentioned assumptions we have obtained for two limits of the load resistance R_L:

for $R_L \to 0$

$$A_0 = \frac{U_{\mathrm{OUT}}}{U_{\mathrm{IN}}} = \frac{\phi^2 Q_m R_L}{\pi Z_0 + \phi^2 Q_m R_L}; \tag{4.9}$$

for $R_L \to \infty$

$$A_\infty = \frac{U_{\mathrm{OUT}}}{U_{\mathrm{IN}}} = \frac{\phi^2 Q_m X_e R_e}{\phi^2 Q_m X_e^2 + \pi Z_0 R_e}. \tag{4.10}$$

Thus

$$\operatorname{tg}\delta_m = \frac{\phi^2 R_L(1-A_0)}{\pi Z_0 A_0} \tag{4.11}$$

and

$$\begin{aligned} \operatorname{tg}\delta_e &= \left(1-k_{31}^2\right)\frac{\phi^2 Q_m X_e - A_\infty \pi Z_0}{\phi^2 Q_m X_e A_\infty}, \\ X_e &= \frac{1}{\omega C_0}. \end{aligned} \tag{4.12}$$

For the determination of the losses it is not sufficient to measure the voltage ratios because in the range of high electric fields the piezoelectric, dielectric and elastic constants change as a function of the field intensity. This effects are due to the domain structure of the ceramics. The domain walls motion changes values of the elastic compliance, the dielectric and piezoelectric constants and the electromechanical coupling coefficient [1, 19, 42].

5. Experimental results

Figure 3 presents the examples of the measured resonance curves of the ceramic rings, Fig. 3 a – input voltage $U_{\mathrm{IN}} = 1\,\mathrm{V}$, Fig. 3 b – $U_{\mathrm{IN}} = 40\,\mathrm{V}$. In Fig. 3 b nonlinear effects described in Sec. 3 (jump effects with hysteresis) are visible. The resonance curve in Fig. 3 a is already slightly asymmetric in spite of low input voltage. The ring was made of hard PZT-type ceramic. Figure 4 presents the output voltage of this transformer as a function of its input voltage. One can see that the output voltage does not increase proportionally to the increase of the input voltage even for small values of U_{IN}.

The values of the ceramic parameters necessary for the calculation of losses were measured. Figure 5 presents the examples of the measured values of s_{11}^E, k_{31}, d_{31} and ϵ_{33}^T as a function of the input electric field for the same ring as above.The changes of the ceramic parameters begin in the same range of the input electric field as the nonlinear effects visible in the $U_{\mathrm{OUT}} = f(U_{\mathrm{IN}})$ curve. These results are similar to the results presented earlier, e.g. in [37, 41]. The density ϱ of this ceramic was equal $7.5 \cdot 10^3\,\mathrm{kg/m^3}$.

Figure 6 presents the examples of the calculated (using the equations given in Sec. 4) dependences of the electrical and mechanical losses on the input electric field. Results presented in Fig. 6 a were obtained for the ceramic ring with the parameters presented in Fig. 5.

We measured also the losses in the ring-shaped transformers made of the ceramic of low quality factor (soft PZT-type ceramic). For these rings the discontinuities of the resonance curves did not appear in the whole measurement range. Therefore it was possible to measure the losses applying Hirose's method [14] described in Sec. 3 and to compare the results obtained using two methods. Figure 7 presents the quality factor Q_m as a function of input electric field for two kinds of ceramics: I – the ceramics of high quality factors (the results of their measurements are presented in Fig. 6), II – the ceramics of low quality factors used for the comparative measurements. The quality factors Q_m were determined as the reciprocals of $\operatorname{tg}\delta_m$ calculated using Eq. (4.11). The values of

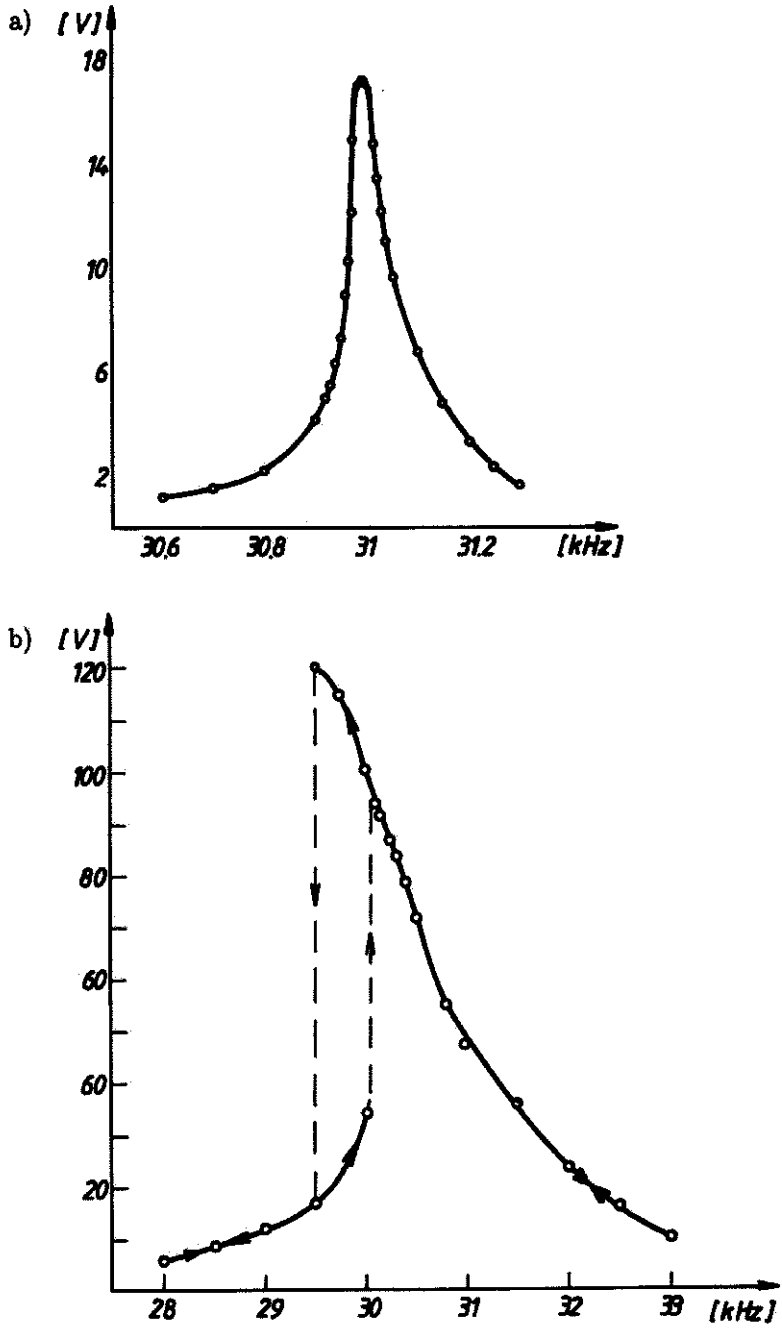


Fig. 3. Resonance curves of the ceramic ring transformer, external diameter $D_{ext} = 38$ mm, internal diameter $D_{int} = 30$ mm, thickness $t = 1$ mm, hard PZT-type ceramic, a) input voltage $U_{IN} = 1$ V, b) $U_{IN} = 40$ V. The arrows show the direction of frequency change during the measurements.

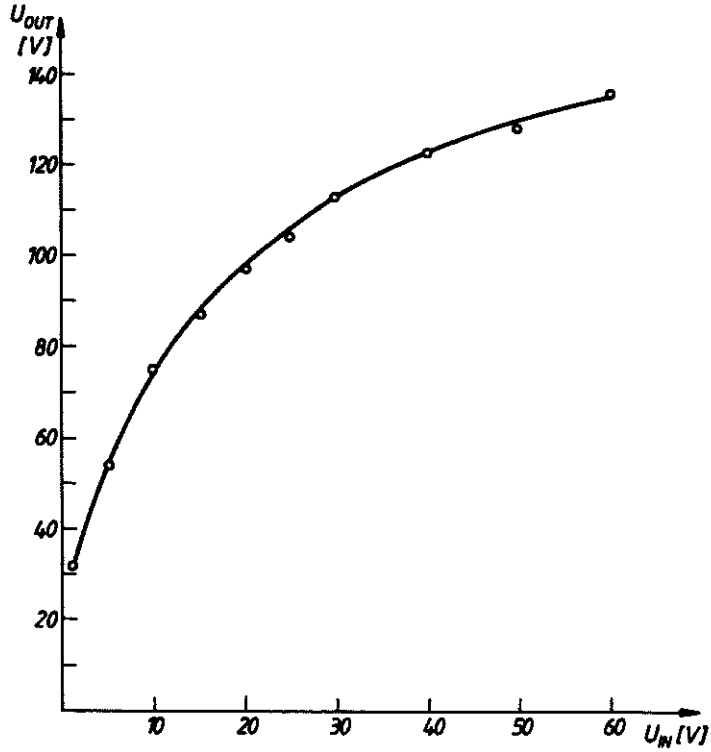
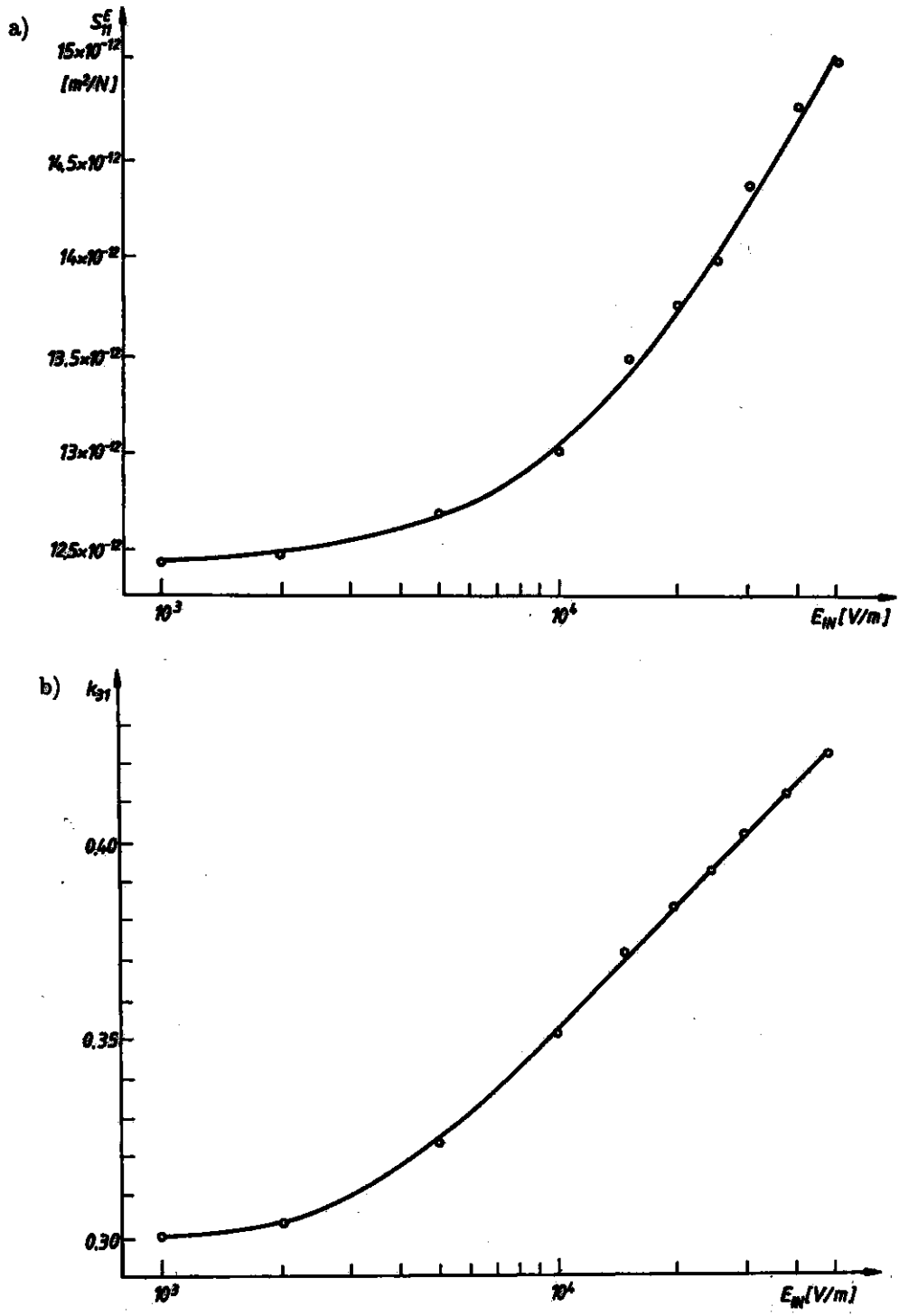


Fig. 4. The output voltage of the ring piezoelectric transformer versus its input voltage, dimensions of the ring as in Fig. 3.

the ceramic parameters were measured as previously. The density of the ceramics II was $\varrho = 7.4 \cdot 10^3\,\mathrm{kg/m^3}$. Figure 8 presents the values of losses determined applying two methods. One can see that the results agree qualitatively. The values of the losses measured using Hirose's method are lower. It may result from the fact that the quality factors Q_A and Q_B have been determined from the widths of resonance curves on the 3 dB-level instead of the perturbation method described in [33]. The curves were deformed, specially at higher input electric fields. The deformations may cause the apparent increase of the measured quality factors. For low input electric fields the differences between the obtained results are small.

During all measurements the rings were air-cooled using a propeller-fan. The losses cause the heating of the ceramic specially in the range of high input electric fields. The temperature increase causes the decrease of quality factors and changes of ceramic parameters [39]. The ageing and ferroelectric fatigue causes changes of the losses with time [26, 31]. The effect of heat generation impedes the measurements of the ceramic properties in high fields. It is possible to avoid this effect applying special, fully computerized, measuring circuits enabling to perform measurement in less than 500 ms [39].

The measurements were done in the frequency range 30 – 44 kHz according to the ring dimensions determining its resonance frequency.


a)
S_{11}^{ε}
15×10⁻¹²
[m²/N]
14,5×10⁻¹²
14×10⁻¹²
13,5×10⁻¹²
13×10⁻¹²
12,5×10⁻¹²
10³
10⁴
E_{IN}[V/m]
b)
k_{31}
0,40
0,35
0,30
10³
10⁴
E_{IN}[V/m]

[Fig. 5a, b]

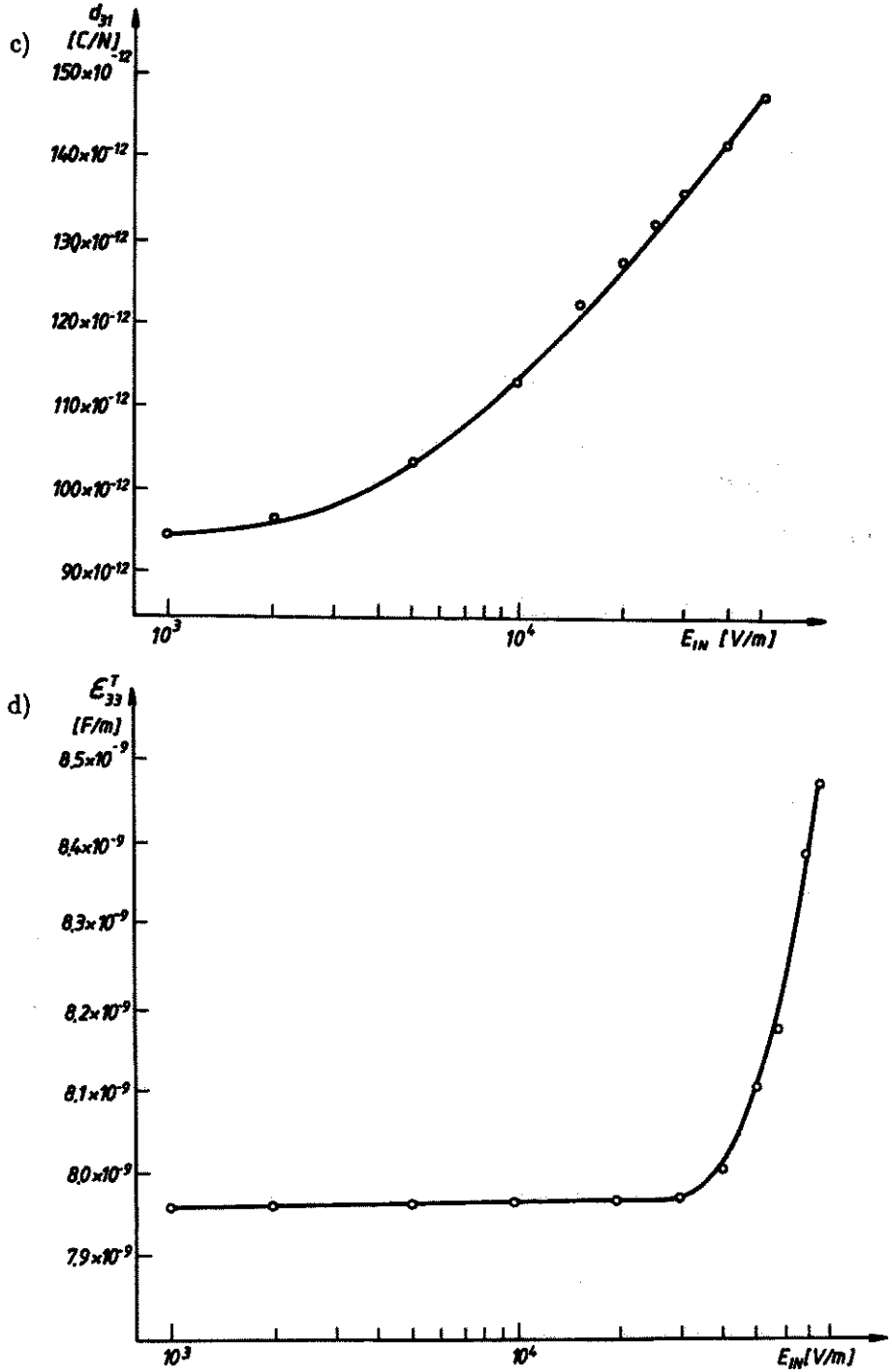


Fig. 5. Examples of measured dependences of the ceramic parameters (hard PZT-type ceramic) on the input electric field, the ceramic ring as in Figs. 3 and 4, a) elastic compliance s^E_{11}, b) electromechanical coupling coefficient k_{31}, c) piezoelectric constant d_{31}, d) permittivity.

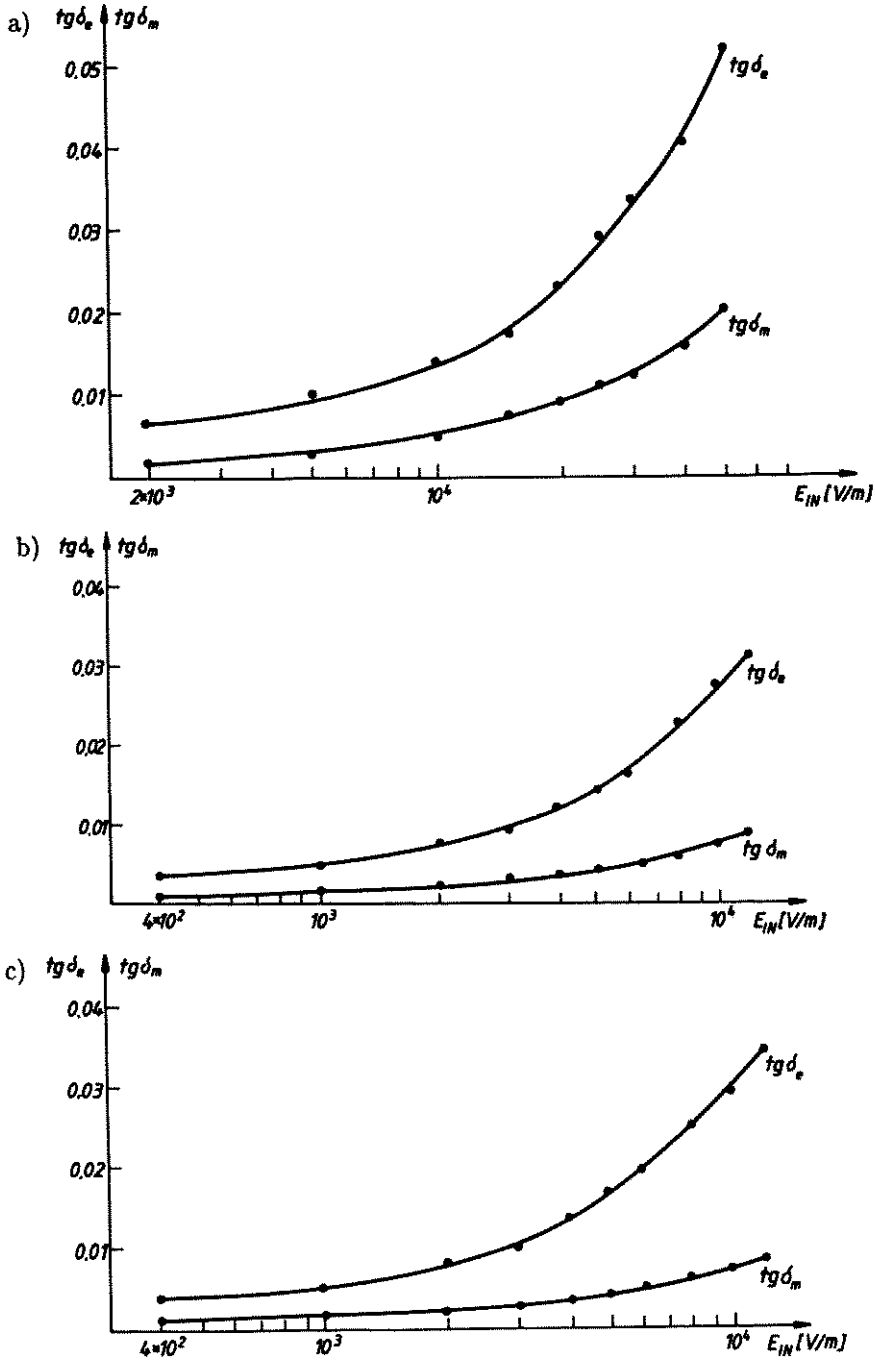


Fig. 6. Electrical ($\mathrm{tg}\,\delta_e$) and mechanical ($\mathrm{tg}\,\delta_m$) losses as a function of input electric field, calculated using Eqs. (4.11) and (4.12) for three piezoelectric ceramic transformers, a) $D_{ext} = 38\,\mathrm{mm}$, $D_{int} = 30\,\mathrm{mm}$, $t = 1\,\mathrm{mm}$, hard PZT-type ceramic, as in Figs. 3–5, b) $D_{ext} = 38\,\mathrm{mm}$, $D_{int} = 28\,\mathrm{mm}$, $t = 5\,\mathrm{mm}$, hard PZT-type ceramic, c) $D_{ext} = 38\,\mathrm{mm}$, $D_{int} = 24\,\mathrm{mm}$, $t = 5\,\mathrm{mm}$, hard PZT-type ceramic.

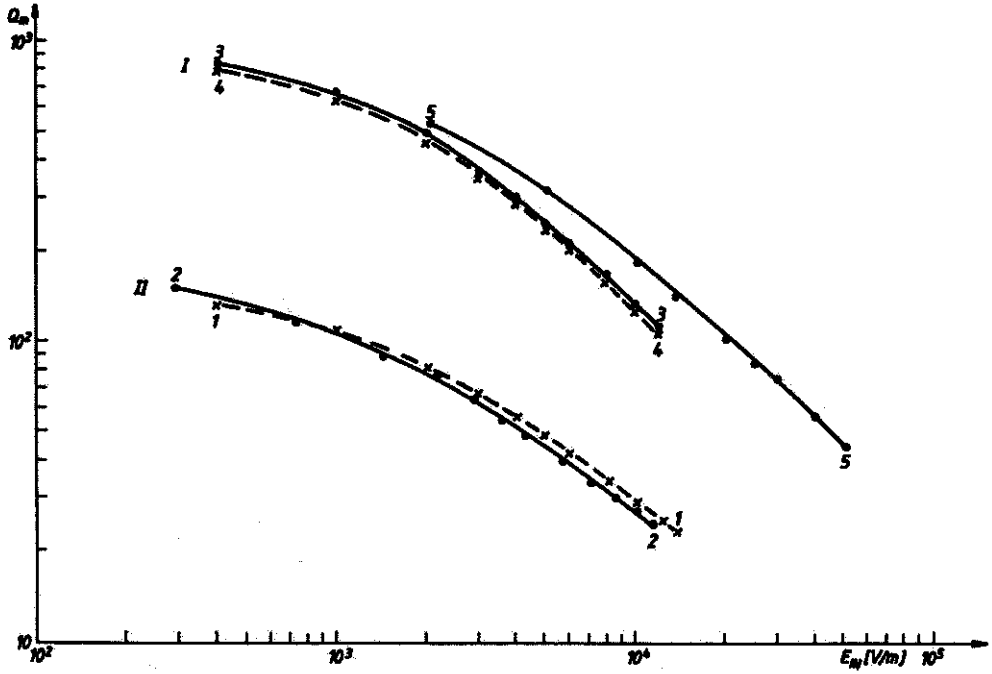
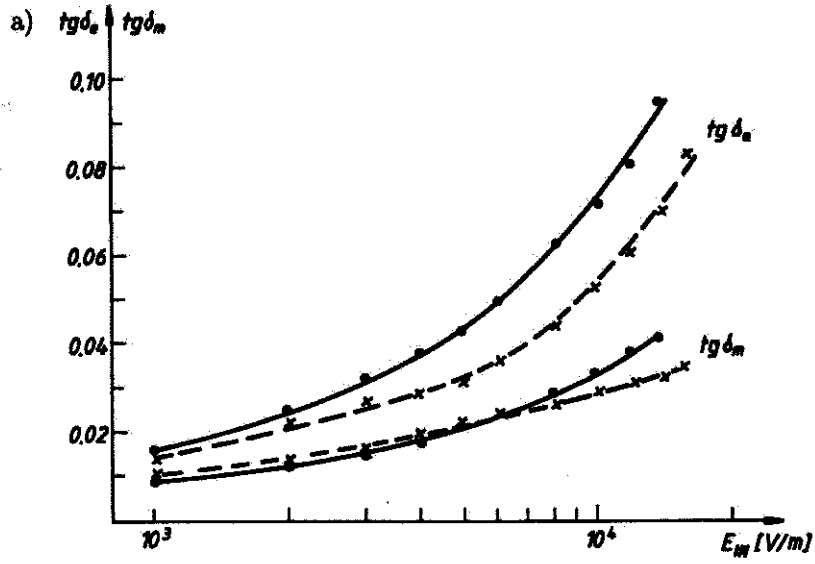


Fig. 7. Mechanical quality factors Q_m of several investigated ceramics as a function of input electric field, *1* – $D_{ext} = 30$ mm, $D_{int} = 16$ mm, $t = 5$ mm, soft PZT-type ceramic, *2* – $D_{ext} = 30$ mm, $D_{int} = 16$ mm, $t = 7$ mm, soft PZT-type ceramic, *3* – ceramic ring as in Fig. 6 c, *4* – ceramic ring as in Fig. 6 b, *5* – ceramic ring as in Fig. 6 a.



[Fig. 8a]

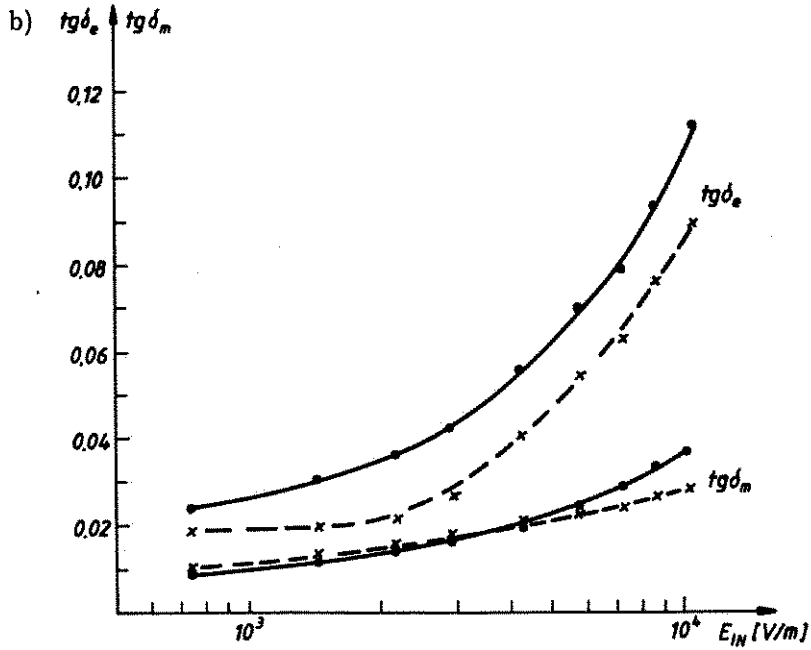


Fig. 8. Comparison of the results obtained for two piezoelectric ceramic transformers applying two methods: solid line – piezoelectric transformer method presented in the paper, dashed line – Hirose's method [14] described in Sec. 3, a) ceramic ring no. 1 in Fig. 7, b) ceramic ring no. 2 in Fig. 7.

6. Conclusions

The results obtained applying the method proposed by the author are qualitatively consistent with the results obtained using the different methods and with the theoretical descriptions of the loss effects. The method presented in the paper does not require to measure the quality factor. The measurements of the quality factor are difficult or even impossible for the ceramics driven by high fields. In above described method the losses are determined from the measurements of the voltage ratios for two kinds of the loading of the piezoelectric transformer. However, the additional measurements of the material constants as a function of the input electric field are necessary. The described method has also certain disadvantage i.e., the sample in the form of the ring transformer is necessary for the loss measurements.

The equations determining the losses have been derived using the simplificating assumptions similarly as in all cited methods of loss analysis. Otherwise the derived equations could not be applied in practice. The simplifications are necessary to obtain the relations permitting to determine the influence of physical parameters of investigated ceramic on the transformer operation and to express the voltage ratios as a function of the mesurable quantities. Further experiments with transformers made of various ceramics with different properties are necessary to estimate quantitatively the accuracy of the described method.

Appendix. Application of KLM equivalent circuit for the calculation of losses

The KLM equivalent circuit is presented in Fig. 2. We calculate the ratios of the output voltage U_{OUT} and the input voltage U_{IN} for two limits of the transformer loading: $R_L \to 0$ and $R_L \to \infty$.

The output current of the inverter J:

$$I = -jB_2U, \tag{A.1}$$

where: U – output voltage of the inverter, B_2 – inverter parameter. The ideal admittance inverter J has also the property that when an admittance Y is connected to one port, the input admittance at the other port is J^2/Y, where $J = |B_2|$ [22].

The voltage at the output of the transmission line:

$$U_2 = \frac{I_1}{\frac{1}{Z_2}\mathrm{ch}\gamma x + \frac{1}{Z_0}\mathrm{sh}\gamma x} \tag{A.2}$$

and

$$U_2 = \frac{U_1}{\mathrm{ch}\gamma x + \frac{Z_0}{Z_2}\mathrm{sh}\gamma x}, \tag{A.3}$$

where: I_1, U_1 – current and voltage at the input of the transmission line, Z_2 – loading impedance at the end of the transmission line, Z_0 – characteristic impedance of the transmission line, $\gamma = \alpha + j\beta$ – propagation constant of the transmission line described by Eq. (4.8), x – length of the transmission line.

The input impedance of the transmission line:

$$Z_{\mathrm{IN}} = \frac{Z_2 + Z_0\mathrm{th}\,\gamma x}{1 + \frac{Z_2}{Z_0}\mathrm{th}\,\gamma x}. \tag{A.4}$$

If αx is small (cf. Sec. 4) then:

$$U_2 = \frac{I_1}{\frac{1}{Z_2}\left(\cos\beta x + j\alpha x \sin\beta x\right) + \frac{1}{Z_0}\left(\alpha x\cos\beta x + j\sin\beta x\right)}. \tag{A.5}$$

In our case Eqs.(A.2) – (A.5) can be considerably simplified because $l_1 = \lambda/2$ in the equivalent circuit in Fig. 2 and the characteristic impedances are identic in all sections of the transmission line. The values of the elements of the equivalent circuit are described by Eqs. (4.1)–(4.7).

Similarly as Rosen we assume that the transformer is driven by the source of ac voltage of constant amplitude. This simplificating assumption allows to neglect the influence of the internal resistance of the voltage source [35].

Using standard relations for the electric circuits and transmission lines [2, 35] we obtain for $R_L \to \infty$:

$$A_\infty = \frac{U_{\mathrm{OUT}}}{U_{\mathrm{IN}}} = \frac{\phi^2 Q_m \left(R_e^2 + X_e^2\right)^{1/2} X_e R_e}{\pi Z_0 \left(R_e^2 + X_e^2\right) + \phi^2 Q_m R_e X_e^2}. \tag{A.6}$$

Since $R_e^2 \gg X_e^2$ then Eq. (A.6) can be simplified:

$$A_\infty = \frac{\phi^2 Q_m X_e R_e}{\pi Z_0 R_e + \phi^2 Q_m X_e^2} \tag{A.7}$$

then

$$R_e = \frac{\phi^2 Q_m X_e^2 A_\infty}{\phi^2 Q_m X_e - A_\infty \pi Z_0} \tag{A.8}$$

as [3]:

$$\operatorname{tg} \delta_e = \frac{(1 - k_{31}^2) X_e}{R_e} \tag{A.9}$$

therefore:

$$\operatorname{tg} \delta_e = \frac{(1 - k_{31}^2)(\phi^2 Q_m X_e - \pi A_\infty Z_0)}{\phi^2 Q_m X_e A_\infty} \tag{A.10}$$

similarly for $R_L \to 0$ we obtain:

$$A_0 = \frac{U_{\mathrm{OUT}}}{U_{\mathrm{IN}}} = \frac{\phi^2 Q_m R_L}{\pi Z_0 + \phi^2 Q_m R_L} \tag{A.11}$$

then Q_m necessary to calculate (A.10) is:

$$Q_m = \frac{\pi Z_0 A_0}{\phi^2 R_L (1 - A_0)} \tag{A.12}$$

and

$$\operatorname{tg} \delta_m = \frac{1}{Q_m} = \frac{\phi^2 R_L (1 - A_0)}{\pi Z_0 A_0}. \tag{A.13}$$

The measurements were done for two values of the transformer loading: a) corresponding to $R_L \to 0$, $R_L = 11\,\Omega \ll R_e$ and b) corresponding to $R_L \to \infty$, in this case the transformer output was loaded by the internal resistance of the voltmeter $R_i \gg R_e$.

The same results of the calculations may be obtained using Mason's equivalent circuit, but the calculations are more time-consuming.

References

[1] G. ARLT, *Domain contributions to piezoelectricity in ceramics*, 1990 IEEE Ultrasonics Symposium Proceedings, 733–742.

[2] G.I. ATABIEKOW, *Theory of linear electric circuits* [in Russian], Sowietskoje Radio, Moskwa 1960.

[3] D.A. BERLINCOURT, D.R. CURRAN and H. JAFFE, *Piezoelectric and piezomagnetic materials*, Physical Acoustics, Vol. IA, W.P. MASON [Ed.], Academic Press, New York 1964, Ch. 3.

[4] R. BRIOT, P. GONNARD and M. TROCCAZ, *Comparison of PZT ceramics based on the frequency dependence of conductance and a grain boundary-domain walls model*, Ferroelectrics, **127**, 155–160 (1992).

[5] J.L. BUTLER, K.D. ROLT and F.A. TITO, *Piezoelectric ceramic mechanical and electrical stress study*, J. Acoust. Soc. Am., **96**, 3, 1914–1917 (1994).

[6] P. CHAMP and P. GONNARD, *Characterization of high power piezoelectric ceramics under high mechanical driving levels*, 1er Congres Français d'Acoustique 1990, 621–624.

[7] A. CHEŁKOWSKI, *Physics of dielectrics* [in Polish], PWN, Warszawa 1993, Ch. 5.

[8] D. DAMJANOVIC, M. DEMARTIN, H.S. SCHULMAN, M. TESTORF and N. SETTER, *Instabilities in the piezoelectric properties of ferroelectric ceramics*, Sensors and Actuators, **A53**, 353–360 (1996).

[9] L. EYRAUD, P. GONNARD and M. TROCCAZ, *Piezoceramics and their characterization for high power ultrasound*, Ultrasonics International 1989 Conference Proceedings, 295–300.

[10] J.J. GAGNEPAIN and R. BESSON, *Nonlinear effects in piezoelectric quartz crystals*, Physical Acoustics, Vol. XI, W.P. MASON and R.N. THURSTON [Eds.], Academic Press, New York 1975, Ch. 5.

[11] J.O. GENTNER, P. GERTHSEN, N.A. SCHMIDT and R.E. SEND, *Dielectric losses in ferroelectric ceramics produced by domain-wall motion*, J. Appl. Phys., **49**, 8, 4485–4489 (1978).

[12] P. GERTHSEN, K.H. HÄRDTL and N.A. SCHMIDT, *Correlation of mechanical and electrical losses in ferroelectric ceramics*, J. Appl. Phys., **51**, 2, 1131–1134 (1980).

[13] D. GUYOMAR, N. AURELLE and L. EYRAUD, *Piezoelectric ceramics nonlinear behavior. Application to Langevin transducer*, J. Phys. III, **7**, 6, 1197–1208 (1997).

[14] S. HIROSE, M. AOYAGI and Y. TOMIKAWA, *Dielectric loss in a piezoelectric ceramic transducer under high power generation; increase of dielectric loss and its influence on transducer efficiency*, Jpn. J. Appl. Phys., Part 1, **32**, 5B, 2418–2421 (1993).

[15] S. HIROSE, *New method for measuring mechanical vibration loss and dielectric loss of piezoelectric transducer under high-power excitation*, Jpn. J. Appl. Phys., Part 1, **33**, 5B, 2945–2948 (1994).

[16] S. HIROSE, M. AOYAGI, Y. TOMIKAWA, S. TAKAHASHI and K. UCHINO, *High power characteristics at antiresonance frequency of piezoelectric transducers*, Ultrasonics, **34**, 2–5, 213–217 (1996).

[17] R. HOLLAND, *The equivalent circuit of an N-electrode piezoelectric bar*, Proc. IEEE, **54**, 7, 968–975 (1966).

[18] T. IKEDA, *Fundamentals of piezoelectricity*, Oxford University Press, 1990, Ch. 6.

[19] B. JAFFE, W.R. COOK and H. JAFFE, *Piezoelectric ceramics*, Academic Press, New York 1971, Ch. 7.

[20] U. KIEFER, TH. FINKLER and K.-D. BECKER, *Piezoelektrische Ultraschallwandler modelliert. Simulation mechanischer und elektrischer Verluste*, Materialprüfung, **36**, 5, 192–195 (1994).

[21] R. KRIMHOLTZ, D.A. LEEDOM and G.L. MATTHAEI, *New equivalent circuits for elementary piezoelectric transducers*, Electron. Lett., **6**, 13, 398–399 (1970).

[22] D.A. LEEDOM, R. KRIMHOLTZ and G.L. MATTHAEI, *Equivalent circuits for transducers having arbitrary even- or odd-symmetry piezoelectric excitation*, IEEE Trans., **SU-18**, 3, 128–141 (1971).

[23] M. LETHIECQ, L.P. TRAN-HUU-HUE, F. PATAT and L. POURCELOT, *Measurement of losses in five piezoelectric ceramics*, IEEE Trans., **UFFC-40**, 3, 232–236 (1993).

[24] W.W. ŁAWRINIENKO, *Piezoelectric transformers* [in Russian], Energia, Moskwa 1975.

[25] I. MALECKI, *Physical foundations of technical acoustics*, Pergamon Press, Oxford 1969, Ch. 14.

[26] W.P. MASON, *Aging of the properties of barium titanate and related ferroelectric ceramics*, J. Acoust. Soc. Amer., **27**, 1, 73–85 (1955).

[27] F.R. MONTERO DE ESPINOZA, J.L. SAN EMETERIO and P.T. SANZ, *Summary of the measurements methods of Q_m for piezoelectric materials*, Ferroelectrics, **128**, 1–4, 61–66 (1992).

[28] K. NAGATA, J. THONGMENG and K. KATO, *Evaluation of the reliability of piezoelectric ceramic transformers*, Jpn. J. Appl. Phys., Part 1, **36**, 9B, 6103–6105 (1997).

[29] K. NEGISHI, *Jump phenomenon in resonance curve of ferroelectric ceramics*, J. Phys. Soc. Jpn., **15**, 3, 534 (1960).

[30] W. PAJEWSKI and M. SZALEWSKI, *Methods of the measurements of the quality factor of piezoelectric resonators with high electrical and mechanical losses*, Arch. Acoust., **20**, 4, 373–385 (1995).

[31] W. PAN, S. SUN and P. FUIERER, *Effects of ferroelectric switching on the dielectric and ferroelectric properties in lead zirconate titanate ceramics and their modelling*, J. Appl. Phys., **74**, 2, 1256–1264 (1993).

[32] D.L. RASTORGUJEW, *Measurements of electromechanical characteristics of piezoelectric materials using piezoelectric transformer* [in Russian], Pribory i Technika Eksperimenta, 1, 155–158 (1997).

[33] S. SAITO, R. AOYAGI and H. SHIMIZU, *A method for automatic measurement of ultrasonic transducer constants*, J. Acoust. Soc. Jpn., **33**, 10, 540–548 (1977).

[34] S. SAITO and H. SHIMIZU, *Influence of elastic loss on measured dielectric-loss-angle of high-coupling piezoelectric material*, Jpn. J. Appl. Phys., **24**, suppl. 3, 145–147 (1985).

[35] *Solid state magnetic and dielectric devices*, H.W. KATZ [Ed.], John Wiley and Sons, New York 1959, ch. 5.

[36] J.J. STOKER, *Nonlinear vibrations in mechanical and electrical systems*, Interscience Publishers, New York 1950, ch. 2 and 4.

[37] S. TAKAHASHI, S. HIROSE, K. UCHINO and K.Y. OH, *Electromechanical characteristics of lead-zirconate-titanate ceramics under vibration-level change*, Proc.9th IEEE Int. Symp. Appl. Ferroel., 377–382 (1994).

[38] S. TAKAHASHI, Y. SASAKI, S. HIROSE and K. UCHINO, *Electromechanical properties of* $PbZrO_3$-$PbTiO_3$-$Pb(Mn_{1/3}Sb_{2/3})O_3$ *ceramics under vibration level change*, Mat. Res. Soc. Symp. Proc., **360**, 305–310 (1995).

[39] S. TASHIRO, M. IKEHIRO and H. IGARASHI, *Influence of temperature rise and vibration level on electromechanical properties of high-power piezoelectric ceramics*, Jpn. J. Appl. Phys., Part 1, **36**, 5B, 3004–3009 (1997).

[40] K. UCHINO, *Materials issues in design and performance of piezoelectric actuators: an overview*, Acta Mater., **46**, 11, 3745–3753 (1998).

[41] VO DUY DAN, *Investigation of piezoelectric ceramic in the range of high mechanical strains* [in Polish], Doctor Thesis, IFTR, Warsaw 1985.

[42] R.S. WOOLLETT and C.L. LE BLANC, *Ferroelectric nonlinearities in transducer ceramics*, IEEE Trans., **SU-20**, 1, 24–31 (1973).

[43] X.D. ZHANG and C.A. ROGERS, *A macroscopic phenonenological formulation for coupled electromechanical effects in piezoelectricity*, J. Intell. Mater. Syst. Struct., **4**, 3, 307–316 (1993).

ARCHIVES OF ACOUSTICS
24, 1, 75–84 (1999)

AN EXPERIMENTAL INVESTIGATION OF THE FINITE AMPLITUDE WAVE

E. KOZACZKA and G. GRELOWSKA

Naval Academy
(81-919 Gdynia, ul. Śmidowicza 71, Poland)

An experimental investigation of the pressure field distribution produced by a plane circular piston in water was carried out by means of two receivers of different size. A transducer of 46-mm-diam and 1.0-MHz center frequency was used as a transmitter. The PVdF needle hydrophone of 1 mm diameter was the first receiver, whereas the second one had a diameter equal to that of the source. All experiments were performed using a high precision computer-controlled tank facility. Measurements of harmonic generation are compared with numerical calculations based on the nonlinear parabolic wave equation.

1. Introduction

Recently there has been a steady growth of interest in the nonlinear problem of finite aperture sound beams in number of disciplines such as acoustic microscopy [10], ultrasound therapy and diagnostics [6, 9], parametric acoustic arrays [8], and the measurements of the nonlinear parameter B/A with finite amplitude effects of sound waves [11, 14].

In many applications the most important are phenomena occurring in the nearfield area. It is caused by the fact that due to the applied measurement set up arrangement and the chosen parameters of used transducers the measured or diagnosed object is often situated within the nearfield area.

Many papers have been published devoted to the problem of the nonlinear distortion of the finite amplitude wave in the nearfield of the source. The fine structure of the nearfield presents difficulties to both: the theoretical and the experimental investigations. One of the first papers devoted to the theoretical description of the nearfield of the finite amplitude source was published in 1971 by INGENITO and WILLIAMS JR. [12]. There the second harmonic field of a piston transducer was calculated by means of the perturbation method. A significant step was made by Norwegian scientist, who solved the KZK equation numerically, using the finite difference scheme [1], which is known as the Bergen code. It accounts for the nonlinearity, absorption and diffraction. The solution is widely used in comparing the measurement results [3, 17] and may be modified according to the measurements conditions, for instance for focused circular sources [4] and for rectangular sources [5, 16].

The paper presents the results of the experimental investigations of the finite amplitude wave radiated by the circular piston. The investigations were carried out close to the transmitter in the nearfield and nearfield-farfield transition area using the high precision positioning device which controlled the movement of the receiver. The distance of occurring the last maximum in the pressure distribution on the beam axis is assumed to be the boundary of the nearfield.

2. Experimental methods

A set up shown in Fig. 1 was used for experimental measurements. The transmitter – a piston of 46-mm-diam – was mounted at one end of the water tank 1.4 m long by 1.2 m wide and 1.2 m deep. The transducer was driven at its center frequency of 1.0-MHz corresponding to $ka = 96$ (k is a wave number, a is the transmitter radius). It was driven with a tone pulse, approximately 50 cycles long and the pulse repetition was of about 8 msec. This gave a quasi-continuous wave in the measuring area field without standing waves. Reflections were eliminated by using a time gate on the receiver pulse.

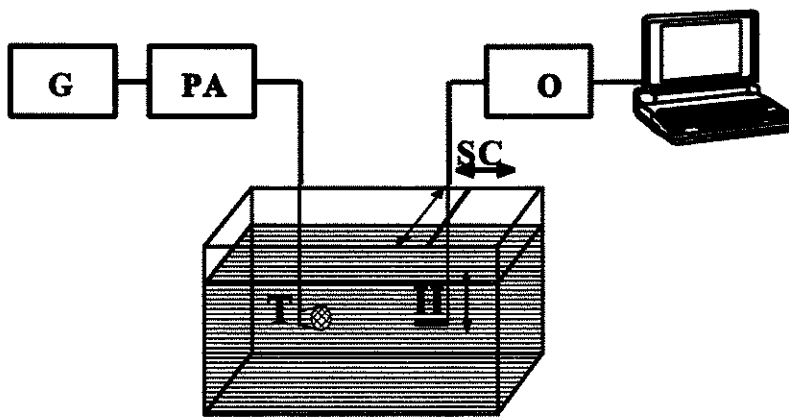


Fig. 1. The measurement set up: G – generator, PA – power amplifier, T – transmitter, H – PVdF needle hydrophone, SC – scanning set up, O - digital storage oscilloscope HP 54503A.

The pressure generated by the transducer was measured using a 1-mm-diam PVdF needle hydrophone. The hydrophone was mounted on a three-dimensional translation stage that allowed to position the hydrophone anywhere in a plane perpendicular to the acoustic axis of the transducer or along this axis. The movement resolution is theoretically equal to 0.0125 mm.

The output from the hydrophone was fed directly into a digital storage oscilloscope (HP 54503A). The oscilloscope was used to capture a middle part of the tone burst. The time waveform was then transferred to the controlling computer and five (or two) cycles underwent the Fourier analysis to extract the harmonic amplitudes. The measured waveforms were averaged before analyzing.

The second receiver used in measurements was of an area equal to that of the transmitter, but it was sufficiently broadband to receive the harmonics appearing as an effect of nonlinear propagation in water. The measuring method of nonlinear distortion in which two coaxial transducers of equal area are used was described by COBB [7].

3. Experimental results

Results of the experiments allow to make a thorough study of the nonlinear distortion growth in the area close to the transmitter. The pressure distribution of the examined plane circular piston transmitter of 46-mm-diam at a distance of 1 mm from the surface is shown in Fig. 2. The amplitude of the pressure p_0 equals to about 157 kPa. In the numerical model that distribution is approximated by the curve shown in Fig. 3.

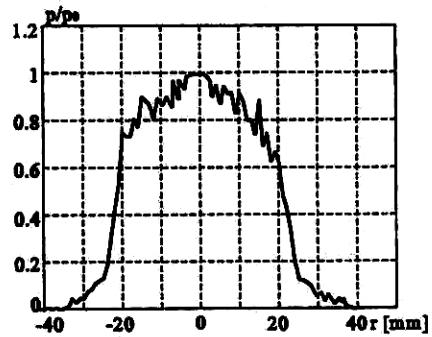


Fig. 2. The pressure distribution measured at a distance of 1 mm from the surface of the examined plane circular piston transmitter.

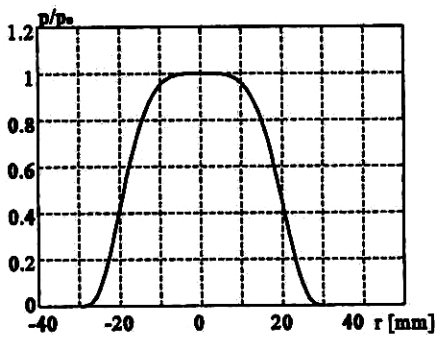


Fig. 3. The curve approximating the measured pressure distribution at the radiating surface.

The growth of the nonlinear distortion can be noticed by observing the changes in the shape of the wave and in the rising of the amplitude of the second harmonic with the increasing distance from the source. Figure 4 shows the shape of the wave measured using 1-mm needle hydrophone at different distances from the source and Fig. 5 shows the shape of the wave measured in the axis of the source at the distance of 70 mm and

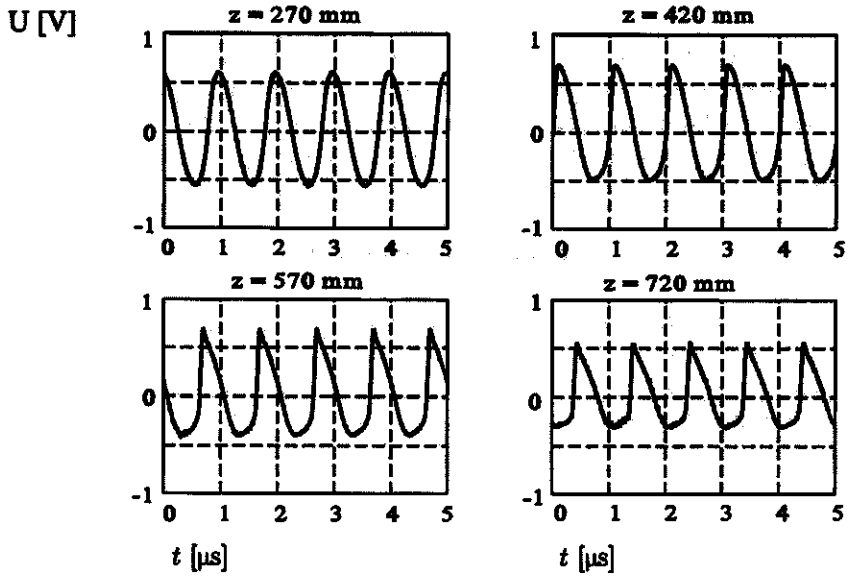


Fig. 4. The changes in the shape of the wave measured on the beam axis with the increasing distance from the source.

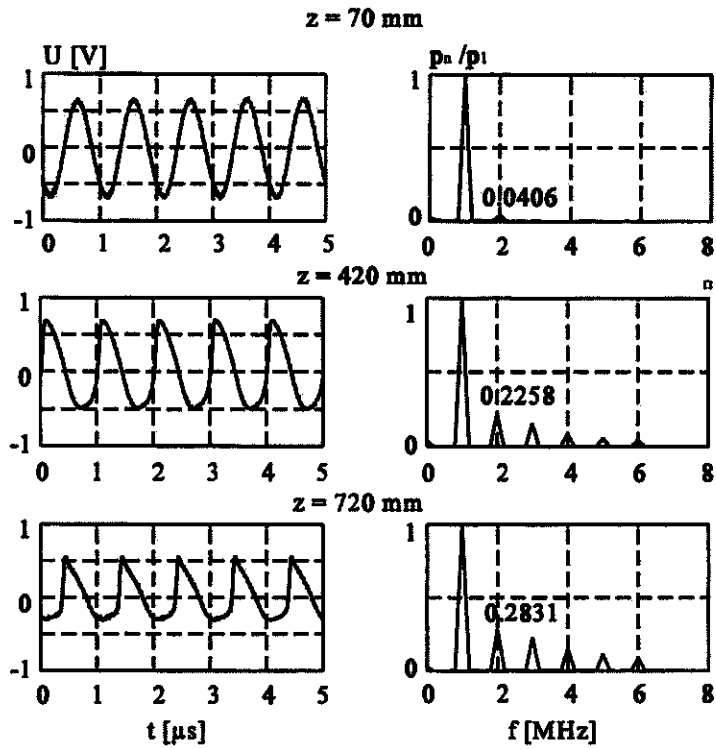


Fig. 5. The shape of the wave measured in the axis of the source at the distance of 70 mm 420 mm and 720 mm and their respective spectra.

720 mm and their respective spectra. In the plane distant 720 mm from the source the distortion is quite great. The amplitude of the second harmonic equals to about 28% of the amplitude of the first harmonic.

The results of measurements allow to obtain the graph (Fig. 6) in which the changes in the amplitude of the pressure harmonics as a function of the distance from the source are presented. The results of experiment (solid line) are shown together with the results of numerical calculation (dashed line). The comparison could be done from the distance come from restriction in validity of the applied numerical model.

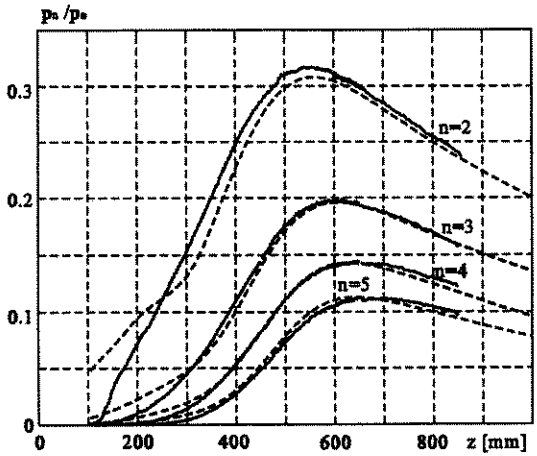


Fig. 6. The changes in the amplitude of the pressure harmonic components on the beam axis as a function of the distance from the plane circular source: (—) obtained experimentally, (- - -) predicted theoretically.

We can also illustrate the nonlinear wave distortion in a form of amplitude pressure distribution in chosen plane. In the next figures are shown distributions of the relative amplitude of the second and higher harmonic components measured in planes placed at the different distance from the source. The growth of the nonlinear distortion can be noticed by observing the rising in the amplitude of the second harmonic component with the distance from the source. It is exemplified in Fig. 7 which shows the growth of the nonlinear distortion expressed in the rising in the relative amplitude of the second harmonic pressure in planes distant 120 mm, 270 mm and 370 mm from the transmitter. In the last of these planes (370 mm from the source) the distortion is so great that higher harmonic components become measurable. In the Fig. 8 are shown the first four pressure harmonics measured in this plane. The distortion rises with the growing of the distance from the source. It is confirmed by the distributions of the relative amplitude of the higher harmonics in the plane distant 720 mm from the source shown in Fig. 9.

Apart from investigations carried out using the needle hydrophone they were made using the receiver of the area as the transmitter. The results of these measurements were compared. The next figures show the results of measurements obtained in the same conditions of the experiment using both of hydrophones at the same distances from the source.

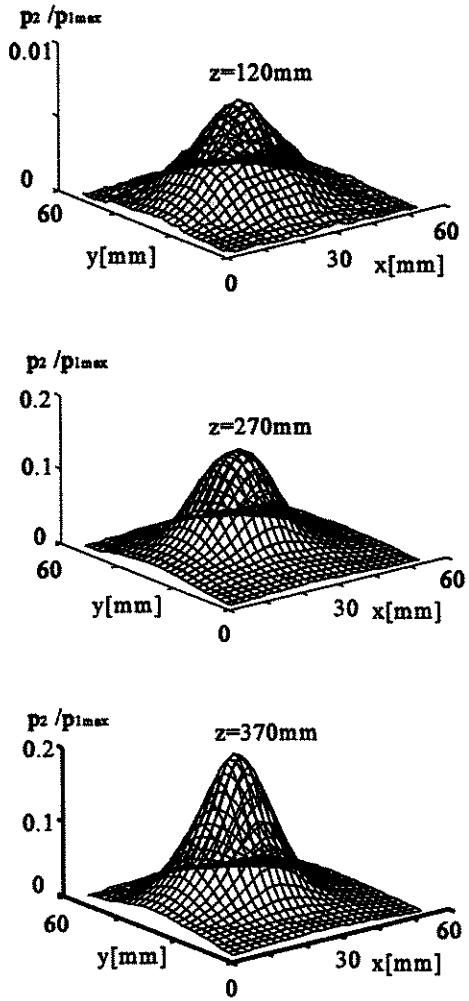


Fig. 7. Relative amplitude of the second pressure harmonic component measured in planes distant 120 mm, 270 mm and 370 mm from the plane circular source.

In Fig. 10 are presented the shape and the spectrum of the wave measured applying the receiver of the same area as the transmitter at two different distances between them. In the successive figure are shown the corresponding shapes and spectra of the wave measured by means of the PVdF hydrophone in the axis of the source (Fig. 11). We can notice that the wave distortion in the axis is more substantial then the one of the wave averaged on the surface of the larger receiver. It is caused by nonuniform distribution of the primary wave pressure across the beam which influences the nonuniform distribution of the higher harmonics. The pressure distribution of the first and second harmonic components at the planes at the same distances as the previous ones can be seen in Fig. 12.

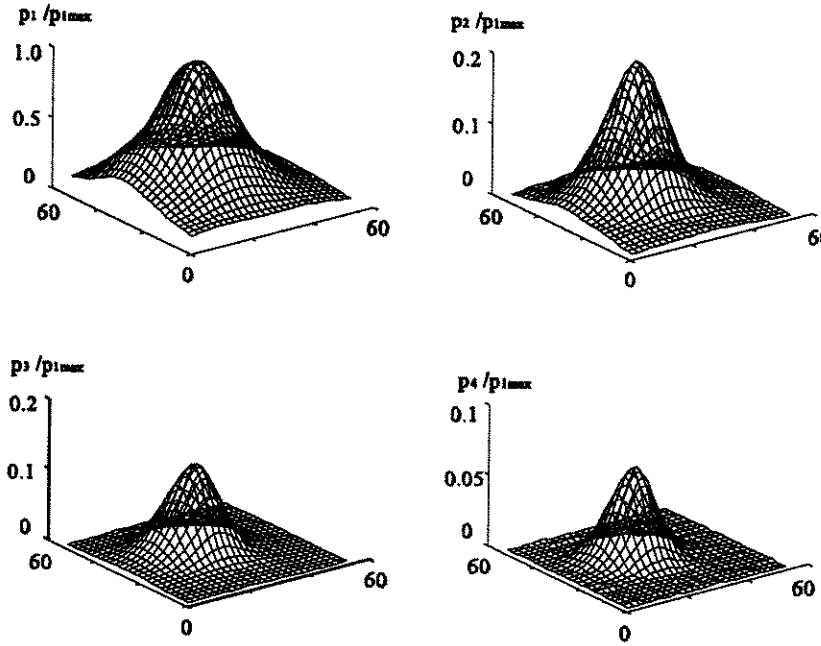


Fig. 8. Relative amplitude of the four first pressure harmonic components measured in the plane placed at the distance of 370 mm from the plane circular source.

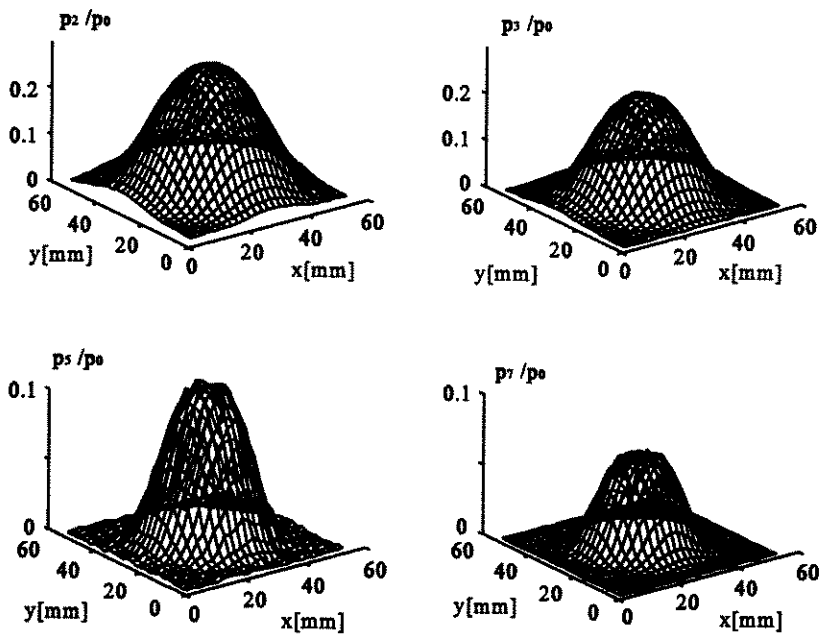


Fig. 9. Relative amplitude of the second, the third, the fifth and the seventh pressure harmonic components measured in the plane distant 720 mm from the source.

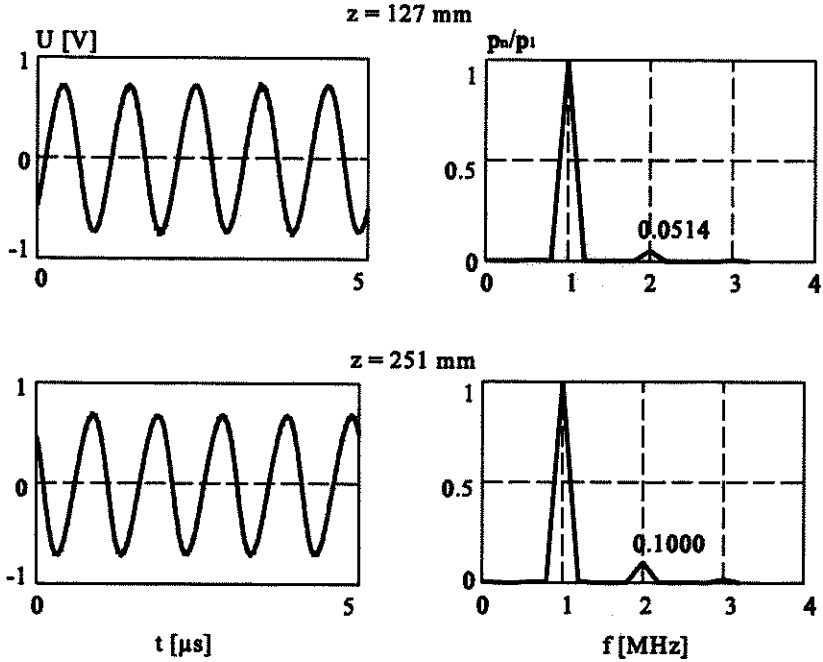


Fig. 10. The shape of the wave measured using the receiver of that area as the transmitter at the different distances and their respective spectra.

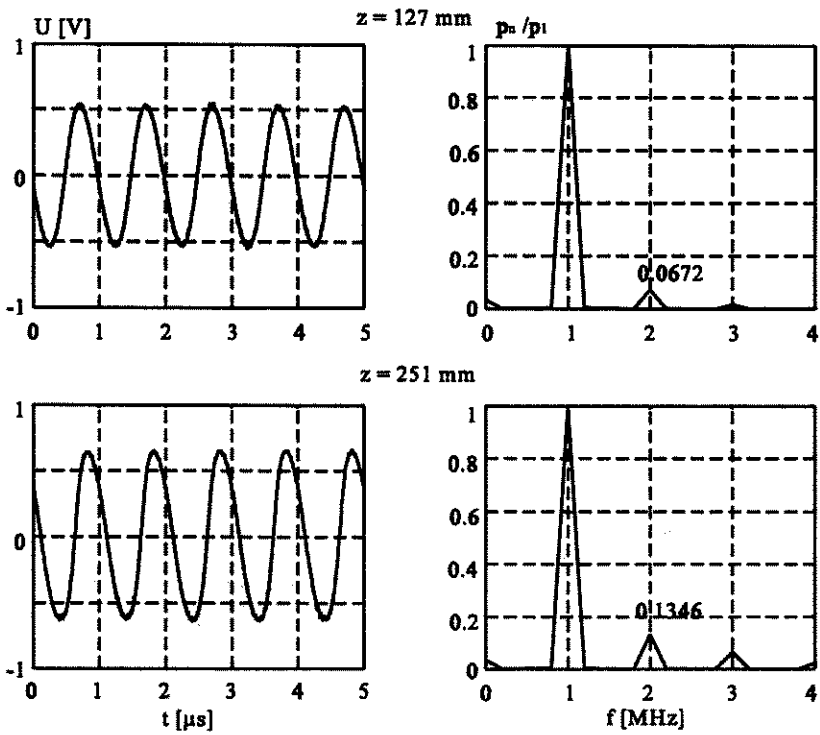


Fig. 11. The shape of the wave in the beam axis measured using the needle hydrophone at the different distances from the source and their respective spectra.

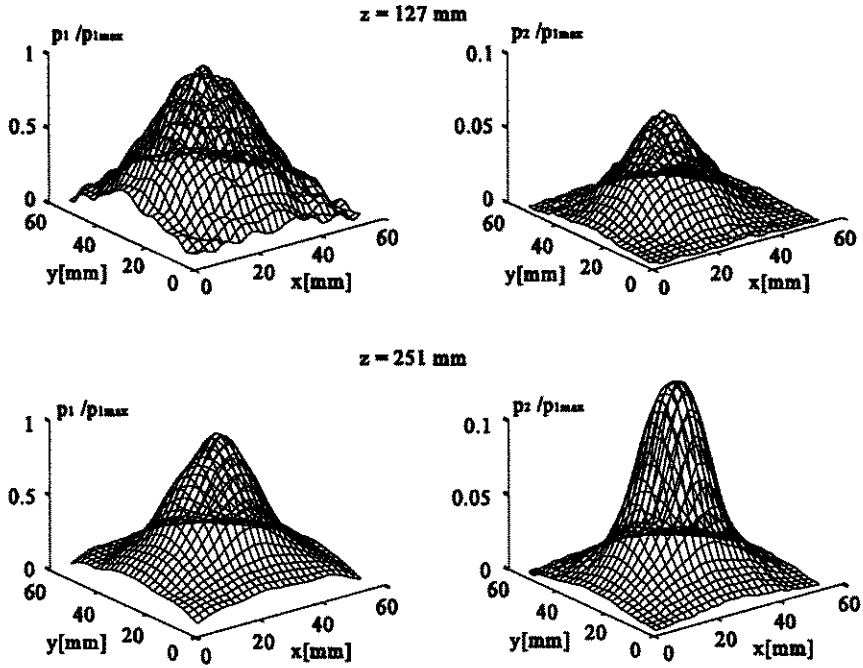


Fig. 12. The pressure distributions of the first and the second harmonics measured in the planes distant 127 mm and 251 mm from the source.

4. Conclusions

The paper presents the results of the experimental investigations of the finite amplitude wave field distribution. The measurement results obtained using the high precision facility which controlled the movement of the receiver were presented as well. The results of measurements together with calculations allow to make a thorough study of the nonlinear distortion growth in the nearfield. They confirm the usefulness of both the elaborated method and the measurement set up for the investigations of the finite amplitude wave source in its nearfield. The use of receivers of an active area dimension greater than a wavelength gives us an averaged value of the measured quantity.

Acknowledgements

The research was supported by the State Committee of Scientific Research (Poland) grant No 839 T07 96 11.

References

[1] I. AANONSEN, T. BARKVE, J. NAZE TJØTTA, S. TJØTTA, *Distortion and harmonic generation in the nearfield of a finite amplitude sound beam*, J. Acoust. Soc. Am., **75**, 749–768 (1984).

[2] K. BEISSNER, *Exact integral expression for the diffraction loss of a circular piston source*, Acustica, **49**, 212–217 (1981).

[3] A.C. BAKER, K. ANASTASIADIS, V.F. HUMPREY, *The nonlinear pressure field of a plane circular piston: Theory and experiment*, J. Acoust. Soc. Am., **84**, 1483–1487 (1998).

[4] A.C. BAKER, *Nonlinear pressure fields due to focused circular apertures*, J. Acoust. Soc. Am., **91**, 713–717 (1992).

[5] A.C. BAKER, A.M. BERG, A. SAHIN, J. NAZE TJØTTA, *The nonlinear pressure field of a plane, rectangular apertures: Experimental and theoretical results*, J. Acoust. Soc. Am., **97**, 3510–3517 (1995).

[6] D. CATHIGNOL, J.Y. CHAPELON, *High energy ultrasound therapy, Part I and Part II*, Advances in nonlinear acoustics, World Scientific, London, 21–35 (1993).

[7] W.N. COBB, *Finite amplitude method for the determination of the acoustic nonlinear parameter B/A*, J. Acoust. Soc. Am., **73**, 1525–1531 (1983).

[8] J. DYBEDAL, *TOPAS: Parametric end-fire array used in offshore applications*, Advances in nonlinear acoustics, World Scientific, London, 264–269 (1983).

[9] L. FILIPCZYŃSKI, J. ETIENNE, M. PIECHOCKI, *An attempt to reconstruct the lithotripter shock wave pulse in kidney: Possible temperature effects*, Med. Biol., **18**, 569–577 (1992).

[10] L. GERMAIN, J.D.N. CHEEKE, *Generation and detection of high order harmonics in liquids using a scanning acoustic microscopy*, J. Acoust. Soc. Am., **83**, 942–949 (1988).

[11] X.F. GONG, X.Z. LIU, *Acoustical nonlinearity parameter and its medical applications*, Advances in nonlinear acoustics, World Scientific, London, 353–357 (1993).

[12] F. INGENITO, A.O. WILLIAMS JR., *Calculation of second harmonic generation in a piston beam*, J. Acoust. Soc. Am., **49**, 319–328 (1971).

[13] E. KOZACZKA, G. GRELOWSKA, *Nonlinearity parameter B/A of the low-salinity seawater*, Arch. Acoust., **19**, 259–270 (1994).

[14] E. KOZACZKA, G. GRELOWSKA, *Investigation of the nonlinearity parameter B/A in the South Baltic Sea*, Nonlinear acoustics in perspective, Nanjing University Press, 88–93 (1996).

[15] S.-W. LI, Z.-X. XU, *The harmonic nearfield of a narrow strip transducer*, Nonlinear acoustics in perspective, Nanjing University Press, 200–205 (1996).

[16] J.A. TENCATE, *An experimental investigation of the nonlinear pressure field produced by a plane circular piston*, J. Acoust. Soc. Am., **94**, 1084–1089 (1993).

ARCHIVES OF ACOUSTICS
24, 1, 85–95 (1999)

THE INFLUENCE OF ORIENTATION ON THE BEHAVIOUR OF THE ACOUSTIC EMISSION IN FACE CENTERED CUBIC METAL SINGLE CRYSTALS COMPRESSED IN A CHANNEL-DIE

A. PAWEŁEK, A. PIĄTKOWSKI, Z. JASIEŃSKI

Aleksander Krupkowski Institute of Metallurgy
and Materials Science
Polish Academy of Sciences
(30-059 Kraków, ul. Reymonta 25, Poland)

S. PILECKI

Institute of Fundamental Technological Research
Polish Academy of Sciences
(00-049 Warszawa, ul. Świętokrzyska 21, Poland)

The influence of the crystallographic orientation on the behaviour of the acoustic emission (AE) in face centered cubic (FCC) metal and alloy monocrystals compressed in channel-die is investigated using five differently oriented copper single crystals ($\{100\}<001>$, $\{011\}<112>$, $\{112\}<111>$, $\{111\}<123>$ and $\{001\}<110>$). The results obtained are also compared to the AE behaviour in silver and to the low-temperature AE behaviour in copper and copper-aluminium alloy single crystals of identical $\{112\}<111>$ orientations. It has been stated that the orientation of crystal affects the final stage of the microstructure evolution (shear bands of the V or X shape or the bands of complex structure), and the orientation dependence of the AE behaviour is only a consequence of the orientation dependence of the deformation mechanisms. In general, however, the AE behaviour is strongly correlated with strain localization related to twinning and shear band formation, and is of universal character since it is similar in all the orientations applied.

The observed correlations between the AE and the strain localization mechanisms are discussed on the basis of the dynamic and nonlinear (solitary wave) properties of dislocations. Consequently, it has been stated that the dynamics of shear band formation is markedly weeker than that in the case of twinning seems to indicate the crystallographic character of the shear band propagation since the non-crystallographic slip should be accompanied by strong acoustic effects.

1. Introduction

Recent investigations of the phenomenon of acoustic emission (AE) are, in general, conducted in two directions. The first one, dealing with the AE phenomenon itself, concentrates mainly on the spectral analysis of the measured AE parameters [1]. However,

the phenomenon of AE is more often used as an experimental method in various investigations in the materials science [2–6]. The AE method, in particular, has been used for some years by the authors of the present study in the investigations of the processes of plastic deformation of metal and alloy single crystals of face centered cubic (FCC) lattice subject to a channel-die compression, where a model of strain in plain state is realized [7–10]. The aim of these investigations was to determine and explain the observed correlations between the behaviour of AE and the strength properties (strain hardening curve) of the material, the evolution of its microstructure and texture, and the mechanisms of strain localization connected with the processes of twinning and formation of shear bands [11, 12]. A particular aim of this work was to establish the influence of crystallographic orientation on the AE behaviour in metal and alloy single crystals of FCC lattice subjected to channel-die compression. The investigations were carried out on the example of copper single crystals of five different orientations (I – {111}<001>, II – {112}<110>, III – {112}<111>, IV – {111}<1123> and V – {001}<110>) and compared with the results of investigations of channel-die compression of single crystals of silver, copper and copper-aluminium alloy (CuAl2) of identical orientations {112}<111>. The results obtained in the present work by means of the AE method have been discussed on the basis of the dynamic and nonlinear (solitary waves) properties of dislocations [13–16]. In particular, these results enabled to formulate a more precise opinion on the still disputable problems referring both to the dynamics of the formation and the possibility of non-crystallographic propagation of shear bands.

2. Investigation methods

Channel-die compression tests were carried out using the tensile testing machine INSTRON-6025, equipped with an additional installation containing a channel-die, which ensured the plastic flow of the metal in the parallel direction only, i.e. along the channel axis, and in the vertical direction, i.e. perpendicular to the channel axis. This is a simple model of the channel-die compression in which the plane state of strain is realized since in the lateral direction (also perpendicular to the channel axis) plastic deformation does not occur. Monocrystalline samples were obtained by means of the Bridgeman's method (in the case of silver and copper) or by the method of zone crystallization at a natural temperature gradient (in the case of copper and copper-aluminium alloy). Samples, having the shape of cubes of a 10 mm edge, were subjected to compression tests by the multi-stage method with the purpose to obtain the final reductions as well as the intermediate reductions of an appropiate value. After each stage, the samples were appropiately cropped so as to make the ratio of the actual value of elongation to that of height not greater than 2. In order to minimize the effects of the sample friction against the channel walls, a teflon foil was used. The speed of the testing machine travers was all the time equal to 0.05 mm/min. Simultaneously with the measurement of the work-hardening curve in the terms of force vs. time, the basic AE parameter in the form of the rate of events $\Delta N_z/\Delta t$ was registered. The number of events ΔN_z was measured in the time interval $\Delta t = 4$ s. A broad-band piezoelectric sensor enabled to

register acoustic pulses of a frequency from some tens to some hundreds of kilohertzs, while suitable filters eliminated the frequency of free vibrations of the testing machine, as well as the higher frequencies connected, e.g. with the noise of broadcasting stations. The contact of the sensor with the sample at the ambient temperature occurred through a steel rail which formed a pad in the channel-die, and at the liquid nitrogen temperature – by means of a wave-guide especially formed from a quartz rod. The total amplification of the apparatus registering the AE signals was 88 dB, and the suitably chosen threshold voltage of the discriminator had a value of 1.20 V. After each compression test, the microstructure was observed using the standard technique of optical microscopy.

3. Investigation results and discussion

In Figs. 1–5 is shown the AE behaviour during the channel-die compression of copper monocrystals of five various crystallographic orientations: I – {100}<001> (Fig. 1), II – {112}<110> (Fig. 2), III – {112}<111> (Fig. 3), IV – {111}<123> (Fig. 4) and V – {001}<110> (Fig. 5). On each of these figures as well as on the following ones, besides the graphs of AE and force (to the left), there are also presented the microstructures (to the right) corresponding to the reductions obtained at the given compression stage. It should be noted that the initial course of AE, illustrating, on each figure, the AE behaviour in the time interval up to about 200–300 s, will not be discussed here since, this interval comprises also beside the transition from the elastic to the plastic state (already fairly accurately examined), effects connected with the mechanical adjustment of the sample to the channel walls in the range of elastic strains. For the purpose of comparison, in Figs. 6–8 there has been presented the behaviour of the AE (for selected compression stages) also in other single crystals of FCC lattice of identical orientation III subjected to channel-die compression at the liquid nitrogen temperature. Consequently, in Fig. 6 is illustrated the AE behaviour in Cu single crystals, in Fig. 7 – in Ag single crystals, and in Fig. 8 – in CuAl2 alloy single crystals. Moreover, in Fig. 7 is illustrated the transition from the deformation mechanism, associated with twinning, to the deformation mechanism connected with the formation of shear bands.

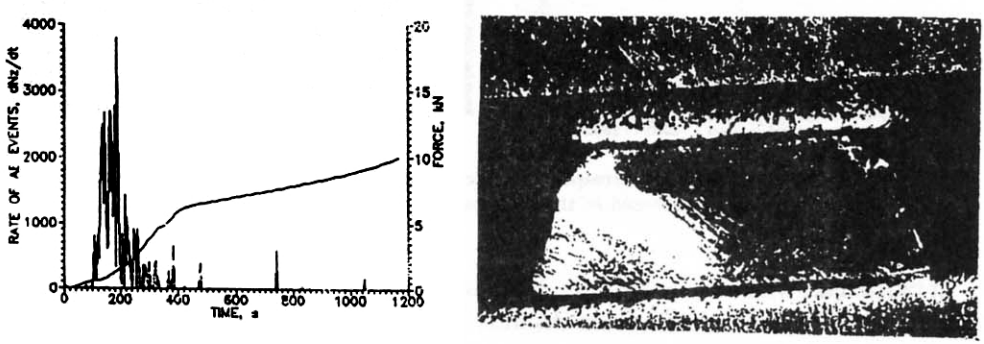


Fig. 1. Behaviour of the acoustic emission (AE) and the external force in copper monocrystals of cubic orientation {100}<001> during a channel-die compression at the ambient temperature. Besides is the corresponding microstructure after a $z = 90.4\%$ reduction; magnified ×31.

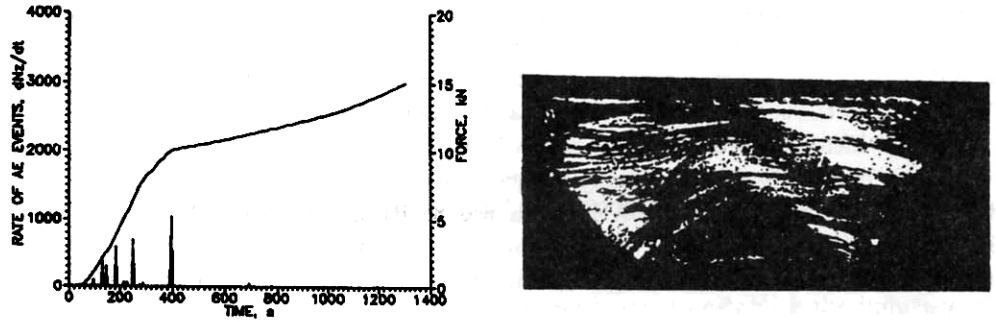


Fig. 2. AE, external force and the corresponding microstructure in Cu monocrystals of {112}<110> orientation compressed in the channel-die ($z = 86.8\%$, $T = 293\,\mathrm{K}$, ×25).

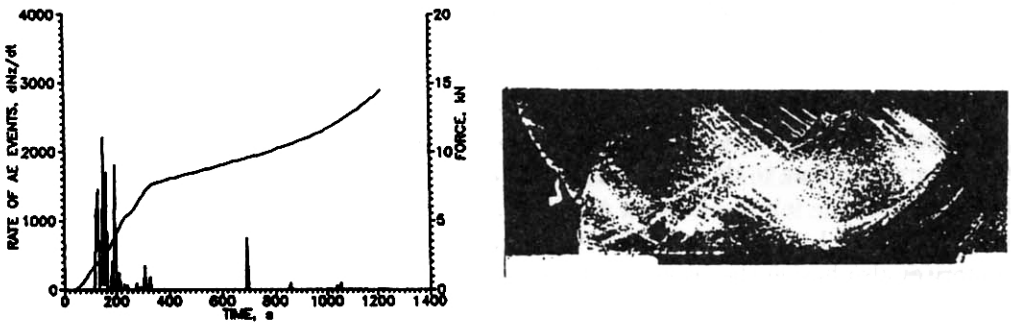


Fig. 3. AE, external force and the corresponding microstructure in Cu monocrystals of {112}<111> orientation compressed in the channel-die ($z = 80\%$, $T = 293\,\mathrm{K}$, ×31).

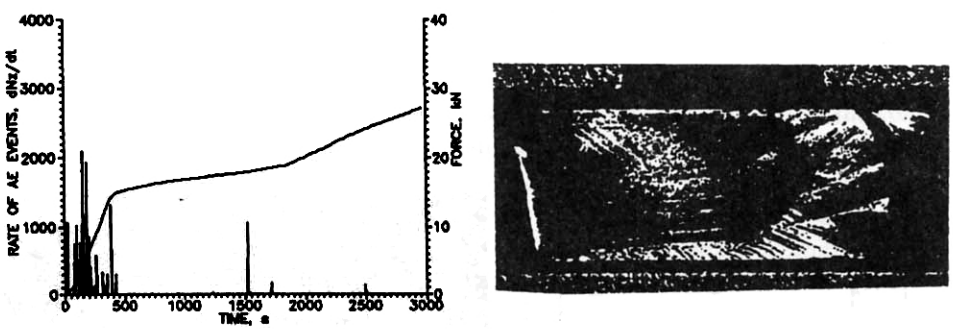


Fig. 4. AE, external force and the corresponding microstructure in Cu monocrystals of {111}<123> orientation compressed in the channel-die ($z = 70\%$, $T = 293\,\mathrm{K}$, ×10).

In particular, when analyzing Figs. 1–5 one can observe that there exists a tendency of universal character indicating that the final stage of the evolution of the microstructure in the process of channel-die compression (high reductions) is the occurrence of a second family of shear bands formed in the secondary activated slip systems; together with the previous family of shear bands formed in the primary activated slip systems, these bands

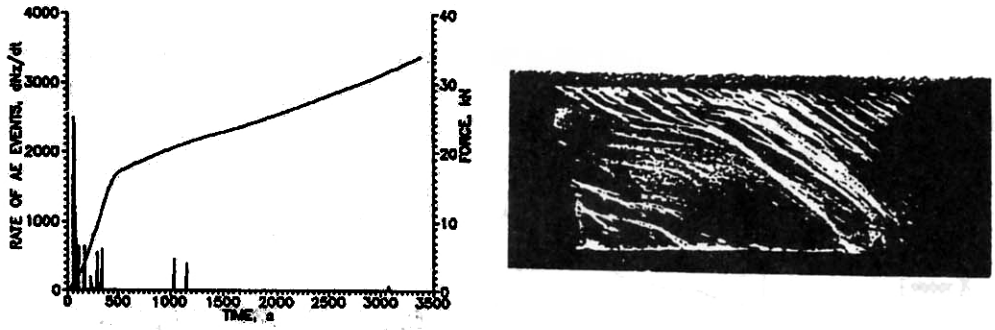


Fig. 5. AE, external force and the corresponding microstructure in Cu monocrystals of {001}<110> orientation compressed in the channel-die ($z = 75\%$, $T = 293\,\mathrm{K}$, ×31).

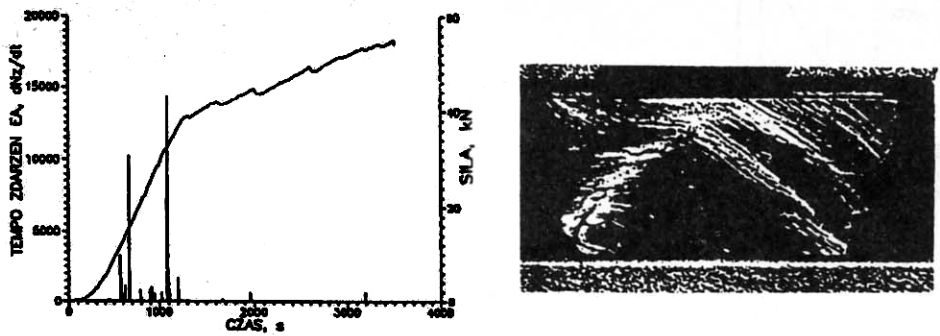


Fig. 6. AE, external force and the corresponding microstructure in CuAl2 single crystals of {112}<111> orientation compressed in the channel-die at the liquid nitrogen temperature ($z = 73\%$, ×20).

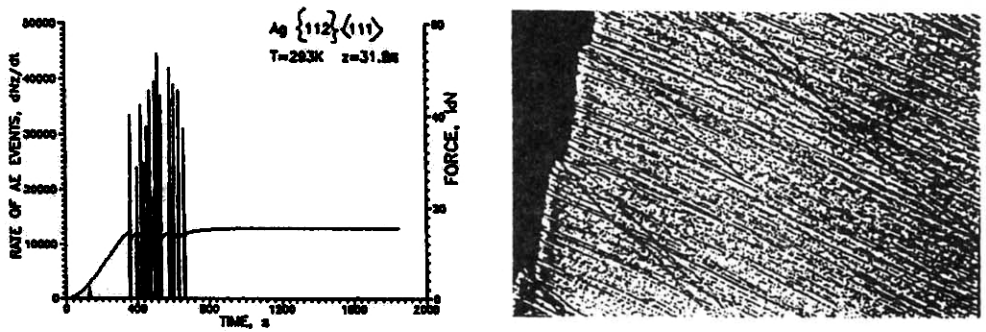


Fig. 7. AE, external force and the corresponding microstructure in silver single crystals of {112}<111> orientation compressed in the channel-die ($z = 31\%$, $T = 293\,\mathrm{K}$, ×200).

have the shape of the letter V (stright or reversed). This can be observed distincly for the orientations I, II and V (Figs. 1, 2 and 5, respectively). In the case of the orientations III and IV (Figs. 3 and 4, respectively), the pattern is more complex due to the influence of the orientation. In the case of the orientation {112}<111>, the bands form a shape resembling the letter X (Fig. 3), and in the case of orientation {111}<123>, the effect

of cooperation of both the tendencies to form bands of V and X type can be observed (Fig. 4). However, in each case (including those shown in Figs. 6 and 8) the formation of shear bands of V type (as they have been called earlier [7, 9]) is always accompanied by a series of a few, more or less regular, AE peaks.

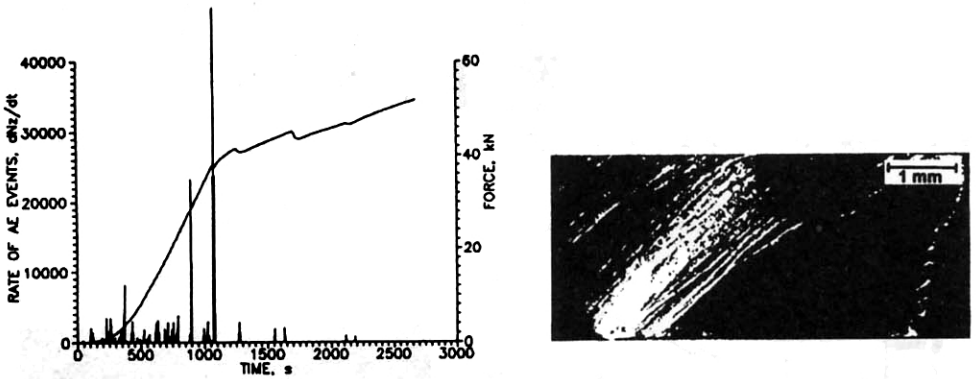


Fig. 8. AE, external force and the corresponding microstructure in Cu single crystals of {112}<111> orientation compressed in the channel-die ($z = 73\%$, $T = 77\,\mathrm{K}$, $\times 20$).

From the presented graphs and microstructures it can be seen that in the case of each orientation there exists a direct correlation between the AE behaviour and the deformation mechanisms, while the influence of orientation is indirect in the sense that the AE depends on the deformation mechanisms determined, in general, by the initial crystallographic orientation.

The above correlations may be interpreted, at least qualitatively, basing on the dynamics of the dislocation configurations (Figs. 9 and 10) and the theoretical concept which seeks the origin of the AE just in the dynamic dislocation phenomena connected chiefly with the acceleration and the processes of dislocation annihilation. It must be taken into consideration that acoustic effects are proportional both to the square of acceleration (including the acceleration in the vibration movement of kinks) and to the square of the relative velocity of annihilating dislocations (see e.g. [14, 15]). In the scheme in Fig. 9 is shown a simple propagation mechanism of a slip band (thus at the first approximation corresponding to the initial phase of shear band formation, Fig. 9a), a wave image of this propagation (solitary wave, Fig. 9b) and the spatial mechanism of the operation of the dislocation sources of the Frank–Read type in the primary activated slip systems (Fig. 9c). In Fig. 10, on the other hand, is illustrated the solitary wave character of the propagation of a group of dislocations generated by a Frank–Read source in a single slip plane (Fig. 10a), a one-dimensional approximation of this propagation (Fig. 10b), and a wave representation (Fig.10c), evidencing the soliton-like character of this propagation (see also [13, 16]). Speaking in general, the observed AE peaks may be the result of a superposition of the effects of acceleration and annihilation of the dislocations both inside a crystal (in the case of the closing of the dislocation loops during the operation of the Frank–Read sources) and also on its free surface (in the case of the formation of the steps on the crystal surface).

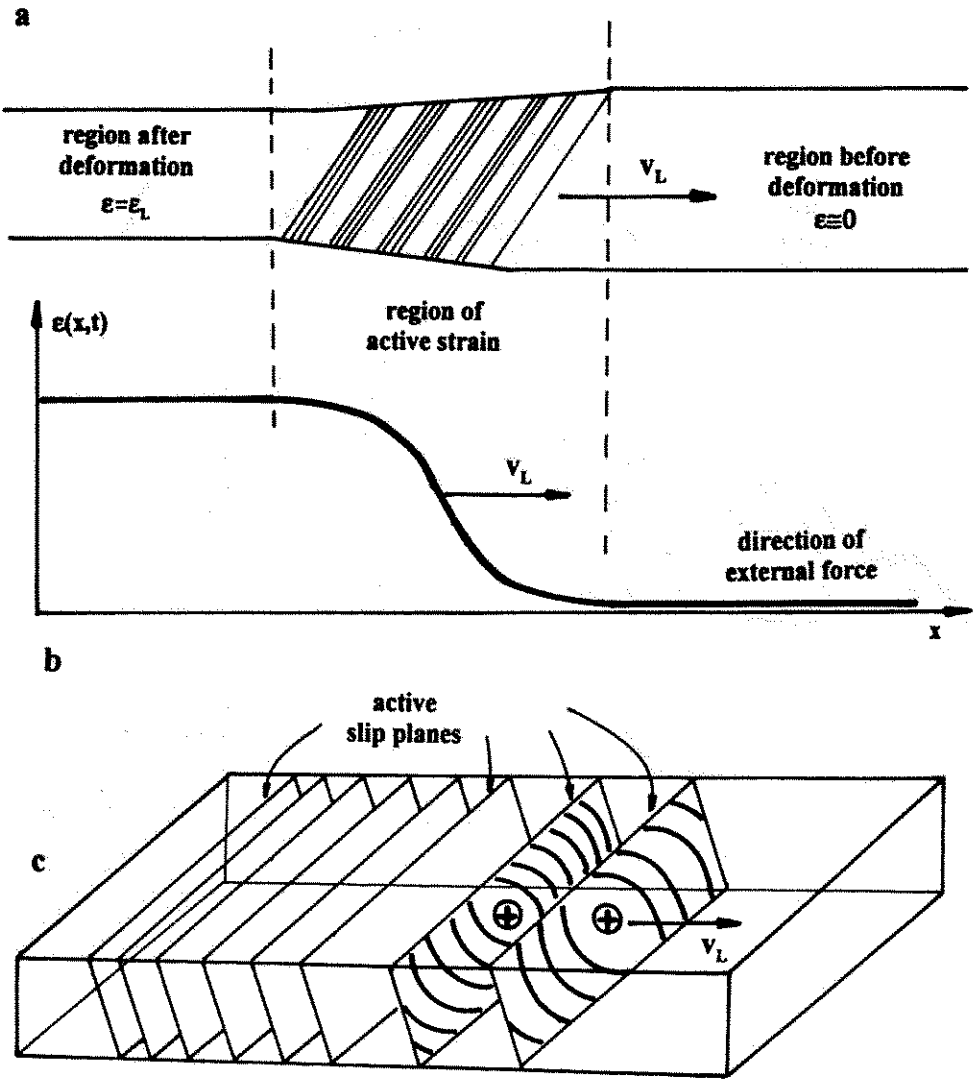
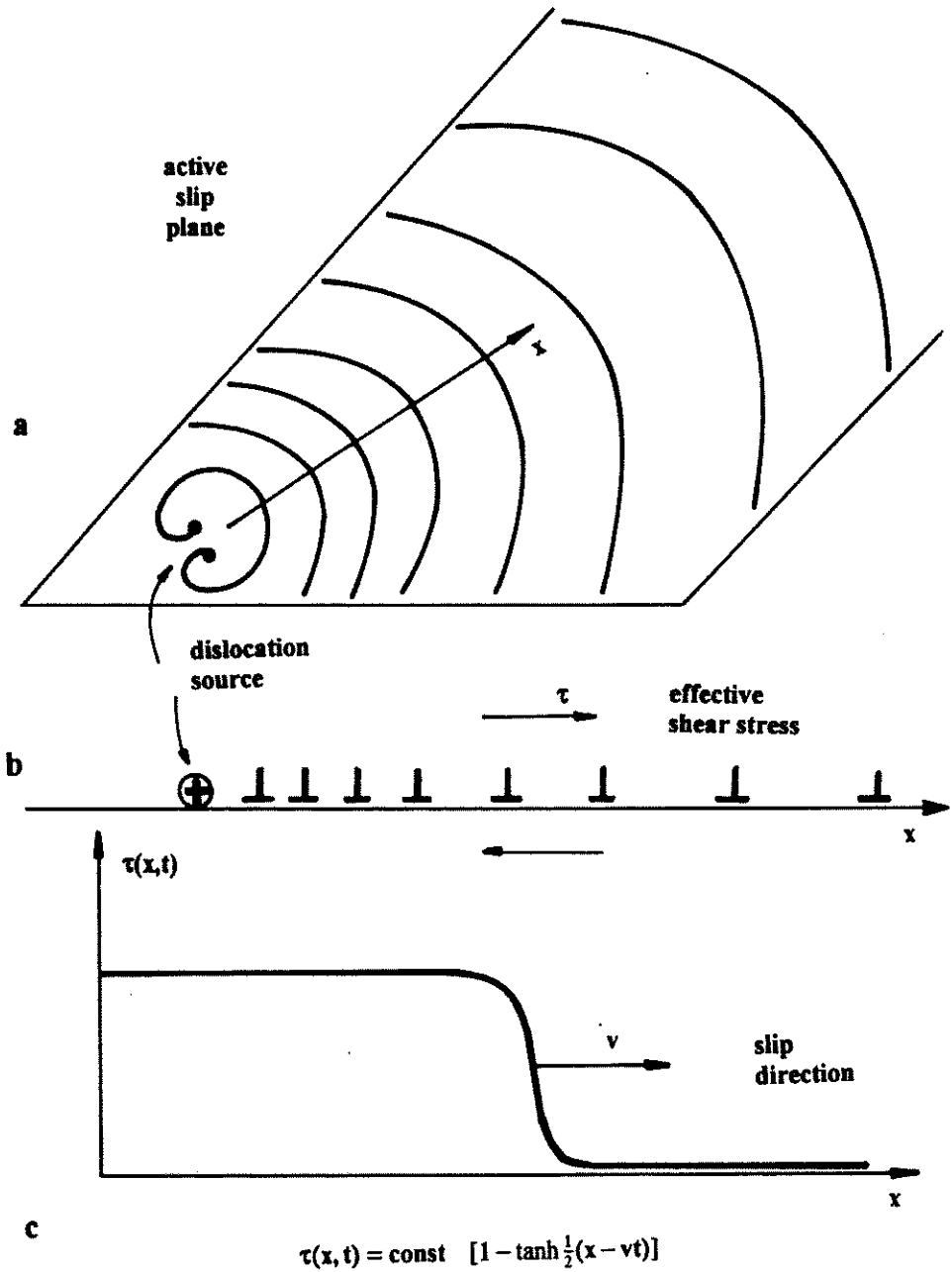


Fig. 9. Schematic illustration of a simple model of shear band propagation (a), its solitary wave representation (b) and the dislocation dynamics in active slip systems (c).

Thus, the high rate of the AE events in the case of twinning, reaching the values of the order of 4×10^4 (in the range from 350 to 650 s in Fig. 7), is due to the fact that the formation of a single twin – assuming the polar mechanism of twinning (see e.g. [15]) – may be associated with the escape to the crystal surface of tens of thousands of dislocations (more accurate estimation has been presented in [17, 18]). The effects of dislocation acceleration may also participate to a great extent in it since it is known that the velocity of twinning dislocations in a given crystal may reach values of the order of the speed of sound. On the other hand, the dynamics of the gliding dislocations (Figs. 9a



$$\tau(x,t) = \text{const} \quad [1 - \tanh \tfrac{1}{2}(x - vt)]$$

Fig. 10. Schematic illustration of the dynamics of dislocation sources in a single slip plane (a), the corresponding configuration of dislocations at the one-dimensional approximation (b), and the soliton-like character of the propagation of dislocations (c).

and 10a) is smaller than that of the twinning dislocations (at least in the sense of the possibility of attaining very high velocities in comparable periods of time). Moreover, shear bands occur in the sample volume considerably less frequently than the twins (Fig. 7). As a result, the shear bands produce much weaker acoustic events of the order of 10^3, i.e. they can be estimated to be at least by one order of magnitude smaller than in the case of twinning.

The considerably more distinct and stronger acoustic effects, accompanying the formation of shear bands at low temperatures (Figs. 6 and 8), can be explained in a similar way. In this case the dynamics of the gliding dislocations is much greater than that occurring at the ambient temperature (higher flow stress since higher dislocation velocity). On the other hand, a quite similar situation occurs when shear bands of type V are formed at the ambient temperature (Figs. 1–5) and at the temperature of liquid nitrogen. The beginning of the formation of shear bands in the secondary activated slip systems (similarly as in the case of the primary activated slip systems) would be associated with the engagement of some tens to some hundreds of dislocation sources (a more accurate assessment is presented also in [17, 18]) operating in parallel slip systems (Fig. 9c); each of them (Fig. 10a) would generate tens or even hundreds of dislocations. In this way each single AE peak, within an advanced range of the work-hardening curve (Figs. 1–5), signalling the beginning of the formation of shear bands and the formation of a step on the crystal surface, would be associated, at the first approximation, with the escape to the sample surface of a number of dislocations of the order from some hundreds (in relation to the ambient temperature) to some thousands (in relation to the low temperature). The number of AE events of this order, observed in Figs. 1–5 and 6 and 8, suggests that in the case of shear bands, independently of the crystallographic orientation, the contribution of the surface annihilation to the AE may be much greater than the that from the internal annihilation and that deriving from the effects of the acceleration of dislocations.

4. Conclusions

1. The observed correlations (very distinct especially at the temperature of liquid nitrogen) between the behaviour of the acoustic emission (AE) and the course of the channel-die compression of monocrystals of metals and alloys of a FCC lattice occur for all the examined orientations being mainly connected with twinning and the formation of shear bands.

2. Single twins, independently of the crystal orientation, are associated with single AE peaks to which the number of AE events of the order of tens of thousands corresponds.

3. The initiation of the formation of shear bands is connected with single AE peaks, to which, however, a smaller number of events correspond – in general, of the order of some hundred to a few thousands – independently on the crystal orientation.

4. Depending on the crystal orientation (although not to a great extent), the final stage of the microstructure evolution (high reductions) are the shear bands, either of type V (most often) or of type X (as in the case of the orientation {112}<111>) or of a

more complex type, being always accompanied by a series of a few, more or less regular, AE peaks.

5. The observed correlations between the AE behaviour and the progress of the compression deformation can be interpreted basing on the dynamic properties of dislocations connected mainly with acceleration and the dislocation annihilation occurring especially on the crystal surface.

6. As a consequence, the dynamics of the formation of shear bands is distinctly weaker than the dynamics of the formation of twins which is an additional evidence of the crystallographic nature of shear bands propagation since a non-crystallographic slip should be accompanied by very strong acoustic effects.

Acknowledgement

The part of this work was supported by the Grant No 7 T07B 043 12.

References

[1] A. VINOGRADOV, M. NADTOCHIY, S. HASHIMOTO, S. MIURA, *Correlation between spectral parameters of acoustic emission during plastic deformation of Cu and Cu-Al single and polycrystals*, Materials Transactions, JIM, **36**, 426–431 (1995).

[2] A. SLIMANI, J. CHICOIS, R. FOUGÈRES, P. FLEISCHMANN, *Etude par emission acoustique de la deformation cyclique de polycristaux d'aluminium de haute purete sonlicites en traction compression*, Mat. Int. Conf. Strength of Metals and Alloys, ICSMA 7, Montreal (Canada) 1985, vol. 2, pp. 1243–1248.

[3] P.M. MUMMERY, B. DERBY, C.B. SCRUBY, *Acoustic emission from particulate-reinforced metal matrix composites*, Acta Metall. Mater., **41**, 1430–1445 (1993).

[4] K.R. THUMMA, V.S. BOIKO, D.N. BESHERS, *Acoustic emission accompanying nonlinear vibrations in the vicinity of a crack*, Scripta Metall. Mater., **28**, 677–681 (1993).

[5] F. CHMELIK, Z. TROJANOVÁ, Z. PŘEVOROVSKÝ, P. LUKÁČ, *The Portevin-Le Chatelier effect in Al-2.92%Mg-0.38%Mn alloy and linear location of acoustic emission*, Mat. Sci. Eng., **A164**, 260–265 (1993).

[6] C.K. MUKHOPADHYAY, K.V. KASIVISWANATHAN, T. JAYAKUMAR, BALDEV RAJ, *Acoustic emission from aging-induced martensite formation in cold worked AISI type 304 stainless steel*, Scripta Metall. Mater., **30**, 303–307 (1994).

[7] A. PAWEŁEK, Z. JASIEŃSKI, A. PIĄTKOWSKI, H. PAUL, A. LITWORA, *Low-temperature acoustic emission during channel-die compression of various oriented silver single crystals*, Proc. XIVth Winter School on Molecular and Quantum Acoustics, Gliwice-Wisła (Poland), Akustyka Kwantowa i Molekularna, **16**, 177–186 (1995).

[8] A. PAWEŁEK, Z. JASIEŃSKI, A. PIĄTKOWSKI, A. LITWORA, H. PAUL, *Acoustic emission and strain localization in shear bands during channel-die compression*, Archives of Metallurgy, **41**, 323–336 (1996).

[9] A. PAWEŁEK, Z. JASIEŃSKI, A. PIĄTKOWSKI, H. PAUL, A. LITWORA, *Acoustic emission during the channel-die compression of the polycrystalline copper and brass with columnar structure of grains*, Archives of Metallurgy, **41**, 465–477 (1996).

[10] H. PAUL, A. PAWEŁEK, Z. JASIEŃSKI, A. PIĄTKOWSKI, A. LITWORA, *Investigation methods of the structure of materials in deformed and in recrystallized states*, Proc. Inter. Conference under auspices of E-MRS „Nowoczesne materiały ceramiczne i metody badań", Białowieża 1996, Polish Ceramic Bulletin 12, Ceramics, **50**, 81–94 (1996).

[11] H. PAUL, Z. JASIEŃSKI, A. PIĄTKOWSKI, A. LITWORA, A. PAWEŁEK, *Crystallographic nature of shear bands in polycrystalline copper*, Archives of Metallurgy, **41**, 337–353 (1996).

[12] H. PAUL, Z. JASIEŃSKI, A. PIĄTKOWSKI, A. PAWEŁEK, A. LITWORA, *Pasma ścinania w nieswobodnie ściskanych monokryształach stopu Cu 2% Al o orientacji* (112)[11$\bar{1}$] [in Polish], Proc. I-st Polish Conf. „Inżynieria Materiałowa' 96", Gdańsk–Sobieszewo 1996, pp. 31-40.

[13] A. PAWEŁEK, A. KORBEL, *Soliton-like behaviour of a moving dislocation group*, Phil. Mag., **B61**, 829–842 (1991).

[14] A. PAWEŁEK, *Possibility of a soliton description of acoustic emission during plastic deformation of crystals*, J. Appl. Phys., **63**, 5320–5325 (1988).

[15] A. PAWEŁEK, I. MALECKI, *Emisja akustyczna a plastyczność kryształów* [in Polish], Prace IPPT, No 22, Warszawa 1993, pp. 1–63.

[16] A. PAWEŁEK, S. PILECKI, Z. JASIEŃSKI, *Propagation of an effective shear stress as a solitary wave and possibility of non-crystallographic slip in polycrystals*, Archives of Mechanics, **47**, 101–116 (1995).

[17] A. PAWEŁEK, Z. JASIEŃSKI, A. PIĄTKOWSKI, H. PAUL, A. LITWORA, *Acoustic emission and dislocation dynamics during channel-die compression of F.C.C. metal and alloy single crystals*, Proc. XXVI-th Winter School on Molecular and Quantum Acoustics, Gliwice – Ustroń 1997, pp. 63–67.

[18] A. PAWEŁEK, A. PIĄTKOWSKI, Z. JASIEŃSKI, *Nonlinear and dislocation dynamic aspects of acoustic emission and microstructure evolution during channel-die compression of metals*, Akustyka Molekularna i Kwantowa, **18**, 321–358 (1997).

ARCHIVES OF ACOUSTICS
24, 1, (1999)

C H R O N I C L E

DISSERTATIONS

Ultrasound scattering for characterization of marine crude oil spills

by STANISŁAW J. POGORZELSKI
Institute of Experimental Physics, Gdańsk University,
ul. Wita Stwosza 57, 80-952 Gdańsk, Poland
e-mail: fizsp@paula.univ.gda.pl

Laboratory and at-sea surface scattering measurements, for wind generated water surfaces covered with crude oil derivative films of well-defined and oceanographically relevant viscoelastic surface properties, were carried out using a directional acoustic system based on high-frequency forward specular scattering. In the light of the "specular point" scattering theory applicable to ultrasounds (and also to laser, sun glitter, and microwave surface probing) the scattering coefficient of acoustic waves in these conditions i.e., for large values of Rayleigh parameter, depends only on the mean square slope of the rough surface. Water wave attenuation by a viscoelastic film is attributed to the Marangoni effect which causes a strong resonance – type damping in the short-gravity-capillary water wave region. The Marangoni damping depends on the physicochemical nature, concentration of the film-forming substance, and the rheological surface film parameters (E – elasticity modulus, P – film pressure, ω_0 – structural diffusion parameter). The scattered signal signatures turned out to be unequivocally related (via the Marangoni damping of wind waves by an elastic film) to the structural and rheological film parameters.

It has been demonstrated in open-sea experiments with artificial crude oil slicks spread over the Baltic Sea surface and a buoy-like free drifting low-weight acoustic system that the film parameters can be recovered from the ratio of the scattered signal low-frequency modulation spectra (with to without films) adopting a theoretical form of the relative spectrum with terms responsible for the Marangoni damping of wind waves, their growth, and spatial film homogeneity (i.e., film-filling factor). The latter parameter, which is wind-speed dependent, plays a principal role in the proper determination of the film properties derived, whereas a variation of the wind waves growth rate, also affected by the film presence, is of secondary importance.

The acoustically determined viscoelastic surface properties of a model slick are in moderate agreement with these simultaneously evaluated "in situ" using a novel film

sampler-elastometer-Langmuir trough system, and are characteristic of a natural slick of biogenic origin and/or "weathered" crude oil spills. A discrepancy between the acoustically derived and theoretically predicted relative spectra of wind waves is believed to result from a simplified model of the wind-wave interaction and a very complex form of fine structures of "real" wind-generated surfaces (parasitic, instability waves known as "cat's paws"), which may not express itself correctly in the signatures of the "specular point" scattering applicable to ultrasounds. The elastic properties of composite sea surfaces likely to be found in polluted coastal waters and consisting of oil spills filled with a surface-active material, floating solid particles (dust), gas bubbles, or drops of a third fluid with their important implications in remote sensing techniques are also discussed.

The statistics of the scattered signal fluctuations are approximated by expanding the Gaussian function into a Gram-Charlier series, taking into consideration statistical moments up to the fourth one. Evolution of the shape, skewness, and kurtosis of the signal distribution reflects an important role played by the film elasticity. Simultaneous analyses of all the statistical parameters could be a starting point for determining the fraction weight of the given spill-forming substance its layer thickness, and finally a form of the pollutant (monolayer, thick layer, or individual dispersed spots). The system allows the passage of the edge of the oil spill to be detected, and can provide significantly better resolution in the study of small spatial-scale air-sea interaction processes taking place at the sea surface than radar, and may be more convenient to operate in long-term continuous monitoring and less costly to use.

[in:] **ENCYCLOPEDIA OF ENVIRONMENTAL CONTROL TECHNOLOGY**, P. Cheremisinoff [Ed.], Vol. 9, Geotechnical Applications Leak Detection, Treatment Options, Chapter 15, Gulf Publishing Company, Houston 1995, pp. 485–538.

Acoustical spectroscopy of cyclic and heterocyclic compounds, ketones and polluted water surface

Doctor Thesis (Dissertation 1997)
by BOGUMIŁ, BOLESŁAW, JULIUSZ LINDE
Institute of Experimental Physics
Acoustics and Solid State Physics Division, Gdańsk University
ul. Wita Stwosza 57, 80-952 Gdańsk, Poland

THE CONTENTS OF THE THESIS

The subject of this thesis was the research of ultrasonic waves interaction with liquid molecules of pure organic compounds, solved in the other ones or spilt on the water surface.

There were two main aims of this thesis:

1) mainly on the basis of own acoustic investigations however enriched with some data taken from literature, concerning ultrasound absorption in ketones, cyclic and heterocyclic compounds, the author would like to established certain strict relations between acoustical absorption (α/f^2) rel and the molecules structure of a given liquid, and moreover to study some pure organic liquids and liquid mixtures with Kneser relaxation;

2) looking for the remote system in order to identify and control the ecological condition of seas as the first step in checking of oil pollution on surface of water.

As a consequence the work was divided into two parts: the first "A" – "Acoustic investigations in cyclic & heterocyclic compounds & their mixtures as well as some ketones", and the second "B" – "Amplitude spectrum statistical analysis of the acoustic wave, reflected from the undulated water surface, for identifying its pollution".

All results of investigations presented in this work are closely related to interactions of ultrasonic waves and the molecules of liquids.

The results presented in the I-st part of the Thesis has allowed the author to make following conclusions:

There exists a strict relation between acoustical absorption, in the frequency range below the Kneser relaxation process, and molecule structure. It closely depends on: numbers of the groups attached to the benzene or pyridine ring, the size of the group attached, the kind of the atom in the benzene ring, saturation of the compounds, numbers of the vibrational degrees of freedom as well as their distribution. The acoustical absorption dependence on aromacity and compressibility is rather ambiguous. There is no special relation between absorption and dipole moment.

The author has found the Kneser acoustical relaxation in six new compounds.

The situation similar to the first case takes place also for the group of compounds in which an isomeric relaxation is observed. The acoustical absorption depends on the size of chemical group attached to the ketone group what is connected with the conjugation of the $C_1 - C_2$ bond and increasing of the rotation barrier. This dependence in saturated ketones is much weaker than in unsaturated ones.

In the mixture of two Kneser liquids of highly-absorbing and low-absorbing abilities, it has been shown that the energy transfer between two different molecules is more

probable than between the same. There appear fast T–V and V–V transfers which fastens the process of molecular deactivation, they shorten the acoustical relaxation time and decrease the absorption coefficient α/f^2.

Referring to the II-nd part of the Thesis one may conclude that follows:

The measurements of the ratio of the measured spectra $D_r(f)$ or the acoustic amplitude pattern may enable the nature of the surface film to be characterised. It is of course difficult to deduce an exact physical and chemical nature of the sea surface films from the intensity and frequency of this characteristic peak, but it does seem feasible in principle to characterize surface films by this method. The direct influence of the surface film should be considered in terms of the chemical structure of the film's hydrophobic part of damping surface waves.

The attenuation of a surface wave propagated on the surface of water covered with an oil film attains values several times higher than those predicted by the classical hydrodynamic theory for a totally clean surface. The attenuation reaches its maximum, equal to $\alpha_{\max} \sim 2\alpha_{\text{imm}}$ (immo-bile), at low surface pressure. The maximum position is consistent with the prediction of the Dorrestein theory, however, it is lower and broader, due to the significant effect of viscosity. When the thickness of the film approaches or exceeds the generally assumed value which determines the depth of penetration of the undulatory motion $d \sim \lambda_b/2$, the values of the coefficient tend to approach the Stokes' expression α_m, exceeding it only by several per cent.

The observed instances of departure of a from the theoretical value are probably caused by the short distance between interfacial surfaces and the excessive viscosity of heavier fractions of crude oil, which leads to the unfulfilment of the applicability condition of the Stokes' approximation for an ideal liquid.

The stabilizing properties of the water surface covered with a monolayer of crude oil derivative and under the influence of an air stream depend on the type of derivative and surface concentration. A decisive role in the stabilisation process is being played by expansion surface elasticity ε_d of the monolayer. For strictly defined values of surface concentrations of the mono-layers of the investigated substances, the stabilisation of flow on the surface attains a maximum value.

Surface concentrations of monolayer, for which a maximum of surface capillary wave attenuation is observed, are very close to the values predicted by the theory for the maximum stabilisation. The agreement in the range of 30–55% indicates a close dependence between the process of wave energy dissipation and their formation under the influence of the air stream on the surface. The process is different for various viscoelastic properties.

The observed differences between both values of Γ results from a simplified and idealised model of the effect of the air stream on the liquid surface. The threshold air stream velocity capable of generating waves on the water surface covered by a monolayer of crude oil derivative is increased by a factor of about 18, while for the maximum achieved surface concentrations – by a factor of 5 – 7.5.

As far as the registration of the water pollution is concerned an acoustic high-frequency, surface scattering system consisting of the directional transducers of a narrow transmitting characteristic, is suitable to investigate the surface changes of an undulated

water caused by the presence of an oil substance layer. The statistical parameters of the echo peak value distribution from the rough water surface covered with the layer are markedly different from those for a clean surface. The changes of the parameters are of the order of several dozen percent for the normalized mean amplitude and fluctuation coefficient and the other parameters change by a larger factor, thus being more sensitive to the surface contamination.

On the basis of the results provided in natural condition (even open-sea measurements) carried out by means of free-floating (buoy-like) acoustic system all the laboratory observations presented here, as well as in many other our papers, were confirmed. Such an analysis of statistical properties of the acoustic field scattered at a rough surface can be a suggestion of a new, contactless method of remote sensing of surface pollution of natural water with oil substances.

Investigations of the finite amplitude wave propagation in water [in Polish]
by ANNA BARANOWSKA
Naval Academy, 22 April 1998,
Contact: Central Library
ul. Śmidowicza 71, 81-919 Gdynia, Poland
Supervisor prof. dr. hab. Stanisław Dobrociński

The aim of the thesis is the numerical analysis of the finite amplitude wave propagation in water. Following the description of the problem itself the physical model is presented and then the mathematical one. Further existing analytical solutions are discussed. The mathematical model is built on the basis of the KZK equation that describes the changes in acoustic pressure along the sound beam. Basing on the assumption that the wave source is circular and the amplitude of harmonic piston distribution is only a function of its radius the problem is considered as an axial symmetric one. The numerical methods must be used to solve the problem. Therefore usefulness of the choice of numerical methods is analysed. To solve this problem harmonic analysis, finite-element method and finite-difference method are applied. All these methods are described in detail. Own computer programs are worked out on the basis of obtained algorithms. Accuracy, convergence, influence of values of physical parameters such as for example nonlinearity parameter or dissipation coefficient of the medium and the fundamental wave distribution on the source (rectangular, polynomial and some different distributions which adequately simulate the real distribution on the source) have been investigated numerically. The calculations were carried out for different values of physical parameters and different values of parameters, which have influence on discrete model. The wave distortion and harmonic amplitudes evaluation as a function of the distance from the source on the beam axis and along the sound beam are studied. Moreover the average acoustic pressure on the receiving transducer is considered. The results of numerical calculations are compared with experimental ones.

On the near field of ultrasonically diffracted light [in Polish]

by GRZEGORZ GONDEK
Institute of Experimental Physics,
Acousto-Optics and Laser Physics Department,
Gdańsk University
ul. Wita Stwosza 57, 80-952 Gdańsk, Poland
e-mail: fizgg@univ.gda.pl
Advisor: Prof. dr hab. Piotr Kwiek, Institute of Experimental Physics,
Gdańsk University
Reviewers: Prof. dr hab. Aleksander Opilski, Gliwice Institute of Physics,
Silesian Technical University
Prof. dr hab. Antoni Śliwiński, Institute of Experimental Physics,
Gdańsk University

Keywords: acousto-optics, near field (Fresnel region), far field (Fraunhofer region), intermediate diffraction regime, self-imaging phenomenon, secondary interference planes (Nomoto planes), KML method, NOA method, additional optical phase shifts.

The work is devoted to theoretical and experimental study of the Fresnel region of ultrasonically diffracted light. However some more general considerations are introduced yet the main interest is focused on the case of plane monochromatic light wave diffraction, incidenting normally a plane sinusoidal sound beam, propagating in transparent, uniform and isotropic medium (distilled water). The Thesis constitutes an attempt at analysis and recapitulation of nowadays state of the near field examinations as well as the experimental verification of the main theories describing the phenomenon of light diffraction by ultrasonic waves. In particular, the so called additional optical phase shifts (namely: initial phase shift and v-dependent phase shift), discovered and described in 1994 by P. Kwiek (Gdańsk University, Poland) and R. Reibold (PTB, Braunschweig, Germany), are under consideration and their influence over the secondary interference planes displacement is discussed.

Moreover, it is shown that while taking into account the complex character of the diffracted light amplitudes one can perfectly explain all the effects that, previously, have been attributed to the existence of phase-amplitude diffraction grating. Besides, a comparison is done of the results of the near field diffraction theory, given in 1976 by B.D. Cook with the ones making use of NOA method, proposed in 1987 by E. Blomme. The same, hitherto existing opinions have been revised, concerning B.D. Cook's theory correctness, by indicating limited range of its applicability.

It was also shown, that within the Klein–Cook parameter range under consideration, i.e. $0 < Q < 2.5$, there were no appreciable discrepancy between experimental data and the ones obtained by means of NOA method. Thus, no doubt that, at any rate within this range of the Klein–Cook parameter values, the NOA method is really powerful tool which provides us an excellent description of ultrasonic light diffraction phenomenon. Besides, it can be easily adopted for numerical calculations which, at least theoretically, can be carried on with unrestricted accuracy.

However, one should keep in mind that NOA method is in fact an approximative one and that at present there are no explicit and unequivocal opinion concerning the range of its validity. The experimental part of the work comprised near field measurements based on three independent techniques: (i) direct one – with light detector placed immediately within the Fresnel region of diffracted light, (ii) by means of an additional lens, transforming the exit plane from the sound wave to the image plane containing photodetector, and finally (iii) the one based on the so called single photon counting technique.

The experimental setup was fully automatized and controlled by a PC computer via GPIB/IEEE488 interface. Adequate configuration of the apparatus used enabled the author to obtain 5% accuracy of the time-dependent near field light intensity distribution measurements. It is worth noting that similar or analogical quantitative results have not been mentioned so far.

The Thesis consists of 7 chapters and introduction describing the historical background of the acousto-optics roots and evolution, giving one also a wide bibliographical review. Chapter I defines the basic notions which are commonly used to describe the phenomenon of light diffraction by ultrasonic waves. It also includes a short characterization of the so called self-imaging phenomenon. In Chapter II a procedure is introduced starting from Maxwell equations and leading to the Raman-Nath equations system, while Chapter III reveals different methods usually used to find solutions to these equations. Chapter IV gives the detailed description of the Fresnel region of light diffracted by ultrasound. In succeeding subsections several diffraction cases are analyzed and periodical character of the near field is discussed. Besides, a sound field reconstruction procedure is mentioned, based on the information contained in the Fresnel region of ultrasonically diffracted light. In Chapter V the so called finite amplitude effects are discussed. Thanks to detailed analysis of considerations presented in this chapter, suitable experimental conditions were found, justifying the negligence of the mentioned nonlinear effects influence over the carried measurements results. Chapter VI contains the detailed experimental setup description, presents the results obtained as well their discussion and comparison with appropriate theoretical predictions. Chapter VII comprises conclusions and final remarks.

Surface acoustic wave propagation in selected phthalocyanine thin layers [in Polish]
by WIESŁAW JAKUBIK
Silesian Technical University, Gliwice
March 1998 Contact: Institute of Physics, Acoustooptics Staff,
ul. Krzywoustego 2, 44-100 Gliwice, Poland
e-mail: murban@zeus.polsl.gliwice.pl
Supervisor: Prof. A. Opilski

The aim of this thesis was a theoretical and experimental study of the surface acoustic wave (SAW) propagation in selected phthalocyanine thin layers from the point of view of their application in toxic gas sensors. The thickness of all the investigated layers was much smaller than SAW wavelength ($h \ll \lambda$). This is a case of small perturbations in the SAW propagation on account of the thin layer placed on a piezoelectric crystal surface.

The theoretical study was based on Auld's perturbation theory using the notion of surface impedance. In the work the influence of the mechanical and the electrical properties (conductivity, diffusion constant and electrical permittivity) of thin semiconducting layers for SAW propagation in a piezoelectric crystal was applied. It has been shown that for SAW propagation the main influence is exerted by slight changes of the surface mass loading and surface conductivity of thin phthalocyanine layers. These two phenomena (i.e. mass and electric loading of the crystal) are the base for designing toxic gas sensors in low-range concentrations with the use of sensitive organic layers (for instance various metalophthalocyanine complexes). The electric effect results from the interaction of the electric field associated with SAW with mobile charge carriers in the phthalocyanine layer. It has been found that there is a strong frequency dependence of the conductivity of thin phthalocyanine films. This phenomenon facilitated the acoustoelectric effect between the SAW electric field and the mobile charge carriers in thin films. The strong frequency dependence of the conductivity of thin phthalocyanine layers has been confirmed by direct electrical measurements using a Tesla VHF Bridge.

The sensor properties of various thin phthalocyanine films have been tested with use of the well-known experimental set of dual-delay lines. On a piezoelectric substrate (i.e. $LiNbO_3$ Y-Z) two identical circuits are formed to facilitate the propagation of the surface wave excited by means of interdigital transducers. Next, by sublimation in a vacuum in one of the paths a thin layer of metalophthalocyanine is formed. The free path of the crystal serves as a reference, permitting easy measurements of the arisen difference of frequencies. Five types of phthalocyanines (PbPc, CuPc, FePc, NiPc and H_2Pc) have been examined at various concentrations of toxic gases (NO_2, SO_2, CH_4 and NH_3). The best results have been achieved for the PbPc and NO_2 low concentrations (below 1 ppm in air at 70°C). The gases were batched with the use of mass flow controllers.

From the physical point of view it is essential that the influence of each above mentioned effect (mass and electrical) should be considered apart from its reaction to the interaction of the layer with gas. For this purpose a special experiment with a lot of measurements has been carried out. The very thin aluminium layer left in one of the acoustic paths reduced the electric potential associated with SAW to zero. The obtained results confirmed the theoretical study in the quality range. It also determined the approximate

mobility of the charge carriers in thin layers. For all the investigated phthalocyanine films the obtained values were in good agreement with other methods for amorphous structures dealt with in literature.