ARCHIVES OF ACOUSTICS
5, 1, 3-30 (1980)

AN ACOUSTIC METHOD FOR THE EVALUATION OF THE STATE OF THE LARYNX SOURCE IN CASES INVOLVING PATHOLOGICAL CHANGES IN THE VOCAL FOLDS*

RYSZARD GUBRYNOWICZ, WŁADYSŁAW MIKIEL, PIOTR ŻARNECKI

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

With a view to discussion of the role of the larynx source in the formation of the acoustic speech signal, this paper presents a method for the evaluation of the state of the larynx for different pathological states of the vocal chords. 143 persons, both healthy and with specific pathological voice changes were investigated. The pathological conditions included unilateral and bilateral paralysis of the larynx nerves, inflammation states and cancerous development, both benign (polyps, papillomas etc.) and malignant. For the analysis of the speech signal an Intonograph-77 was used, which together with a MERA-300 minicomputer provides an integral system for automatic measurement and registration of the values of successive periods of the larynx tone, currently determined, during the articulation of longer texts. The evaluation of the state of the larynx source was performed on the basis of the analysis of the phonation time and the statistical distributions of the values of periods of the larynx excitation and their percentage changes in time. A final evaluation was performed, using a diagnostic model in which fuzzy relations were used for the classification of voices into two sets-normal and pathological voices. The mean classification error did not exceed 8%, with a lack of decision occurring in about 20% of the casses.

1. Introduction

For a number of years an increasingly widespread tendency to use more or less automated measurement systems, permitting an objective evaluation of the state of specific organs of the human organism, could have been observed in medicine. Development of suitable low-cost methods and diagnostic apparatus, permitting mass screenings to be conducted, mainly for preventive purposes and the early detection of disease, is a particularly important problem. In the literature on the investigation of human speech organ there are relatively few papers concerning acoustical methods of voice evaluation, although, compared

* The paper was written under the problem 10.4.3 — Biocybernetic systems and elements.

with many traditional methods used in phoniatry, they show many advantages. First of all, they permit the taking of measurements fully or partly automatically, using microprocessors or minicomputers. In addition, they do not require the introduction of devices or sensors into a speech organ or their application outside the organ. When traditional methods are used, e.g. endoscopic, stroboscopic, bioelectrical, radiological etc., the voice examination is carried out under abnormal articulation conditions, which causes considerable uncontrolled variation in the measurements. On the account of the great discomfort to the patient in this type of examination, they are limited to voice evaluation on the basis of an analysis of some vowels uttered in isolation, and do not usually permit an investigation of the action of a speech organ during the utterance of longer texts, which gives more credible and representative results.

The above-mentioned faults of traditional methods for voice analysis are particularly conspicuous in the investigation of the activity of the larynx tone. This investigation is nearly always performed under abnormal articulation conditions and is usually limited to the evaluation of phonation during the articulation of short speech segments. Probably this is one of the reasons why these methods fail to detect early pathological changes occurring in the larynx while they are still small.

The above-mentioned limitations do not apply to the acoustical methods of voice investigation which analyze a complex speech signal measured using a microphone placed at a certain distance from the mouth of the speaker. In addition the investigation of voice using these methods can be performed on the basis of a signal recorded on an audio tape, which facilitates the storage of measurement data and the repetition of analyses as measurement and analytical methods improve.

This paper presents a method for the investigation of larynx phonation based on the analysis of the phonation time, the variation of the frequency of the fundamental tone, and the statistical distributions of the period of the larynx tone and its percentage changes. It also describes a diagnostic model assumed in the investigations for the evaluation of the state of the larynx source, which was verified for specific cases of pathological changes in the vocal chords.

2. Larynx excitation in the case of normal voices

One of the basic anatomical parts of the speech organ, which participate directly in the creation of an acoustical speech signal, is the larynx which is the upper end of the trachea through which the air pressed out of the lungs passes into the mouth cavity. The air stream which flows through the larynx is modulated in the glottal slit formed between the cyclically opening and closing vocal folds. Due to the motion of the vocal folds, air pulses are created, which excite vibrations of the air mass in the throat and mouth cavity. The behaviour of variations of the glottis opening area (and thus the similar variations of the volume velocity of the air stream) is shown schematically in Fig. 1.

Larynx excitation is the dominating kind of excitation of the vocal cavities and is for the most part the primary source of acoustic energy in a speech signal. In addition, the occurrence or lack of a voiced characteristic in physiological speech has phonetical significance, and changes in the frequency of the larynx tone or its amplitude have linquistic significance. Thus in articulation the larynx plays two important functions: (a) a phonetic function which permits sounds to be distinguished as voiced and unvoiced, and (b) a prosodic function, since changes in the intensity of speech and its fundamental frequency serve to stress chosen syllables and to create different melodic contours.

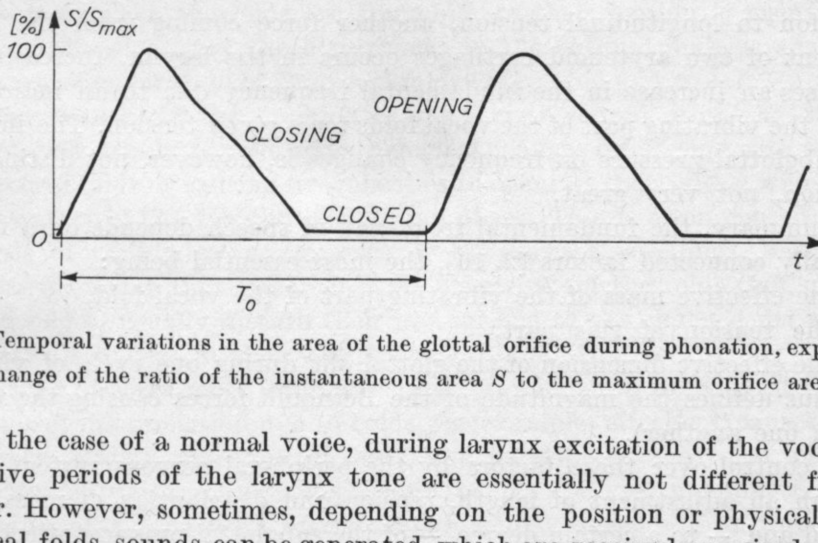


Fig. 1. Temporal variations in the area of the glottal orifice during phonation, expressed by the change of the ratio of the instantaneous area S to the maximum orifice area $S_{\max}$

In the case of a normal voice, during larynx excitation of the vocal tract, successive periods of the larynx tone are essentially not different from one another. However, sometimes, depending on the position or physical state of the vocal folds, sounds can be generated, which are received as voiced, although successive periods of the fundamental tone differ essentially one from another. This can occur in two ways. Most often successive periods are alternately shorter and longer. The periods occurring alternately become increasingly closer to each other (nearly identical) than to the adjacent ones. However, steplike changes occur relatively less often in the duration and shape of the larynx pulses.

As was already mentioned, the larynx source plays a dominant role in the formation of the prosodic features of a speech signal, which are expressed in the behaviour of the melodic contour and in changes of intensity. A third factor affecting prosodic characteristics, duration of the individual speech sounds, which is physiologically conditioned, primarily by the velocity of the supraglottal articulation elements, depends directly on the larynx excitation. The intensity of the output signal is conditioned only partly by the intensity of the larynx source, which depends mainly on the magnitude of the subglottal pressure. If, for example, this pressure increases by a factor of two, the level of larynx excitation increases by about 12 dB [2]. A number of investigations have shown that the most important prosodic characteristic is the behaviour of the melodic contour, while an exact control of changes in the fundamental frequency is essential for the creation of an utterance with a predetermined intonation.

The melodic contour is defined by the variations of the fundamental frequency which is physiologically dependent primarily on the eigenfrequency of the vocal folds. The value of this frequency is affected by such physical parameters of the folds as their length, thickness, mass of vibration, and tension. Thus tension is directed along the vocal ligaments and the frequency of the chord vibrations increases with increasing tension in the ligaments. Also, the lengthening of the folds causes the fundamental frequency to increase, due to a decrease in their thickness, and thus in the mass that is vibrating [8]. In addition to longitudinal tension, another force coming from the central attachment of two arytenoid cartilages occurs in the larynx. Increasing this force causes an increase in the fundamental frequency due to an indirect decrease of the vibrating part of the vocal folds for a given tension. The influence of the subglottal pressure on frequency changes is, however, not distinct and, in addition, not very great.

In summary, the fundamental frequency of speech depends on a number of mutually connected factors [2, 16], the most essential being:

1. the effective mass of the vibrating part of the vocal fold,
2. the tension of this part,
3. the effective dimension of the glottal slit during one cycle of vibration (the radius defines the magnitude of the Bernoulli forces causing the folds to approach one another).

The control over these factors by the peripheral nervous system of man is through an adjustment of length, tension and distribution of mass in the vocal fold and — to a certain degree — of the subglottal pressure affecting the magnitude of the Bernoulli forces.

In pathological cases involving the larynx source, disturbances in their phonation function can occur due to anatomical changes (e.g. the occurrence of polyps) or due to damage in the system which controls its work, e.g. paralysis of the recurrent laryngeal nerves.

3. Disturbances in the function of the larynx source, due to pathological changes in the vocal chords

The investigations presented in this paper involved disturbances of voice caused only by pathological changes in the larynx. Disturbances in its phonation function due to changes in the endocrine glands, in the central nervous system, and personality and emotional changes were not considered.

Most often the faulty operation of the larynx source is a result of a failure of the glottis when periodically closing, to do so in a complete manner, or of asynchronous motion of the vocal folds. The failure of the glottis to close totally and the absence of a rapid break in the flow of the air stream cause the higher harmonics of the larynx tone to have a lower level than in the case where full closing occurs. This is due to the fact that variations in the volume velocity of the stream with incomplete closing of the glottis are considerably smaller than when there is a full closing of the chords, which is necessary for a suitable increase in the subglottal pressure to occur. In addition when, at the moment of the maximum proximity of the folds to each other, the area of the glottal slit is sufficiently large, turbulence can occur in the air flow. The absence of a closing phase for the vocal folds usually causes the voice to be weakened, and often coarse. It is often accompanied by a shortening of phonation time and, in more serious cases when the folds are immobile or their motion is insufficient, by aphonia. Almost always the disturbance in the closing phase is accompanied by asynchronous motion of the vocal folds [11].

Disturbances in the function of the larynx source, from the anatomical point of view, can be caused by pathological changes interfering with the motion of the vocal folds or causing irregularities to occur at their edges, which prevent exact contact of the folds in the closing phase. This is often caused either by paralysis of certain larynx muscles or by growths, swellings, scars, polyps. It is worth noting that larger growths occurring on the folds (not necessarily at their edges) usually disturb their motion due to the changed mass distribution and differences in their elasticity [11], for example.

Of the diseases most often affecting the phonation activity of the larynx all states of inflammation due to colds, for example, effusions (caused by excessive vocal effort), cancerous changes, and paralysis of the vocal folds should be mentioned. Cancerous changes, both benign (polyps, papillomas etc.) and malignant, affect the work of the larynx source, depending on their size and location. When they are not very large and occur on the edges of the folds, the changes in the voice which they cause, can be unnoticed. This is particularly critical in the case of malignant tumours, which can be successfully cured, provided they are detected early enough. Otherwise, it is usually necessary to perform a partial or total laryngoctomy. Adduction and also obduction of the vocal chords can also be limited as a result of paralyses or paresis of the larynx muscles. These paralyses, which are usually unilateral (less often bilateral), are most frequently caused by the damage to larynx nerves during surgical operations on the thyroid. It is a rather complicated operation and even faultless operational technique does not sometimes protect the patient from the damage to larynx nerves which run very close to the thyroid gland, usually in an untypical way.

The work of the larynx source in cases of a unilateral damage depends predominantly on the position of the paralysed vocal fold. If it is close to the central line of the glottis (in the so-called medial position), it can function in an acceptable manner, since the healthy fold reaches the damaged one without difficulty and no changes can be observed in the voice. In the case of a more distant position of the damaged fold, the healthy fold cannot cross the centre line and a full closing cannot thus occur [11].

In general, from the acoustic point of view, malfunction of the larynx source can result in time variations of the fundamental tone, the intensity of the speech signal, the duration of the voiced excitation and the voice quality.

The pitch is determined by the frequency of the larynx pulses which, in the case of pathological changes in the vocal folds, can be subject to considerable random variations, as a result of a decrease in the tension and flexibility of the folds and also of changes in the spatial distribution of their mass. These factors decisively affect the speed of motion of the folds and the glottal closing time. The energy generated in each successive larynx pulse has a dominating effect on the behaviour of the level of speech intensity with time; thus, for example, asynchronous motion of the ligaments or factors affecting the degree of glottal closing play an essential role here, since the excess subglottal pressure generated just before the closing of the folds depends on these. Also, the phonation time, which in the case of incomplete closing of the glottis decreases due to an increase in the rate of flow of air through the larynx, depends on the degree to which the glottis closes. Voice quality, which is defined in a rather subjective manner, depends on the factors mentioned above. In particular, asymmetric motion of the folds is reflected in a hoarseness, weaker sonority and an unnatural sound of the voice, although the closing may be nearly complete.

4. Methods for the investigation of the phonation activity of the larynx source

In laryngology and phoniatry generally accepted clinical methods for voice investigation exist, which can be divided into two basic groups: (a) audio monitoring and (b) visual monitoring. Since structural changes in the larynx are often accompanied by changes in the voice, the performance of suitable articulation tests can supply the physician with essential information [4]. The main fault of audio monitoring methods is their subjective character, which results in more difficult cases when changes in the voice are not distinct, in the possibility of even experienced phoniatrists giving different diagnoses for the same patient. In addition, there is a lack of quantitative standards for audio evaluation of the state of the larynx source.

Visual methods, e.g. endoscopic, stroboscopic, electromyographic, radiological etc., enable the physician to perform a fairly reliable evaluation of the operation of the larynx source in cases with advanced pathological changes. These methods, however, mainly because they are time-consuming and expensive, are not useful for mass examinations which are becoming increasingly necessary for the early detection of disease, particularly those of the cancerous type [5]. In addition, when using these methods it is, in general, impossible to perform a long-term investigation of the phonation activity of the larynx source during the utterance of longer texts, mainly because these must occur under abnormal articulation conditions which are distorted by devices and alien bodies introduced into the speech organ and by the performance of additional tiring activities. These investigations are, therefore, usually limited to the observation of the larynx during the articulation of isolated vowels and, in many cases, when structural changes in the larynx are slight, disturbances in phonation can occur in a random and rather infrequent manner, so that they may pass unnoticed in this shorter observation time.

Acoustical methods, which are free from the faults mentioned above, investigate a speech signal recorded by a microphone placed at a fixed distance from the speaker's mouth. In the case of an investigation of the larynx source, a special measurement methodology must be developed, which permits information to be obtained from the analysis of such a complex signal as the acoustical speech signal. The final form of the signal is influenced not only by the frequency characteristics of the larynx source, but also by transmission characteristics of the throat-mouth (and possibly nasal) tract and by the mouth radiation characteristic.

A basic acoustic parameter that is analysed in the evaluation of the state of the larynx source is the behaviour of the fundamental frequency in the speech. LIEBERMANN [12], who was the pioneer in the investigation of this type behaviour for cases with pathology of the larynx, analysed the distribution function of the differences between successive periods of the larynx tone and introduced a synthetic index – the so-called *pitch perturbation factor* – defining the probability of occurrence of differences in excess of 0.5 ms. Liebermann showed that this index is larger for cases with pathological changes of the larynx than for physiologically normal states. KOIKE [10] and DAVIS [5] used an evaluation method of a similar type, although slightly modified, for the assessment of the state of the larynx source.

The variation in the intensity of the harmonic excitation in a speech signal is another parameter which, it seems, should be more widely used for evaluation of the state of the larynx source. However, the measurement is very difficult and complicated. Davis [5] was one of the first investigators to use an analysis of the intensity variation of the excitation for the evaluation of the phonation activity of the larynx. He investigated the distribution of the amplitudes of the larynx pulses by inverse filtering a signal recorded during a lengthened articulation of an isolated vowel /a/. He found that in the cases with pathology of the larynx, the distribution is usually more diffuse and has a smaller kurtosis than that for normal voices. This is most probably due to the fact that additional larynx noises, caused by turbulent air flow through the glottial slit, often occur in pathological cases.

Recently investigators have attempted to evaluate the pathology of the larynx source by an analysis of the whole spectral structure of the speech signal, and not only chosen parameters connected directly with the source [5, 6, 10]. It appears [5], however, that the analysis of the behaviour of the frequency and amplitude of the larynx excitation gives most information on the phonation activity of the larynx.

Acoustical methods, which are being developed for the evaluation of the voice forming activity of the speech organ, should be assessed from two points of view. They must first of all permit a differentiation between normal and pathological states; and subsequently must provide a quantitative measure of the changes occurring in the voice during the course of the curing and rehabilitation procedures. Further development of these methods will probably be towards the provision of an objective evaluation of the degree of advancement of the pathological states, with particular emphasis on their early detection.

5. Bases of the diagnostic model used in the investigations

The primary task of medical diagnostics is first of all to establish if the person presenting himself for examination is healthy or not. It is only later that the type of disease is identified and a decision can be made about the proper treatment. In the course of therapy, diagnosis serves to follow the advancement of the treatment and permits it to be ended at the proper time.

Automation or, at least, computer-aiding of the diagnostic process, is becoming more and more attractive for physicians and scientists. Most of investigations performed in this field are based, however, on statistical decision models, and the results obtained have not been very promising so far from the point of view of medical applications [19]. The criterion that is most often used in decision-making is Bayes' criterion based on the assumption that the best set of patterns is the one which gives least probability of classification error. In practice, the creation of such a set requires an exact knowledge of the multidimensional distributions of conditional probabilities, which involves the need to possess a sufficiently large data base, the dimensions of which are, more often than not, unattainable.

Thus, it is often necessary to use approximate distributions which usually deviate considerably from the real distributions. In addition, all attempts at expanding the model to include new diseases require the recalculation of all the conditional probability distributions, (based on a wider data base) and a complete change of the decision-making rules. In practice, the precision of classification with different statistical decision-making models is unsatisfactory, and when tested on an open set of data does not usually exceed 70 % [15]. It seems that further development of statistical decision-making models will not bring an essential improvement and it is thus necessary to use a completely different approach to the diagnostic process.

In a classical diagnostic technique used in medicine, the state of the human organism or a particular organ is evaluated by the physician on the basis of a number of symptoms in a manner which is, to a dominating degree, subjective (based on experience acquired in the professional practice). Even those of symptoms that can be measured, e.g. temperature, are described in an approximate manner, using designations: normal, slightly increased, increased, high, very high etc. temperature. Individual symptoms together form a characteristic set and diagnosis consists in selection of the disease most typical for the given set of symptoms of disease. The use here of statistical classification methods necessitates the formation of an exact quantitative description of the disease phenomena investigated, which are naturally described in a more or less approximate way.

In the diagnostic model used in this work it was assumed that the analysis of individual symptoms (features) is expressed in mutually opposing fuzzy sets [20], i.e. less formally in sets, where a sharp change from full assignment to nonassignment to the set does not occur, but where intermediate assignment stages can also exist. In this approach, the degree of assignment (so-called *grade of membership*) is expressed by a pair of numbers in the range [0, 1], the former indicating nonassignment, and the latter assignment to a given set.

It is assumed that N is a subset in the space X. A function $\mu_N(x)$ is defined, called the *characteristic function* of the set N, giving for each element x in the space X the degree of its assignment to the subset N in terms of the number $\mu_N(x)$ in the range [0, 1]. When $\mu_N(x) = 1$, the element x belongs to N, while the condition $\mu_N(x) = 0$ corresponds to the fact that x does not belong to N.

If each element x is determined, using a series of measurable properties $C = C_1, C_2, \ldots, C_n$, it is assigned to a suitable value of $\mu_N(x)$ on the basis of an analysis of variations of the individual parameters, followed by a so-called composite inference rule [20], based in this case on the linguistic characteristics of the individual features.

An example of voice pathology will serve to give a brief explanation of this method of feature description. As was mentioned in Section 2, in most cases of larynx pathology an decrease in the phonation time can be observed with respect to normal voices. If the phonation time is short, the hypothesis that pathological changes occur in the larynx cannot be rejected. If, however, the time is long, it can be assumed that the larynx source functions normally. In the present paper, the ratio of the total duration of voiced excitation to the whole duration of the utterance was used. Theoretically, this ratio can take any value in the range [0, 1], depending on the type of text uttered, the emotional state of the speaker or the type of disease. In the case of aphonia, for example, it is close to zero. In practice, for a normal voice and phonetically balanced spoken text, this ratio does not usually exceed a value of 0.85.

As an example, Fig. 2 shows the behaviour of the assignment function $\mu_L(A)$ for a fuzzy subset of "low values", which takes the following values:

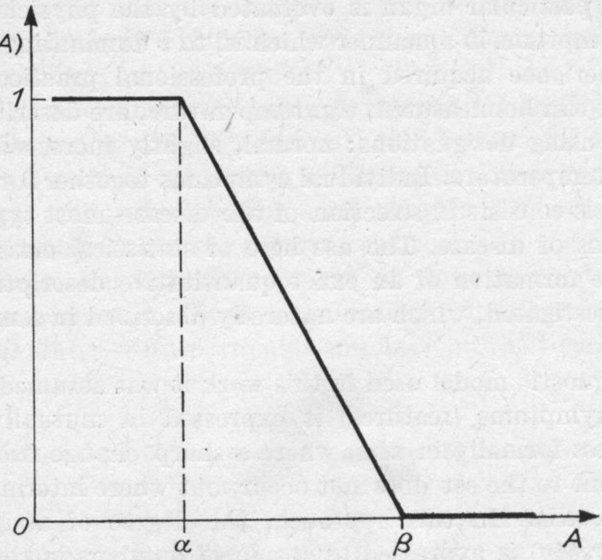


Fig. 2. An example of the behaviour of the assignment function $\mu_L(A)$ of a fuzzy subset "low values" of A

$$\mu_L(A) = \begin{cases} 1 & \text{for } 0 \leqslant A \leqslant \alpha, \\ (\beta - A)/(\beta - \alpha) & \text{for } \alpha < A < \beta, \\ 0 & \text{for } \quad A \geqslant \beta. \end{cases} \tag{1}$$

Thus, in the case of aphonia, $\mu_L(A) = 1$. In an analogous manner, the subset of "high values" can be defined, for which the assignment function $\mu_H(A)$ is connected with $\mu_L(A)$ by the following relation:

$$\mu_H(A) = 1 - \mu_L(A). \tag{2}$$

Determination of the characteristic points α and β of the assignment function $\mu(A)$ is performed as a result of a learning process, based on utterances by normal voices, with long phonation times, and by pathological voices, with short phonation times. The quantity α is defined as that value of the parameter A, below of which the phonation time is low, while that of β is that value above of which its value is certainly high. A similar description is made for the remaining features, with the assignment function of type μ_H or μ_L being determined for each separately. In this manner, the state of the larynx source can be described, using a fuzzy relation of its features, and the degree of assignment of the ith investigated voice to the set of pathological voices or the set of normal voices can be determined, based on the formulae

$$\mu_P(i) = \bigvee_j \mu_{R_p(C_j)}(a_{ij}), \qquad \mu_N(i) = \bigvee_j \mu_{R_n(C_j)}(a_{ij}), \tag{3}$$

where $\mu_P(i)$ and $\mu_N(i)$ define the degree of assignment of the ith voice, to the set P of pathological voices and set N of normal voices, respectively. $R(C_j)$ denotes the fuzzy limit of the value a_j of the parameter A_j, defining a subset of values $\{a_j\}$ having a common property, according to the assumed description (e.g. A_j is *low*); $\bigvee_j$ is the operator of determining the upper limit with respect to j.

The values of the degree of assignment μ_p and μ_N are the criteria for making the decision as to which set an analysed voice will be assigned. The analysed voice is assigned to the set whose degree of assignment is greater; with an additional condition of the form $(\mu \geqslant d) \vee (\mu < 1-d)$, where $d = 0.07$, being introduced before a decision can be made. The final form of the decision-making rule is thus the following:

$$\begin{aligned} P_d &= \{i \,|\, (\mu_p(i) \geqslant d) \wedge (\mu_N(i) < 1-d)\}, \\ N_d &= \{i \,|\, (\mu_p(i) < 1-d) \wedge (\mu_N(i) \geqslant d)\}. \end{aligned} \tag{4}$$

When these conditions are not satisfied, the voice analysed was assigned to a "no-decision" set.

6. Measuring system

The frequency of the period of the larynx tone can be measured in many ways and for normal voices (with individual exceptions) the measurement is not usually a very difficult technical problem. In pathological cases, however, the measurement is not easy (see [3]), routine digital methods of larynx tone extraction fail totally and it is necessary to use tiring, semi-automatic procedures. This was one of the reasons why an analogue method, which imposes less limitations on a signal and at the same time permits real time measurements to be made, was used to extract the larynx tone from the speech signal.

In the investigations a measuring system was used which permitted measurement of the following running value of changes in the period (frequency) of the larynx tone, the percentage variation recorded for each period, and the relative duration of voiced excitation (expressed as the ratio of the duration of larynx excitation to the duration of total utterance). A schematic block diagram of the system is shown in Fig. 3.

An Intonograf-77, designed and produced in Acoustics Laboratory of the Institute of Fundamental Technological Research of the Polish Academy of Sciences, whose design and operation has been discussed previously [14], was used to extract and to measure the period of the larynx tone. Two filter systems were used in the device to extract the fundamental tone. One was fixed with a pass band of 78 to 500 Hz, and the other was used as a low pass tracking filter, which retuned automatically so that the fundamental frequency of the speech signal was always slightly above its limiting frequency. The period of a signal thus filtered is determined in a continuous manner by measurement of time between two succesive unidirectional zero crossings of the signal (the measuring precision of a period is 0.05 ms). In the case of the absence of voiced excitation, the signal quantization time automatically assumes a constant value of 12.5 ms, and the register of the device is zeroed.

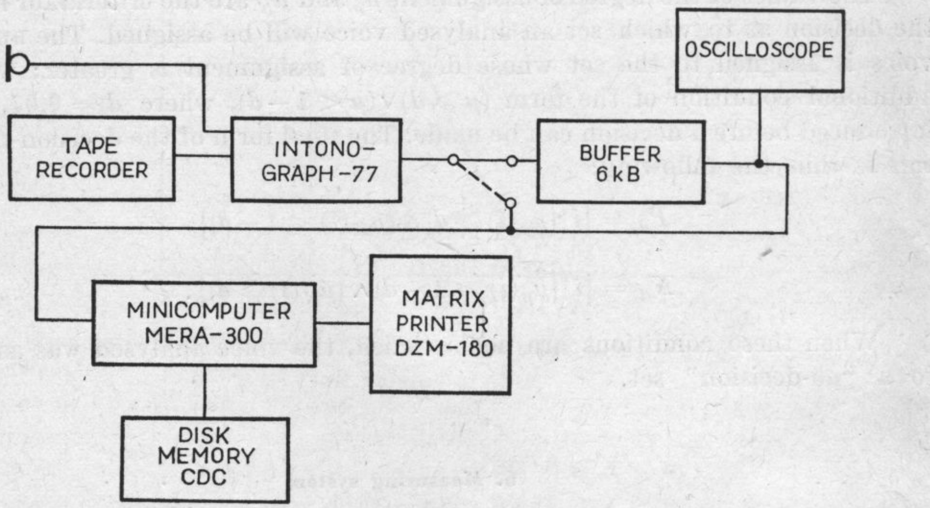


Fig. 3. A schematic block diagram of the measuring system used in the investigation

In the present measuring system, the Intonograf is connected on-line to a Mera-300 minicomputer which has 8-bit words and a memory cycle time of about $1\mu s$. A 7502 Brüel and Kjaer Digital Event Recorder was used in the system for buffering the digital data from the Intonograf. It was thus possible to check the data visually on an oscilloscope screen before sending it to the minicomputer. The software of the measuring system consists of a series of programmes for plotting the frequency variation of the larynx tone with time, for plotting statistical distributions of the values of the periods of larynx excitation and of the values of the relative differences between successive periods, and for the determination of distribution parameters such as the mean values, various dispersion measures, coefficients of asymmetry and kurtosis, modes, medians etc. The relative differences between successive periods of the larynx tone were determined according to the formula

$$\Delta TO = \frac{2(TO_i - TO_{i+1})}{TO_i + TO_{i+1}} \tag{5}$$

for $TO_i \wedge TO_{i+1} \neq 0$, where i is the number of the (successive) periods measured.

The distribution of the values of ΔTO can be considered, as an approximation, to be equal to the distribution of the values of the derivative of the function $TO(t)$ with respect to time.

To date this has not been discussed in the literature.

Part of the programmes was written in assembler, while the remainder (concerning statistical analysis) were written in a limited version of Fortran adapted for the minicomputer. The plot of the variation of F_0 was made immediately at the end of a measuring cycle, and the statistical processing of the data lasted about 5 min. An example of the printout of the results of the statistical calculations and the distributions of TO and ΔTO [%], obtained for a voice with polyp changes of a vocal fold, is shown in Fig. 4.

7. Investigation material

The investigations were performed on 143 persons, including 54 with normal voices (26 men and 28 women), 54 persons (including 37 women) with paralysis of the larynx, caused by injury to the recurrent nerves following operation on the thyroid gland, and 18 persons (including 13 women) with polyp changes on the vocal folds. The other disease states included chronic laryngitis, angiomas, larynx cancers, etc. In addition, for 12 of the subjects, several measurements were made at different stages of the treatment process and in the course of rehabilitation. Prior exact laryngological and phoniatric examinations were performed on all patients to exclude all other causes (outside the larynx) which could essentially affect the respiratory and voice-making functions of the larynx.

Healthy subjects also underwent exact phoniatric examinations.

The material used for measurement and analysis was a speech signal recorded on audio tape, of the individual utterance of two texts, successively: (a) a text of three pairs of sentences with the same phonetic structure and differing only in the intonation pattern (statement — question), and (b) a so-called newspaper text, i.e. part of a press article of adequate length to be phonetically balanced. The reading time of the second text lasted about 1 minute.

8. Analysis of the measurement results and the selection of parameters for a description of the condition of the larynx source

On the basis of the results obtained previously [7], the present investigations concentrated on the analysis of the intonation contours in questions and on the statistical distributions of TO and ΔTO, determined for a newspaper text.

```
LEOKADIA S          L.51
POLIPOWATY TWOR PRZY LEWYM FALDZIE NALEWKOWYM
NAGR. 30.05.1975

                    CZAS (MS)   (%)    PROBKI  (%)
    CALA PROBA      44862.54  100       6144  100
    BRAK SYGNALU    22124.99   49.3     1770   28.8
    ODRZUCONE          50.89    0.1       21    0.3
    SYGNAL          22686.64   50.5     4353   70.8
    ODRZUCONO PROBKI DLA T <  1.69 MS   ORAZ DLA T > 12.69 MS

    PARAMETRY ROZKŁADU TØ

                     SREDNIA =   5.21 (MS)    (191.87 HZ)
                   WARIANCJA =   2.25
     ODCHYLENIE STANDARDOWE =   1.50 (MS)
                   NORMOWANE =   0.2878

    PARAMETRY CENTRALNE        ASYMETRIA =   2.1401
                                  EKSCES =   6.8302

    PARAMETRY POZYCYJNE        MODA 1 :   4.52 %  DLA   4.89 MS
                               MODA 3 :   4.38 %  DLA   4.89 MS
                               MODA 5 :   4.20 %  DLA   4.94 MS
                   MEDIANA :   4.94 MS  (202.02 HZ)
       KWANTYLE        P.10 :   4.04 MS
                       P.25 :   4.64 MS
                       P.75 :   5.34 MS
                       P.90 :   6.29 MS
                      ROZPROSZENIE =   0.34 MS
                         NORMOWANE =   0.0707
             ASYMETRIE   P.10-P.90 =   0.1999
                         P.25-P.75 =   0.1428
                            EKSCES =   0.1555


    PARAMETRY ROZKŁADU DTØ

    PRZETWARZANO WARTOSCI DTØ W ZAKRESIE OD  -150  DO   150  %
                     SREDNIA =  0.13 (%)
                   WARIANCJA = 668.42
     ODCHYLENIE STANDARDOWE =  25.85 (%)

    PARAMETRY CENTRALNE        ASYMETRIA =  -0.2591
                                  EKSCES =   5.3507

    PARAMETRY POZYCYJNE        MODA 1 :   4.43 %  DLA DTØ =    0
                               MODA 3 :   4.31 %  DLA DTØ =    1
                               MODA 5 :   4.18 %  DLA DTØ =    1
                   MEDIANA :   0.00  %
       KWANTYLE        P.10 : -22.99
                       P.25 :  -6.99
                       P.75 :   7.99
                       P.90 :  22.99
                      ROZPROSZENIE =   7.50
             ASYMETRIE   P.10-P.90 =   0.0000
                         P.25-P.75 =   0.0666
                            EKSCES =   0.1630
```

Fig. 4a

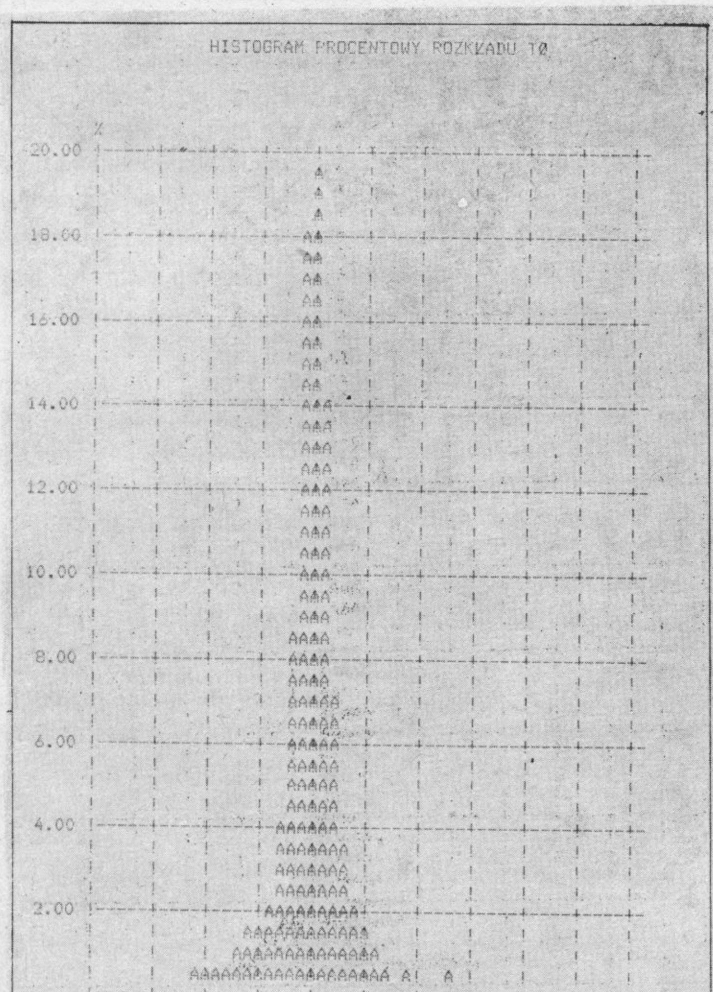

HISTOGRAM PROCENTOWY ROZKŁADU TØ
%
20.00
18.00
16.00
14.00
12.00
10.00
8.00
6.00
4.00
2.00
1.25 2.5 3.75 5 6.25 7.5 8.75 10 11.25 12.5 (MS)
800 400 250 200 160 125 100 80 (HZ)

Fig. 4b

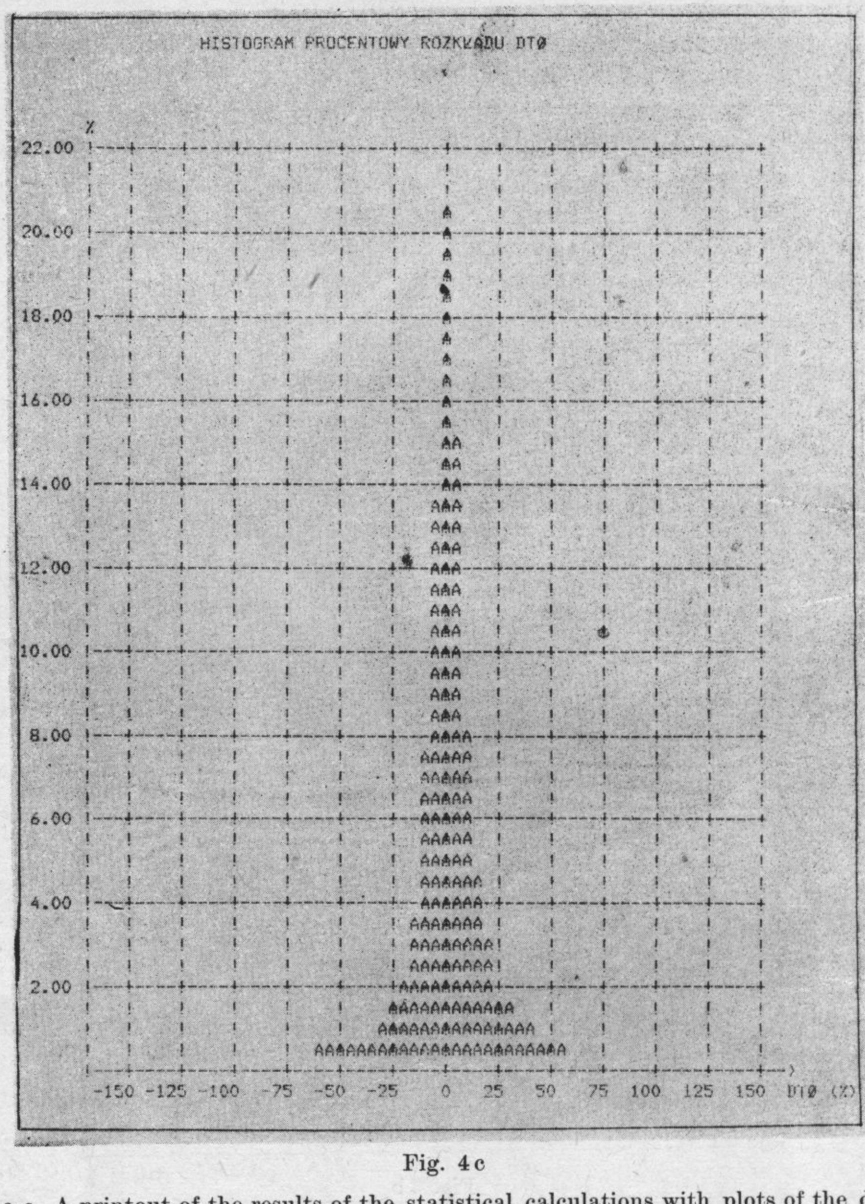


Fig. 4c

Fig. 4a-c. A printout of the results of the statistical calculations with plots of the distributions of *TO* and *TO* obtained for a voice with polyp changes on the left-hand vocal fold

For a general evaluation of the state of the larynx source, both at the time of reporting and in the course of the patient's rehabilitation, an analysis of the melodic contour of recorded questions can give the phoniatrist much information. Firstly, it can easily be estabilished the regularity ("smoothness") of tem of the measured larynx tone. Secondly, it can readily be observed the degree to which the patient is able to raise the intonation at the end of a sentence. In the first case, the occurrence of relatively large steplike changes in the frequency F_0 can indicate asynchronous vibration of the vocal folds (e.g. due to paresis, polyps ets.). In the second case, the inability to raise the intonation at the end of a question results from an unnaturally increased tension in the muscles of the vocal folds, which, as a rule, occurs in all the more serious pathological changes of the larynx.

As an example, Fig. 5 shows the variation of the frequency of the larynx tone in a word "domu" uttered in the question: "Czy idziemy do domu?" /t͡ʃi id͡zemi do domu/ measured on a female voice with a paralysed right vocal fold in the near-centre position. This variation was recorded during the patient's first visit (65a) and one month later, in the course of rehabilitation (65b). In both tests the patient was able to raise the intonation at the end of the phrase, with the increase being more distinct in the second test (about 50 % as opposed to 30 %). However, in the case considered it is more significant that immediately after the initial increase in the frequency F_0, a smooth fall in the final part of the phrase, together with numerous, rather large steplike changes (of the order of 10-15 %) between successive periods of the larynx tone occurred. Thus, the poor condition of the larynx source was only evident at the end of phonation.

Analysis of this type of behaviour is very convenient, since it gives a phoniatrist a general image of the state of the larynx source, viewed from the point of the acoustical signal generated. Current visualization of frequency changes in the larynx tone can be a valuable aid for the rehabilitation of the voice, for learning the correct formation of phrases with suitable melodic contours with preset intonation. However, when an unambiguous quantitative evaluation of the state of the larynx is required, this type of behaviour is less convenient and requires tedious analyses, which excludes the possibility of using it for mass examination.

For this reason, the analysis of the statistical distributions of selected physical quantities, connected in an unambiguous manner with the character of larynx excitation, is more suitable for the quantitative evaluation of the state of the larynx source. These quantities may be, for example, the fundamental frequency connected directly with the vibration frequency of the vocal chords and the amplitude of the larynx pulses, which, however, is very difficult to determine unambiguously for a continuous speech signal. Such a procedure is easier in the case of vowels spoken in isolation (cf. [6]). However, the deter-

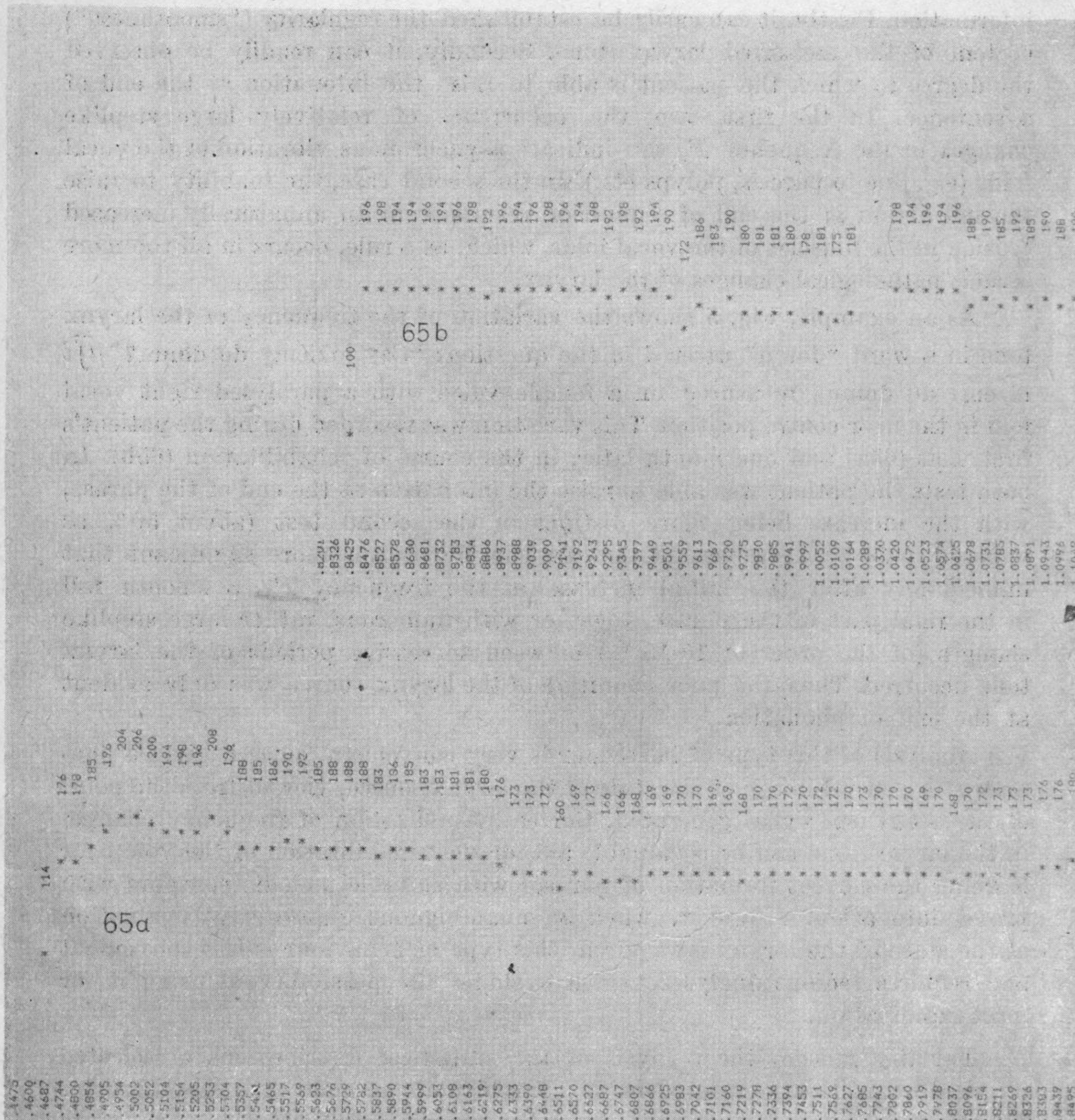
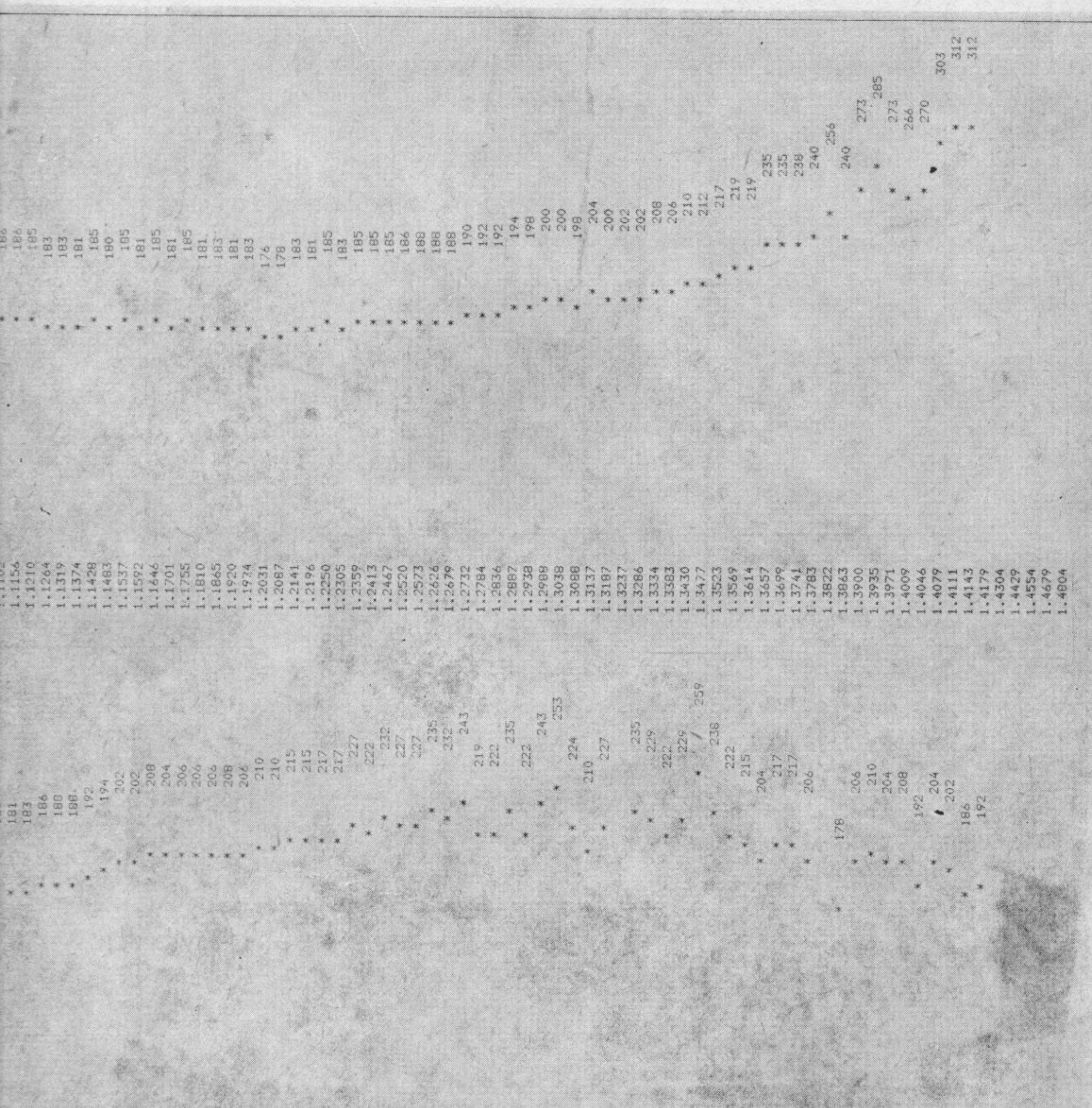
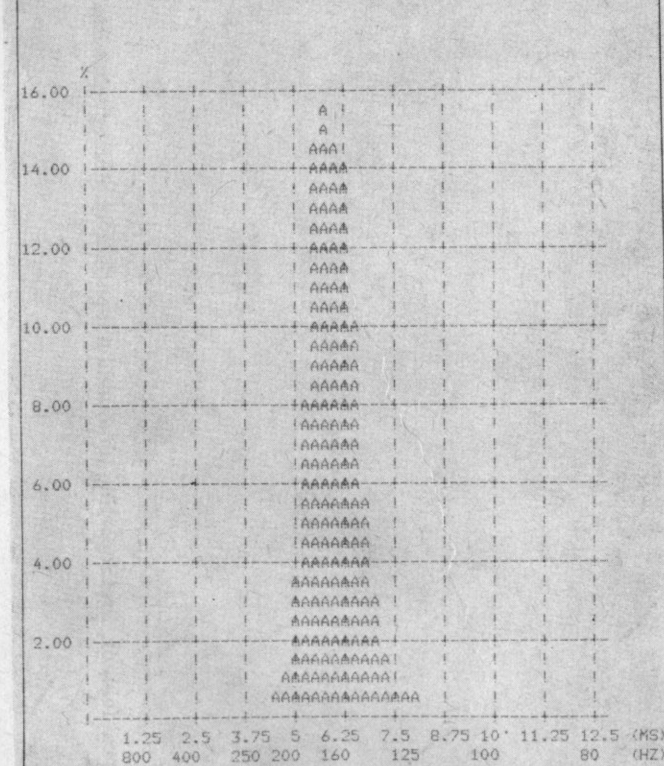


Fig. 5. Temporal variations of F_0 in the word /domu/ uttered by a female voice with a paralysed

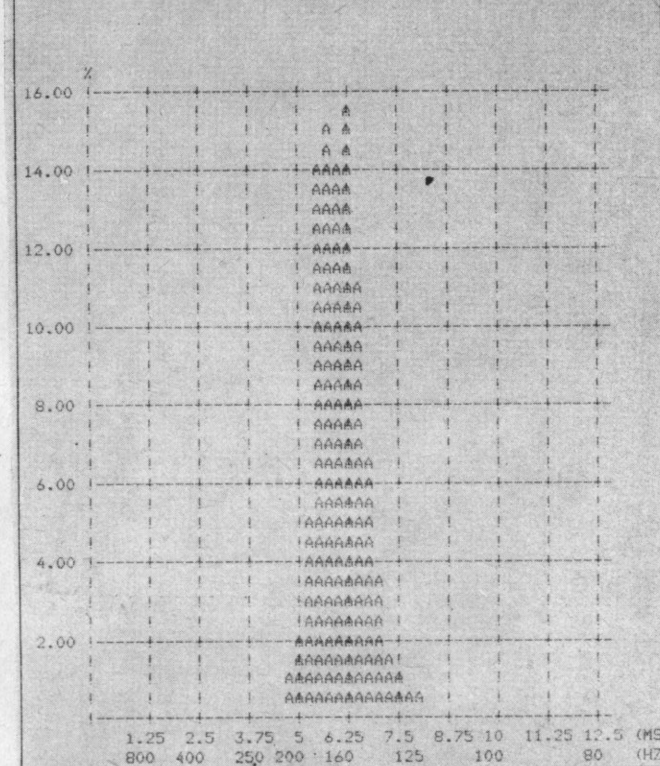


right vocal fold, on first reporting (65a), and a month later, during rehabilitation (65b)

41a


HISTOGRAM PROCENTOWY ROZKŁADU TØ
%
16.00
14.00
12.00
10.00
8.00
6.00
4.00
2.00
1.25 2.5 3.75 5 6.25 7.5 8.75 10 11.25 12.5 (MS)
800 400 250 200 160 125 100 80 (HZ)

41b


HISTOGRAM PROCENTOWY ROZKŁADU TØ
%
16.00
14.00
12.00
10.00
8.00
6.00
4.00
2.00
1.25 2.5 3.75 5 6.25 7.5 8.75 10 11.25 12.5 (MS)
800 400 250 200 160 125 100 80 (HZ)

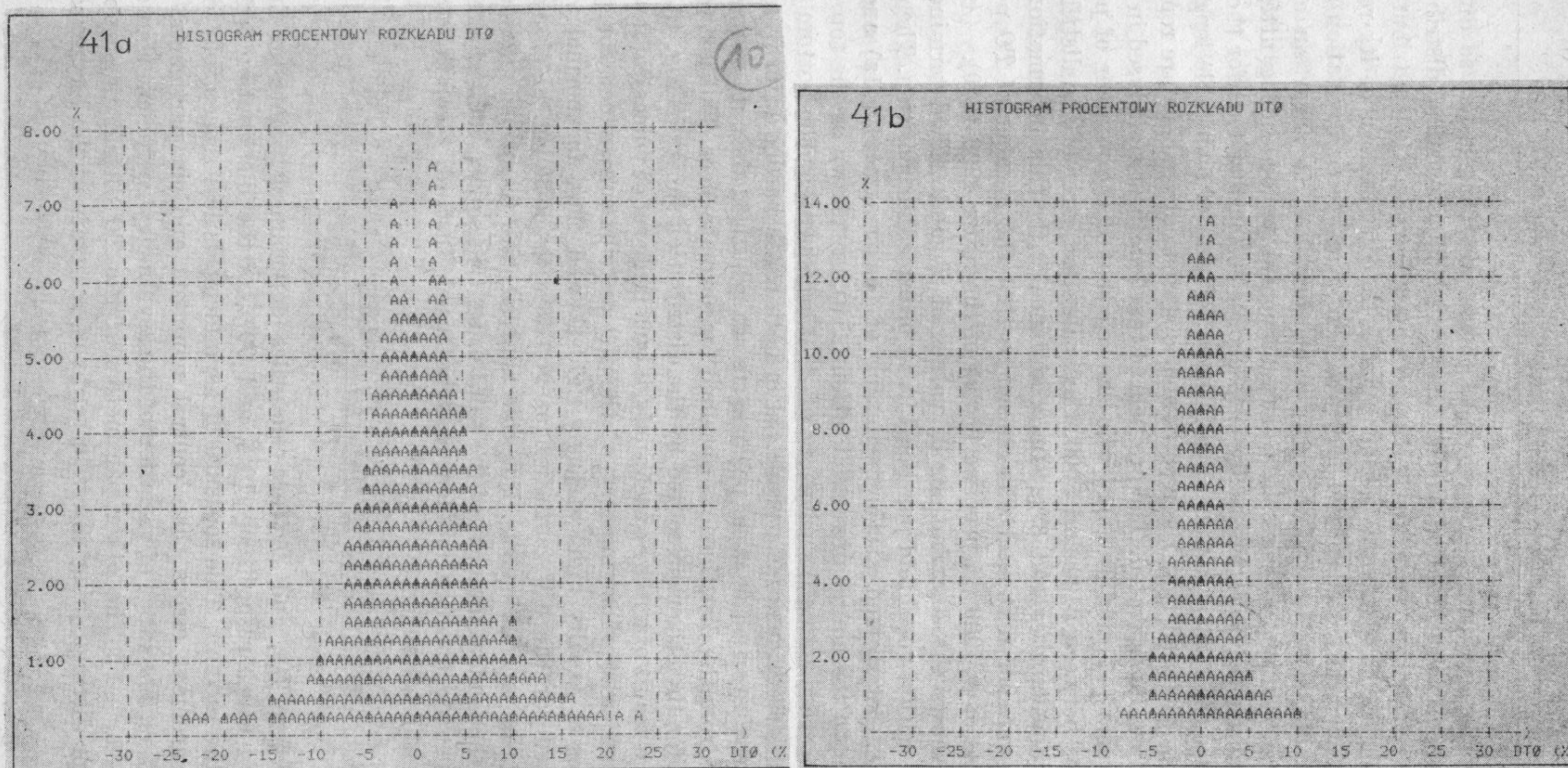


Fig. 6. The distributions of the values of TO and ΔTO for a male voice with a paralysed left vocal fold, determined at the beginning of treatment (41a) and five months later (41b)

mination, of the frequency (or period) of the larynx tone is not so difficult.

This paper is thus limited to the analysis of the statistical distributions of only one parameter connected with the phonation, i.e. the distribution of the values of the fundamental period T_0 and the distribution of its relative changes in time, ΔT_0, determined for each period of voiced excitation according to formula (5). The second distribution describes, to a certain extent, the stationarity of the changes in the fundamental period during utterance. It is essential during statistical analysis to choose the proper text for the utterance recorded, so that it is phonetically balanced and of a suitable length for it to be safely assumed that the results of the statistical analysis are representative for voices investigated. The problem of text length was analysed in detail by Ramishvilli [17] and Boë [1] who showed that if the number of periods measured lies between 2400 and 5000, the evaluation error of the distribution parameters does not exceed 3%. In accordance with this information, the length of newspaper text read was chosen so that the number of *TO* periods measured was over 4000 (in practice it was of the order 4200-4600).

Fig. 6 shows the distributions of values of T_0 and ΔT_0 determined for a male voice with a vocal fold paralysed in an intermediate position. The distributions of T_0 determined at the beginning of the treatment (41a) and five months later, in the course of rehabilitation (41b), are shown at the top of the figure. The corresponding distributions of ΔTO are shown at the bottom of the figure. The distributions of T_0, presented in both cases, hardly differ from each other and have nearly the same parameters. In addition, they do not deviate significantly from the distributions obtained for normal voices. This last remark is also valid for the case represented by the distribution shown in Fig. 4.

Only as a result of a more exact analysis on a large set of voices did it appear that the coefficient of asymmetry of the *TO* distribution (which for pathological voices is usually more strongly right-skewed than for normal voices) could be used for the differentiation of voices with pathological changes in the larynx from normal voices. For this analysis, the traditional measure of the ratio of the central moment of the third order to a square root of the third power of the second order moment was not used. Instead, a coefficient defined [18] by the formula

$$As = \frac{D_{0.9}+D_{0.1}-2\mathrm{ME}}{D_{0.9}-D_{0.1}}, \tag{6}$$

was used, where $D_{0.1}$ and $D_{0.9}$ are quantiles determining points on the *TO* axis below of which respectively 10% and 90% of the distribution is found, and ME is the median. This measure of asymmetry is less susceptible to incidental measurement errors than central moments and, in particular, those of higher orders. The range of variation of the asymmetry coefficient thus determined is $[-1, +1]$. The other parameters of the *TO* distribution appeared less useful for distinguishing voices in terms of the state of the larynx source.

However, the distribution of ΔTO gives a great deal of essential information on the phonation function of the larynx. A relatively large dispersion of the distribution suggests unstable behaviour in $T_0(t)$ and that numerous, steplike changes of large amplitude occur in it. Also, the asymmetry of this distribution can be used for differentiation of voices. In the example in Fig. 6 it can be observed that the dispersion of the ΔTO distribution after five months of rehabilitation has distinctly decreased (by nearly 50%). The standard deviation or the coefficient defined from the following formula [18] can be used as a measure of the dispersion:

$$R = \frac{D_{0.75} - D_{0.25}}{2}, \tag{7}$$

where $D_{0.25}$ and $D_{0.75}$ are the quartiles of distribution, defined in an analogous way to the quantiles occurring in formula (6). The investigations showed that the R coefficient is slightly better for the differentiation of voices than the standard deviation. This coefficient does not usually exceed a value of 0.05 for normal voices. In the case considered the asymmetry also decreased a little, although it is less visible on the plots shown.

Three parameters connected with distributions analysed, i.e., the asymmetry coefficients for the TO and ΔTO distributions and the coefficient R of dispersion in the distribution ΔTO, were finally used for the description of the state of the larynx source. In addition, another parameter, the relative phonation time τ_{ph}, which has been discussed more fully in Section 5, was used for the evaluation.

9. Classification of voices according to the diagnostic model assumed

The model decribed in Section 5 was assumed for the classification of voices into normal and pathological, in terms of the state of the larynx. Each voice was defined, using the four parameters described in Section 8. For each of them separately, the assignment function was heuristically determined for the set of low values (see Fig. 2), at the same time assuming that for normal voices the dispersion R is small, the relative phonation duration τ_{ph} is long, and that the right-skewed asymmetry coefficients of both distributions are small. The values of the parameters corresponding to points α and β were determined for each assignment function, based on the analysis of data from normal voices and voices with paralysis of the larynx. It was found necessary to determine these functions for male and female voices separately. These values are given in Table 1.

On the basis of the behaviour determined for the assignment function, the corresponding degrees of assignment of a given voice to a set of voices with pathological changes in the larynx were determined from the values of the individual parameters. Thus each utterance was described using four degrees of assignment.

The process of classification had two stages. In the first stage only two parameters, R and τ_{ph}, were considered, with the decision-making rule defined by formula (4) being used. When the conditions of the assignment to one of the sets were not satisfied, the two remaining parameters were added to the decision-making rule. The classification precision thus obtained was 72 % while a lack of a classification decision for the assignment of a given voice to one of the sets occurred in about 20 % of cases. Detailed results of the classification of the voices are shown in Table 2.

Table 1. Values of α and β for the assignment functions of the individual parameters

voices		parameter			
		R	τ_{ph}	As_{TO}	$As_{\Delta TO}$
male	α	0.01	0.5	-0.05	-0.6
	β	0.07	0.7	0.3	0.6
female	α	0.03	0.5	0	-0.8
	β	0.1	0.7	0.4	0

Table 2. Results of the classification of voices from the point of view of the condition of larynx

voices	standard			paralyses			total pathology		
	n.d.	i.	c.	n.d.	i.	c.	n.d.	i.	i.
male	19 %	–	81 %	15 %	15 %	70 %	16 %	19 %	65 %
female	21 %	–	79 %	26 %	5 %	69 %	24 %	7 %	69 %

n.d. – no decision, i – incorrect classification, c. – correct classification

It can be seen from the results shown in Table 2 that better classification results were obtained for female voices. This is particularly visible for the pathological voices and is caused by the limitations imposed by the Intonograf whose lower measuring frequency is 78 Hz, while in many cases the instantaneous value of the frequency of the larynx tone of males reaches as low as 40 Hz and this does not only occur for pathological voices (see [9]). For this reason the investigations did not include very low voices and voices which had a distinct "cut-off" on the right-hand side of the TO distribution. The lower limitation of the measuring frequency in these cases caused a decrease in the dispersion in both distributions and a decrease in the right-skew of the TO distribution. Nevertheless, in some cases errors of this type could not be avoided and, probably for this reason, the pathological voices were classified as normal voices.

As was already mentioned, assignment functions were determined heuristically, based on the results of measurements taken of the parameters of the distributions and τ_{ph}, obtained for normal voices and voices with paralysis of the larynx. The same functions were used for the evaluation of the state of the larynx source for other kinds of pathology of the vocal chords, e.g. polyps, papillomas, cancers etc. From the data in Table 2 it can be seen that the expansion to include these other cases of pathological changes in the vocal folds caused a slight increase in the classification errors. This occurred mainly in the cases of polyp changes for which the phonation function of the larynx depends decisively on the size of the polyp and its position on the vocal chord. If it occurs on the outer side of a vocal fold and is not too large, in practice no changes in the voice and in the behaviour of the larynx tone can be observed.

The present method for the description of the condition of the larynx source is also very convenient for the objective evaluation of changes in the voice occurring in the course of treatment and rehabilitation. As an example, Table 3 shows the results of measurements on a female voice with a paralysed

Table 3. Measurement results for the chosen parameters of the distributions of TO and ΔTO and the relative time τ_{ph}, at different stages in the treatment and rehabilitation of a female voice with unilateral paralysis of the vocal fold

recording No.	parameter				
	$\bar{F}_0$ [Hz]	R [%]	τ_{ph}	As_{TO}	$As_{\Delta TO}$
I	184	4.5	0.46	0.3	−0.04
II	236	10.5	0.64	0.3	−0.09
III	244	7.5	0.65	0.44	−0.15
IV	240	7.5	0.69	0.29	−0.2

right vocal fold in an intermediate position. Measurements were taken when the patients first reported (I), after an operation consisting of a "Teflon" injection into the paralysed vocal chord (II), and twice during the rehabilitation of the voice: three months after the operation (III) and a year later (IV). Table 4 shows the values of the degrees of assignment of the voice to the group of pathological voices, calculated for the individual parameters according to formula (1) for the values of coefficients α and β in Table 1.

It can be seen from the results given in Tables 3 and 4 how the phonation function of the patient's larynx gradually improved in the course of treatment and rehabilitation, although it never reached a perfect condition and the process of voice rehabilitation had to be continued. A distinct improvement occurred first of all in the control of the air flow through the larynx during phonation, particularly after the "Teflon" injection into the paralysed vocal fold, due to which a full closing was attained during larynx excitation. Conclusions concerning changes of this type can be drawn on the basis of an increase in the value of the parameter τ_{ph} obtained on the basis of an increase in the value of the parameter τ_{ph} (and, at the same time, a decrease in the degree of the assignment to the pathological voices). However, the other parameters did not improve as much after this operation, and some even became worse. Audio evaluation

Table 4. Values of the degrees of assignment to a set of pathological voices for a female voice with unilateral paralysis of the vocal fold, determined at different stages of treatment and rehabilitation, based on the data in Table 3

recording No.	parameter			
	τ_{ph}	R	As_{TO}	$As_{\Delta TO}$
I	1	0.22	0.75	0.95
II	0.3	1	0.75	0.89
III	0.25	0.64	1	0.81
IV	0.05	0.64	0.72	0.75

confirmed bad phonation of the patient who after the operation (II) spoke with more effort, and in subsequent trials the voice was a little hoarse and only a slight improvement was observed. It is worth noting that, as a rule, after operations on the vocal folds (e.g. to remove polyps) a worsening of the phonation can be observed in the first stage (speaking with greater effort, hoarseness etc.), which improves only as a result of a very long voice rehabilitation. However, improvement in the closing of the folds occurs very soon after the operation and this is reflected in an increase of the relative phonation duration τ_{ph}.

10. Conclusions

The method suggested for the analysis of the distributions of values of the periods of the larynx tone TO, its relative percentage changes ΔTO, and the relative phonation time τ_{ph}, can be used for initial, objective evaluation of the state of the larynx source. In addition, it can also be used for the observation of changes occurring in it during the processes of treatment and rehabilitation. The introduction of fuzzy relations into the description of the state of the larynx source permits easy assessment of its function.

It is an advantage of the present analytical method that a possible change in the decision-making rules in the case of expansion to include other pathological entities is very easy and does not require the laborious calculations.

The method presented is fully automatic and useful for mass examinations based on the recordings of voices on a tape recorder.

The precision of the classification of voices into normal and pathological was about 72 %, with the mean classification error not exceeding 8 %. It is worth noting here that all healthy voices were properly classified.

The measuring system presented can serve to assist phoniatric diagnosis, and visualisation of the variations of the larynx tone in time during the utterance of sentences with a specific type of intonation can be used not only when a diagnosis is made, but also during the process of the rehabilitation of the voice and speech. The patient can then directly observe the intonation of his utterance and compare it with the given one, chosen for exercises.

Acknowledgement. The recordings of voices used in the investigation were made in the Otolaryngological Clinic Phoniatric Laboratory of the Medical Academy, Warsaw, headed by doc. dr. hab. med. W. Tłuchowski. The authors wish to express their gratitude for the possibility of their use together with medical documentation.

References

[1] L. J. Boë, M. Contini, H. Rakotofiringa, Etude statistique de la fréquence laryngiènne, Phonetica, **32**, 1-23 (1975)

[2] D. J. Broad, *Phonation*, in *Normal aspects of speech hearing and language*, ed. F. D. Minifie, T. J. Hixon, F. Wiliams, Prentice-Hall, Englewood Cliffs 1973.

[3] M. J. Cheng, L. R. Rabiner, A. E. Rosenberg, C. A. McGonegal, *Some comparisons among several pitch detection algorithms*, Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Philadelphia Pa. 1976, 332-334.

[4] F. L. Darley, D. C. Spriestersbach, *Diagnostic methods in speech pathology*, 2nd ed., Harper and Row Publ., New York 1978.

[5] S. B. Davis, *Computer evaluation of laryngeal pathology based on inverse filtering of speech*, SCRL Monograph 13, Speech Communications Research Laboratory, Inc., Santa Barbara, 1976.

[6] B. Frøkjaer-Jensen, S. Prytz, *Registration of voice quality*, Technical Review Brüel & Kjaer, **3**, 3-17 (1976)

[7] R. Gubrynowicz, W. Mikiel, P. Żarnecki, *Acoustical analysis for evaluation of laryngeal dysfunction in the case of vocal cords paralysis*, in *Speech analysis and synthesis*, ed. W. Jassem, vol. 5 (in press).

[8] K. Ishizaka, J. L. Flanagan, *Synthesis of voiced sounds from two-mass model of the vocal cords*, Bell System Techn. J., **51**, 1233-1268 (1972).

[9] W. Jassem, M. Steffen-Batóg, S. Czajka, *Statistical characteristics of short-term average F_0 distributions as personal voice features*, in *Speech analysis and synthesis*, ed. W. Jassem, vol. 3, PWN, Warsaw 1973, 209-225.

[10] Y. Koike, *Application of some acoustic measures for the evaluation of laryngeal dysfunction*, Studia Phonologica, **7**, 17-23 (1973).

[11] H. von Leden, P. Moore, R. Timcke, *Laryngeal vibrations. Measurements of the glottic wave.* Part III; *The pathologic larynx*, A.M.A. Archives of Otolaryngology, **71**, 1, 16-35 (1960).

[12] P. Lieberman, *Some acoustic measures of the fundamental periodicity of normal and pathologic larynges*, JASA, **35**, 244-353 (1963).

[13] W. Mikiel, R. Gubrynowicz, W. Hagmajer, *Intonograf – system for measurement and visualisation of changes in intensity and melody of speech* (in Polish), Proc. XXIV Open Sem. on Acoustics, vol. I, Gdańsk–Władysławowo 1977, 120-123.

[14] A. Niedźwiecki, W. Mikiel, *Digital measurement and coding of fundamental frequency in a harmonic speech signal*, in *Speech analysis and synthesis*, ed. W. Jassem, vol. 4, PWN 1976, 293-300.

[15] E. Patrick, *Review of pattern recognition in medical diagnosis and consulting relative to a new system model*, IEEE Trans. on System, Man, Cybernetics, **SMC-4**, 1 (1974).

[16] W. H. Perkins, *Speech pathology. An applied behavioural science*, Mosby, Saint Louis 1971.

[17] G. S. Ramishvili, *Rechevoy signal i individual'nost' golosa*, Mecniereba, Tbilisi 1976.

[18] L. Sachs, *Angewandte Statistik; Planung und Auswertung, Methoden und Modelle*, Springer Verlag, Berlin 1974.

[19] H. Wechsler, *A fuzzy approach to medical diagnosis*, Int. J. Bio-Medical Computing, **7**, 191-203 (1976).

[20] L. A. Zadeh, *A fuzzy-algorithmic approach to the definition of complex or imprecise concepts*, Int. J. Man-Machine Studies, **8**, 249-291 (1976).

Received on May 21, 1979

ARCHIVES OF ACOUSTICS
5, 1, 31-46 (1980)

AUTOMATIC DETERMINATION OF THE FREQUENCY BEHAVIOUR OF THE LARYNX TONE BY THE METHOD OF LINEAR PREDICTION

ANDRZEJ DZIURNIKOWSKI

ASW (02-656 Warszawa)

This paper presents a method of automatically tracing the frequency of the larynx tone, based on a numerical implementation of the method of linear prediction on a digital computer. It also discusses the algorithm for the practical implementation of this method, derived by the author, and presents the results obtained from the analysis of a continuous speech signal with a duration of approximately 2 seconds.

1. Introduction

The frequency of the larynx tone F_0, also called the *fundamental frequency*, is one of the basic accoustic parameters in the investigation of a speech signal. It is generally known that the intonation of continuous speech is connected with the parameter F_0. If the speaker is in the state of emotional excitement, the frequency F_0 of his larynx tone usually increases. From the acoustical point of view it is connected with changes in the duration of the vibration of the vocal chords which depends on the air pressure in the larynx and the physiological excitation of the chords [6]. The frequency F_0 can be measured by different methods: being determined by the analysis of a speech signal in the time domain in real time using the PPD method [8] or by the analysis of the parameters of a speech signal in the frequency domain, e.g. by the calculation of the cepstrum [1]. The objective of the present paper is to present a method for automatic determination of the parameter F_0 in a continuous speech signal, based on the method of linear prediction and using an electronic digital computer.

Developed in the 1950's, the technique of linear prediction was first used to analyse and synthesize speech by SAITO and ITAKURA [9] in the 1970's. It is based on a linear model of speech production which was created by FANT in the late 1950's [4, 6]. This model reflects, with certain simplifications, the most essential phenomena occurring during the course of the process of speech signal generation by man. This process can be roughly described in the following way. An acoustic speech wave consists of a variable acoustic pressure whose source is a positive pressure in the lungs. During articulation the air flows through the larynx and the pharyngo-oral cavity. The following parameters affect the characteristics of the acoustic wave at the outlet from the vocal channel that is formed by the lips and, in particular, its pressure: the pressure of the air exhaled from the lungs, the velocity of vibration of the vocal chords (depending on their mass, elasticity etc., which is a characteristic feature of the speaker), and the anatomical structure of the vocal tract, particularly those of its elements which participate in the dynamic changes of the parameters of the tract (e.g., lips, tongue, soft palate) [6, 11]. During generation of non-nasal sounds, the soft palate closes the nasal cavity, separating it from the other part of the vocal tract. Plosive sounds in turn occur when the mouth is violently opened after previous accumulation of air in the vocal tract. Unvoiced sounds are created when the vocal chords remain open and the air flows freely through the larynx. From the point of view of the present considerations we are, however, interested first of all in generation of such speech sounds as are characterized by the quasiperiodicity of the vocal wave signal (sonorous, voiced fricatives), i.e. those whose generation involves the vocal chords.

The vibration frequency of vocal chords can be measured by the acoustical method as the inverse of the time interval of the wave periods observed in the acoustic pressure at the outlet of the vocal tract. This frequency is defined as the F_0 frequency. The linear model of speech generation given by Fant (Fig. 1)

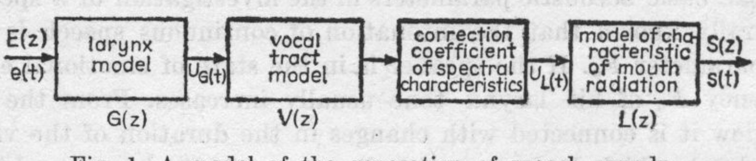


Fig. 1. A model of the generation of speech sounds

permits voiced sounds, voiced fricatives and other sounds of quasiperiodic structure to be represented under the assumption of the vocal tract being excited by an impulse function or by white noise [6]. It is thus possible to estimate the F_0 frequency in the speech signal based on this model.

Excitation with pulses of amplitude σ, separated in time by $T_0 = IT$, where I is a positive whole number and T – a standardized sampling interval (for the sake of simplification it can be assumed that $T = 1$), can be expressed in the z-domain by the formula

$$E(z) = \sigma \sum_{n=0}^{\infty} (z^{-I})^n = \frac{\delta}{1-z^{-I}} \quad \text{for } |z| > 1. \tag{1}$$

In the case of quasiperiodic sounds this excitation is applied at the outlet of the larynx model $G(z)$, represented in Fant's model by a low-pass bipolar filter. The signal $U_G\,(t)$, obtained at the outlet of the filter $G(z)$, is in turn supplied to the model of the vocal tract $V(z)$ which has the form of step-like resonators connected with each other. The number of resonators is connected with the number of the resonance frequencies of formants assumed in a given implementation of the model [6]. In practical research the number of formants assumed is relatively small, resulting in an imprecise representation of the spectral characteristic over the low frequency range. In order to equalize the imprecision of the representation of the spectrum over this range, Fant introduced in his model the so-called *correlation factor of the spectral characteristic* [6] which, however, in numerical representations of this model is not usually taken into consideration [7]. The signal $U_L(t)$ which is formed at the output of the model is supplied to the system $L(z)$ representing the frequency characteristic of the mouth radiation. It is the last element of the linear model of speech sound generation. The model given by Fant can be written as a z-transform using the equation

$$X(z) = E(z)G(z)V(z)L(z), \tag{2}$$

where $X(z)$ is the z-transform of the discrete signal $X(nt)$. On account of the fact that in the present model excitation may be assummed to be with either periodic pulses or noise and that a limited fixed number of formant frequencies and formant bandwidths are assumed, Markel [6] suggests that this model may be directly used for the representation of vowel and fricative sounds. Based on the analytical model (Fant's model), the investigation of quasiperiodic speech segments can be performed in order to define successive periods of larynx excitation and, at the same time, of successive values of the frequency of the larynx tone.

With simplification, the model of speech sound analysis can be written as

$$E(z) = X(z)A(z), \tag{3}$$

where

$$A(z) = \frac{1}{G(z)V(z)L(z)} \tag{4}$$

is the inverse filter. It can be expressed, with simplification, by the following equation in the z-domain:

$$A(z) = \sum_{i=0}^{M} a_i z^{-i}, \quad a_0 = 1, \tag{5}$$

where a_i are the coefficients of the filter $A(z)$, which in addition to the order of the filter M and the signal $X(z)$ define completely the analytical model given by formula (3). In the (discrete) time domain equation (3) can be written as

$$e(n) = \sum_{i=0}^{M} a_i x(n-i) = x(n) + \sum_{i=1}^{M} a_i x(n-i), \tag{6}$$

which means that the M-th coefficient of the inverse filter, otherwise known as the *coefficient of prediction* of the sample $x(n)$, requires a linear combination of the previous M samples. In order to define the coefficients of the inverse filter (i.e. the prediction coefficients), in order to minimize $e(n)$, M linear equations must be solved [5, 6]. In practice, this is achieved by the use of covariance or autocorrelation methods of analysis of the speech signal samples. This means that on the basis of the model of linear prediction, the parameters of the model of speech analysis can be determined in a relatively easy manner which is essential for the further considerations.

2. Determination of the F_0 parameter using the method of linear prediction

The model of speech sound analysis, based on the technique of linear prediction, is a model which can, in a natural manner, due to its relatively easy numerical implementation, be used in speech analysis performed with an electronic digital computer. The methods for the estimation of the frequency of the larynx tone which have been used so far, used the idea of flattening (smoothing) of the spectrum of the signal analyzed, in which the phases of the individual harmonic components of the spectrum were reduced to zero. The signal obtained through the inverse transformation had a structure of the form of pulses repeating at intervals equal to a period of the larynx tone (in the case of a voiced signal) or had a random character (in the case of an unvoiced signal) [10].

The technique of linear prediction and, in particular, the method of analysis by which we obtain the sequence $\{e(n)\}$ with parameters described by formula (6), can give similar effects. The sequence $\{e(n)\}$, obtained in the analysis by the method of linear prediction, called in this method the *error signal*, is a signal with distinct pulses occurring at successive moments of larynx tone generation. The distance between them is the period of this tone (if a voiced quasiperiodic signal is analysed). The signal $\{e(n)\}$ has a random character for all other signal classes. This is in agreement with the assumption of the linear model of speech generation, which is the basis for linear prediction. There is, however, a certain "fault" in the characteristics of the error signal obtained for a quasiperiodic signal, which consists in a rather significant decrease in the amplitude of pulses and changes of sign within the sounds analyzed or in transitions from phoneme to phoneme. Fig. 2 shows the error signal obtained with the present model for a fragment of a word "pokoju" contained in the phrase "w pokoju paliła się

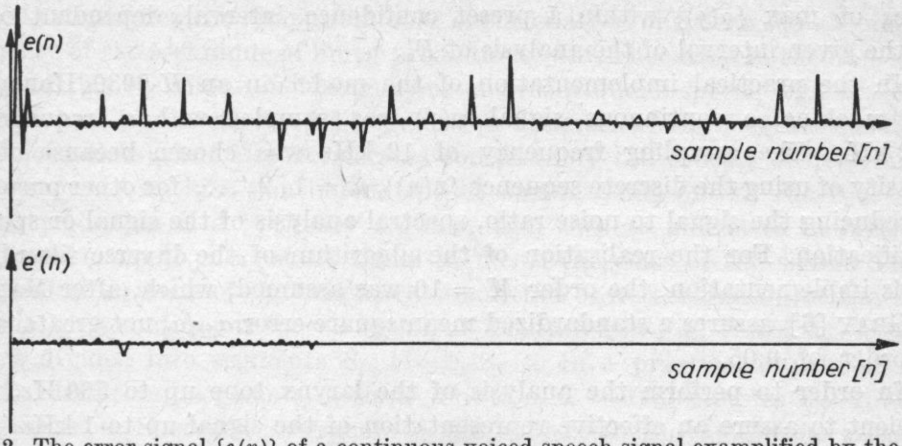


Fig. 2. The error signal $\{e(n)\}$ of a continuous voiced speech signal examplified by the word "pokoju" ($\delta = 1$)

słaba żarówka". This fault can be overcome by adequate operation on the coefficient σ, for example by making it dependent on the amplitude of the real signal. In spite of the above remarks, it can be stated that this approach provides a basis for the development of an effective method for estimation and extraction of the parameter F_0 in a signal of continuous speech, using the technique of linear prediction. A general functional diagram for implementing the estimation of the frequency of the larynx tone is shown in Fig 3. At the

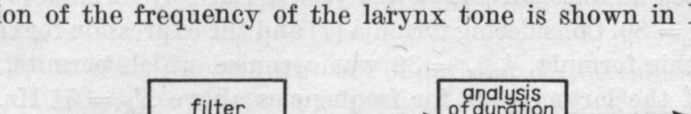


Fig. 3. A functional model of the estimation of the frequency of the larynx tone

output of the model in Fig. 3 we obtain the sequence $\{F_{0i}\}$ of values of the frequency of the larynx tone $F_0 = 1/T_0$ corresponding to successive periods of larynx excitation in a quasiperiodic signal of voiced human speech. An analysis of the length of the period of the larynx tone T_0 can be performed using an autocorrelation analysis of the sequence of the error signal $\{e(n)\}$, described by the following expression:

$$\varrho(j) = \sum_{n=0}^{N-1-j} e(n)e(n+j). \tag{7}$$

In the estimation of F_0 the sequence of values obtained for the autocorrelation $\{\varrho(j)\}$ is in practice calculated for $j \leqslant N/2$. The choice of the values of N depends on the sampling frequency of the analogue signal f_p and affects the effect of averaging the values of the sequence $\{F_{0i}\}$. The sequence $\{\varrho(j)\}$ obtained serves directly for determination of duration of the period T_0 by finding the values of max $\{\varrho(j)\}$ within a preset confidence interval, dependent on f_p and the given interval of the analysis of F_0.

In the practical implementation of the model on an H-6030 Honeywell Bull machine, a continuous signal $x(f)$ was sampled with a frequency $f_p = 12$ kHz. The sampling frequency of 12 kHz was chosen because of the necessity of using the discrete sequence $\{x(n)\}$, $n = 1, 2, \ldots$, for other purposes, e.g. reducing the signal to noise ratio, spectral analysis of the signal or speaker identification. For the realisation of the algorithm of the inverse filter $A(z)$ in this implementation, the order $M = 10$ was assumed, which, after MARKEL and GRAY [6], assures a standardized mean square error a_M/a_0 not greater than the order of 0.05.

In order to perform the analysis of the larynx tone up to 500 Hz, it is sufficient to assure an effective representation of the signal up to 1 kHz. This means that the effective sampling frequency should be at least $f_p = 2$ kHz. To achieve this, for a signal sampled at $f_p = 12$ kHz, it is sufficient to use the procedure of initial filtering consisting in the selection of samples with numbers increasing by a constant value K assumed in the relation $f_p = f_p' K$. This approach additionally ensures a considerable reduction in the number of elements of the sequence $\{x(n)\}$ to be analyzed, which significantly affects the speed of signal processing. In the present situation $K = 6$ was assumed. 480 samples of the primary signal $\{x(n)\}$ were included in each analytical step, which permitted an autocorrelation analysis, defined by formula (7), of a signal of duration $N = 80$. Considering formula (7) and the expression for the constant j occurring in this formula, $j_{\max} = 39$ was assumed which permits, in practice, an analysis of the larynx tone for frequencies above $F_d = 51$ Hz. The upper frequency limit of the analysis is fixed by practical considerations of searching for the absolute maximum in the sequence $\{\varrho(i)\}$.

In other words, we search for the absolute maximum value of the sequence $\{\varrho(i)\}$ in a certain preset interval Q, subsequenly called the *confidence interval*. In the algorithm implemented here the lower boundary of the confidence interval was taken as $l = 6$, and the upper boundary as $k = 37$, which permits an estimation of the frequency of the larynx tone up to $F_g = 333$ Hz and involves the assumption of such a value of i that the following condition would be satisfied within the voiced signal:

$$\max_{l \leqslant i \leqslant k} \{\varrho(i)\} > \varrho(l). \tag{8}$$

3. Algorithm for determining F_0 by the method of linear prediction

The algorithm presented for using the method of linear prediction served for its numerical implementation in the form of a programme in Fortran.

It was assumed from the beginning that the analysis would be performed only within the class of voiced signals of continuous speech. This assumption not only eliminates the "imperfections" of the analysis of speech signals inherent in the use of the technique of linear predictions, which occurred in the implementation of this technique in the analysis of a continuous speech signal without separation of signal classes [6] (e.g. within a signal which does not belong to the voiced class), but also accelerates considerably the process of determining the parameter F_0 in a continuous speech signal. Thus one of the first steps in this algorithm is a "primary segmentation" whose aim is to define, with the precision up to a constant signal segment, the class of the signal. This is an essential element of the schematic algorithm of determination of the parameter F_0 in a continuous speech signal, as illustrated in Fig. 4. A discrete signal $x(n)$ is divided into segments S_l, $l = 1, 2, \ldots$ In a practical implementation of the algorithm, segments with a duration $T = 128$ samples (~ 10.6 ms) were assumed. Successive segments S_l are subject to the amplitude criterion. The segments are considered as silent if the maximum amplitude of a signal within a given segment does not exceed β times the threshold value of β_1. The threshold value β_1 is determined for each segment S_l separately as 15% of the maximum amplitude of the signal $x(n)$ from 8 successively analyzed segments S_l; but it cannot take a value less than 10 (with 128 levels of quantization of the amplitude of the input signal).

Let $a_{1,l}$ denote the factor by which the threshold β_1 is exceeded, where

$$a_{\beta_{1,l}} = n\{X_i\colon X_i \geqslant \beta_1 \wedge X_i \in S_l\}, \tag{9}$$

$X_i \in S_l$ being the values of the l-th segment of duration T (the values of the signal sampled and quantized), while $n\{\cdot\}$ is the size of the set of the values of the amplitude of the samples of the signal exceeding the threshold value β_1 in the segment S_l. The value of $a_{\beta_{1,l}}$ is the basis for determining the class of the signal represented by the segment S_l. A segment is assigned to the class of "no signal" if the following condition is satisfied:

$$a_{\beta_{1,l}} < \beta. \tag{10}$$

$\beta = 30$ was assumed for this algorithm. Only those segments S_l for which $a_{\beta_{1,l}} \geqslant \beta$ were subject to further analysis. Segments with a quasiperiodic character are recognized as all those l-th segments of duration T for which $a_{\beta_{2,l}} \geqslant \beta_0$, where β_2 is the threshold value of the amplitude, and $\beta_0 = 40$. The parameter β_2, similarly to β_1, is dynamically determined as 80% of the value of the maximum amplitude (its value cannot be less than 42). All the remaining segments undergo further analysis based on the frequency criterion. They are initially assigned to the class of "noise" signals, to the class of "voiced quasiperiodic" signals or to the class of signals conditionally recognized as "noise" signals ("conditional noise"). The frequency criterion is based on the calculation of the zero-crossing parameter. The segments for which the number of zero-crossings is greater than $\beta_3 = 42$ (which for $f_p = 12$ kHz corresponds to a frequency of about 2 kHz) are recognized as "noise" signal. However, the segments with a number of zero-crossings over $\beta_4 = 25$ (~ 1.2 kHz) and a maximum amplitude satisfying the condition

$$\max\{X_i\} < \beta_2, \qquad X_i \in S_l, \tag{11}$$

are considered as "conditional noise". The segments of "conditional noise" can be considered as segments of voiced signal or as "noise". If in the direct neighbourhood of the segment recognized as a "noise" segment, a group of

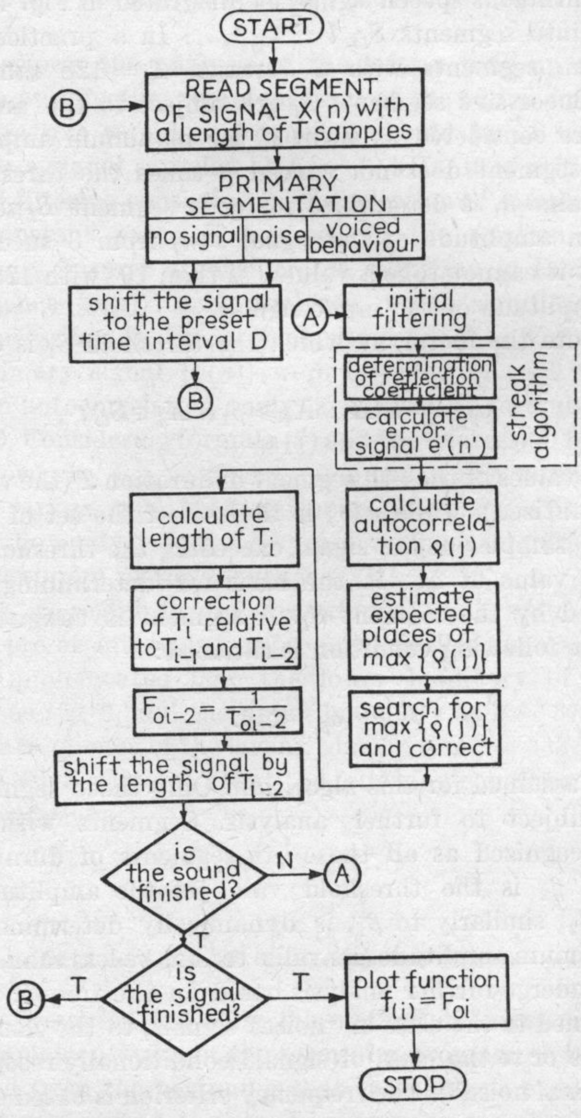


Fig. 4. A schematic diagram of the algorithm for determination of the parameter F_0 in a continuous speech signal; N – no, T – yes

"conditional noise" segments occurs, the whole group is in its entirety considered to be "noise". In the opposite case the "conditional noise" segments are assigned to the class of voiced segments.

A fragment of the algorithm for the determination of the parameter F_0 in a continuous speech signal, concerning the primary segmentation, is shown in Fig. 5. Sections of the speech signal with a duration of about 85.3 ms were

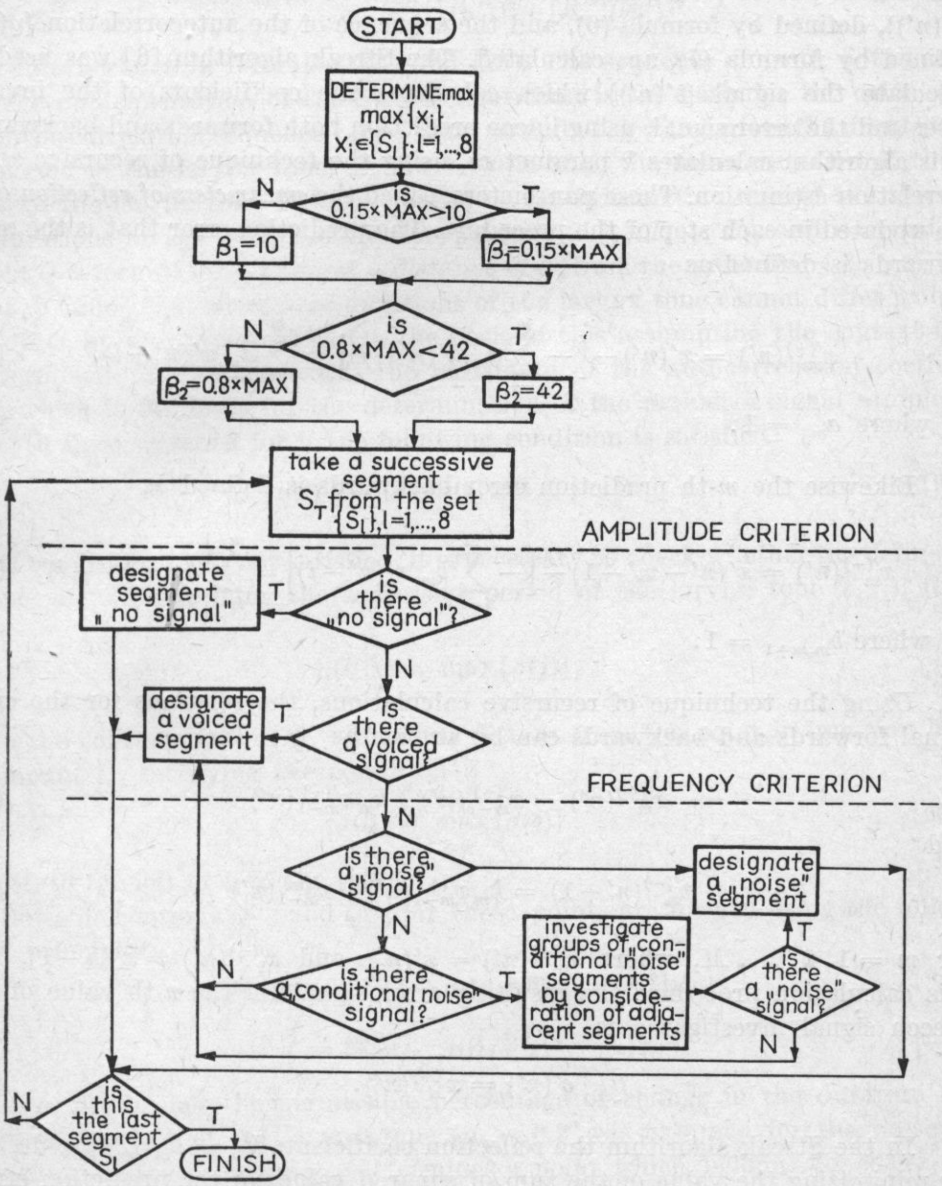


Fig. 5. The algorithm of the primary segmentation for a group of eight segments S_l
N – no, T – yes

subject to a single primary segmentation which corresponded to two tape blocks with 512 samples of the signal sampled at $f_p = 12\,\text{kHz}$. Within the segments assigned to the voiced class, further denoted by $\{x(n)\}_D$, the initial filtering (described in Section 2) was conducted, as a result of which we obtained the signal $\{x'(n')\}_D \subset \{x(n)\}_D$, where $n' = Ki$, $i = 1, 2, \ldots$ and $n = 1, 2, \ldots$ According to the foregoing discussion, the size of the set $\{x'(n')\}_D$, taken for analysis in one step of the algorithm, is 80. For this signal the error signal $\{e'(n')\}$, defined by formula (6), and the sequence of the autocorrelation $\{\varrho(j)\}$, defined by formula (7), are calculated. The Streak algorithm [6] was used to calculate the signal $\{e'(n')\}$ which calculates the coefficients of the inverse filter and the error signal, using linear prediction both forwards and backwards. This algorithm calculates k parameters, using the technique of recursive autocorrelation estimation. These parameters, called the *parameters of reflection* of k, are updated in each step of the procedure. The prediction error that is the m-th forwards is defined as

$$x_m'^{(+)}(n') = x'(n') - \Big[-\sum_{i=1}^{m} a_{m_i} x'(n'-i)\Big] = \sum_{i=c}^{m} a_{m_i} x'(n'-i), \tag{12}$$

where $a_{m_0} = 1$.

Likewise the m-th prediction error backwards is defined as

$$x_m'^{(-)}(n') = x'(n'-m'-1) - \Big[-\sum_{i=1}^{m} b_{m_i} x'(n'-i)\Big] = \sum_{i=1}^{m+1} b_{m_i} x'(n'-i), \tag{13}$$

where $b_{m,m+1} = 1$.

Using the technique of recursive calculations, the relations for the error signal forwards and backwards can be written as

$$x_m'^{(+)}(n') = x_{m-1}'^{(+)}(n') + k_m x_{m-1}'^{(-)}(n') \tag{14}$$

and

$$x_m'^{(-)}(n'+1) = k_m x_{m-1}'^{(+)}(n') + x_{m-1}'^{(-)}(n') \tag{15}$$

for $m = 1, 2, \ldots, M$, where $x_0'^{(+)}(n') = x'(n')$ and $x_0'^{(-)}(n') = x'(n-1)$. For this calculation procedure the resulting error signal for the n-th value of the speech signal investigated is

$$e'(n') = x_M'^{(+)}(n'). \tag{16}$$

In the Streak algorithm the reflection coefficients $k_m = k_m(n')$ are defined by minimizing the value of the sum of squared values of the prediction errors forwards and backwards:

$$[x_m'^{(+)}(n')]^2+[x_m'^{(-)}(n'+1)]^2 \\ = 2k_m x_{m-1}'^{(+)}(n')x_{m-1}'^{(-)}(n')+(1+k_m^2)\{[x_{m-1}'^{(+)}(n')]^2+[x_{m-1}'^{(-)}(n')]^2\}. \tag{17}$$

This minimization gives, in effect, the expression

$$k_m = k_m(n') = \frac{-2x_{m-1}'^{(+)}(n')x_{m-1}'^{(-)}(n')}{[x_{m-1}'^{(+)}(n')]^2+[x_{m-1}'^{-1}(n')]^2}, \tag{18}$$

taken into account in formulae (14) and (15). The sequence $\{e'(n')\}$, obtained from the implementation of the Streak algorithm, is the basis for the calculation of the correlation sequence $\{\varrho(j)\}$ using formula (7), which gives the duration of a period of the larynx tone. In order to do this, the mean duration of the period of the larynx tone T_s is calculated as the arithmetic mean of the three last durations of the period of the larynx tone T_{i-2}, T_{i-1}, T_i, and a certain context Q is formed for a point at a distance of T_s from the preceding $j_{\max}\{\varrho(j)\}$. It was assumed that successive durations of the larynx tone cannot differ from each other by more than 30 %. On the basis of this assumption the context Q is determined. In the context Q the maximum of the autocorrelation coefficient, which is the basis for the determination of the period of signal samples of length l_s, is searched for if the following condition is satisfied:

$$\bigvee_{l\in Q} [\varrho(l) = \max\{\varrho(j)\};\, 6 \leqslant j \leqslant 37]. \tag{19}$$

If condition (19) is not satisfied, it is necessary to check whether the defined distance l_{i+1} (determining the successive period of the larynx tone T_{i+1}), for which

$$\varrho(l_{i+1}) = \max_j\{\varrho(j)\},$$

lies in the contexts $Q^{(+)}$ or $Q^{(-)}$ of points distant from the point in question by an amount l_s, satisfying the condition

$$\varrho(l_s) = \max_{j\in Q}\{\varrho(j)\} \tag{20}$$

by $0.5l_s$ to the left or l_s to the right.

Suitable contexts $Q^{(-)}$ and $Q^{(+)}$ for these points are formed using the following principles:

$$Q^{(-)} \equiv [0.5l_s(1-\mu_1),\ 0.5l_s(1+\mu_1)], \tag{21}$$

$$Q^{(+)} \equiv [2l_s(1-\mu_1), 2l_s(1+\mu_1)], \tag{22}$$

where μ_1 establishes the permissible percentage of change in the duration of successive periods of the larynx tone ($\mu_1 = 0.2$ was assumed for the present algorithm). If the distance l_{i+1} determines a point which belongs to either of the contexts $Q^{(-)}$ or $Q^{(+)}$, then the distance l_s is accepted only as the duration of a successive period of the larynx tone, recognizing that an error of a double decrease or increase of the distance occurred in the course of the analysis. Examples of these type of errors are reflected in Figs. 6 and 7. In both cases

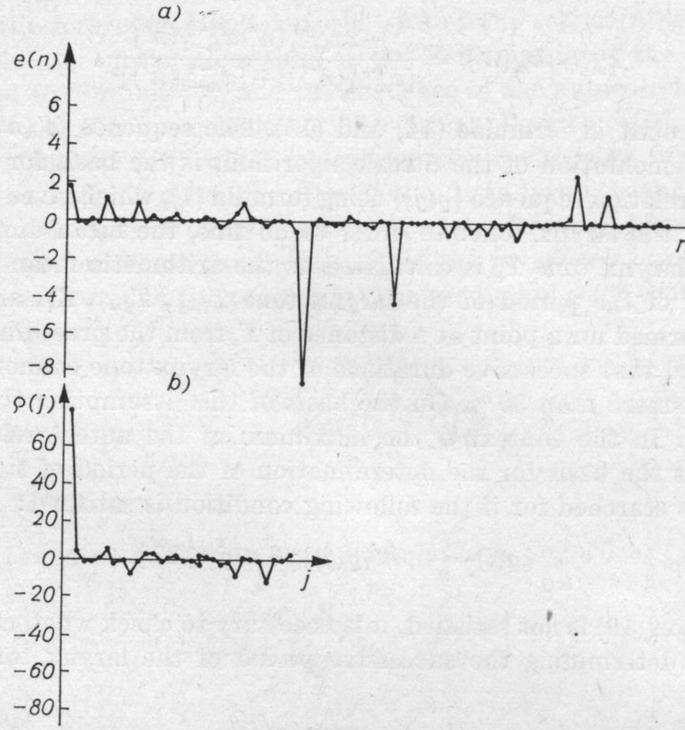


Fig. 6. The error signal $\{e(n)\}$ of the final phase of the duration of the sound /u/ in the word "pokoju" − (a), and the corresponding sequence of the values of the autocorrelation $\{\varrho(j)\}$ − (b)

they apply to points in the real signal $\{x(n)\}$, where the values of the signal samples are lower than the others in a specific segment of the signal, e.g. in the final phase of the duration of a sound in the word "pokoju" in the phrase whose signal is presented in Fig. 8 (Fig. 6), or when the amplitude of the signal is decreased and its structure is changed, e.g. in the transition between sounds (a) and (l) in the word "paliła" represented by the signal in Fig. 8 (Fig. 7). If, however, the distance determined does not belong to the contexts $Q^{(-)}$ or $Q^{(+)}$, we have

$$\varrho(l_s) \geqslant u\,\varrho(T_{i+1}), \tag{23}$$

where u is the threshold coefficient, which in the present algorithm took the value $\mu = 0.75$. If condition (23) is satisfied, then l_s is assumed as the duration of a successive period of the larynx tone. If not, the duration of this period is established as l_{i+1}.

Thus, errors in the definition of the larynx tone which were due to a large variation in the values of the sequence $\{e'(n')\}$ for successive periods, and hence in the sequence $\{\varrho(j)\}$, were eliminated (cf. Figs. 6 and 7).

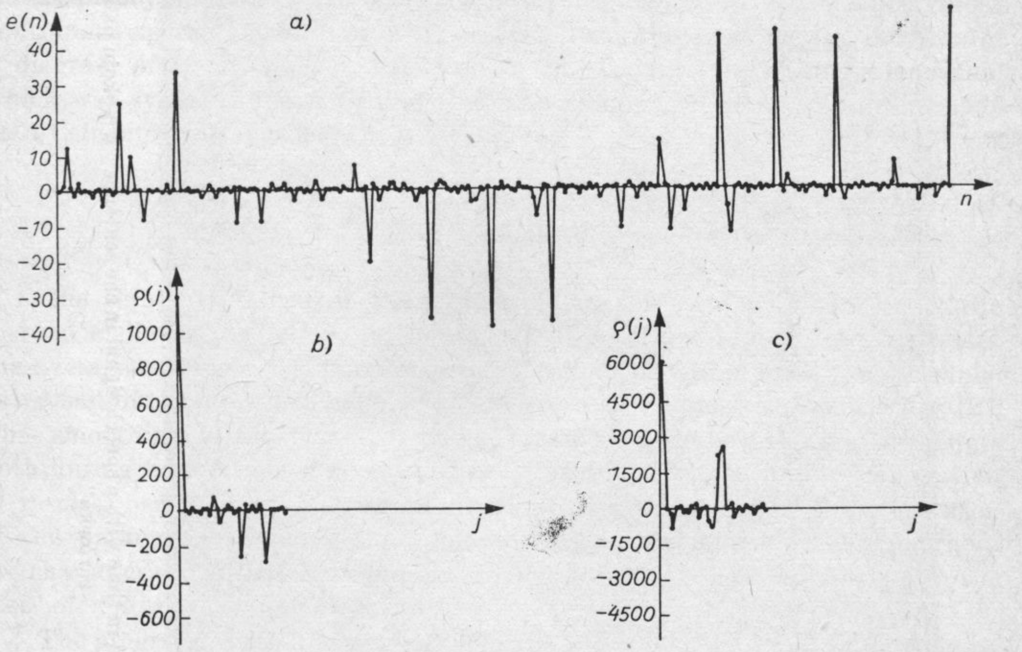


Fig. 7. The error signal $\{e(n)\}$ calculated for the transition between the sounds /a/ and /l/ in the word "wypaliła" $-(a)$, and synchronised therewith in terms of the origin of the sequence of values of the autocorrelation $\{\varrho(j)\}-(b)$ and (c)

The calculated duration of the period of the larynx tone T_{i-1} and T_i is subsequently corrected in order to eliminate errors resulting from large changes in the amplitude of the signal $e(n)$ (for $\delta = \text{const.}$ see formula (2)). In the correction stage the three successive values of the length of the period of the larynx tone T_{i-1}, T_i and T_{i+1}, determined in the preceding steps of the algorithm, are considered. Taking the trend of these values and approximating them by linear extrapolation (such extrapolation was assumed for the present algorithm after the extrapolation given by Markel [6]), new values for the duration of the larynx tone are determined, which effectively "corrects" those values which deviated from the extrapolated ones and were caused by the errors mentioned earlier.

In this way, in the i-th step of the algorithm the final frequency of the larynx tone, determined in the $(i-2)$-nd step, is calculated from the period T_{i-2}.

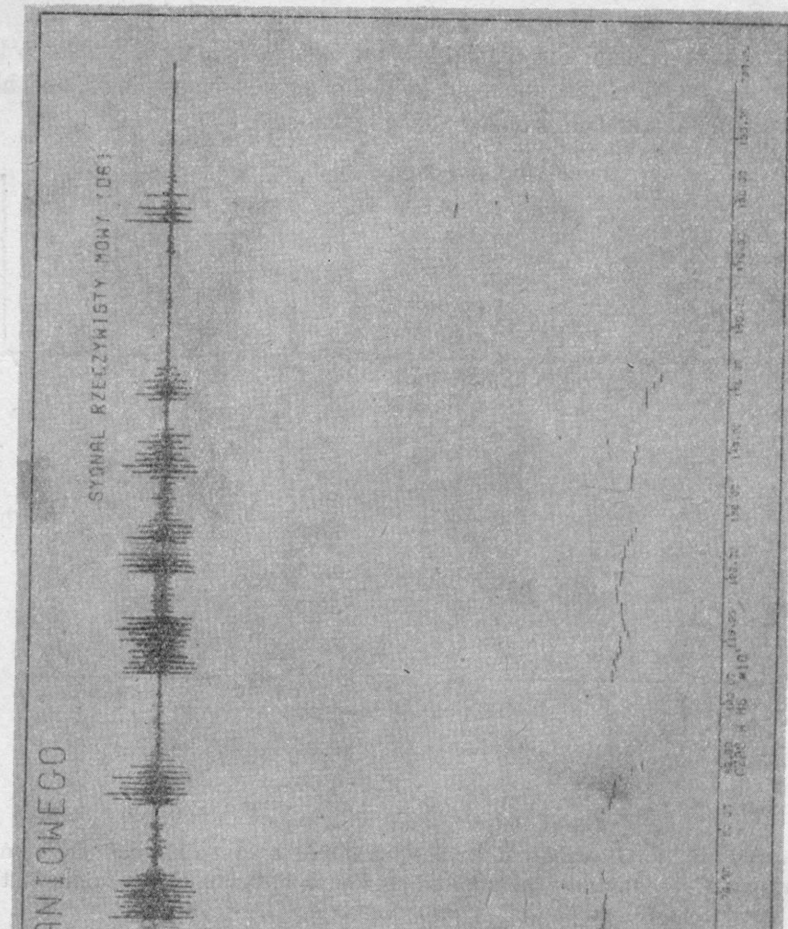


Fig. 8. A diagram of the larynx tone synchronised in time with a speech signal in the sentence “w pokoju paliła się słaba żarówka”

In the present algorithm this frequency is calculated according to the formula

$$F_{0i-2} = \frac{12000}{(T_{i-2}-1)\cdot 6} \tag{24}$$

and is one of the values of the sequence $\{F_{0i}\}$ corresponding to the values of the frequency of the larynx tone dynamically calculated for the class of voiced continuous speech signal. The sequence $\{F_{0i}\}$ was the basis for the creation of a diagram of the frequency variation in the larynx tone during phonation, which was synchronised with real time. An example of such a diagram made on a Calcomp plotter is shown in Fig. 8.

4. Conclusions

The results obtained of the estimation of the parameter F_0, using the method of linear prediction, in the algorithm described above, fully represent the averaged (smoothed) signal of the real values of F_0 obtained, for example, as a result of the implementation of an algorithm of "primary segmentation" [2]. This smoothing is the result of the averaging of the values of parameters both during calculation of the autocorrelation function and the approximation of results, permitting a partial elimination of disturbances in the estimation of the parameters. Thus the parameters $\{F_{0i}\}$, determined by this method, are characterised with a considerably smaller scatter of their values than in the case of "primary segmentation".

The program, written in Fortran 6000 for a H-6030 Honeywell Bull machine, developed by the author, was not optimized either in terms of memory occupied or processing time, and in the present form occupies about 15 K words of memory (together with Fortran subprogrammes and programmes for plotter operation). Thus as an example the implementation time of a programme for a 2-second continuous speech signal (recorded in 50 tape blocks, 512 words-signal samples each) was $\cong 57.9$ s (processor time). It is a relatively long time and excludes the implementation of investigations in real time. It should, however, be noted that the H-6030 machine was designed for data processing and not for scientific and technical calculation, and with a calculation speed of the order of about 200000-300000 operations per second naturally offered no such possibilities.

The mean number of wrongly estimated values of the sequence $\{F_{0i}\}$, using the method of linear prediction, calculated for a continuous speech signal with a duration of 80 s, did not exceed 2.5 % of the total number of values estimated. The errors occurred particularly in those places of the signal where plosives appeared, which could be classified as voiced quasiperiodic signals. These can be eliminated by the introduction of a more precise way of determining the boundaries of the segments of voiced sounds in the stage of signal classification. In spite of those imperfections, the results obtained can be used in further stages of human speech analysis. While the results obtained from "primary segmentation" [2] served for a speech recognition oriented spectral analysis synchronised with the larynx tone [3], the results obtained from the implementation of the programme, based on the method of linear prediction, are not suitable for this type of analysis. They may, however, still be used in statistical investigations of the parameter F_0 with a view to identification or verification of a speaker on the basis of the statistical characteristics of the distribution of the frequency F_0.

References

[1] J. W. Bayless, S. J. Campanella, A. J. Goldberg, *A survey of speech digitization techniques*, Proceedings of Carnahan conference on electronic crime countermeasures, 80-81 (1972).

[2] A. Dziurnikowski, *Microphonemes as fundamental segments of a speech wave, Primary segmentation – automatic searching for microphonemes*, Advence Papers of the IV IJCAI – 75, Vol. 2, Tbilisi 1975.

[3] A. Dziurnikowski, *Primary sequentation of speech sound signals in the SUSY system* (in Polish), Reports JJ UW, **52** (1976).

[4] G. C. M. Fant, *Acoustic theory of speech production*, Mouton and Co., s' – Gravenhage, The Netherlands, 1960.

[5] K. Jaszczak, *Digital modelling of speech signal using linear prediction*, JJ UW, **62** (1977).

[6] J. D. Markel, A. M. Gray, *Linear prediction of speech*, Springer Verlag, Berlin – Heidelberg – New York 1976/1977, 132-134, 190-206.

[7] L. R. Rabiner, *Digital-formant synthesizer for speech synthesis*, JASA, **43**, 822-828 (1968).

[8] D. R. Reddy, *Pitch period determination of speech sound*, Communications of the ACM, **10**, 6, 343-348 (1967).

[9] S. Saito, F. Itakura, *The theoretical consideration of statistically optimum methods for speech spectral density*, Report No. 3107, Electr. Commun. Lab., N.T.T., Tokyo 1966.

[10] M. M. Sondhi, *New methods of pitch extraction*, IEEE Trans., AU-16, 262-266 (1968).

[11] B. Wierzchowska, *Polish pronunciation*, PZWS, Warszawa 1971.

Received on February 23, 1979.

ARCHIVES OF ACOUSTICS
5, 1, 47-58 (1980)

ANALYSIS OF THE CONFIGURATION OF THE ACOUSTIC FIELD OF LOUDSPEAKER SYSTEMS EXCITED WITH A SINUSOIDAL SIGNAL

EDWARD HOJAN

Acoustics Department, UAM (60-769 Poznań)

The paper discusses the accuracy of the approximation of the acoustic pressure distribution radiated by loudspeaker systems with the pressure distribution radiated by a flat circular membrane. The far and near field ranges of loudspeaker systems were determined and the effect of the irregularity of the acoustic pressure distribution in the near field of the systems on their amplitude and frequency characteristics was investigated.

1. Introduction

Determination of the acoustic potential function at a given measurement point of the acoustic field of the sources of disturbances can be performed using the Kirchhoff integration formula [5, 6]. In the case of radiation from a plane or a limited field on a plane, this formula becomes identical to the Huygens' integration formula.

Considering a vibrating element placed in a rigid baffle of infinite extent, the acoustic pressure at a point P (Fig. 1) can be determined from the formula

$$p(P) = \sigma \frac{\partial}{\partial t}\left[\frac{1}{2\pi}\iint_F \frac{e^{-ikr}}{r}\frac{\partial \psi}{\partial n}\,dF\right], \tag{1}$$

where σ is the density of the medium, r — the distance between the measurement point and an element of the source, k — the wave number, and ψ — the potential on the surface of a source of area F.

A detailed analysis [7] of formula (1) shows the necessity of distinguishing between the near field and the far field in the radiation field of the source of the disturbance.

In the far field the particle velocity of the medium is in phase with the acoustic pressure, and in the near field it is shifted by $\pi/2$. Accordingly, the near field does not contribute to the energy transmitted by the source to the medium. The imaginary component of the acoustic field characterizes the streaming of the medium near the source, where the medium not only moves in the direction of the radiation but also tangentially to the vibrating surfaces.

Determination of the analytical conditions differentiating the far field and the near field [7] leads to two inequalities which define the range of the far field:

$$R^2 \gg a^2 = x^2+y^2, \tag{2}$$

$$R^2 \gg \pi a^2/\lambda. \tag{3}$$

In the near field only the inequality

$$R^2 \gg a^2 \tag{4}$$

must be satisfied, with the condition $R \gg \lambda$ being valid for both near and far fields.

The radiation region that is directly adjacent to the vibrating surface, where inequalities (2) and (3) are not satisfied, has been called the *range of the very near field*.

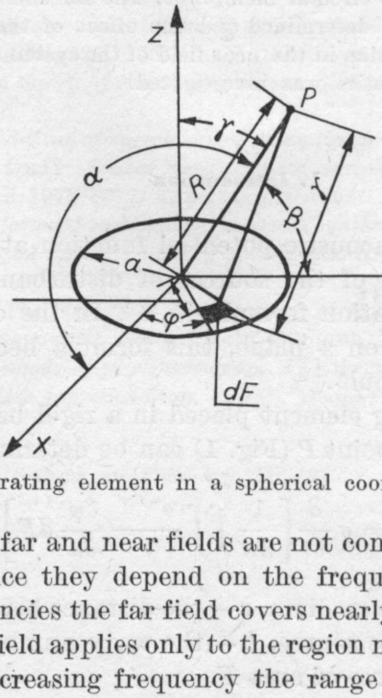


Fig. 1. A vibrating element in a spherical coordinate system

The ranges of the far and near fields are not constant quantities for a given vibrating element, since they depend on the frequency of the acoustic wave emitted. At low frequencies the far field covers nearly the whole of the radiation region, while the near field applies only to the region near the surface of the vibrating element. With increasing frequency the range of the near field expands. Fig. 2 shows the curves illustrating the acoustic pressure distribution along the axis of a circular membrane of a radius $a = 9$ cm, under the assumption that it performs piston-like simple harmonic motion in a direction perpendicular to the surface of the source. The calculations were made using formula (1).

It follows from the curves presented that for $(R/a^2)\lambda \geqslant 0.75$ they decrease monotonically, being proportional to $1/R$. The point for which $(R/a^2)\lambda = 0.75$

separates (for successive curves whose parameter is the ratio a/λ in the range 0.24-5.00) the region, where the phase and interference phenomena which occur affect the uniformity of the acoustic pressure distribution of the signal, from the region, where these phenomena do not occur.

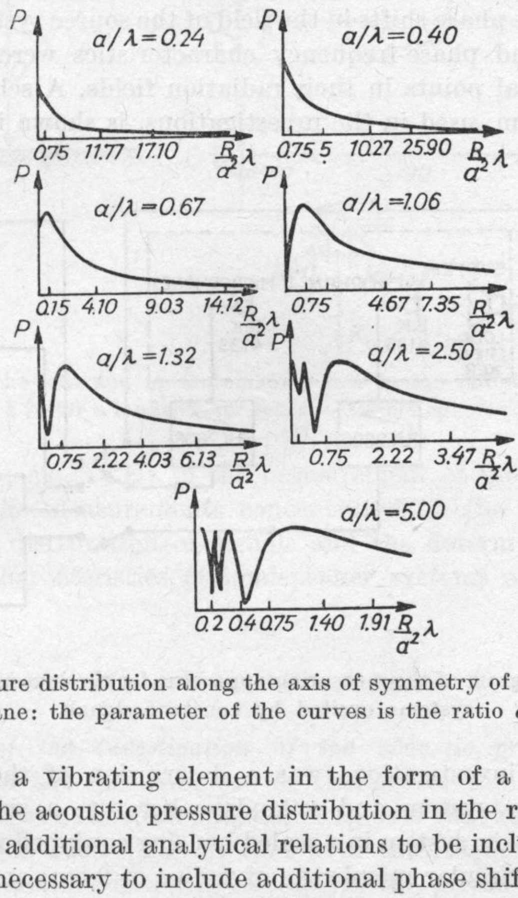


Fig. 2. Acoustic pressure distribution along the axis of symmetry of a circular piston membrane: the parameter of the curves is the ratio a/λ

In the case of a vibrating element in the form of a conical membrane, the calculation of the acoustic pressure distribution in the radiation field of the membrane requires additional analytical relations to be included in formula (1). In particular, it is necessary to include additional phase shifts due to the length and angle of inclination of the ruling of the cone. These additional phase shifts are also dependent on the form of the division of the loudspeaker membrane. For a loudspeaker set these problems become more complicated because of the parallel operation of several loudspeakers with different conical membranes.

Recognizing that there appeared to be no previous attempt at an analysis of all problems connected with a description of the acoustic field of a loudspeaker system over the whole radiation region, experimental investigations were begun which were aimed at the evaluation of the accuracy of the approximation of this type of system by a simple vibrating element (Fig. 1). At the same time the effect of the irregularities occurring in the acoustic pressure distribution in the radiation field of systems (very near field, near field) on the amplitude and frequency characteristics registered in these regions was determined.

2. Experimental method and measuring apparatus

The investigation of the configuration of the acoustic field of simple and complex loudspeaker systems was performed for sinusoidal excitation signals. In the framework of these investigations the acoustic pressure distribution and the magnitude of the phase shifts in the field of the source system were measured. Both amplitude- and phase-frequency characteristics were measured for the systems at individual points in their radiation fields. A schematic diagram of the measuring system, used in the investigations, is shown in Fig. 3.

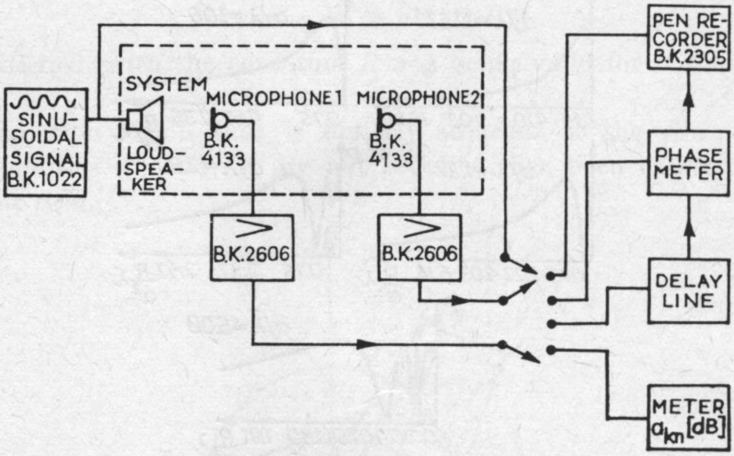


Fig. 3. A schematic diagram of the measuring apparatus for the investigation of loudspeaker systems excited by constant signals

The object of investigations was a description of the acoustic field of a single loudspeaker system and a loudspeaker set: a single Isophon type PSL 203/25 loudspeaker system in a sealed housing equipped with one broadband loudspeaker with a circular membrane of radius of 9 cm; and a Klein-Hummel loudspeaker set consisting of two identical broadband loudspeakers and one high-tone loudspeaker, installed in a sealed housing. The spacing of measurement points in the acoustic field of both systems is shown in Fig. 4 (the investigation was performed in an anechoic chamber). Each loudspeaker system was excited by acoustic generators. A signal from the loudspeaker system was simultaneously received by two microphones, with microphone M-2 being always at the same place (the signal received from this microphone was used as a reference signal), and microphone M-1 changing its position in the course of the measurements. Signals from the microphones were supplied, through amplifiers, to recording systems for further processing.

The excitation level of the system, corresponding to a specific value of SPL, was established at a measurement point (a, $R = 180$ cm) for a signal at

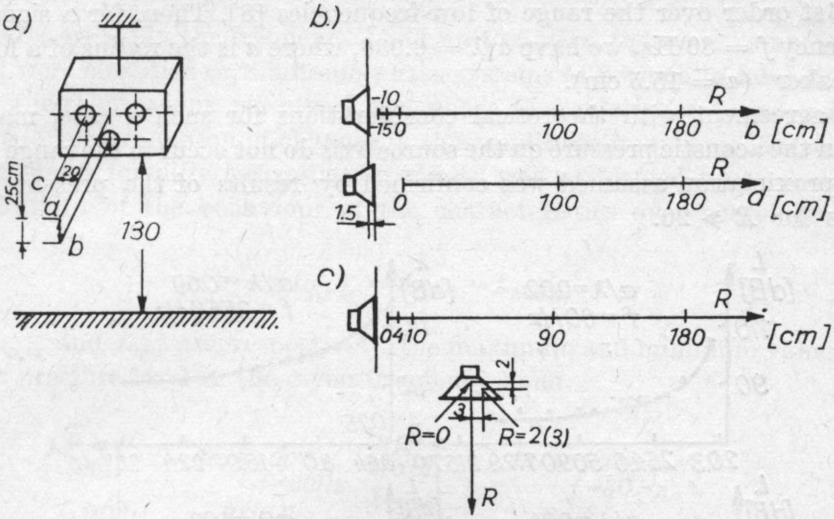


Fig. 4. Position of the axis and measurement points in the radiation fields of loudspeaker systems: *a, b* — for a loudspeaker set, *c* — for a single loudspeaker system

a given frequency, e.g. 1 kHz in the measurement of the amplitude-frequency characteristics. The measurements concentrated on the determination of the acoustic pressure distribution of signals and the determination of the amplitude frequency characteristics of loudspeaker systems along the loudspeaker axes.

3. Analysis of results of the experimental investigations

(a) *Results of the investigation of the acoustic pressure distribution of signals.* Results of the measurement of changes in the level of the acoustic pressure of a signal as a function of distance R, over the radiation field of a single loudspeaker system, show (Fig. 5) that over the frequency range, where the ratio $a/\lambda \ll 1$, there is a uniform decrease in the pressure level for $R \geqslant 10\text{-}20$ and for an exciting voltage giving SPL $= 80$ dB (an increase in the excitation of the single loudspeaker system to a level corresponding to SPL $= 90$ dB caused significant changes in this behaviour, in particular, for frequencies $f > 150$ Hz and $R < 70$). When the ratio $a/\lambda \geqslant 1$, the greatest irregularities occur for small values of R. A uniform decrease in the acoustic pressure level then occurs for $R \geqslant 40\text{-}60$.

In the evaluation of variation of the acoustic pressure level L as a function of R for a loudspeaker set (Fig. 6), an essential difficulty occurs in the determination of ranges of the near and far fields, connected with the fact that a loudspeaker set is a complex radiating system in which several sources operate at the same time. Neglecting all complex phase and interference phenomena in the first approximation, such a system can be considered as a spherical source of the 1st order over the range of low frequencies [8]. Then, for a signal with a frequency $f = 80$ Hz, we have $a/\lambda = 0.036$, where a is the radius of a low-tone loudspeaker ($a = 15.5$ cm).

In agreement with theoretical considerations for such a case, maximum values in the acoustic pressure on the source axis do not occur in the range $R \gg a$. The approximation assumed was confirmed by results of the present measurements for $R \gg 20$.

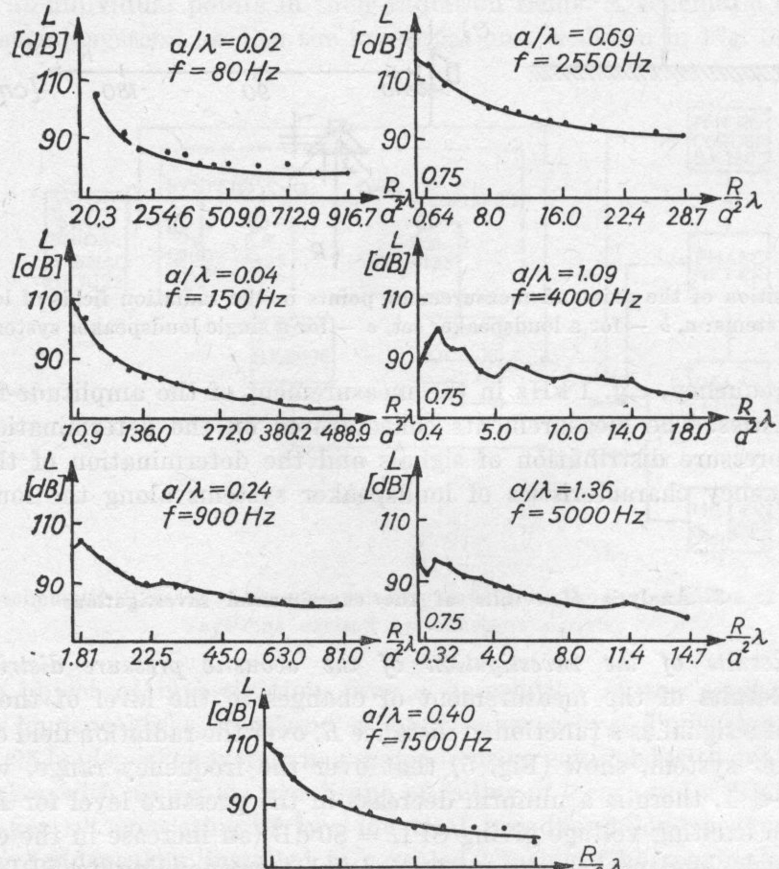


Fig. 5. Relative acoustic pressure level L of a sinusoidal signal as a function of the distance R of the measurement point from the radiation plane of a single loudspeaker system; SPL = 80 dB

This simplification is not valid for signals of higher frequencies ($f = 1000$ Hz and $f = 2000$ Hz). In this case the system should be considered as a group of sources with individual directional characteristics during the calculation of the ratio a/λ, and a is the linear size of the housing of loudspeaker set.

(b) *The results of the investigation of the amplitude-frequency characteristics.* Based on the comparison of the behaviour of the amplitude- and phase-frequency characteristics [1] in the frequency band up to 5 kHz, the systems under investigation were classified as minimum phase systems [4], which in further evaluation of their behaviour permitted the phase characteristics to be neglected. A quantitative evaluation of changes in the amplitude-frequency characteristics of the systems for different radiation ranges was performed by measuring the nonuniformity of the behaviour of the characteristics over chosen frequency ranges,

$$\Delta L_{f_1-f_2} = L_{\max} - L_{\min}, \tag{6}$$

where $L_{\max}$ and $L_{\min}$ are, respectively, the maximum and minimum values of the acoustic pressure level in the given frequency band.

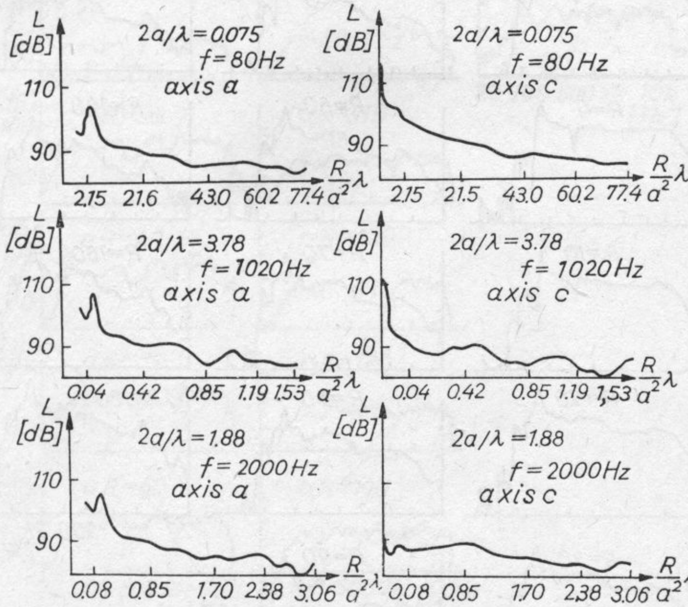


Fig. 6. Relative acoustic pressure level L of a sinusoidal signal as a function of the distance R of the measurement point from the radiation plane of a loudspeaker set; SPL = 80 dB

It was found that the maximum changes of $\Delta L_{f_1-f_2}$ with increasing distance R of the measurement point from the radiation plane occur in the frequency band $f_1-f_2 = 3\text{-}5$ kHz. In the acoustic field of a single loudspeaker system these changes are largest for $R \leqslant 10$, when the excitation level of the system corresponding to an output SPL = 80 dB and may reach 17-25 dB (Fig. 7). For the lower frequency band the ranges of R, where the changes in $\Delta L_{f_1-f_2}$ occur, are the same as before with the numerical value of $\Delta L_{f_1-f_2}$ being lower than the maximum noted above.

An increase in the distance of the measurement point from the axis of symmetry of the system for low R ($R = 2/3$) gave a visible decrease in the absolute value of $\Delta L_{f_1-f_2}$ in the highest frequency band (closely approximating the values of $\Delta L_{f_1-f_2}$ for high R), while this was not observed at low frequencies. In the evaluation of these changes in the radiation field of the loudspeaker set (Fig. 8,9) it can be noted that they are not only a function of the distance R, but also of the position of the measurement axis.

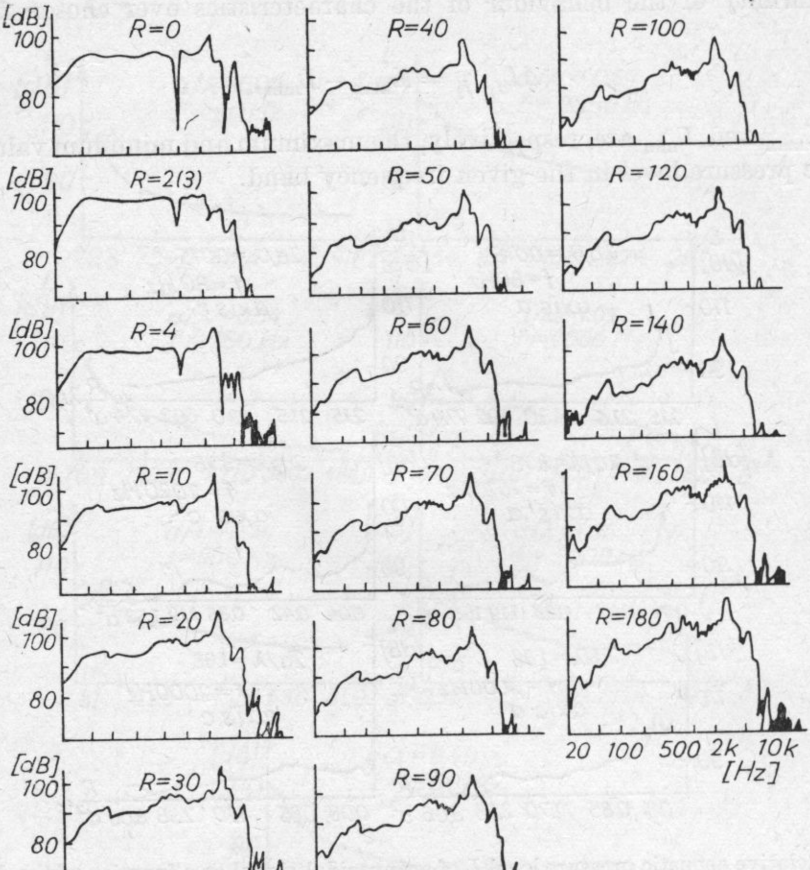


Fig. 7. Amplitude-frequency characteristics of a single loudspeaker system for different values of R; SPL = 80 dB

The greatest differences in the behaviour of the characteristics occurred for a shift of the measurement axis with respect to the axis of symmetry of the system in a vertical plane, and the smallest ones for a shift in a horizontal plane. With the location of the measurement points along the axis of symmetry of the set (the a-axis), it was found that the values of $\Delta L_{f_1-f_2}$ in the low frequency band are practically constant functions of the distance R. At higher frequencies the values of $\Delta L_{f_1-f_2}$ become constant for $R \geqslant 30$. Changes in the values of $\Delta L_{f_1-f_2}$ along the b- and c-axes are strongly dependent on frequency; they also

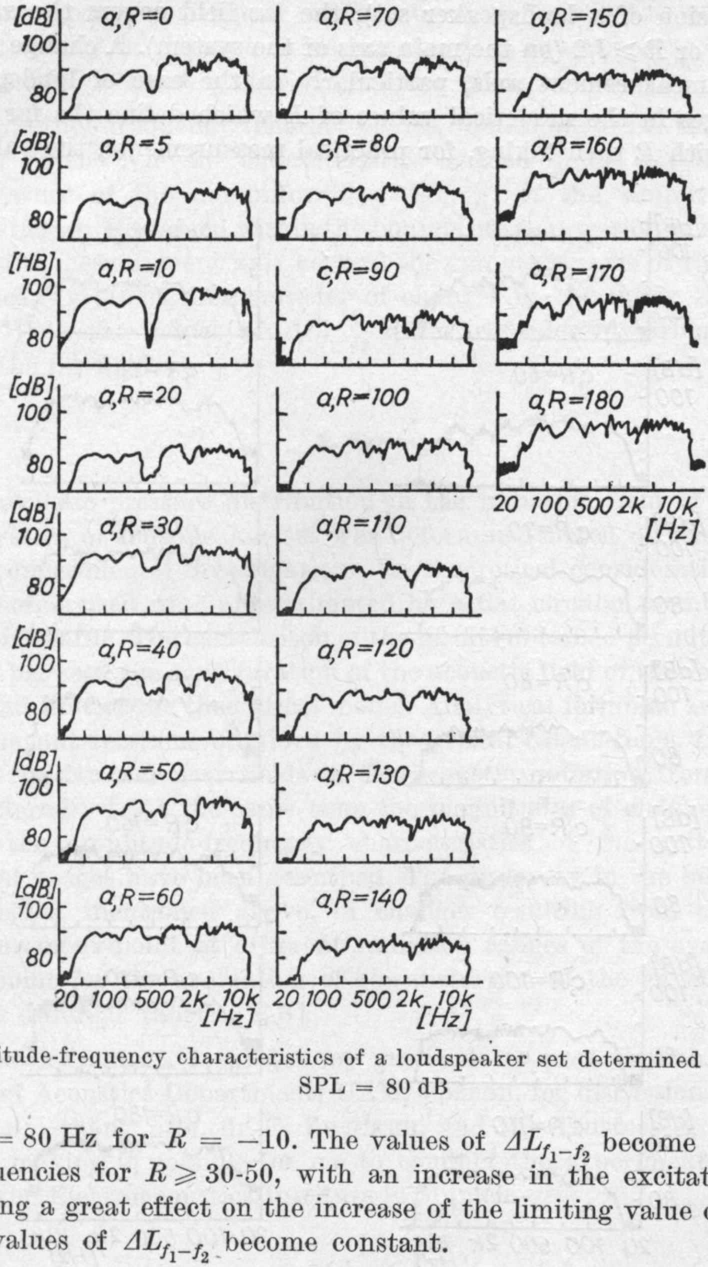


Fig. 8. Amplitude-frequency characteristics of a loudspeaker set determined on the a-axis; SPL = 80 dB

occur at $f = 80$ Hz for $R = -10$. The values of $\Delta L_{f_1-f_2}$ become constant at higher frequencies for $R \geqslant 30$-50, with an increase in the excitation level of the set having a great effect on the increase of the limiting value of R beyond which the values of $\Delta L_{f_1-f_2}$ become constant.

4. Conclusions

1. The following rules, which specify the relations given in formulae (2) and (3), can be used to define the far field of loudspeaker systems:

at low frequencies, for which the inequalities $a < 0.50\,\lambda$ or $l < \lambda$ are valid (a – the radius of membrane in a single loudspeaker system, l – the maximum linear dimension of a loudspeaker set), the far field covers the range R for which $R \geqslant a$ or $R \geqslant l/2$ (on the main axis of the system). A change in the position of the measurement axis, particularly in the case of loudspeaker sets, causes changes in the numerical values of R which define the far field range to change, with R then taking, for practical measurements, the values $R \geqslant 2a$ or $R \geqslant l$;

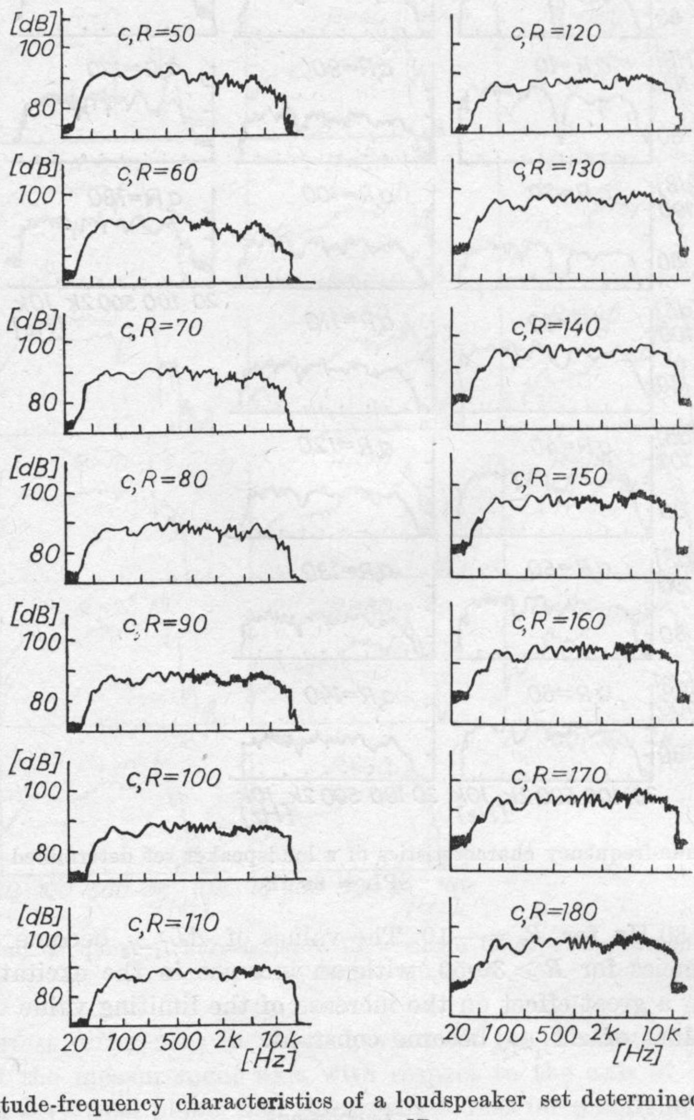


Fig. 9. Amplitude-frequency characteristics of a loudspeaker set determined on the c-axis; SPL = 80 dB

at high frequencies, for which the inequality $a \geqslant 0.50\lambda$ or $l \geqslant \lambda$ is satisfied, the far field covers the range where $R \geqslant \pi a^2/\lambda$ for a single loudspeaker system and $R \geqslant 0.75 \cdot (l/2)^2/\lambda$ for a loudspeaker set (on the main axes of the systems).

A change in the position of the measurement axis in the systems investigated, particularly in the case of loudspeaker sets, causes changes in the numerical values of R defined by the inequalities given above.

2. Amplitude-frequency characteristics, determined in the near fields, are similar to the relevant characteristics defined in the far fields. Changes in the measure of the nonuniformity, $\Delta L_{f_1-f_2}$, of the amplitude-frequency characteristics are contained within the bounds of the precision of measurements. A shift of the measurement axis beyond the symmetry axis of the system does not significantly affect the character of changes in the index $\Delta L_{f_1-f_2}$. In the very near field the values of $\Delta L_{f_1-f_2}$ differ considerably from the relevant values in the far field.

5. Summary

The acoustic pressure distribution in the radiation field of a single loudspeaker system or lousdpeaker set was determined based on theoretical equations and experimental investigations. In theoretical considerations the loudspeaker systems used were approximated by a flat circular membrane installed in an infinite baffle. The comparison of the results obtained permitted differences to be seen between the configuration of the acoustic field of the real loudspeaker systems and that of the theoretical model. Analytical formulae are given, which specify relevant relations obtained by theoretical calculations, from which the ranges of the far and near fields in the acoustic radiation from loudspeakers can be determined. At the same time the magnitudes of changes in the behaviour of the amplitude-frequency characteristics of the systems measured at different ranges have been described. The discovery in the behaviour of the characteristics, mentioned above, of changes resulting from the position of the measurement point at different radiation ranges of the systems, was the starting point for the evaluation of the distortion of the envelopes of pulsed signals at different ranges [2, 3].

Acknowledgement. I express my gratitude to prof. dr hab. H. Ryffert, the head of Acoustics Department, UAM, Poznań, for discussion on the present paper. I also thank prof. dr E. Zwicker, and the Trustees of the Humboldt Fund for making it possible for me to conduct the experimental work in the paper at the Electroacoustics Institute in Munich.

References

[1] E. Hojan, *Analysis of the deformation of transient signals connected with the configuration of the acoustic field of a given system of sources* (in Polish), Wyd. UAM, Seria Akustyka, **6**, Poznań 1979.

[2] E. Hojan, *Detectability of pulse distortion in an acoustic loudspeaker field,* Archives of Acoustics, **2**, 4, 49-66 (1977).

[3] E. Hojan, *Distortion of transients over the radiation range of loudspeaker systems* (in Polish), ibidem (in press).

[4] B. P. Lathi, *Theories of signals and telecommunication systems* (in Polish), PWN, Warszawa 1970.

[5] I. Malecki, *Theory of waves and acoustic systems* (in Polish), PWN, Warszawa 1964.

[6] P. Morse, K. Ingard, *Theoretical acoustics*, Mc Graw Hill, New York 1968.

[7] H. Stenzel, O. Brosze, *Leitfaden zur Berechnung von Schallvorgängen*, Springer Verlag, Berlin 1958.

[8] B. Vinogradova, *Blizhnee pole lineynoi izluchayushchei gruppy*, Akusticheskii Zhurnal, **12**, **1**, 7-11 (1966).

Received on May 23, 1978; final version on September 10, 1979.

ARCHIVES OF ACOUSTICS
5, 1, 59-72 (1980)

PSYCHOACOUSTIC INVESTIGATION OF LISTENING WITH QUADROPHONIC EARPHONES

TOMASZ ŁĘTOWSKI, WOJCIECH MAKOWSKI

Department of Musical Acoustics, State Higher Musical School (00-368 Warszawa)

Despite the existence of a variety of designs of quadrophonic earphones, there have been no papers concerned with the psychoacoustic phenomena accompanying quadrophonic earphone projection. The present paper presents the results of seven experiments in the field of the psychoacoustics of hearing under conditions of earphone projection involving the use of four transducers, and also presents the technical principles of the formation of binaural images. The results obtained permit the conclusion to be drawn that quadrophonic or even stereophonic earphone projection involving four transducers enables the sound sources to be located outside the listener's head. It seems interesting that the listening to stereophonic signals with four transducers gives effects incomparably better than those for the projection with two transducers and is of a quality comparable with the monitoring of quadrophonic signals.

1. Introduction

The development of phonic techniques, permitting the multichannel recording of sound has caused an increasing interest in multichannel systems of sound projection which permit the formation of complex spatial images with the signals emitted from many electroacoustic sources placed in a given space. The best known of systems of this type is that of quadrophonic projection consisting of sound emission from four sources located around the listener and controlled by the signals: *LF* (left front), *RF* (right front), *LR* (left rear), and *RR* (right rear). Various multichannel systems of sound projection are met at present in shows and cinema halls, in the outdoor spectacles of the "Son et Lumière" type, in variety shows, and even in apartments. In all cases the sound is emitted from sets or columns of loudspeakers.

The spreading of multichannel techniques of sound projection caused the first multichannel earphones, the so called quadrophonic earphones, to appear in the early 1970's. Typical quadrophonic earphones are equipped with four transducers — placed in pairs in enclosures around the ears — permitting independent emission to the listener's ears of signals from the front (F) and rear (R) channels. The idea of quadrophonic earphones — permitting the formation of spatial sound impressions similar to the impressions caused by multichannel systems of loudspeaker projection, and at the same time permitting the technical requirements to be simplified and the cost to be lowered considerably — has long fascinated designers, while at same time raising many doubts. The difficulties in the evaluation of the possibility of building coherent three-dimensional spatial impressions with earphone projection were due primarily to the lack of fundamental research in the field of multichannel audio monitoring. The possiblity of the listener being able to separate and locate the signals from the front and rear channels seemed the most controversial question. In general, there is a lack of papers concerning, for example, the influence of the ear lobe on the sound impressions produced under conditions of earphone projection.

The investigations reported here were aimed at the evaluation of the resolution of signals radiated by the front and rear transducers under different projection conditions and the qualitative comparative evaluation of quadrophonic earphones in the stereo and quadrophonic projection of test signals. The investigations were performed on three models (A, B, C) of quadrophonic earphones designed in KAM-PWSM (Department of Musical Acoustics, State Higher Musical School) [6]. Each of the models was equipped with four Tonsil SN 60 transducers placed in pairs in the right and left earphone enclosures. All the transducers were in the same horizontal plane, producing a system which was symmetrical with respect to the plane which included the centres of the listener's ears. In models A and B, the individual pairs of transducers were placed parallel to each other (side by side) at distances of 30 and 45 mm, respectively. In model C the distance between the transducers was $l = 45$ mm, as in model B, but the planes of the transducer membranes were at an angle of 135° to each other. Model C was designed in two versions (C-1 and C-2) differing only in the form of the earphone pads. The C-2 version, which was equipped with deeper and more elastic pads, was investigated in the later phase of the experiments.

In addition, three standard foreign-made earphones were included in the investigations: DT 204 of Beyer, DH 32Q of Hosiden, and 5844 of JVC Nivico.

Also, in the comparative evaluation phase the investigations included a pair of broadband stereophonic ZWG Tonsil SN 60 earphones.

The audio monitoring group consisted of 15 persons chosen from among the staff and students of the Sound Recording Department of State Higher Musical School in Warsaw. All the tests were made on each participant individually.

The experimental investigations performed consisted of 7 tests. The tests were concerned with:

(a) the recognition of signals radiated by individually activated front and rear transducers;

(*b*) the discrimination between mono and stereophonic projections from a set of transducers operating in a common casing;

(*c*) the discrimination between stereophonic signals radiated by the front and rear pairs of transducers;

(*d*) the discrimination between two-transducer and four-transducer stereophonic projections by quadrophonic earphones;

(*e*) the discrimination between mono, stereo and quadrophonic projections;

(*f*) a comparative evaluation of the quality of quadrophonic earphones;

(*g*) a comparison of the properties of quadrophonic monitoring with earphones and with loudspeakers.

2. Recognition of signals radiated by individually activated transducers

The aim of this phase of the investigations was the determination of the possibility of the listener recognition of the signals emitted by the individual transducers. The earphones were activated individually with a tone at a frequency of 1000 Hz or with white noise. The test signals were reproduced from a tape recorder through a switch system permitting the selection of one of eight possibilities: *FR*-tone, *FR*-noise, *LF*-tone, *LF*-noise, *RR*-tone, *RR*-noise, *LR*-tone, and *LR*-noise.

A single test consisted of 16 tasks (two tasks of each type) in random succession. The duration of the test tasks was 5 s, with pauses for the listener's answer lasting about 3 s. The duration of the test for one earphone was about 2 min, with the duration of the test for all the earphones being about 20 min. The task of the experts was to determine the direction of the signal received: *RF*, *LF*, *RR*, or *LR*.

The results obtained for the different signals are shown in Table 1.

Table 1. Recognition of signals radiated by individually activated front and rear transducers

Kind of signal	Percentage of correct answers P				
	A	*B*	$C-1$	DT 204	DH 320
Tone 1000 Hz	57	87.5	68.1	78.8	58.8
White noise	98.8	98.8	100	100	97.6

The results obtained show that with the earphones having four transducers, in addition to the obvious possibility of "left-right" direction differentiation, it is possible also to differentiate between transducers in the "front-rear" configuration. The resolution of the transducers is closely correlated with the spectral structure of the signal emitted. In the testing of the individual transducers with a signal having a wide spectrum (white noise), the sample investigation showed nearly 100 percent agreement between the responses and the real projection conditions. However, for the emission of a sinusoidal signal, the number of correct answers decreased to between 60 and 90 %, depending on the specific design of the earphone (the possibility of identification is represented by $P \geqslant 75\,\%$). In the interpretation of the results obtained it should be stressed that the possibility of identifying the transducers is not equivalent to the specific directional location of the sound impressions received. The basic criterion permitting the resolution of the signals emitted by the front and rear transducers was the difference in the timbre of the impressions received, resulting from the transformation of the spectrum of the signals received, due to the influence of the ear lobes and hearing channels of the observer. In the case of broadband signals, a characteristic distortion of timbre occurring on the signals emitted in the rear channels (a decreased content of higher frequency components) enabled the evaluators to perform the test tasks correctly. However, at the same time – apart from whether the front or rear transducer was operating – the listeners retained the impression that the sound source was located directly in the excited ear.

In case of narrowband signals when the screening effect of the ear lobe was evidenced only by slight changes of the loudness of the signal emitted by the front or rear transducer, the discrimination between the projection conditions was considerably worse.

The above observations permit the conclusion to be drawn that in quadrophonic earphone projection, the amount of directional information on the location of the sources in the lateral spaces is very limited. The situation is analogous to the uniaural reception od signals reaching the listener as a result of quadrophonic loudspeaker projection [2]. In both cases the decrease in the amount of information on the spatial distribution of the sources results from the lack of binaural perception of the signals carrying the information.

3. Discrimination between mono- and stereophonic projection for a set of transducers working in a common casing

The aim of the investigation was to determine the possibility of listener discrimination between mono and stereophonic projection of signals to one ear. The sound material of the test was a stereophonic tape recording of the song "Aquarius" from the musical "Hair". A sound signal was transmitted to a pair of transducers placed in a common encasing and reproduced mono- or stereo-phonically in ten-second fragments with a one-second pause between successive tasks. Altogether, the test contained eight tasks in random order. Each test was preceded by four examples demonstrating mono and stereophonic projections in a M-S-M-S regime (4×5 s). The duration of the test for one earphone was about 2 min. and it took about 10 min. to test all the earphones. The task of the experts was to determine whether the fragment of the song presented was reproduced mono-or stereophonically.

The results obtained are shown in Table 2.

Table 2. Results of the discrimination between mono- and stereophonic projections for one ear

Earphone	A	B	C	D	E	F	G	H	J	K	$\bar{x}$	σ	$P[\%]$
KAM/A	4	4	3	5	5	3	4	4	6	5	4.3	0.95	53.8
KAM/B	7	7	8	4	7	7	6	6	7	6	6.5	1.08	81.2
KAM/$C-1$	7	8	3	8	5	5	6	5	6	7	6.0	1.56	75.0
DT 204	6	4	6	8	5	7	8	4	6	7	6.4	1.43	80.0
DH 32 Q	5	4	6	8	5	8	3	8	6	7	5.7	1.70	71.2

$\bar{x}$ – arithmetic mean of results obtained for 10 listeners
σ – standard deviation of results
P – percentage of correct answers

The results obtained show that with two-transducer emission of signals to one ear it is possible to distinguish a monophonic projection from a stereophonic one ($P \geqslant 75\,\%$). The criteria permitting such a distinction are not, however, unambiguous, since none of the earphones investigated gave a 100 percent correct response and none of the experts avoided mistakes in the solution of the test tasks.

The results of the experiment performed show a significant correlation with some of the results obtained in the first experiment (see Table 1 under a continuous tone test signal at 1000 Hz). Classifying the earphones tested in terms of the percentage of correct answers, both tests gave the same rank order: KAM, DT 204, KAM $C-1$, DH 32Q, KAM/A. The results of the Students t – test showed that, in addition, the differences between the first four earphones of the above series were statistically insignificant. However, the differentiation of the type of projection for earphone KAM/A was significantly worse (at a level of $\alpha = 0.05$). In the interpretation of the above results it should also be observed that the first three earphone designs are among those with a relatively long distance between transducer centres ($l = 45$ mm), while in model KAM/A the centre spacing was considerably shorter ($l = 30$ mm). The design of the DH 32Q earphones is essentially different from the other designs. Three transducers (two high-tone and one low-tone) are placed in each enclosure of the earphones, with the signals from both F and R channels being supplied – through electrical switches – to a common low-tone transducer in the central position facing the inlet of the ear channel, and to two high-tone transducers deeper placed inside the enclosure, with their radiation axes facing each other. As a result, for both mono and stereo emissions the low-tone signals retain the central location, and the signals emitted by high-tone transducers mix in the common part of the inner earphone chamber; thus the directional information contained in these signals becomes less distinct.

4. Discrimination between stereophonic signals radiated by the front and rear pairs of transducer

As before, the sound material used in the experiment was a stereophonic recording of the song "Aquarius" from the musical "Hair". The sound signal was supplied to the front and rear pairs of transducers in ten-second fragments divided by one-second pauses. The test consisted of eight sound tasks presented in random order, preceded by four examples demonstrating the projection by the front and rear pairs of transducers — in the regime $F-R-F-R$ (4×5 s). The duration of the test was about 2 min. for one earphone. The task of the experts was to determine which pair of the transducers was reproducing the song fragment heard.

The results obtained are shown in Table 3.

Table 3. Results of discrimination between of stereophonic signals radiated by the front and rear pairs of transducers

Earphone \ Expert	A	B	C	D	E	F	G	H	J	K	$\bar{x}$	σ	$P\%$
KAM/A	7	8	5	4	6	6	7	8	8	8	6.7	1.42	83.75
KAM/B	8	7	8	7	8	8	8	8	8	8	7.8	0.42	97.5
KAM/$C-1$	8	8	6	7	8	7	6	8	8	8	7.4	0.84	92.5
DT 204	8	8	7	8	8	8	8	8	8	8	7.9	0.32	98.75
DH 32Q	6	7	6	7	6	7	7	4	6	4	6.0	1.15	75

$\bar{x}$ — arithmetic mean of results obtained for 10 listeners
σ — standard deviation of results
P — percentage of correct answers

The results of the experiment performed show the possibility of resolving the projection of stereophonic signals emitted from the front and rear pairs of the transducers of quadrophonic earphones ($P \geqslant 75\,\%$). The very high mean values for the percentage of correct answers for most of the earphones tested, show that the correct solution of the test tasks was not difficult for the participants in the experiments. The imposing evaluation criterion was a different sound perspective for the work of the front and the rear channels. Changes of the spectrum of the signals radiated by the rear transducers, resulting from the screening effect of the ear lobes, were interpreted by the monitoring group as fading and flattening of the subjective space of the sound events. The effect was greatest for the earphones with a long distance between the transducer centres (DT 204, KAM/B, KAM/$C-1$), and was less noticeable for earphones KAM/A, and, in particular, for DH 32Q (a difference significant at a level of $\alpha = 0.02$).

5. Discrimination between two-transducer and four-transducer stereophonic projections by quadrophonic earphones

The aim of the following experiments was to establish subjective criteria which would permit the listeners to distinguish between two-transducer and four-transducer stereophonic projections (the possibility of such a distinction was shown indirectly in the investigations discussed above).

The investigations were performed on the three pairs of earphones, which in the previous experiments showed the highest resolution of the signals emitted by the individual transducers, i.e. *DT* 204, KAM/*B*, KAM/*C*-1, also on KAM/*C*-2 (a model similar in design to KAM/*C*-1, but equipped with different ear-pads) and on JVC 5844 (of Nivico). The JVC 5844 earphone represented in terms of design a type similar to the KAM/*C* models, with transducers inclined at an angle, of centre-spacing $l = 50$ mm, placed in spacious, deep, earphone enclosures.

Two experiments were made within the framework of the evaluation. In the first experiment a stereophonic recording of popular music was switched by the operator to the transmission channels *LF* and *RF*, or $LF+LR$ and $RF+$ $+RR$, respectively. The test contained 10 tasks presented in random order (five tasks for each projection variation), preceded by four examples demonstrating two-transducer and four-transducer projection. The duration of each task was 10 s., the duration of the pause between successive tasks was 1 s. In view of the higher pressure level produced by the earphones in four-transducer emission, and also because of the varying efficiency of the earphones tested, the investigations were preceded by a regulation of the testing levels so that during the evaluation, the listeners had the impression that the individual signals were of the same loudness.

Five experts participated in the audio monitorings. The task they were given was to determine whether the signal they heard was emitted from two or four transducers. As was expected, the results of the evaluation showed 10 percentage agreement between the answers and the real projection conditions.

The same group of listeners participated in the second part of the experiment. The experts' task was to compare the features of the projection discussed for all the earphones investigated, and subsequently to formulate observations on the variation in the sound timbre, directional location and spatial character of each of the projection variants. The monitoring was performed according to the principles of free rate of testing and free selection [5].

The experimental material permitted the following hypotheses to be formulated, concerning esthetic criteria for differentiating between two-transducer and four-transducer stereophonic conditions.

The essential differences, as stressed by all the evaluators, concern the imaginative space of the sound impressions received. In the case of two-transducer projection the sources of the sound image were located, in the listeners' opinion, inside the head, along a line in the form of a flat arc connecting the ears. With four-transducer projection the sources of the sound image were lowered or retreated, forming a space of sound action which expanded behind the listener's head, with decreasing resolution of the sound impressions obtained.

The fascination with spatial effects – unobtainable in the two-transducer transmission – caused the listeners' attention to concentrate to a lesser degree on the timbre of the sound. Nevertheless, the impressions obtained in four-transducer projection were found to be more natural, and closer to the subjective standards established in the direct listening.

In the general evaluation of the quality of the monitoring the experts' preference was the same: the four-transducer projection was considered to be decisively better.

An attempt at relating the regularities observed to the objective transmission conditions occurring in two-transducer and four-transducer projection of the signals leads to the observation that the basic cause of the difference in the sound impressions for the two-transducer and four-transducer stereophonic transmission is the asymmetry of the outer ear, which introduces different spectral transformations of the signals from the front and rear transducers.

The present asymmetry – apart from the modification of the transmission characteristics for the individual channels – also introduces time delays between the signals reaching the inlet of the ear channel from the front and the rear transducers. The combined influence of the factors mentioned above causes the four-transducer stereophonic projection to give a stronger impression (with respect to the conventional two-transducer projection) of the deepening of the sense of space of the sound image and the location of the sound sources outside the listener's head.

It is worth noting that in stereophonic earphone projection the time delays between the signals reaching the listener's ears are the basic cause of the effect of the sound sources being located outside the head. The time delays can have a twofold character

(a) as a delay in the time arrival of the same signal from one ear to the other,

(b) ambiphonic delays in signals reaching the same ear. In the case of four-transducer stereophonic projection the second factor plays an essential role. The time delay between the ears for the same signal (immitating the monitoring conditions occurring with loudspeaker projection) can also lead to the location of the sound sources outside the listener's head; however, at present there is a lack of adequate data concerning the conditions necessary to obtain a natural spatial effect in this manner.

6. Discrimination between four-transducer mono, stereo and quadrophonic projections

The aim of the present investigation was to determine the possibility of the listeners differentiating between the mono, stereo and quadrophonic projection of signals by quadrophonic earphones. The investigations were performed for five types of quadrophonic earphones, i.e. DT 204, JVC 5844, KAM/*B*, KAM/$C-1$ and KAM/$C-2$. The sound material used consisted of quadrophonic recordings of popular and symphonic music from the demonstration tape of the SANSUI firm, reproduced on a quadrophonic tape recorder M 2406 QD. Signals from the tape-recorder were supplied to the earphones through a mono-stereo-quadro switch permitting the combination of signals in accordance with the desired projection conditions.

Ten experts participated in the evaluation, each performing 45 test tasks (9 tasks for each earphone). Test tasks for each earphone were randomly ordered. The duration of each test task was 10 s., with a one-second pause between successive tasks. Each monitoring series was preceded by three examples: of mono, stereo and quadrophonic projections, respectively.

The results obtained are shown in Table 4.

Table 4. Discrimination between mono-, stereo-, and quadrophonic projections by quadrophonic earphones

Projection conditions	M	S	Q	P [%]
M	147	3	0	98
S	1	86	63	57.3
Q	0	67	83	55,3

The results of the investigation show, under the conditions of the experiments performed, that the essential qualitative difference permitting a correct identification of the variants of the projection presented occurred for the comparison of "mono" and "stereo" monitorings; while "stereo" and "quadro" projections were, in general, indistinguishable.

The results of the investigations complement the conclusions drawn earlier that stereophonic transmission through quadrophonic earphones leads to a considerable increase in the quality of the monitoring compared to conventional monitoring involving two transducers. As has been shown by the present investigations, the use of four independent transmission channels does not — under earphone conditions — appear to make the listener's sound impressions any more attractive.

On the basis of the conclusions and observations made during the investigations performed a general classification can be established of earphone systems of sound projection, in terms of the possibility of forming an imaginary space for the sound events (Table 5).

7. Qualitative evaluation of earphones with quadrophonic projection

The evaluation of the quadrophonic monitoring was made for all the seven pairs of quadrophonic earphones used in the previous experiments. A stereophonic pair SN 60 was also included in the investigations, its transducers being supplied with signals LF and RF, respectively.

Table 5. Earphone projection classification

No	Signal	Projection system	Observations
1	Monophonic	One-, two-transducer to one ear	Location of sound directly in the stimulated ear
2	Monophonic	Two-, four-transducer to both ears	Location of sound "in the middle" of the listener's head
3	Stereophonic	Two-transducer to both ears	Location of sound "inside" the head on a line between the ears
4	Stereophonic Quadrophonic	Four-transducer to both ears	Location of sound "outside" the listener's head*

* The effect of the location of the sound outside the listener's head can also be achieved during conventional binaural two-transducer projection. However, it involves the use of specialised recording or reproduction conditions for the sound signals, including the "artificial head" technique [3], two-channel projection of coded signals of a matrix of quadrophonic signals [1], or the technique of electronic mixing and mutual delaying of the *L*- and *R*-signals [7].

The evaluation was made, based on the method of ordering the earphones, with each of the 10 experts giving 7 points to the earphones he found best, while the other earphones were assigned 6, 5, 4, 3, 2, 1 and 0 points. The timbre of sound, sense of space, location of the sources and listening comfort were taken as the basic parameters giving jointly a general qualitative evaluation of the earphones compared.

Before the monitoring the levels of signals controlling the individual earphones were individually adjusted. A quadrophonic test recording of popular music by Polish Radio and Television was used as the sound material. The monitoring was performed at the listeners' own rate and with free selection. The total duration of the listening was about 30 min. for each listener.

The results obtained are shown in Table 6.

Table 6. Results of general evaluation of the quality of the earphones during quadrophonic projection

Earphone	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*J*	*K*	Σ	$\bar{x}$	σ
KAM/*A*	2	3	1	4	3	3	5	1	4	2	28	2.8	1.31
KAM/*B*	3	5	3	6	6	6	7	6	7	6	55	5.5	1.43
KAM/*C* – 1	5	4	5	3	2	2	1	3	2	4	31	3.1	1.37
KAM/*C* – 2	6	5	2	7	4	4	4	4	5	5	46	4.6	1.35
DT 204	7	0	6	2	7	7	6	5	3	3	46	4.6	2.46
JVC 5844	4	7	7	5	5	5	3	7	6	7	56	5.6	1.43
DH 32 *Q*	0	1	4	0	0	1	2	2	0	0	10	1.0	1.34
SN 60	1	2	0	1	1	0	0	0	1	1	7	0.7	0.68

The results obtained show a lack of conformity of opinion among the experts participating in the monitoring. Two earphones were outstanding in the general evaluation: JVC 5844 and KAM/B. The second group — from the point of view of the preferences recorded — included earphones KAM/$C-2$ and DT 204, another group included KAM/$C-1$ and KAM/A, while DH 32Q earphones and the stereophonic earphones SN-60 were found to be the worst. The existence of significant differences was determined, using the Students t—test at a level $\alpha = 0.05$, between the first and the third groups and between the fourth and the first three groups, respectively.

A random examination of the listeners' opinions showed that earphones JVC 5844 and KAM/B were found to be best from the point of view of timbre, JVC 5844 and KAM/$C-2$ were found best from the point of view of the sense of sound space, and earphones KAM/B and JVC 5844 were found best from the point of view of the spatial location of the sources. These opinions are confirmed by the objective properties of the earphones discussed [6]. Earphones JVC 5844 and KAM/B showed the flattest frequency transmission characteristics of the transducers as measured by the $B-K$ type 4133 "artificial head" method. At the same time these earphones, for the group under investigation reproduced high tones best, and were thus ranked highly from the point of view of the differences in the locations of sound sources. In turn, earphones JVC-5844 and KAM/$C-2$, the designs with the deepest earphone chambers and with the planes of the transducers inclined at an angle, gave the best impression of the sound space [6].

The very low ranking of the SN-60 stereophonic earphone (which was found worst in the group) confirmed an earlier thesis that four-transducer earphone monitoring is decisively better than two-transducer monitoring.

8. Comparison of the properties of quadrophonic monitoring by earphones with loudspeaker monitoring

A comparative evaluation of quadrophonic monitoring by earphones and by loudspeakers was performed using the following criteria: a general evaluation, the sense of the sound space, the quality in differentiating the phantom sources, sound quality and naturalness of the phantom sources located at the back, and the distinctness of the sound impression. A quadrophonic JVC 5844 earphone and four cased matched loudspeaker sets $GK-132$ (from "Fonia") were selected for the investigations.

The comparison was performed in the audio monitoring room of the State Higher Musical School in Warsaw. A test recording of popular music by Polish Radio and Television was used as the sound material. Fifteen experts participated in the investigations. They compared — individually — the two variants of quadrophonic sound projection.

The results of the experiment showed considerable differences of opinion from all the participants in the monitoring. Some experts, considering the presented projection variants as incomparable, declined to give their opinions in terms of the qualitative criteria suggested. The preferences of the remaining listeners presented a very indistinct picture of the relation between the two systems of quadrophonic sound projection compared (Table 7).

Table 7. Results of quality evaluation of quadrophonic monitoring by loudspeakers and by earphones

Criterion	Preferred monitoring system			No. opinion
	Loudspeakers	Earphones	Comparable quality	
General evaluation	9	5	—	1
Sense of space	6	2	2	5
Spatial stability	1	3	—	11
Quality in differentiating the phantom sources	2	2	3	8
Naturalness of the sources behind the listener's head	3	1	—	11
Distinctness of sound impression	2	4	2	7

The results obtained show that the sound impressions occurring under quadrophonic sound projection by loudspeakers and by earphones belong to essentially different aesthetic categories, the preference of one of the two projection variants depending primarily on the predispositions and likings of the individual listeners.

9. Conclusions

(1) With two-transducer emission of signals into one ear, the listener is able to recognize the projection conditions when only one transducer is being operated, and to distinguish mono from stereophonic projection during the simultaneous operation of the transducers. The degree of recognition of the projection conditions is closely related to the spectral structure of the signals emitted (the wider the bandwidth of the signal, the better the recognition of the projection conditions becomes).

(2) Four-transducer earphones assure the posibility of distinguishing between signals coming from the front and the rear pairs of transducers.

(3) It follows from the experiments performed that differences in timbre and the mutual time delays of the signals coming from the front and rear transducer pairs determine the directional and spatial impressions during four-transducer earphone monitoring.

(4) Four-transducer earphones permit monitoring of quadrophonic signals and also of stereophonic signals emitted by two-and four-transducer systems.

(5) Quadrophonic earphone monitoring permits location of the sound sources outside the head and considerable deepening of the spatial sense of the sound image. However, the possibility of locating directions over the whole range of space angles is not to be expected.

(6) Four-transducer two-channel monitoring represents a very great improvement in quality with respect to traditional two-transducer monitoring (considerably greater than the step between the four-transducer two-channel monitoring and quadrophonic monitoring).

(7) Four-transducer projection of stereophonic signals permits location of the sound outside the listener's head and compared to other, special techniques of earphone projection which give the effect mentioned above, seems to be the most natural and at the same time the most attractive from the economic point of view.

(8) The quality of the sound image and the listening comfort achieved under conditions of four-transducer binaural projection depend primarily on:

(a) the quality of the earphone transducers used,

(b) sufficiently large earphone enclosures,

(c) the distance between the transducers in a pair, and their position relative to the ear lobe,

(d) the quality of contact between the earphones and the listener's head.

(9) In comparing quadrophonic earphone monitoring with quadrophonic loudspeaker monitoring it should be noted that the sound impressions occurring under the two monitoring conditions belong to basically different aesthetic categories, preference towards one or other of the two sound projection variants depending primarily on the individual predispositions and likings of the listeners.

References

[1] B. B. Bauer, S Q *dichophony-quadrophonic earphone listening*, JAES, **24**, 5, 387 (1976).

[2] J. Blauert, *Räumliches Hören*, S. Hirzel Verlag, Stuttgart 1974.

[3] R. Litwin, T. Łętowski, A. Rakowski, *The possibility of using an "artificial head" microphone system for music recordings* (in Polish), Technika R i TV (in press).

[4] T. Łętowski, W. Makowski, *Properties of sound images* (in Polish), Technika R i TV, **1**, 16, (1977).

[5] T. Łętowski, K. Szlifierski, *General principles of subjective evaluation of electroacustic equipment* (in Polish), Technika R i TV, **4**, 9 (1974).

[6] W. Makowski, T. Łętowski, W. Hamer, *A quadrophonic earphone model – documentation of a problem developed in* KAMPWSM *for ZWG Tonsil*, Września, Warszawa 1976.

[7] M. V. Thomas, *Improving the stereo headphone sound image*, JAES, 25, 7/8, 474 (1977).

Received on December 18, 1978; final version on June 27, 1979.

ARCHIVES OF ACOUSTICS
5, 1, 73-80 (1980)

MUTUAL ACOUSTIC IMPEDANCE OF CYLINDRICAL SOURCES FOR A SPECIFIC CASE

WITOLD RDZANEK

Higher Pedagogical School (65-069 Zielona Góra)

The paper analyzes the imaginary component of the mutual impedence of two pulsating cylindrical rings placed on an infinitely long and stiff circular cylinder. The specific case of the radius of the cylinder being considerably shorter than the wave lengths of the acoustic waves radiated is considered. The imaginary component of the mutual impedence was calculated using the HILBERT transformation, based on the approximate expression for the real component of the mutual impedance given by ROBEY. Compared to earlier results the formulae have a simple form, and are thus convenient for numerical calculations. These calculations are illustrated graphically.

1. Introduction

In the consideration of the acoustical properties of a system of sources, significant interaction between sources should, in general be included. The theoretical calculation of these interactions consists in the evaluation of the mutual impedance of the sources.

It appears that the number of papers on the mutual acoustical interactions of cylindrical sources is relatively small.

An exact expression for the mutual acoustic impedance of a radiating system of cylindrical sources placed on the surface of an infinitely long and stiff circular cylinder, was given by Robey [10]. He assumed that the sources of the acoustic field consist of pulsating cylindrical rings of finite length. He used a Green's function for the calculation of the impedance. On account of the axial symmetry of the radiating sources he used a Green's function in a cylindrical coordinate system which is independent of the angular variation. For sources on a cylinder whose diameter is short compared to the wave length, he obtained an approximate formula for the mutual resistance.

The acoustic pressure and the mutual radiation impedance of rectangular pistons placed on the surface of a stiff and infinitely long circular cylinder has been investigated by GREENSPON and SHERMAN [2]. The pressure distribution on the baffle and the surface of the sources was obtained using the Neumann boundary condition problem in a cylindrical coordinate system. The knowledge of the pressure distribution was used to determine the mutual impedance of the two sources. They showed that the expression for the mutual impedance of two rectangular pistons becomes in the limiting case the same as the formula for the mutual impedance between two rings, which was given by Robey.

In paper [8] RDZANEK and WYRZYKOWSKI investigated the radiation of a system of slits of finite length placed symmetrically on the surface of a stiff circular cylinder. In order to derive the formulae for the mutual impedance of two slits they used the Green's function method, with dependence on all three cylindrical variables [7].

The problem of the mutual reactance of two pulsating cylindrical rings placed on a common cylindrical baffle was considered by Greenspon [1] and THOMPSON [12]. Referring to Robey's investigations, they calculated the mutual reactance of two sources with the assumption that the sources are placed on the surface of a cylinder whose diameter is shorter than the acoustic wave length. The final approximate formulae obtained for the mutual reactance were expressed by a single integral and an infinite series containing the axial function, which in the case of numerical calculations requires the use of a digital computer.

The use of the Hilbert transformation in the solution of the present problem permitted the final formulae in this paper to be obtained in a form more convenient for the performance of numerical calculations. The approximate expression for the real component of the mutual impedance, given by Robey, was assumed as the starting point. In addition to the trigonometric functions, only integral sines occur in this expression. Using the known real component of the mutual impedance, the imaginary component of the mutual impedance was calculated and expressed in terms of trigonometric functions and integral cosines. The numerical calculations are also illustrated graphically.

Notation

a — the radius of the cylinder
c — the wave propagation velocity
$Ci(x)$ — the integral cosine (21)
i — $\sqrt{-1}$
J_n — cylindrical Bessel function of the nth order
k — wave number
K_n — cylindrical MacDonald function of the nth order
l — half the source length
l_{ms} — the distance between the surface centres of the mth and the sth sources
N_n — cylindrical Neumann function of the nth order
p_{ms} — acoustic pressure from the sth source, acting on the mth source
v — the normal component of the particle velocity

v^x – the complex conjugate of the complex particle velocity v
v_0 – the amplitude of the particle velocity
$Si(x)$ – the integral sine (17)
Z_{ms} – the mutual mechanical impedence between the sth and the mth sources
θ_{ms} – relative acoustic resistance (4)
θ^0_{ms} – relative acoustic resistance of sources vibrating in phase
χ_{ms} – relative acoustic reactance (4)
χ^0_{ms} – relative acoustic reactance of sources vibrating in phase
ζ_{ms} – relative acoustic impedance (3)
δ_m – the initial phase of the particle velocity at the mth source
ϱ – the density of the medium
σ – the surface area

2. Assumptions of the analysis

It is assumed that there is a system of N harmonically vibrating sound sources in a liquid medium on the surface of an infinitely long circular cylinder of radius a, which acts as an ideal stiff baffle. The sound sources are pulsating rings of the cylinder (pistons in the shape of cylindrical rings), each having a length $2l$, and radius a (Fig. 1). The areas of the sources are equal and are

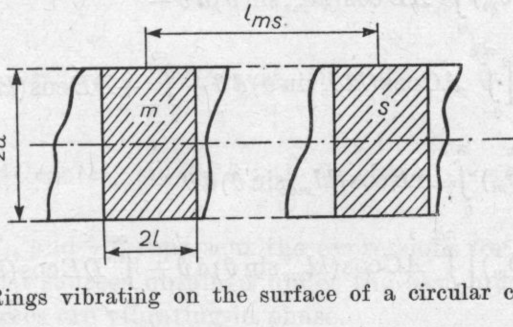


Fig. 1. Rings vibrating on the surface of a circular cylinder

$\sigma_m = \sigma_s = \sigma = 4\pi al$. The distance between the surface centres of the mth and sth sources is l_{ms}. It is assumed that all the points of the source vibrate in phase with a constant amplitude in the radial direction, with a shift of the phase of vibration occuring between individual extended sources. The normal component of the amplitude of the particle velocity (the radial component) of the mth source is equal to

$$v_m = v_0 e^{-i\delta_m} \tag{1}$$

where v_0 is the amplitude of the particle velocity for the initial phase on the mth source, σ_m, equal to zero.

The mechanical mutual impedance of the sth and mth sources is [6]

$$Z_{ms} = \frac{1}{v_{0s} v^*_{0m}} \int_{\delta m} p_{ms} v^*_m d\sigma, \tag{2}$$

where p_{ms} is the sound pressure from the sth source acting on the mth source, while v_m^x is the complex conjugate of the particle velocity v_m. From relation (1), $v_{0s} = v_{0m} = v_{0m}^x = v_0$.

Relating the mechanical impedance to the characteristic resistance of the medium, ϱc, and the area ofthe source, σ, we obtain the relative acoustic impedance

$$\zeta_{ms} = \theta_{ms} + i\,\chi_{ms}, \tag{3}$$

where quantities

$$\Theta_{ms} = \frac{\mathrm{Re}(Z_{ms})}{\varrho c\,\sigma}, \quad \chi_{ms} = \frac{\mathrm{Im}(Z_{ms})}{\varrho c\,\sigma} \tag{4}$$

are the mutual resistance and the mutual reactance, respectively.

3. Solution of the problem

Exact expressions for the mutual acoustic impedance of two sound sources under the given assumptions, can be written in the following way [2, 9, 10]:

$$\Theta_{ms} = \cos(\delta_s - \delta_m) \int\limits_0^{\pi/2} AB\cos(kl_{ms}\sin\vartheta)\,d\vartheta -$$

$$-\sin(\delta_s - d_m)\Big[\int\limits_0^{\pi/2} AC\cos(kl_{ms}\sin\vartheta)\,d\vartheta + \int\limits_0^{\infty} +DE\cos(kl_{ms}\cos h\psi)\,d\psi\Big], \tag{5}$$

$$\chi_{ms} = \sin(\delta_s - \delta_m) \int\limits_0^{\pi/2} AB\cos(kl_{ms}\sin\vartheta)\,d\vartheta +$$

$$+\cos(\delta_s - \delta_m)\Big[\int\limits_0^{\pi/2} AC\cos(kl_{ms}\sin\vartheta)\,d\vartheta + \int\limits_0^{\infty} DE\cos(kl_{ms}\cos h\psi)\,d\psi\Big]; \tag{6}$$

Using the following notation:

$$A = \frac{2}{\pi kl}\,\frac{\sin^2(kl\sin\vartheta)}{\sin^2\vartheta}, \tag{7}$$

$$B = \frac{2}{\pi ka}\,\frac{1}{\cos\vartheta\,[J_1^2(ka\cos\vartheta) + N_1^2(ka\cos\vartheta)]}, \tag{8}$$

$$C = \frac{J_0(ka\cos\vartheta)J_1(ka\cos\vartheta) + N_0(ka\cos\vartheta)N_1(ka\cos\vartheta)}{J_1^2(ka\cos\vartheta) + N_1^2(ka\cos\vartheta)}, \tag{9}$$

$$D = \frac{2}{\pi kl}\,\frac{\sin^2(kl\cos h\psi)}{\cos h^2\psi}, \tag{10}$$

$$E = \frac{K_0(ka\sin h\psi)}{K_1(ka\sin h\psi)}, \tag{11}$$

where $k = 2\pi/\lambda$ is the wave number, J_n represents a Bessel function, K_n — a MacDonald function, and N_n — a Neumann function (of the nth order in each case) [4, 13].

In the calculation of the mutual impedance from (5) and (6), when $ka \ll 1$, it is convenient to use the properties of cylindrical functions which have simple expressions for values of the argument. Thus the factors (8) and (9) can be reduced to expressions containing elementary functions, and subsequently the integrals over the range $(0, \pi/2)$ which occur in formulae (5) and (6), (see [9, 10, 12]), can be calculated. This property cannot be applied to the factor (11) which contains MacDonald functions with the argument $ka \sin h\psi$, since integration over the variable ψ in formulae (5) and (6) is over the infinite limits $(0, \infty)$ (see [1]).

To avoid the difficulty of the calculation of the integral over infinite limits $(0, \infty)$, another method will be used. Formulae (5) and (6) can be written in the following way:

$$\Theta_{ms} = \cos(\delta_s - \delta_m)\Theta^0_{ms} - \sin(\delta_s - \delta_m)\chi^0_{ms}, \tag{12}$$

$$\chi_{ms} = \sin(\delta_s - \delta_m)\Theta^0_{ms} + \cos(\delta_s - \delta_m)\chi^0_{ms}, \tag{13}$$

where

$$\Theta^0_{ms} = \int_0^{\pi/2} AB\cos(kl_{ms}\sin\vartheta)\,d\vartheta, \tag{14}$$

$$\chi^0_{ms} = \int_0^{\pi/2} AC\cos(kl_{ms}\sin\vartheta)\,d\vartheta + \int_0^{\infty} DE\cos(kl_{ms}\cos h\psi)\,d\psi. \tag{15}$$

The quantities Θ^0_{ms} and χ^0_{ms} represent the expressions for the mutual impedance of the cylindrical sources obtained under the assumption that $\delta_s = \delta_m$, and thus that the sources are vibrating in phase.

The real component of the mutual impedance θ^0_{ms}, when $ka \ll 1$ and $l_{ms} \geqslant 2l$, is known and is [10].

$$\Theta^0_{ms} = \frac{a}{l}\left\{\frac{1}{4}[k(l_{ms}+2l)]\mathrm{Si}[k(l_{ms}+2l)] + \frac{1}{4}[k(l_{ms}-2l)]\mathrm{Si}[k(l_{ms}-2l)] - \right.$$
$$\left. -\frac{klms}{2}\mathrm{Si}(kl_{ms}) - \sin^2 kl\cos(kl_{ms})\right\}, \tag{16}$$

where [3]

$$\mathrm{Si}(x) = \int_0^x \frac{\sin t}{t}\,dt. \tag{17}$$

The imaginary component of the mutual impedance χ^0_{ms} is found by a Hilbert transformation [5], [6], using the known real part of the mutual impedance (16).

The Hilbert's transform, in the notation assumed here, is written as

$$\chi^0_{ms}(k) = \frac{1}{\pi} \int\limits_{-\infty}^{+\infty} \frac{\Theta^0_{ms}(y)}{y-k}\, dy, \tag{18}$$

where we take the principal value of the integral. Observing that the real part of the mutual impedance Θ^0_{ms} in (16) is an even function of the wave number k, the transformation (18) can be reduced to the form

$$\chi^0_{ms}(k) = \frac{2k}{\pi} \int\limits_0^{\infty} \frac{\Theta^0_{ms}(y)}{y^2-k^2}\, dy. \tag{19}$$

To calculate the integral in formula (19), we refer to reference [11]; we have

$$\int\limits_0^{\infty} \frac{y\,\mathrm{Si}(ay)}{y^2-k^2}\, dy = \frac{\pi}{2}\,\mathrm{Ci}(ka) \tag{20}$$

for $a > 0$ and $k > 0$, where [3]

$$\mathrm{Ci}(x) = \int\limits_{\infty}^{x} \frac{\cos t}{t}\, dt. \tag{21}$$

To show that

$$\int\limits_0^{\infty} \frac{\sin^2(ay)\cos(by)}{k^2-y^2}\, dy = \frac{\pi}{2k}\sin^2(ka)\sin(kb), \tag{22}$$

the equality

$$\sin^2(ay)\cos(by) = \frac{1}{2}\cos(by) - \frac{1}{4}\cos(b+2a)y - \frac{1}{4}\cos(b-2a)y \tag{23}$$

and [11] formula

$$\int\limits_0^{\infty} \frac{\cos(by)}{k^2-y^2}\, dy = \frac{\pi}{2k}\sin(kb), \tag{24}$$

for $b > 0$ and $k > 0$, should be used. Inserting (20), (22), (23) and (24) into (19) we obtain

$$\chi^0_{ms} = \frac{a}{l}\left\{\frac{1}{4}[k(l_{ms}+2l)]\,\mathrm{Ci}[k(l_{ms}+2l)] + \frac{1}{4}[k(l_{ms}-2l)]\,\mathrm{Ci}[k(l_{ms}-2l)] - \right.$$
$$\left. - \frac{1}{2}kl_{ms}\,\mathrm{Ci}(kl_{ms}) + \sin^2(kl)\sin(kl_{ms})\right\} \tag{25}$$

for $kl_{ms} > 0$, $l_{ms} > 2l$, which is the expression for the imaginary part of the mutual impedance of the sources for small values of the parameter $ka\,(ka \ll 1)$.

If $l_{ms} = 2l$, i.e. when the sources are contiguous (see Fig. 1), we obtain

$$\Theta^0_{ms} = \frac{a}{l}\,[kl\,\mathrm{Si}(4kl) - kl\,\mathrm{Si}(2kl) - \sin^2(kl)\cos(2kl)] \tag{26}$$

and

$$\chi^0_{ms} = \frac{a}{l}\,[kl\,\mathrm{Ci}(4kl) - kl\,\mathrm{Ci}(2kl) + \sin^2(kl)\sin(2kl)]. \tag{27}$$

4. Conclusion

The calculation of the mutual acoustic reactance of two pulsating cylindrical rings on a stiff circular cylinder, performed using a Hilbert transform gives a formula which is simple in form. The formula permits the analysis of the acoustic interactions between cylindrical sources to be performed for small values of the parameter ka.

From formulae (16) and (25) the mutual impedance of cylindrical rings vibrating in phase can be calculated. The knowledge of the impedance permits the mutual impedance to be determined in the more general case when the sources do not vibrate in phase. Formulae (12) and (13) can be used for this purpose. The formulae for the mutual impedance take a particularly simple form when the sources are contiguous.

The dependencies of the mutual acoustic impedance on kl and kl_{ms}, as calculated from the formulae given in the present paper, are plotted in Figs. 2 and 3.

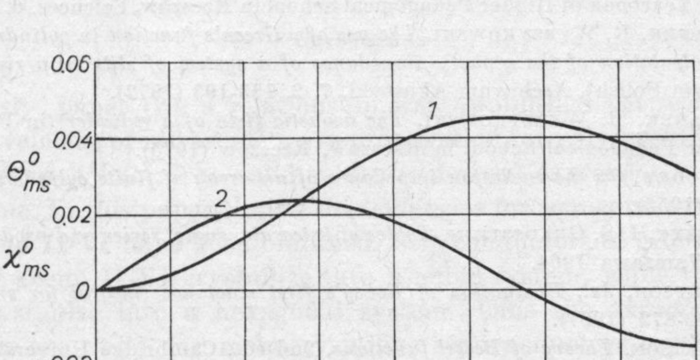


Fig. 2. The plot of the mutual impedance of two contiguous rings ($l_{ms} = 2l$), against the parameter $kl = 2\pi l/\lambda$

curve *1* – resistance, curve *2* – reactance. It is assumed that $a/l = 0.1$

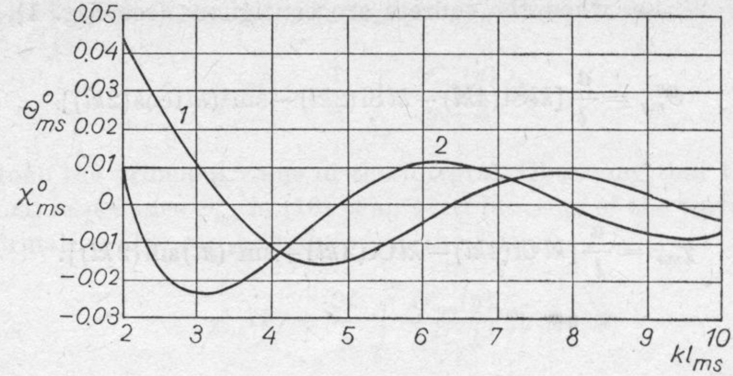


Fig. 3. The plot of the mutual impedance against the parameter $kl_{ms} = (2\pi/\lambda)l_{ms}$ curve *1* – resistance, curve *2* – reactance. It is assumed that $kl = 1$, $a/l = 0.1$

References

[1] J. E. Greenspan, *An approximation to the remainder of Robey's reactive impedance integral*, JASA, **33**, 1428-1429 (1961).

[2] J. E. Greenspan, C. H. Sherman, *Mutual radiation impedance and nearfield pressure for pistons on a cylinder*, JASA, **36**, 1, 143-153 (1964).

[3] E. Janke, F. Emde, F. Lösch, *Tafeln höherer Funktionen* BG, Teubner Verlagsgesellschaft, Stuttgart 1960.

[4] N. W. McLachlan, *Bessel functions for engineers* (in Polish), PWN, Warszawa 1964.

[5] P. M. Morse, H. Feshbach, *Methods of theoretical physics*, Vols. 1 and 2, McGraw – Hill, New York 1953.

[6] W. Rdzanek, *Mutual and total acoustic impedance of a system of sources with a variable surface distribution of particle velocity* (in Polish), Reports of Higher Pedagogical School in Zielona Góra, Zielona Góra (1979).

[7] W. Rdzanek, *The acoustic Green's function for a cylinder* (in Polish), Scientific and Didactic Yearbook of Higher Pedagogical School in Rzeszów, Sciences, 6, 103-129 (1969).

[8] W. Rdzanek, R. Wyrzykowski, *The use of a Green's function in cylindrical coordinates for the calculation of the acoustic impedance of a system of slits in a ring of a circular cylinder* (in Polish), Archiwum Akustyki, 7, 2, 183-193 (1972).

[9] W. Rdzanek, R. Wyrzykowski, *The acoustic field of a cylinder* (in Polish), Reports of Higher Pedagogical School in Rzeszów, Rzeszów (1975).

[10] D. H. Robey, *On the radiation impedance of an array of finite cylinders*, JASA, **27**, 4, 706-710 (1955).

[11] I. M. Ryżyk, I. S. Gradsztejn, *Tables of integrals, sums, series and products* (in Polish), PWN, Warszawa 1964.

[12] W. Thompson, Jr., *Evaluation of Robey's first reactance integral for small ka*, JASA, **42**, 4, 870-872 (1967).

[13] G. N. Watson, *Theory of Bessel functions*, 2nd ed., Cambridge University Press, 1966.

Received on October 28, 1978; final version on June 21, 1979.

ARCHIVES OF ACOUSTICS
5, 1, 81-90 (1980)

RELATION BETWEEN THE ENERGY GAP AND ULTRASONIC WAVE PROPAGATION VELOCITIES IN SEMICONDUCTORS

L. OPILSKA

Institute of Theoretical Mechanics Silesian Technical University (44-100 Gliwice)

A. OPILSKI

Institute of Physics, Silesian Technical University (44-100 Gliwice)

Z. LATACZ

Department of Solid State Physics, Polish Academy of Sciences (41-800 Zabrze)

In this work the existence of the relation between the energy gap and ultrasonic wave velocities for semiconductors crystallizing into cubic and hexagonal systems has been shown. Data from the literature confirmed the validity of the relations obtained. Using the acoustical method the variation of the energy gap with the percentage composition of the solid solutions has also been obtained.

Introduction

Previously, paper [2], a relationship was established between the energy gap and the velocity of ultrasonic wave propagation for polycrystalline semiconductors of group $A^{II} B^{V}$. The validity of the derived expression was confirmed by experiment. In this paper similar dependencies for semiconductors of groups IV, III-V, and II-VI have been obtained. Semiconductors of groups IV, III-V, and some of group II-VI crystallize into a cubic system, while the remainder of II-VI crystallize into a hexagonal system. Thus the expressions derived in this work are valid for monocrystalline semiconductors crystallizing into both systems. The relation of these results to the question of chemical bonds in semiconductors was considered. As is well known, semiconductors of group IV have valent bonds and semiconductors of groups III-V, II-VI have valent-ionic bonds.

Theory

In our considerations we start with the well known thermodynamic relation

$$\left[\frac{\partial^2 U}{\partial^2 V}\right]_T = -\left(\frac{\partial p}{\partial V}\right)_T + T\left[\frac{\partial}{\partial T}\left(\frac{\partial p}{\partial V}\right)_T\right]_V, \tag{1}$$

where U is the internal crystal energy, T – the temperature, V – the volume and p – the pressure.

The bulk modulus $K = -V(\partial p/\partial V)_T$ may be expressed as a function of the propagation velocities of ultrasonic waves in the following manner:

$$-V\left(\frac{\partial p}{\partial V}\right)_T = \varrho\left(C^2_{l[111} - \frac{4}{3}C^2_{t[100]}\right) \tag{2}$$

for a cubic system; and

$$-V\left(\frac{\partial p}{\partial V}\right)_T = \varrho\,\frac{C^2_{l[001]}(C^2_{l[100]} - C^2_{t[100]}) - (2C^2_{l[101]} - C^2_{l[100]} - 2C^2_{t[001]})^2}{C^2_{l[001]} + C^2_{l[100]} - C^2_{t[100]} - 2(2C^2_{l[101]} - C^2_{l[100]} - 2^2_{t[001]})} \tag{3}$$

for a hexagonal system, where ϱ is the density of material, $c_{l[\ldots]}$ and $c_{t[\ldots]}$ are the velocities of longitudinal and transverse waves propagating in the given directions. The semiconductors of group IV are characteristically valent, and of the remaining groups valent-ionic. In general, the potential energy of atomic interaction may be written in the following form:

$$\Phi = Ae^{-R/a} - \frac{B}{R^n} - \frac{B_1}{R}. \tag{4}$$

In this formula the first part describes the forces of repulsion, the second, the forces of attraction, and the last, the ionic forces of attraction.

From formula (4) it may be derived that

$$\frac{\partial^2 \Phi}{\partial R^2} = A_1\Phi + A_2\Phi_j, \tag{5}$$

where

$$A_1 = \frac{n[a(n+1) - R_0}{aR_0(an + R_0)}, \quad A_2 = \frac{[an^2 + R_0(n^2 + n - 2)}{R_0^2(an + R_0)},$$

where Φ is the potential energy of atomic interaction, and Φ_j is the ionic bond component.

Using the relation

$$\frac{\partial^2 \Phi}{\partial V^2} = \frac{R_0^2}{9V^2}\,\frac{\partial^2 \Phi}{\partial R^2}$$

and the fact that the kinetic energy does not directly depend on the volume, from equations (1), (2), (3), and (5) we obtain

$$\Phi = D_1 w^2 + D_2 \Phi_j, \tag{6}$$

where

$$w^2 = C^2_{l[111]} - \frac{4}{3} C^2_{t[100]} \tag{6a}$$

for a cubic system, and

$$w^2 = \frac{C^2_{l[001]}(C^2_{l[100]} - C^2_{t[100]}) - (2C^2_{l[101]} - C^2_{l[100]} - 2C^2_{t[001]})^2}{C^2_{l[001]} + C^2_{l[100]} - C^2_{t[100]} - 2(2C^2_{l[101]} - C^2_{l[100]} - 2C^2_{t[001]})} \tag{6b}$$

for a hexagonal system, with

$$D_1 = \frac{1}{9} a^2 R_0^2 A_1 \quad \text{and} \quad D_2 = -\frac{1}{9} A_2.$$

It is also known that the dependence of the energy gap on the potential energy of atomic interaction [1, 4] has a linear character:

$$E_g = a\Phi + b, \tag{7}$$

where a and b are constants for particular groups of semiconductors.

Associating the above dependence with the expression obtained above between the velocity of ultrasonic wave propagation and the energy gap (6) we obtain

$$E_g = a_1 w^2 + a_2 \Phi_j + b, \tag{8}$$

where w^2 is given by equations (6a) and (6b) for cubic and hexagonal systems respectively, a_1, a_2, b – constants characteristic for the particular semiconductor groups.

Expression (8) should be analysed with respect to the chemical bonds in the semiconductors. $\Phi_j = 0$ can be put equal to zero in formula (8) for semiconductors of group IV, which, as is well known, have pure valent bonding. Thus the dependence of the energy gap on the combination of ultrasonic wave velocities for this group is one straight line. For semiconductors of groups III-V and II-VI the share of the ionic component in the bond is different. In these compounds, the dependence of the energy gap on the ionic component has to be determined. If the coefficient of proportionality between E_g and Φ_j was equal for the whole group of semiconductors considered then in the dependence $E_g = f(w^2)$ all the compounds of types III-V and II-VI would lie on one straight line. However, because of the fact that the coefficient of proportionality between E_g and Φ_j is constant only for a given series of compounds, the dependence $E_g = f(w^2)$ will not be the same for the whole group of these compounds, but only for a given series. For the semiconductors having valent-ionic bonding, the relation between the energy gap and the velocities of ultrasonic wave propagation is

$$E_g = aw^2 + \beta, \tag{8a}$$

where a and β are constants, characteristic for a given series of semiconductors of groups III-V and II-VI.

For group III-V semiconductors three series have been distinguished:

MX_i-InSb, InAs, InP,
NX_i-GaSb, GaAs, GaP,
PX_i-AlSb, AlAs, AlP.

For these semiconductors three straight lines have been obtained for the dependence $E_g = f(w^2)$.

For group II-VI semiconductors two series have been distinguished:

RY_i-ZnSe, ZnTe, ZnS
TY_i-CdSe, CdTe, CdS.

Thus, for these semiconductors two straight lines have been obtained.

The dependence of the energy gap on the ionic component of the bond for semiconductors of groups III-V and II-VI is similar.

These relations for the solid solutions of particular groups of semiconductors may also be generalized. For this purpose the relationship between the velocity of ultrasonic wave propagation, the percentage composition, and the propagation velocities in the pure components [3] has been used. The following dependence has thus been obtained for solid solutions with cubic and hexagonal systems:

$$E_g = a\left[\frac{\mu M_1}{M} w_1^2 + \frac{(1-\mu)M_2}{M} w_2^2\right] + \beta, \tag{9}$$

where w_1, w_2 are the expressions given in formulae (6a, b) for the first and second components respectively, μ is the molecular fraction, M_1, M_2 are the masses of components, and $\bar{M}$ is the average mass, given by

$$\bar{M} = \frac{\mu_1 M_1 + \mu_2 M_2}{\mu_1 + \mu_2}.$$

Results

In Figs. 1 and 2 the dependence of the energy gap on the ultrasonic wave velocities, for semiconductors of groups IV, III-V, II-VI, respectively, are presented graphically. These dependences have been drawn on the basis of data from the literature. From these data the corresponding constants in equation (8) were determined to be:

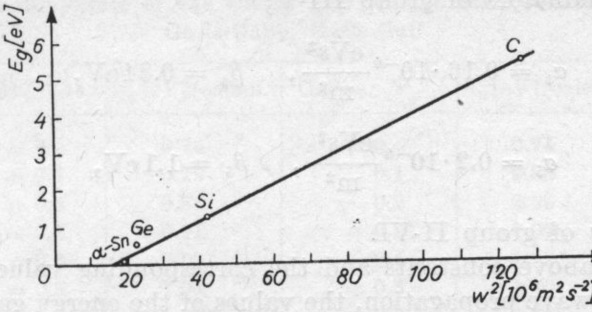


Fig. 1. Dependence of the energy gap on the combination of ultrasonic wave velocities for semiconductors of group IV

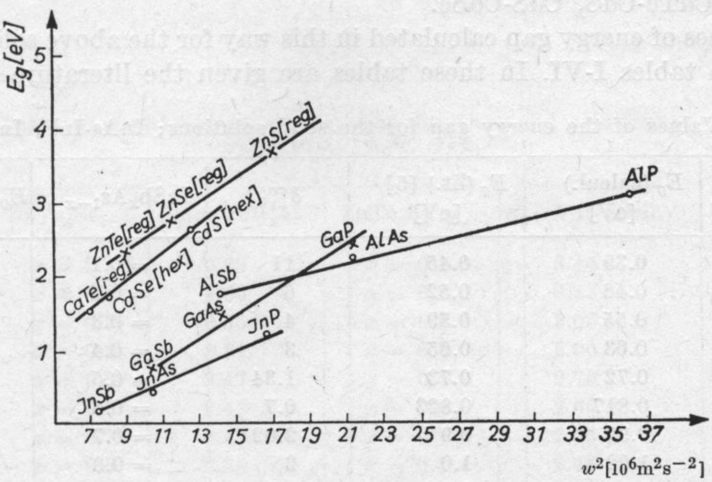


Fig. 2. Dependence of the energy gap on the combination of ultrasonic wave velocities for semiconductors of groups III-V and II-VI

$$a_1 = 0.0485 \cdot 10^{-6} \frac{\mathrm{eVs}^2}{\mathrm{m}^2}, \quad b = -0.882\,\mathrm{eV},$$

for semiconductors of group IV;

$$a_1 = 0.115 \cdot 10^{-6} \frac{\mathrm{eV \cdot s}^2}{\mathrm{m}^2}, \quad \beta_1 = -0.69\,\mathrm{eV},$$

$$a_2 = 0.14 \cdot 10^{-6} \frac{\mathrm{eVs}^2}{\mathrm{m}^2}, \quad \beta_2 = -0.6\,\mathrm{eV},$$

$$a_3 = 0.062 \cdot 10^{-6} \frac{\mathrm{eVs}^2}{\mathrm{m}^2}, \quad \beta_3 = 0.82\,\mathrm{eV},$$

— for semiconductors of group III-V;

$$\alpha_4 = 0.16 \cdot 10^{-6} \frac{\mathrm{eVs}^2}{\mathrm{m}^2}, \quad \beta_4 = 0.84\,\mathrm{eV},$$

$$\alpha_5 = 0.2 \cdot 10^{-6} \frac{\mathrm{eVs}^2}{\mathrm{m}^2}, \quad \beta_5 = 1.1\,\mathrm{eV},$$

for semiconductors of group II-VI.

Knowing the above constants and the corresponding values of the velocities of ultrasonic wave propagation, the values of the energy gap for the solid solutions were calculated according to equation (9). The following solid solutions were considered: InAs-InP, InAs-InSb, InSb-InP, GaAs-GaP, GaAs-GaSb, GaSb-GaP, AlAs-AlP, AlAs-AlSb, AlSb-AlP, ZnTe-ZnSe, ZnTe-ZnS, ZnS-ZnSe, CdTe-CdSe, CdTe-CdS, CdS-CdSe.

The values of energy gap calculated in this way for the above solid solutions are shown in tables I-VI. In these tables are given the literature values and

Table 1. Values of the energy gap for the solid solutions; InAs-InP, InAs-InSb

$InAs_{1-x}P_x$	E_g (calcul.) [eV]	E_g (lit.) [5] [eV]	δ [%]	$InSb_xAs_{1-x}$	E_g [eV] (calcul.)
$x = 0$	0.39	0.45	11	$= 0.1$	0.21
$x = 0.1$	0.46	0.52	6	$= 0.2$	0.23
$x = 0.2$	0.55	0.59	4	$= 0.3$	0.25
$x = 0.3$	0.63	0.65	3	$= 0.4$	0.28
$x = 0.4$	0.72	0.73	1.34	$= 0.5$	0.31
$x = 0.5$	0.817	0.823	0.7	$= 0.6$	0.35
$x = 0.6$	0.92	0.9	2.02	$= 0.7$	0.38
$x = 0.7$	1.02	1.0	2	$= 0.8$	0.40
$x = 0.8$	1.13	1.1	2.7	$= 0.9$	0.44
$x = 0.9$	1.25	1.2	3.9		
$x = 1.0$	1.37	1.31	4.5		

Table 2. Values of the energy gap for the solid solutions; InSb-InP, GaAs-GaP

$InSb_{1-x}P_x$	E_g [eV] (calcul.)	$GaAs_{1-x}P_x$	E_g (calcul.)	E_g (lit.) [5]	δ [%]
$= 0$	0.10	$x = 0.1$	2.47	2.59	4.6
$= 0.1$	0.17	$x = 0.2$	2.32	2.44	4.9
$= 0.2$	0.24	$x = 0.3$	2.17	2.31	6.0
$= 0.3$	0.34	$x = 0.4$	2.04	2.17	5.9
$= 0.4$	0.44	$x = 0.5$	1.92	2.04	5.8
$= 0.5$	0.55	$x = 0.6$	1.81	1.91	5.2
$= 0.6$	0.68	$x = 0.7$	1.71	1.79	5.0
$= 0.7$	0.80	$x = 0.8$	1.60	1.66	3.8
$= 0.8$	0.95	$x = 0.9$	1.51	1.55	2.5
$= 0.9$	1.12				
$= 1.0$	1.30				

Table 3. Values of the energy gap for the solid solutions: GaAs-GaSb, GaSb-GaP

$GaSb_{1-x}As_x$	E_g [eV] (calcul.)	$GaSb_{1-x}P_x$	E_g [eV] (calcul.)
= 0	0.74	= 0	0.74
= 0.1	0.79	= 0.1	0.83
= 0.2	0.85	= 0.2	0.86
= 0.3	0.91	= 0.3	1.09
= 0.4	0.97	= 0.4	1.23
= 0.5	1.04	= 0.5	1.40
= 0.6	1.105	= 0.6	1.56
= 0.7	1.18	= 0.7	1.78
= 0.8	1.25	= 0.8	2.02
= 0.9	1.34	= 0.9	2.32
		= 1.0	2.63

Table 4. Values of the energy gap for the solid solutions: ZnTe-ZnSe, ZnTe-ZnS

$ZnTe_xSe_{1-x}$	E_g (calcul) [eV]	$ZnTe_xS_{1-x}$	E_g [eV] (calcul.)
$x = 0.1$	2.65	$x = 0.1$	3.44
$x = 0.2$	2.60	$x = 0.2$	3.23
$x = 0.3$	2.55	$x = 0.3$	3.05
$x = 0.4$	2.51	$x = 0.4$	2.90
$x = 0.5$	2.47	$x = 0.5$	2.76
$x = 0.6$	2.43	$x = 0.6$	2.66
$x = 0.7$	2.40	$x = 0.7$	2.56
$x = 0.8$	2.36	$x = 0.8$	2.45
$x = 0.9$	2.33	$x = 0.9$	2.32

Table 5. Values of the energy gap for the solid solutions: ZnS-ZnSe, CdTe-SdSe

$ZnSe_xS_{1-x}$	E_g (calcul.) [eV]	$CdTe_xSe_{1-x}$	E_q [eV] (calcul.)
$x = 0.1$	3.55	= 0.1	1.52
$x = 0.2$	3.42	= 0.2	1.53
$x = 0.3$	3.30	= 0.3	1.55
$x = 0.4$	3.20	= 0.4	1.56
$x = 0.5$	3.10	= 0.5	1.58
$x = 0.6$	2.99	= 0.6	1.60
$x = 0.7$	2.92	= 0.7	1.62
$x = 0.8$	2.82	= 0.8	1.64
$x = 0.9$	2.76	= 0.9	1.66

Table 6. Values of the energy gap for the solid solutions: CdTe-CdS, CdS-CdSe

$CdTe_xS_{1-x}$	E_g [eV] (calcul.)	$CdSe_xS_{1-x}$	E_g [eV] (calcul.)
$x = 0.1$	1.56	$x = 0.1$	2.53
$x = 0.2$	1.64	$x = 0.2$	2.41
$x = 0.3$	1.72	$x = 0.3$	2.30
$x = 0.4$	1.81	$x = 0.4$	2.20
$x = 0.5$	1.90	$x = 0.5$	2.09
$x = 0.6$	2.01	$x = 0.6$	1.98
$x = 0.7$	2.13	$x = 0.7$	1.9
$x = 0.8$	2.26	$x = 0.8$	1.81
$x = 0.9$	2.41	$x = 0.9$	1.76

the percentage difference between these values, and those determined in this work by acoustic and other methods. On the basis of this data the dependence of the energy gap on the percentage composition of the above solid solutions has been made graphically and is shown in Figs. 3 and 4.

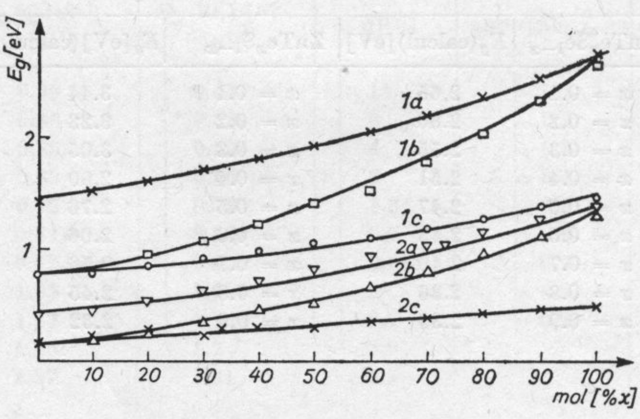


Fig. 3. Dependence of the energy gap on the percentage composition of solid solutions of group III-V

1a – GaAs-GaP, *1b* – GaSb-GaP, *1c* – GaSb-GaAs, *2a* – InAs-InP, *2b* – InSb-InP, *2c* – InSb-InAs

Conclusions

The relations obtained theoretically are in agreement with the experimental ones. This fact enables us to determine the energy gap indirectly with knowledge of the velocities of ultrasonic wave propagation for the pure semiconductors and their solid solutions. The results obtained by this method for some solid solutions have been compared to the data obtained by other methods i.e. (opti-

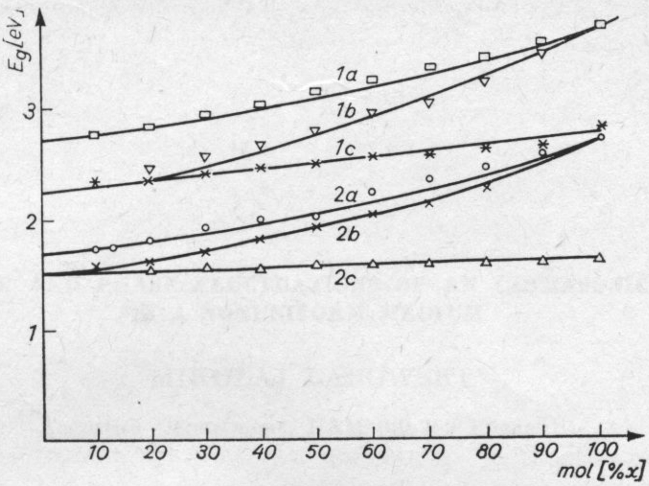


Fig. 4 Dependence of the energy gap on the percentage composition of solid solutions of group II-VI

1a – ZnTe-ZnS, *1b* – ZnSe-ZnS, *1c* – ZnSe-ZnTe, *2a* – CdSe-CdS, *2b* – CdTe-CdS, *2c* – CdTe-CdSe

cal and electrical). The maximum error between the values of the energy gap determined by the other methods and by the acoustic method is 11%. To determine the energy gap for semiconductors which have a cubic structure it is necessary to measure the velocity of longitudinal waves in the [111] direction and transverse waves in the [100] direction. For semiconductors with a hexagonal structure, it is necessary to measure the following velocities: of longitudinal waves in the directions [001], [100], [101] and of transverse waves in the directions [100], [001].

It is worth mentioning that the ultrasonic wave velocity is a parameter which may be easily measured. The accuracy of its determination is of the order of 0.2%.

References

[1] P. Manca, Energy gap and bonds in semiconductors, J. Phys. Chem. Solids, **20**, 268 (1961)

[2] L. Opilska, A. Opilski, Relations between the energy gap and the velocity of ultrasonic wave propagation in $A^{II}B^{V}$ group semiconductors, Phys. Stat. Solidi, *a* **35** *K*, 183 (1976).

[3] L. Opilska, W. Zdanowicz, A. Kalińska, Ultrasonic wave propagation in polycrystalline solid solutions of semiconductors of groups $A^{II}B^{V}$; $Cd_3As_2-Zn_3As_2$, $Cd_3AS_2-Cd_3P_2$, Acta phys. polonica, *A* **50** (1976).

[4] N. N. Sirota, Khimicheskaya svyaz v poluprovodnikah i termodynamika, Mińsk 1966.

[5] A. G. Thomson, J. C. Woolley, Energy gap variation in mixed III-V alloys, Canada Rec. **27** (1966).

Received on November 7, 1978; final version on June 10, 1979.

ARCHIVES OF ACOUSTICS
5, 1, 91-112 (1980)

AMPLITUDE AND PHASE FLUCTUATIONS OF AN ULTRASONIC WAVE IN A NONUNIFORM MEDIUM

MIKOŁAJ ŁABOWSKI

Acoustics Department, UAM (60-769 Poznań)

The paper presents the method of investigation of the results obtained for the fluctuation of quantities of the wave field in CO_2 close to the critical state. The method used the phenomenon of light diffraction by an ultrasonic wave. The analysis of the experimental results was performed by a correlation method, in which the autocorrelation functions and the power spectra of the fluctuation were determined for individual thermodynamic conditions and various frequencies of the ultrasonic wave.

1. Introduction

An ultrasonic wave, propagating in a non-uniform medium, (e.g. in a liquid-liquid or liquid-vapour system under critical conditions) where large fluctuations in concentration or density may occur, undergoes certain deformations. The deformations are seen in the fluctuation of characteristic quantities of the wave field due to the superimposition of the primary wave and the waves scattered by the heterogeneities of the medium.

If a medium with a fluctuation in density is characterized by the refraction index which is subject to fluctuation, a definite relation can be found between the fluctuations of the coefficient and the characteristic quantities of the wave field.

The quantities of the wave field which are subject to fluctuation are primarily amplitude and phase.

Chernov [1] gives two methods for deriving a general expression for the amplitude and phase fluctuations of a wave propagating in a non-uniform medium. One is the method of small perturbations, the other, a more general approach, is Rytov's method.

The basic method permitting the evaluation of diffraction effects in the case of large heterogencities (i.e., $\lambda \ll a$, where a is the size of heterogencity) is the method derived by Rytov [2], and developed by Obukhov [3], Chernov [4] and others. The formulae that they derived give the dependence of the phase and amplitude fluctuations of a wave propagating in a non-uniform medium on the statistical properties of the medium characterised by the function $\mu(\xi, \eta, \zeta)$. The statistical process of the fluctuation of the refraction coefficient was considered to be stationary, so that the dependence of μ on time could be neglected. Describing such a process with a correlation function and knowing its form, the mean squares of the fluctuations of the phase $\overline{S_1^2}$ and of the amplitude $\overline{B^2}$ of the wave can be calculated.

2. Use of the phenomenon of light diffraction by an ultrasonic wave for the investigation of the amplitude and phase fluctuations of the wave propagating a non-uniform medium

Pictures of the diffraction of light by an ultrasonic wave can for such a medium give information on the heterogeneity of the medium and also on the deformation of the ultrasonic wave itself. The phenomenon of light diffraction by an ultrasonic wave propagating in a non-uniform medium is considerably more complicated [5] than the well-known phenomenon in a uniform medium [6-14].

The phenomenon of light diffraction by an ultrasonic wave in a medium where large fluctuations in density or concentration occur is such that the resulting diffraction image is also subject to fluctuation [15-20]. The light intensity fluctuates in the individual diffraction orders, and also does the angle of diffraction of the light wave. The character of the phenomenon is different from that of the RAMAN-NATH theory [6]. ŚLIWIŃSKI [5] attempted to consider the effect of the fluctuations of the medium on the diffraction image. Fluctuations in the density of the medium cause amplitude and phase fluctuations of the ultrasonic wave which causes an additional change in the coefficient of light diffraction. Considering these changes he derived an expression for the instantaneous distribution of the light diffracted over the plane of departure from the ultrasonic layer.

The mean value of the angle of diffraction that is subject to fluctuation is given by the expression [5]:

$$\operatorname{tg} a_p = \frac{p}{k}\left(K + \frac{\beta}{2\gamma}\right), \tag{1}$$

where $k = 2\pi/\lambda$, $K = 2\pi/\Lambda$, p is the order of the spectrum, with β, γ being functions of the mean squares of fluctuation in phase $\overline{S^2}$ and amplitude, $\overline{B^2}$, of the ultrasonic wave, given by the following formulae:

$$\beta = \frac{R_{BS_1}}{2\sqrt{\overline{B^2}}\sqrt{\overline{S_1^2}(1-R_{BS}^2)}}\,; \quad \gamma = \frac{1}{2\overline{S_1^2}(1-R_{BS_1}^2)}.$$

R_{BS_1} is the correlation coefficient between B and S_l which is defined by

$$R_{BS_1} = \frac{\overline{BS_1}}{\sqrt{\overline{B^2} \cdot \overline{S_1^2}}}.$$

2.1. The method of measurement of the fluctuations in the ultrasonic wave. The light intensity in the individual diffraction orders and the angles of diffraction of the light waves are subject to fluctuations. Owing to the fluctuation of the diffraction angles the widths of individual diffraction lines change. The mean width of these lines, defined as the "half" width of the curve of the light intensity distribution over a line, is a measure of the diffraction angle which in turn may be related to the fluctuations of the amplitude and the phase of the ultrasonic wave.

It is possible to develop a method of defining the magnitude of the fluctuations of the amplitude and the phase of the ultrasonic wave propagating in a medium close to the critical point, based on the measurement of the diffuseness of the width of diffraction lines of the light diffracted by the wave [28]. A schematic diagram of the light diffracted by a scattered ultrasonic wave is shown in Fig. 1.

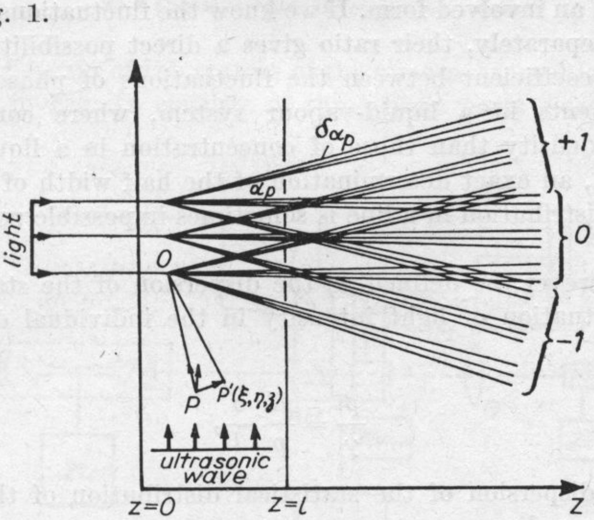


Fig. 1. A diagram of light diffraction by a scattered ultrasonic wave, according to [5]

Relation (1) can be written for small angles in the form

$$\operatorname{tg}(\alpha_p + \delta\alpha_p) \approx \sin(\alpha_p + \delta\alpha_p) = p\left(\frac{\lambda}{\Lambda} + \frac{\beta}{2\gamma}\right), \tag{2}$$

whence

$$\alpha_p + \delta\alpha_p = p\left(\frac{\lambda}{\Lambda} + \frac{\beta}{2\gamma}\right). \tag{3}$$

Since $a_p = p\lambda/\Lambda$, and $\delta a_p = \Delta p/l$, $\delta a_p = p\beta/2\gamma$, where p is the order of the spectrum, δa_p is the fluctuation in the diffraction angle, β and γ are functions of the mean squares of the fluctuations of the amplitude and phase of the ultrasonic wave [6, 13, 21, 15], λ is the wavelenght of the light, Λ is the ultrasonic wavelength, Δ_p is the mean width of the pth line, and l is the distance of the objective from the recording photographic film (or photo multiplier).

The ratio β/γ can be expressed in the following way [22]:

$$\frac{\beta}{\gamma} = \frac{\overline{B \cdot S_1}}{\overline{S_1^2}} = R_{\beta S_l}\sqrt{\frac{\overline{B^2}}{\overline{S_1^2}}}. \tag{4}$$

From a determination of the width of the diffuseness of a line and with knowledge of the distance l, the ratio β/γ which is defined in terms of the mean squares of the fluctuations of amplitude and phase of the ultrasonic wave, can be found. This ratio can be assumed to be the measure of the fluctuation of an ultrasonic wave propagating in a medium close to the critical state.

From the width of the diffuseness of the diffraction lines, information can be obtained on both the phase and the amplitude fluctuations of the ultrasonic waves, albeit, in an involved form. If we know the fluctuations of the phase and the amplitude separately, their ratio gives a direct possibility of determining the correlation coefficient between the fluctuations of phase and amplitude.

In experiments in a liquid-vapour system, where considerably larger fluctuations in density than those of concentration in a liquid-liquid system can be observed, an exact determination of the half width of the curve of the light intensity distribution in a line is sometimes impossible using photographic methods.

It may, however, be defined as the dispersion of the statistical distribution of the fluctuation of light intensity in the individual diffraction orders [5, 15]. Then

$$\frac{\beta}{\gamma} = \frac{K}{p}\frac{\sigma}{l},$$

where σ is the dispersion of the statistical distribution of the fluctuation of light intensity in a line.

This paper presents the results of the investigation of fluctuations of quantities of the ultrasonic wave field in CO_2 close to the critical state, using a method involving the diffraction of light by an ultrasonic wave. Due to a linear relation between the intensity (not too high an intensity*) of the ultrasonic wave, and the intensity of the light in the diffraction figures, this method permits the fluctuations of quantities of the wave field to be investigated through measurement of the fluctuation of the light diffracted by this wave.

---

* When only lines of the ± 1 orders, and none of higher orders occur.

Fluctuations of the intensity of the light diffracted by an ultrasonic wave at a few chosen frequencies (0.8, 1.2, 3.51, 5.76, 27.8 MHz) propagating in CO_2 which was held in a state close to the critical point where distinct fluctuations in density occur, were investigated.

The analysis of the experimental results for a liquid-vapour system was performed by a correlation method, in which the autocorrelation functions and the power spectra of the fluctuation for specific thermodynamic conditions and various frequencies of the ultrasonic wave [20] were determined.

It should be noted that the object of the investigation was the slow ("coarse grained") fluctuations connected with fluctuations of the thermodynamic state close to the critical point, which have also been investigated using the photographic method, by Śliwiński [23] and mentioned by Brown [24]. For purely technical reasons the fast ("fine grained") fluctuations of light diffracted by the ultrasonic wave (which also occur when this wave is absent), which cause the well-known phenomenon of optical opalescence, have not been investigated so far.

3. Apparatus and investigation method

A general diagram of the system of the apparatus used is shown in Fig. 1a. The apparatus in Fig. 1a consists of three basic parts closely linked to each other:

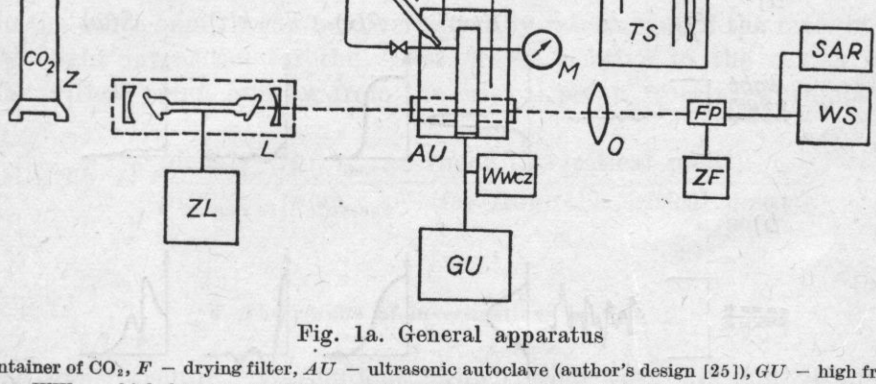


Fig. 1a. General apparatus

Z – container of CO_2, F – drying filter, AU – ultrasonic autoclave (author's design [25]), GU – high frequency generator, WW_{cz} – high frequency voltmeter, TS – Höppler thermostat, T_1, T_2 – thermometers, M – manometer, L – He-Ne laser, ZL – laser power supply, O – objective, FP – photomultiplier, ZF – photomultiplier power supply, WS – voltmeter pen recorder, SAR – statistical distribution analyser

(1) A system for generating adequate thermodynamic conditions, i.e. an adequate ultrasonic autoclave [25] with a thermostat combined with pressure and temperature meters.

(2) A system propagating an ultrasonic wave into the autoclave.

(3) An optical system with photographic or photoelectrical recording.

A laser beam of 2.5 mm diameter passes through the ultrasonic autoclave where there is a hetereogeneous medium maintained under suitable thermodynamic conditions and where an ultrasonic wave generated by a piezoelectric transducer propagates in a direction parallel to the laser beam.

The diffracted waves after being collected by the objective O give a diffraction image which can be recorded with a photoelectic system or a photographic method.

The starting point was an exact positioning of the optical system and checking of the linearity of the photoelectrical system, which plays an extremely important role in these cases.

This linearity was confirmed by the agreement of the spectral distributions obtained in liquid CO_2 under thermodynamic conditions far from the critical point with the Raman-Nath distribution [6]. The intensity of the ultrasonic wave close to the critical point was determined using the method described in reference [26].

The recording of the fluctuation of a given diffracted line (of the zeroth or ± 1-st order) was made by recording the photoelectrical current. It was essential to place the inlet hole through which a beam of a chosen order fell on the photomultiplier in the correct position. The hole was fixed and thus fluctuations of the diffraction angle caused part of the beam to be extinguished.

A decrease in the light intensity due to fluctuation of the diffraction angle is shown schematically in Fig. 2.

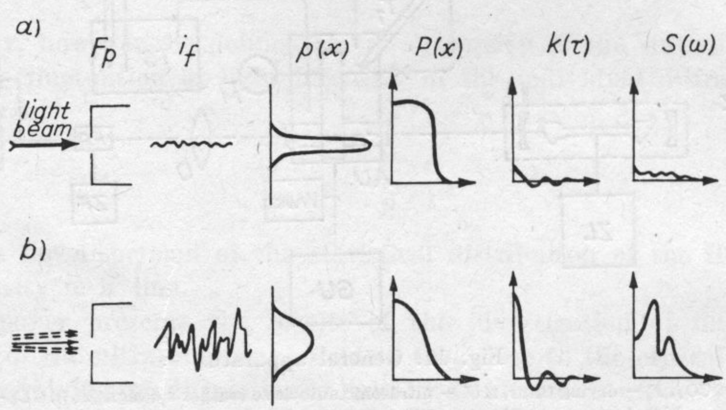


Fig. 2. A schematic representation of the process of recording and analysis of the measurement results

In addition to a decrease due to fluctuation of the diffraction angle, the photomultiplier also recorded fluctuations due to fluctuations of brightness connected with fluctuation of the amplitude of the diffracted light. The process recorded a given diffraction line is a combined phenomenon: of the fluctuation of the diffraction angle (formula (11)) which depends on the ratio, and (thus on the fluctuation of both the amplitude and the phase of the ultrasonic wave), and also on the fluctuation of the light amplitude, which depends also on the fluctuation of the amplitude of the ultrasonic wave, and thus on the quantities α, β, and γ defined by the formulae given in reference [15].

Fig. 2. shows schematically successive phases of recording and analysis of this combined process. The apparatus used permits an intensity proportional to the photoelectrical current i_f to be recorded, and a statistical distribution $p(x)$ and the distribution function $P(x)$ to be simultaneously determined automatically. Further processing of the statistical data was performed by calculations based on the fluctuation recording in which the autorrelation function $k(\tau)$ and the power spectrum of the fluctuation process $S(\omega)$, were determined.

The following quantities were used to characterize the fluctuation process: σ (the dispersion of the statistical distribution of the changes of light intensity in a line), τ_0 for $k(\tau) = 0$ (the first zero crossing of the function on the time axis), $S(\omega)_{\max}$ and $\omega_{\max}$ for which $S(\omega) = S(\omega)_{\max}$.

Based on the data obtained from measurements taken near the critical point and far from the critical point, the quantities γ and β which are related to the phase and amplitude of the ultrasonic wave, can be determined. Fluctuations of the diffraction angle can be expressed by the β/γ ratio which is related to σ in the following way:

$$\frac{\beta}{\gamma} = \frac{k}{p}\frac{\sigma}{l}.$$

On the other hand, γ can be determined by calculation of the ratio of most probable light intensities (in the $+1$st orders relative to the zeroth order) near the critical point, and far from the critical point, from the formula [28]:

$$e^{-1/2\gamma} = \frac{p(x)_{\max+1}/p(x)_{\max 0} \quad \text{(near the critical point)}}{p(x)_{\max+1}/p(x)_{\max 0} \quad \text{(far from the critical point)}}. \tag{5}$$

4. The results of investigations in CO_2

The experimental investigations, which were very difficult technically and laborious, permitted a relatively large number of recording to be obtained of the fluctuations of the intensity of light in diffraction figures from an ultrasonic wave under different thermodynamic conditions (far from and near the critical point). It was only partly possible to systematize the experiments, since the thermodynamic states obtained could not be, for purely technical reasons, planned beforehand and subsequently established with precision.

Furthemore for some thermodynamic states, the occurrence of a large absorption of the ultrasonic wave, prevented investigations over longer distances. In order to compare the results of measurements on lines of the same order (0 or +1) under different thermodynamic states, a rather severe selection of results was made and only a limited number of them underwent further manipulation and analysis.

The measurement results selected for this paper are shown in Table 1.

In order to explain the method used for the analysis of results, and also to present the differences between the fluctuation processes near and far from the critical point, Figs. 3 and 4 show, as an example, two full sets of results of all the intermediate phases (of. Table 1).

Figs. 3 and 4 present: a) original recordings of the fluctuation processes (for two different paper speeds —10 and 30 mm/s), b) their statistical distributions, $p(x)$, c) their distribution functions $P(x)$, and also d) the autocorrelation function, $k(\tau)$ (continuous line) and e) the density functions of the spectral power fluctuations $S(\omega)$ (discontinuous line), of the processes analyzed. The autocorrelation function has been deliberately given in an unnormalized form, since such a form was used for the calculation of $S(\omega)$.

Table 1 shows the characteristic quantities calculated for five different frequencies of the ultrasonic wave, corresponding to different thermodynamic conditions. The table also includes categories representing the order of the line and the distance from the quartz. In individual groups (of three and four lines) the values are ordered so that the first line corresponds to thermodynamic states close to the critical point; and succeeding lines to states further and further from the critical point. Thus the last line of each group corresponds to thermodynamic states very far from the critical state, for which the fluctuations are already relatively small and it can be assumed that the Raman-Nath distribution is valid for the light diffraction lines. The values in these lines gave a reference for the calculation of γ from formula (5).

All the recorded fluctuations of the light diffracted by an ultrasonic wave at specific frequencies for different thermodynamic conditions were analyzed in the manner mentioned above.

Figs. 5-8 show as an illustration the final calculations of the functions of spectral power of the fluctuation process for four selected cases.

For two of the frequencies of the ultrasonic wave used, (5.76 MHz and 27.8 MHz, Figs. 6 and 8), the behaviour of $S(\omega)$ is characterized by very sharp maxima at low pulsation rates. In particular, close to the critical point (continuous curves) these maxima have large values. It can also be noted that at a frequency of 27.8 MHz, the number of characteristic maxima becomes larger as the thermodynamic conditions diverge further from those of the critical point (discontinuous curves).

For the lines of the 1st order these very characteristic maxima are, considerably shifted (compared with the zeroth order) towards lower pulsation rates.

It can be seen from the diagrams and tables that for all the fluctuations of the ultrasonic wave investigated, the mean value of the power spectrum of the fluctuations becomes higher, as the thermodynamic state approaches the critical point. Also the closer the state comes to the critical point, the greater the shifts from the mean fluctuation power become. These deviations have components lying in the range of low pulsation rates and are thus connected with the occurrence of the "coarse grained" fluctuations.

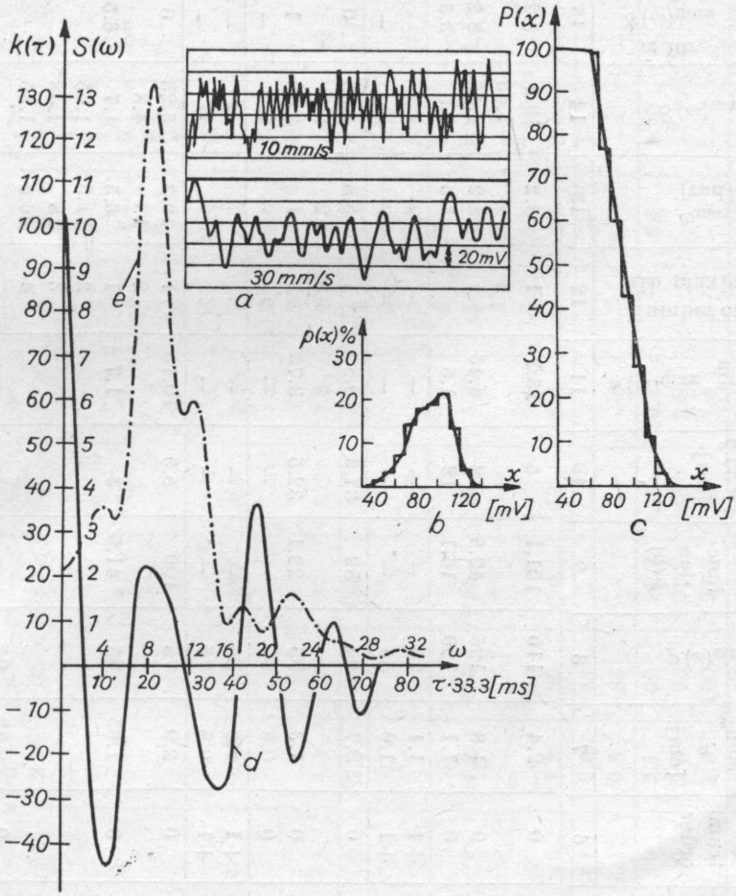


Fig. 3. An example of successive steps in the analysis of the fluctuation of the light diffracted by an ultrasonic wave at a frequency $f = 0.8$ MHz near the critical point ($t = 31.4$°C, $p = 71$ at) for the single line of the zeroth order

a – original fluctuation recordings (for two paper speeds 10 and 30 mm/s), b – statistical distribution $p(x)$, c – distribution function $P(x)$, d – autocorrelation function $k(\tau)$ (continuous line), e – fluctuation power spectrum $S(\omega)$ (discontinuous line)

When the thermodynamic conditions are farther from the critical point, the values of the maxima of the fluctuation components in the power spectrum decrease until they vanish completely when no fluctuations of the light diffracted by the ultrasonic wave can be observed.

Table 1

Ultrasonic wave frequency [MHz]	Ultrasonic wave intensity $\left[\frac{\text{Watt}}{\text{cm}^2}\right]$	Thermodynamic conditions t [°C]	Thermodynamic conditions p [at]	Distance from wave source x [cm]	Spectrum order	Dispersion σ [cm]	$p(x)_{\max}$*	Autocorrel. function $k(0)$	$\tau \times 33.3$ [ms]	Fluct. power spectrum $S(\omega)_{\max}$	Number of the nth maximum	$\omega_{\max_n}$ [rad/s]	$S(\omega)_{\max_n}$	Pulsation rate: ω for $S(\omega)_{\max}$ [rad/s]	Pulsation rate: $\omega_{\max}$ [rad/s]	β/γ	γ	β
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
0.8	0.1	31.4	71	1.3	0	2.4	110	101.1	5	13.3	1	8.2	13.3	8.2	36	0.14	4.01	0.55
											2		1.65					
	0.1	34.5	75	1.3	0	1.8	125	43.8	12	4.45	1	2.2	4.45	2.2	32	0.10	—	—
	0.1	27.4	62	1.3	0	1.1	120	15.1	12	1.5	1	3.5	1.5	3.5	30	0.06	—	—
	0.1	31.6	71	1.3	+1	1.1	70	—	—	—	—	—	—	—	—	0.05	—	—
	0.1	27.6	63	1.3	+1	1.0	55	—	—	—	—	—	—	—	—	—	—	—
	0.1	31.2	72.5	6.5	0	2.7	95	82	37.5	10.8	1	5.8	6.95	0	32	0.15	3.52	0.54
											2	12	2.3					
	0.1	35.1	76	6.5	0	1.3	95	22.1	32.5	3.75	1	1	3.75	1	30	0.07	—	—
	0.1	26.3	61.5	6.5	0	0.8	95	—	—	—	—	—	—	—	—	0.04	—	—
	0.1	32	73.5	6.5	+1	2.8	65	—	—	—	—	—	—	—	—	—	—	—
	0.1	27	62.5	6.5	+1	0.8	75	—	—	—	—	—	—	—	—	—	—	—
	0.1	32.1	73	1.3	0	2.9	75	100	5.3	10.15	1	10.3	5.25	0	40	0.16	6.25	1.03
											2	18	4.8					
	0.1	30	68	1.3	0	1.65	85	41.5	3	1.7	1	6.5	1.7	6.5	45	0.09	—	—
											2	11.5	1.65					
											3	18	1.48					
											4	26.5	1.35					
	0.1	28.6	60.5	1.3	0	0.82	95	—	—	—	—	—	—	—	—	0.05	—	—
	0.1	26.4	59.5	1.3	0	0.8	95	—	—	—	—	—	—	—	—	0.04	—	—
	0.1	31.4	71.5	6.5	0	2.4	75	91	5	6.05	1	6.2	6.05	6.2	43	0.15	12	1.85
											2	11.6	5.06					
											3	15.8	3.45					

											4	19.3	2.75					
	0.1	29.6	67	6.5	0	1.8	80	31.2	8	3.20	1	4.8	3.2	4.8	26	0.09	—	—
1.2											2	9.7	2.65					
	0.1	28	60	6.5	0	1.2	85	27.2	32.5	2.5	—	—	—	0	40	0.09	—	—
	0.1	31.2	71.5	6.5	+1	0.9	55	—	—	—	—	—	—	—	—	—	—	—
	0.1	28.2	60	6.5	+1	0.8	65	—	—	—	—	—	—	—	—	—	—	—
	0.1	32.2	73.5	1.3	+1	2.7	55	185.5	6	17.8	1	6.8	13.2	0	44	—	—	—
											2	13	13.5					
											3	16	5.1					
	0.1	30.3	69	1.3	+1	1.62	45	85.2	15.5	9	1	1.8	9	1.8	38	—	—	—
											2	5.3	6.7					
											3	14	3.35					
											4	21.7	0.78					
	0.1	29.4	63	1.3	+1	1.55	55	38	4.5	2.25	1	6.8	2.25	6.8	34	—	—	—
											2	14.6	1.95					
											3	23	0.75					
	0.1	27.1	60.5	1.3	+1	1.1	75	31	3.5	1.6	1	12	1.6	12	40	—	—	—
	0.1	31.5	73	1.3	0	3.3	95	136.7	15.5	13.15	1	2.2	13.15	2.2	40	0.19	2.6	0.49
											2	5.2	11.1					
											3	14.8	6.25					
											4	27.5	0.8					
	0.1	33.4	74.5	1.3	0	1.9	115	55	30	6.32	1	1.4	6.32	1.4	36	0.11	—	—
											2	8.5	3.15					
3.51											3	18.6	1.8					
	0.1	35.8	80	1.3	0	1.2	120	20	13	2.4	1	2.6	2.4	2.6	32	0.07	—	—
	0.1	31.2	73	1.3	+1	2.65	55	—	—	—	—	—	—	—	—	—	—	—
	0.1	36	79	1.3	+1	0.8	85	—	—	—	—	—	—	—	—	—	—	—
	0.1	32.8	75.5	6.5	0	3.1	90	89.2	40	20.6	1	6.2	5.73	0	34	0.18	8.1	1.43
	0.1	34.1	76	6.5	0	2.8	95	39.7	48	8.8	—	—	—	0	30	0.16	—	—
	0.1	35.6	79.5	6.5	0	1.9	100	37.3	63	7.6	—	—	—	0	34	0.11	—	—
	0.1	34.1	70.5	1.3	0	3.2	85	117.3	5.5	10.86	1	5.6	8.8	10.8	38	0.18	3.57	0.65
											2	10.8	10.86					
											3	16	2.5					
	0.1	30.1	67	1.3	0	2.0	85	59.3	8	4.6	1	2.6	4.2	7.2	38	0.10	—	—
											2	7.2	4.6					

Table 1 ctd.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
											3	19.6	1.35	—				
5.76	0.1	27.2	61	1.3	0	1.05	85	9.7	66.5	1.97	1	0	1.97	0	20	0.04	—	—
	0.1	34.2	70	1.3	+1	2.65	65	139.5	6.5	17.1	1	5.6	17.1	5.6	36	—	—	—
											2	10.6	11.03					
	0.1	29.4	66.5	1.3	+1	1.9	55	48.8	4.5	3.83	1	9.4	2.30	14.2	48	—	—	—
											2	14.2	3.83					
											3	23.6	1.05					
	0.1	27.4	61.5	1.3	+1	1.1	75	17.3	29.5	2	1	—	—	0	36	—	—	—
	0.1	32.6	73.5	2	0	3.6	85	152.2	5.5	16.5	1	11.2	15.5	11.2	40	0.20	6.25	1.28
	0.1	30.1	69	2	0	2.9	95	63.8	5	4.05	1	3	4.05	3	40	0.16	—	—
											2	11.4	4					
											3	17.6	2.55					
	0.1	27.5	62	2	0	1.3	85	26	18	2.35	1	4	2.35	4	36	0.07	—	—
27.8											2	12	1.63					
											3	18	0.63					
	0.1	32.5	74.5	2	+1	3.1	65	120	6.5	—	—	—	—	—	—	—	—	—
	0.1	30.6	69.5	2	+1	2.8	70	72.2	5	7.7	1	3.2	3.73	9.6	40	—	—	—
											2	9.6	7.7					
											3	15.2	3.45					
	0.1	36.9	63	2	+1	0.9	70	23.2	75	4.8	—	—	—	0	32	—	—	—

* $p(x)_{\max}$ denotes the most probable value of the photoelectric current (i.e. light intensity).

The distribution of the light intensity in the diffraction spectrum then agrees with the Raman-Nath theory. The characteristic maxima in the fluctuation power spectrum for a line of the zeroth order at a distance of 6.5 cm are shifted towards lower values of pulsation rate, compared to those at a distance of 1.3 cm. A similar shift occurs at the same distance for a line of the 1st order, compared with a line of the zeroth order. The occurrence of specific maxima in the fluctuation power spectrum for a line of the zeroth order under critical conditions shows that the behaviour of the fluctuations is "quasiperiodic". In the case when several characteristic maxima occur in the fluctuation power spectrum, it can be observed that they occur at nearly equal intervals on the scale of pulsation rate ω. Thus the spectrum is "quasiharmonic".

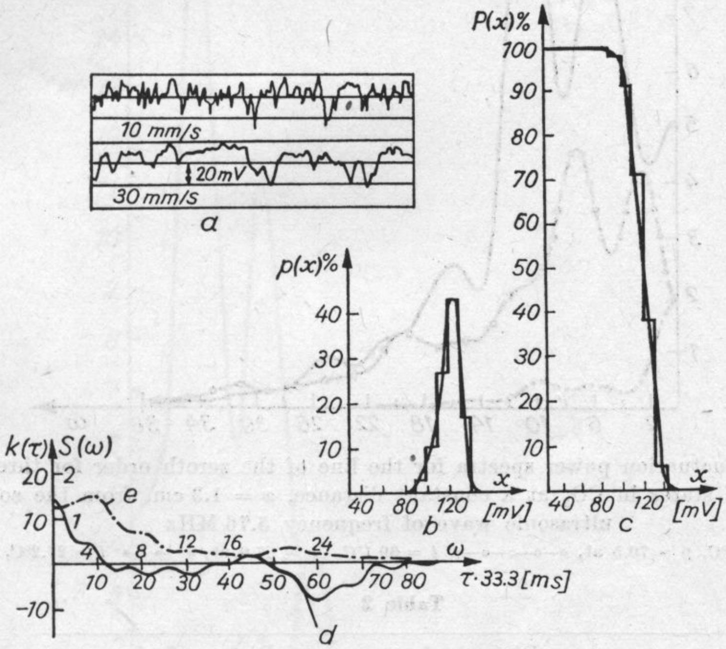


Fig. 4. An example of successive steps in the analysis of the fluctuations of the light diffracted by an ultrasonic wave at a frequency of $f = 0.8$ MHz relatively far from the critical point ($t = 27.4°C$, $p = 62$ at), for the single line of the zeroth order

a – original fluctuation recordings (for two paper speeds – 10 and 30 mm/s), b – statistical distribution $p(x)$, c – distribution function $P(x)$, d – autocorrelation function $k(\tau)$ (continuous line), e – fluctuation power spectrum $S(\omega)$ (discontinuous line)

The β/γ ratio and subsequently γ itself, were calculated, on the basis of previously determined values of the dispersion (δ) of the statistical distribution of the fluctuations of the intensity of the light diffracted by the ultrasonic wave under different thermodynamic conditions. From these results β could be determined. Values corresponding to these quantities are shown in Table 1.

For all frequencies of the ultrasonic wave used, the value of β/γ decreases as the deviation of the thermodynamic conditions from the critical point decreases; thus showing that fluctuations in quantities of the wave field become increasingly smaller. It can be also observed that the β/γ ratio rises sligtly close to the critical point as the frequency of the ultrasonic wave increases. This

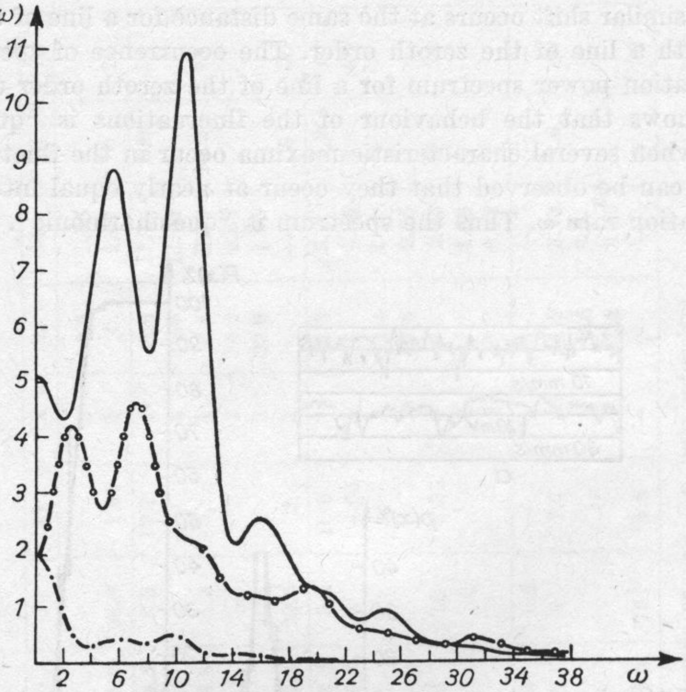


Fig. 5. Light fluctuation power spectra for the line of the zeroth order for three different thermodynamic states in CO_2 at a constant distance, $x = 1.3$ cm, from the source of an ultrasonic wave of frequency 5.76 MHz

—— $t = 34.1$ 1°C, $p = 70.5$ at, o—o—o—o—o $t = 30.1$°C, $p = 67.0$ at, •—•—• $t = 27.2$°C, $p = 61.0$ at

Table 2

Frequency f[MHz]	Distance from source of ultrasonic wave 1.3 [cm]		Distance from source of ultrasonic wave 6.5 [cm]	
	values of γ	values of β	values of γ	values of β
0.8	4.01	0.55	3.52	0.54
3.51	2.6	0.49	8.1	1.43
5.76	3.57	0.65	—	—
27.8	6.25	1.28	—	—

results from the fact that a stronger interaction between the ultrasonic wave and the fluctuations of the medium occurs close to the critical point.

The values of γ and β for five frequencies and two distances from the source of the ultrasonic wave are shown in Table 2.

It follows from the statistical distributions of the light intensity fluctuations of the diffraction spectrum recorded (at the shortest distance from the quartz), that for a line of the zeroth order these distributions are generally symmetrical (differing slightly from normal distributions), while for lines of the 1st order they are very asymmetrical and irregular. As an illustration,

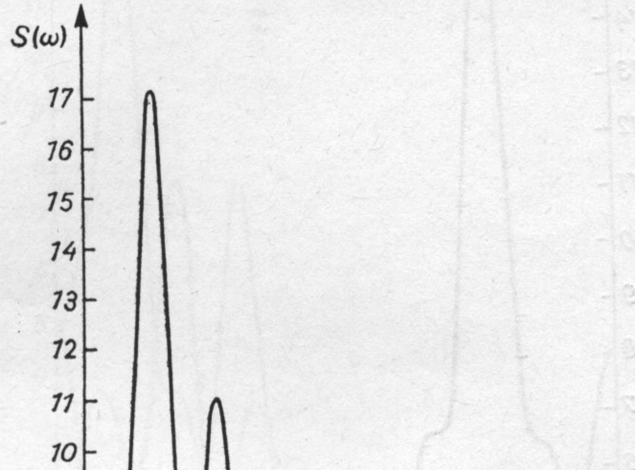


Fig. 6. Light fluctuation power spectra for a line of the first order for three different thermodynamic states in CO_2 at a constant distance, $x = 1.3$ cm, from the source of an ultrasonic wave of frequency 5.76 MHz

——— $t = 34.2$°C, $p = 70.0$ at, o—o—o— $t = 29\ 4$°C, $p = 66.5$ at, •—•—• $t = 27.4$°C, $p = 61.5$ at

Fig. 9 shows two distributions recorded at an ultrasonic frequency of 5.76 MHz close to the critical point, for the two cases. At longer distances from the wave source, deviations from the normal distribution are considerably greater, which means that deformation of the acoustic wave increases with distance for a given intensity.

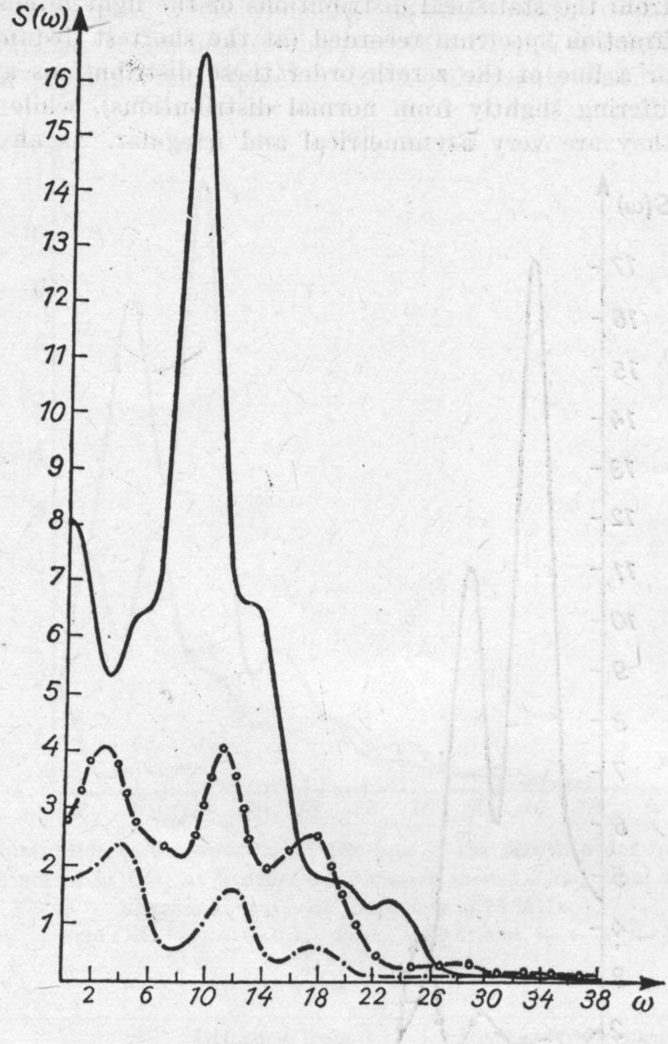


Fig. 7. Light fluctuation power spectra for the line of the zeroth order for three different thermodynamic states in CO_2 at a constant distance, $x = 2$ cm, from the source of an ultrasonic wave of frequency 27.18 MHz

——— $t = 32.6°C$, $p = 73.5$ at, o—o—o $t = 30.1°C$, $p = 69.0$ at, •—•—• $t = 27.5°C$, $p = 62.0$ at

5. Comparison of results for liquid-liquid and liquid-vapour systems

In both a liquid-liquid system [21] and a liquid-vapour system [20] the β/γ ratio takes increasingly higher values as the critical point of the medium is approached, which shows that deformation of a propagating ultrasonic wave increases as the critical point is approached.

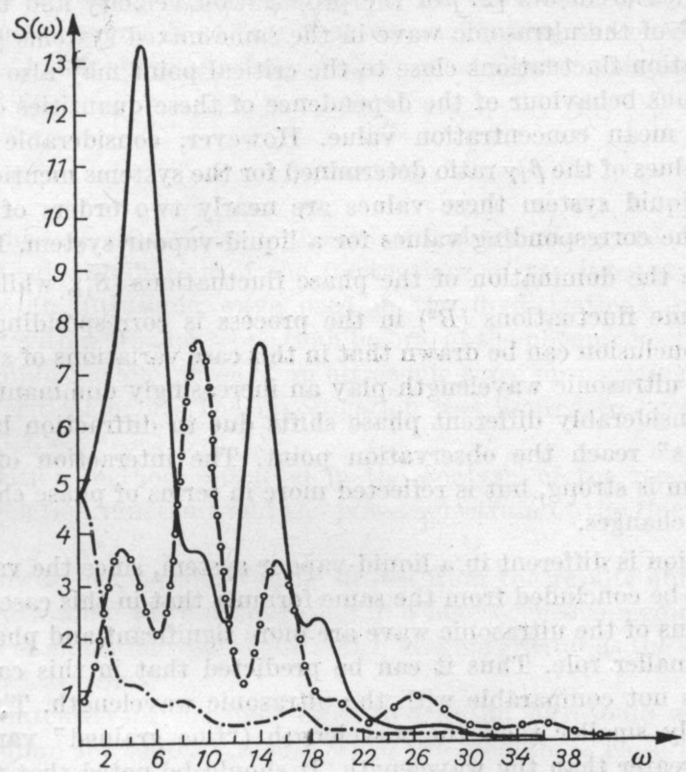


Fig. 8. Light fluctuation power spectra for a line of the first order for three different thermodynamic states in CO_2 at a constant distance, $x = 2$ cm, from the source of an ultrasonic wave of frequency 27.8 MHz

—— $t = 32.5$°C, $p = 74.5$ at, ○—○—○ $t = 30.6$°C, $p = 69.5$ at, •—•—• $t = 26.9$°C, $p = 63.0$ at

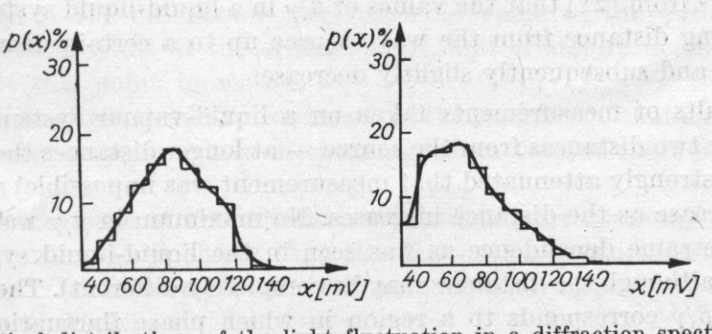


Fig. 9. Statistical distributions of the light fluctuation in a diffraction spectrum near the critical point ($t = 34.1$°C, $p = 70.5$ at)

= for the line of the zeroth order, b — for a line of the first order

Parallel measurements [27] of the propagation velocity and the attenuation coefficient of the ultrasonic wave in the same mixed systems [21] showed that concentration fluctuations close to the critical point may also be a source of the anomalous behaviour of the dependence of these quantities on temperature and the mean concentration value. However, considerable differences occur in the values of the β/γ ratio determined for the systems mentioned above. For a liquid-liquid system these values are nearly two orders of magnitude greater than the corresponding values for a liquid-vapour system. High values of β/γ indicate the domination of the phase fluctuations $(\overline{S_1^2})$, while the share of the amplitude fluctuations $(\overline{B^2})$ in the process is correspondingly smaller. The indirect conclusion can be drawn that in this case variations of sizes comparable with the ultrasonic wavelength play an increasingly dominant role. Thus waves with considerably different phase shifts due to diffraction by the fluctuation "grains" reach the observation point. The interaction of the wave and the medium is strong, but is reflected more in terms of phase changes than in amplitude changes.

The situation is different in a liquid-vapour system, since the values of β/γ are low. It can be concluded from the same formula that in this case the amplitude fluctuations of the ultrasonic wave are more significant and phase fluctuations play a smaller role. Thus it can be predicted that in this case the size of variations is not comparable with the ultrasonic wavelength. They are either considerably smaller than the wavelength ("fine grained" variations) or considerably greater than the wavelength. It should be noted that the sizes of the variations change considerably close to the critical point, and in each thermodynamic state there is a very wide statistical distribution of sizes, in which only the most probable sizes can be distinquished. The ratio β/γ which is determined is the effect of a combined interaction of the ultrasonic wave and variations of different sozes and the remarks above refer only to interactions with the variations of the dominant size in a given distribution.

It follows from [21] that the values of β/γ in a liquid-liquid system increase with increasing distance from the wave source up to a certain distance, reach a maximum, and subsequently slightly decrease.

The results of measurements taken on a liquid-vapour system (although taken at only two distances from the source — at longer distances the ultrasonic wave was so strongly attenuated that measurement was impossible) show a tendency to increase as the distance increases. No maximum in β/γ was observed. However, the same dependence as was seen in the liquid-liquid system could be expected (although the distances may be completely different). The maximum in the ratio β/γ corresponds to a region in which phase fluctuations become comparable with amplitude fluctuations, $\overline{S_1^2} = \overline{B^2}$ (assuming that the statistical distributions of the fluctuations of these quantities are normal [1]). The region is related with the transition from the so called Fresnel region (near field) to the Fraunhofer region (far. field). Phase fluctuations play an essential role in the near field and a less significant role in the far field. In the far field the distance from the source is so long that waves reaching the observation point have only slight shifts in phase. Amplitude fluctuations are then more dominant.

5. Summary

The method of simultaneous photoelectrical recording and determination of the statistical distribution of the fluctuations of the intensity of the light diffracted by an ultrasonic wave used in the investigations, permitted the investigation, with a broader scope than before, of the phenomena of the fluctuation of the amplitude and phase of an ultrasonic wave propagating in a medium close to the critical point. The investigation was limited to "coarse grained" fluctuations.

A correlation method was used in the analysis of the records obtained. The autocorrelation functions and the power spectrum of the fluctuations were determined.

The following conclusions can be drawn from the results obtained:

(1) An ultrasonic wave propagating in a heterogeneous medium undergoes deformations of amplitude and phase which become greater as the system approaches the critical point.

(2) The intensity of the light diffracted by an ultrasonic wave propagating in a medium fluctuates strongly when the medium is close to the critical point.

(3) Statistical distributions of the fluctuations of the light intensity in the zeroth order are approximately symmetrical (being only slightly different from normal distributions).

(4) Statistical distributions of the fluctuations of the light intensity in diffraction lines of the first order are asymmetrical and irregular.

(5) The dispersion of the statistical distributions of the fluctuations of the light intensity in the diffraction spectrum (the value of the autocorrelation function $\tau(o)$) is greatest at the critical point, and gradually decreases as the distance from this point increases.

(6) It follows from the power spectrum of the fluctuation that characteristic components with pulsation rates in the range from 0 to 35 rad/s occur in the process.

(7) The occurrence of characteristic pulsations (corresponding to maxima in the spectrum of $S(\omega)$) shows that the fluctuation processes observed are "quasiperiodic". In some cases the spectrum of $S(\omega)$ appears to be "quasiharmonic" in character.

(8) The mean values of the power spectrum ($S(0)$) reach their highest values for thermodynamic conditions close to the critical point.

(9) Deviations from the mean value of the fluctuation power spectrum at a long distance from the source of the ultrasonic wave (for a line of the zeroth

order) are much smaller than those at a shorter distance and are in a lower range of pulsation rates (up to about 15 rad/s).

(10) A shift of the characteristic maxima towards lower pulsation rates can be observed in the power spectra of the fluctuations of the light intensity in a line of the 1st order.

(11) The methods of determinating the ratio β/γ, and γ itself, given in this paper, permit the value of β to be calculated. This represents an improvement on the investigations of Śliwiński who gave only the method for determinating the ratio β/γ.

(12) The apparatus used to record the fluctuations permitted only "coarse grained" fluctuations to be investigated. The use of the correlation method in the analysis of the results obtained illustrates its value. The correlation method could also be successfully used for the analysis of processes connected with "fine grained" fluctuations. This would require recording devices with very small time constants to be used. This is a separate problem which could be investigated by an analogous method when suitable apparatus is available.

I wish to express my gratitude to Prof. dr hab. A. Śliwiński from Gdańsk University for much valuable advice and discussion during the performance of the work and the preparation of the results obtained.

References

[1] L. A. Chernov, *Rasprostranienije woln w sriedie so sluchajnymi nieodnorodnostjami*, Izd. AN SSSR, Moskwa 1958, 10-16, 58-145.

[2] S. M. Rytov, *Diffraktsia sveta na ultrazvukovikh volnakh*, Izd. AN SSSR, Ser, Fiz., 2 (1937).

[3] A. M. Obukhov, *O vliani slabikh neodnorostei atmosfery na rasprostranenie zvuka i sveta*, Izd. ANSSSR, (seria geofizicheskaya) 2, 155 (1963).

[4] L. A. Chernov, *Korrelatsia fluktuatsii amplitudy i fazy pri rasprostranenii volny v srede so swuchaynimi neodnorostiami*, Akust Zhurn., **1**, 1, 89 (1955).

[5] A. Śliwiński, *Interaction of light and ultrasonic waves in a medium with density fluctuations* (in Polish), Prace Wydz. Mat.-Fiz.-Chem. UAM, Poznań (1964).

[6] C. V. Raman, N. S. Nagendra Nath, *The diffraction of light by high frequency sound waves*, Proc. Indian Acad. Sci., **A2**, 413 (1939).

[7] L. Bergman, *Der Ultraschall und seine Anwendung in Wissenschaft und Technik*, Zürich 1956.

[8] R. Lucas, P. Biquard, *Propriétés optiques des ondes élastiques ultra-sonores*, J. Phys. Radium, **3**, 464 (1932).

[9] O. Nomoto, K. Negishi, *Diffraction of light by ultrasonic waves of finite amplitude*, Acustica, **15**, 4, 223-235 (1966).

[10] G. W. Willard, *Criteria for normal and abnormal ultrasonic light diffraction effects*, JASA, **21**, 2, 101-108 (1949).

[11] V. F. Nozdrev, *Primenie ultraakustiki v molekularnoi fizike*, Izd. Nauka, Moskva 1958.

[12] E. H. Wagner, *Zur Beugung paralleler Lichtbündel an laufendenden Ultraschallwellen*, Zs. Phys., **141**, 5, 604-621 (1955).

[13] J. Satyanarayana Murty, *Theoretical investigation on the diffraction of light by superposed ultrasonic waves*, JASA, **26**, 4, 970-974 (1954).

[14] L. E. Hargrove, E. A. Hiedemann, R. Martens, *Diffraction of light by two spatially separated parallel ultrasonic waves of different frequency*, Zs. Phys. **167**, 3, 326-336 (1962).

[15] A. Śliwiński, *A statistical method for determining the deviation of the fluctuation distribution of lines of light diffracted by an ultrasonic wave from a normal distribution in a heterogeneous medium* (in Polish), Arch. Akust, **1**, 91-101 (1966).

[16] A. Śliwiński, A. E. Brown, *Ultrasonic light diffraction patterns in ethylene in the critical region*, Acustica, **16**, 5, 312-323 (1965-66).

[17] A. Śliwiński, M. Łabowski, *Some aspects of laser light diffraction by an ultrasonic wave in an inhomogeneous medium*, The 6th International Congress on Acoustics, Tokyo, Japan, August 21-28, (1968).

[18] A. Śliwiński, M. Łabowski, *Fluctuations of patterns of light diffracted by an ultrasonic wave in a medium near the critical point, considered as the result of hypersonic processes*, Seventh International Congress on Acoustics, Budapest 1971.

[19] M. Łabowski, *Laser light diffraction by ultrasound in media near the critical point* (in Polish), Arch. Akust, **5**, 2, 239-242 (1970)

[20] M. Łabowski, A. Śliwiński, *Correlation analysis as applied to the observation of fluctuations of laser light diffracted by an ultrasonic wave in an inhomogeneous medium*, Seventh International Congress on Acoustics, Budapest 1971.

[21] A. Śliwiński, M. Łabowski, *Measurement of amplitude and phase fluctuations of an ultrasonic wave by optical method in critical mixtures methanol-hexan, methanol-cyclohexan, phenol-water* (in Polish), Arch. Akust., **4**, 2, 323-328 (1967).

[22] W. S. Pugacher, *Teorie swuchaynikh funktsii*, Fizmatgiz, Moskwa 1960.

[23] A. Śliwiński, *Light diffraction patterns from ultrasonic waves in a liquid-vapour system and a critical solution* (in Polish), Prace Wydz. Mat.-Fiz.-Chem. UAM, Fizyka 1, Poznań 1961.

[24] A. E. Brown, *A note on scattering of sound waves in a binary mixture near its critical solution point*, Acustica, **18**, 3, 169-172 (1967).

[25] M. Łabowski, *Universal ultrasonic autoclave for investigation of liquid-vapour systems near the critical point* (in Polosh), Arch. Akust., **2**, 5, 271-274 (1970).

[26] M. Łabowski, *Determination of the statistical distribution of the ultrasonic wave intensity near the critical point of carbon dioxide by an optical method*, Proceedings of the Xth International Conference on Ultrasound, Prague, September 1972.

[27] M. Łabowski, J. Średzińska, *Propagation velocity and attenuation of ultrasound in binary critical mixtures; methanol-hexan and phenol-water* (in Polish), Arch, Akust. **4**, 2, 329-338 (1967).

[28] M. Łabowski, *Investigation of the amplitude and phase fluctuations of an ultrasonic wave in a heterogeneous medium by laser light*, Doctoral thesis — UAM Poznań 1970.

Received on February 1, 1979

CHRONICLE

INTER-NOISE 79

A successive Inter-Noise conference was held in Warsaw on September 11 - 13, 1979. Over 500 participants from 28 countries took part in it.

Sections

A. Mechanical noise
B. Aerodynamic noise
C. Noise measurements methods
D. Machinery noise
E. Engine noise control
F. Reduction of in – plant noise
G. Designing and planning for industrial noise control
H. Transportation noise abatement
I. Preventive shipboard noise control
J. Aircraft and airport noise
K. Environmental noise
L. Personal noise exposure
M. Noise control engineering in buildings.

At plenary sessions 5 general papers and 54 poster – form papers were delivered. The other part of the debates took place in four parallel sections, including 22 invited papers and 86 contributed papers.

Before the closing of the sessions, panel discussions were held, which summed up discussions in respective sections. Moreover, there was a special debate on the subject of training in the field of noise and vibration.

Conference materials consisting of 944 pages were published in two-volume proceedings.

List of papers delivered

General papers

R. W. B. Stephens, Education in acoustics.
S. Czarnecki, Noise control in Poland.
P. E. Doak, Progress in aerodynamic noise research and its implications for development of noise control measures.
T. Kihlman, Goals and means in industrial noise control.
W. W. Lang, G. Maling, Computer systems for noise control engineering.

Section A — Mechanical noise

Chairmen: W. Schirmer, Z. Engel

Invited paper

W. Schirmer, Engineering methods for computing of the radiation ratio of machine parts.

Contributed papers

M. E. Westcott, Impulsive noise from machinery: validation of prediction methods.

A. Brański, R. Panuszka, M. Zabawa, Sound radiation by a rectangular plate represented by a model in form of a system of rectangular pistons.

Poster-form papers

Z. Engel, H. Jaworowski, S. Kasprzyk, Experimental investigations of impact noise.

N. Kosińska, A. Śliwiński, Z. Trumpakaj, T. Zaleski, Numerical calculations of eigenfrequencies of circular plate with damping concentrically clamped with free periphery.

S. Czarnecki, Z. Engel, R. Panuszka, Estimation of equivalent surface area for determination of the acoustic power of a circular plate.

Section B — Aerodynamic noise

Chairmen: E. A. Müller, W. M. Jungowski.

Invited papers

J. Němec, Problems of similarity of fan noise.

W. M. Jungowski, G. B. Sobieraj, Influence of the outer shape of Hartmann-Sprenger generator on the radiated sound spectrum.

Contributed papers

B. Barsikow, W. Neise, Influence of nonuniform inflow conditions on centrifugal fan noise.

B. Michel, H. Arbey, M. Sunyach, Radiation from subsonic rotor operating in a non-uniform incident flow.

Y. N. Chen, U. Bolleter, Thermo-driven self-sustained vibration and noise of a supercritical reduction valve.

W. Neise, Flow noise level at microphones in flow ducts – determination via comparative measurement with noise cone and slit-tube.

P. Anglesio, E. Cirillo, Acoustic characterisation for combustion heads of forced draft burners.

Poster-form papers

S. Czarnecki, M. Czechowicz, T. Sobol, Aero-vibroacoustic feedback for the edgetone production.

W. C. Selerowicz, Relations between boundary conditions and acoustic effects of the free oscillating gas jets.

K. J. Witczak, Noise generated by annular gas jet at various geometry of its surroundings.

S. Stecco, E. Carnevale, A new pendulum detuner to reduce in-plant turbomachinery vibrations.

J. R. Piechna, Influence of operating conditions of a reciprocating compressor on the spectra of pressure pulsation and noise.

W. M. Jungowski, W. C. Selerowicz, W. J. Stojanowski, B. M. Niewczas, Prevention of the noise generated by the release of gas into the atmosphere.

Section C — Noise measurement methods

Chairmen: H. G. Leventhall, P. V. Brüel

Contributed papers

J. Tichy, Progress in the reverberation room method of sound power measurements.

M. J. M. Jessel, Some evidences for a general theory of active sound absorption.

H. G. Leventhall, K. Eghtesadi, Active attenuation of noise: dipole and monopole systems.

V. Chalupová, Calculation conditions of noise energetic spectra.
E. Eichler, Transmission loss revisited.
J. Stěnička, Impact testing aiming at the determination of structureborn noise transfer parameters.
M. Yanagida, D. G. Stuart, O. Kakusho, Noise separation by generalized inverse of matrices.
J. P. Sirieys, D. Commins, Comparative field study on noise monitoring systems.
T. Suzuki, A transmission coefficient in a spherical incidence wave.
I. L. Vér, M. M. Myles, Evaluation of the acoustic performance of anechoic chambers utilizing pulse response and transient spectrum.
M. Papp, On a new type of infrasound chambers.
E. Zwicker, Advantages of a precise loudness meter.
J. M. Lambert, Application of a modern intensity-meter to industrial problems: example of in-situ sound power determination.
K. Obermayr, Calibration of directional microphones in a sound measuring duct.
W. F. King III, On microphone arrays used to locate sound sources on moving vehicles.
J. E. K. Foreman, Instrumentation and procedures in the measurement and statistical analysis of attitudinal response of people to audible noise from high voltage transmission lines.

Poster-form papers

R. B. Randall, N. Thrane, Impulse analysis using a real-time digital analyzer.
E. Kozaczka, Z. Dolny, Application of the power cepstrum to determine the differences in arrival time between pulses in a duct.
M. Vogt, H. G. Leventhall, Elimination of secondary wave feedback in dipole active attenuators.

Section D — Machinery noise

Chairmen: J. Němec, C. Puzyna.

Contributed papers

M. L. Abrahams, J. S. Willims, D. J. Johns, Paper corrugating machinery: noise sources, measurements, analyses, and cures.
J. Motylewski, Measurement of noise emitted by moulding machines.
K. Szabo, Noise power level measurement at cement works.
V. Irmer, K. Stinshoff, Possibilities of noise reduction for lawn mowers.
V. Alić, M. Jevremović, Models of cybernetic approach to noise control in vacuum systems for water cooling below wet bulb temperature.
M. Makomaski, Influence of some constructional features on the noise from electric arc furnace.
T. Zieliński, Experimental and theoretical investigations of a barking drum model.
P. L. Timár, Technique of reducing the noise of electrical machines.

Poster-form papers

Ir. W. C. du Pre, Noise reduction of weaving looms.
A. G. Herbert, J. M. Burrows, Noise in blanking and piercing.
J. S. Anderson, I. S. Turfanda, Factors affecting the noise in piercing and blanking operations with a small punch press.
A. G. Herbert, Noise in press shops.
H. K. Tönshoff, H.-D. Raschke, A. Scherger, R. Westphal, Noise reduction on cutting processes with high-speed tools of different structure-sound excitation mechanisms.

J. Bednarowski, Influence of the geometrical parameters of a circular saw on the form of noise spectrum in the process of cutting of tubes and profiles.
P.-A. Berg, G. Lagerberg, Are pneumatic tool noise data useful for predicting working noise in steel structures.
J. Kaźmierczak, Some aspects of the non-stationarity of acoustic signals in the case of electric arc furnaces.
H. Łopatowa, Examination of acoustic field generated by the use of vibrotamper for sand moulding.
D. Augustyńska, Results of measurements of infrasounds emitted by piston compressors.
A. Wnuk, Methods and means of noise reduction in textile machinery.

Section E — Engine noise control.

Chairmen: D. Johns, A. Śliwiński

Invited papers

M. J. Crocker, Identifying sources of noise in engines and vehicles.
S. D. Haddad, Piston slap simulation rig – a fundamental tool to study piston slap induced vibration and noise in diesel engines.

Contributed papers

L. Łukaszek, Influence of exhaust conditions on the noise of four-stroke diesel engines.
N. Lalor, E. C. Grover, Mechanical noise associated with the crank mechanism of an engine.
A. F. Seybert, Identification of piston impact noise using a coherence method.
D. M. Waters, Light hovercraft noise control.

Section F — Reduction of in-plant noise

Chairmen: J. Tichy, W. Rostafiński.

Contributed papers

M. Bartenwerfer, On noise radiation from centrifugal fan casings.
M. Mirowska, Influence of structure on sound absorption properties of inorganic fibrous materials.

Poster-form papers

A. M. Gołaś, J. A. Zalewski, E. M. Zalewska, Use of the Monte-Carlo method for identification of acoustical field distribution in industrial interiors.
R. Janczur, S. Czarnecki, Influence of the shape and kind of acoustical treatment of ceiling on distribution of sound energy in a shallow room.
E. Kotarbińska, Influence of acoustic barrier on the field of refracted waves in a flat room.
R. Pompoli, Insertion loss of acoustic barriers in industrial halls.
J. Sułocki, K. Nogalski, N. Kosińska, An example of noise level reduction accompanying a modernization process in a factory.

Section G — Designing and planning for industrial noise control

Chairmen: P. François, R. Pompoli.

H. Bley, E.-L. Noe, Specific computer use for noise abatement in planning factories.
Y. Okamoto, H. Taira, A new programming system for noise prediction in a large industrial plant.

A. H. FRITZ, G. NEUGEBAUER, Predetermination of noise distribution in industrial halls to permit lay-outs with low-noise job positions.

H. HÄGGKVIST, Methods of constructional noise control in Finnish industry.

J. ROBERTS, Designing and planning for noise control in coal preparation plants.

G. ALOCCI, A. BLADACCONI, Industrial halls generated noise: state of art according to Italian legislation.

E. AZZARETTO, I. BARDUCCI, Criteria underlying noise control measures for single-blade disk saws used in woodworking.

Poster-form papers

B. A. KUGLER, A model for the assessment and analysis of industrial noise problems.

J. DMOCHOWSKI, A. BAŃKA, Method determining the influence of design-failures at successive stages of the plant investment process on acoustic environment.

J. SCHOLZE, Sound attenuation of ventilation apertures in exterior walls and roofs of wide-span workshops.

Section H — Transportation noise abatement

Chairmen: A. LAWRENCE, P. M. da SILVA.

Invited paper

A. W. LAWRENCE, Variation in individual vehicle noise attenuation provided by a building facade.

Contributed papers

G. BLÜCHER, A. SKOLD, Joint Nordic computing model for predicting road traffic noise levels.

O. ABDEL ALIM, Road traffic noise as a stochastic process.

N. WANG, A noise control conoid model for traffic noise simulation.

D. J. MARTIN, Low frequency traffic noise and building vibration.

J. KRAGH, Motorways noise propagation and screening under varying metereological conditions.

A. MOERKERKEN, H. J. L. van WIJK, Meteorological influences on the transmission of traffic noise.

C. LARSSON, S. ISRAELSSON, Influence of the metereological parameters on the sound propagation from a traffic road.

U. SANDBERG, A road surface for reduction of tire noise emission.

S. SATO, H. MATSUHISA, Wheel-rail noise reduction of rail rapid transit.

V. KUNZL, Acoustic pulse — a useful tool of vehicle internal noise investigation and control.

S. KURRA, A computer model for predicting sound level distribution behind the buildings which are to be used as barriers against traffic noise.

Poster — form papers

R. J. BERACHA, Prediction of the traffic noise in residential areas.

H. S. AL-SAMARRAI, D. M. WATERS, Noise from interrupted traffic flows.

S. CZARNECKI, E. SŁAWIŃSKA, J. SZUBA, Highway noise control inside a residence by using the barriers.

D. C. HOTHERSALL, R. R. K. JONES, Computer simulation of road traffic noise in the region of roundabouts.

T. E. GRANQUIST, Development of the national traffic noise abatement programme in Norway.

R. CASTELLI, S. MANICARDI, Some aspects of noise emissions of Grand Prix motorcycles.

Y. Ugai, S. Suzuki, Infrasound from centrifugal fans used in subway ventilating equipment, and infrasound control.

J. Miazga, K. Janicka, Road traffic noise on major highway outlets of Warsaw.

Section I — Preventive shipboard noise control

Chairmen: J. H. Ødegaard-Jensen, J. Verheij.

Contributed papers

J. H. Janssen, Simultaneous shipboard noise-and-vibration annoyance rating.

J. Buiten, H. Aarsten, Simplified method for predicting sound level A in accomodation spaces aboard sea-going motorships.

J. L. Goldberg, N. H. Clark, J. A. Lund, A study of noise-generating mechanisms in fast 16 – metre patrol craft.

E. Brubakk, Noise control on small ships – with special reference to supply ships and tugs.

J. Ødegaard-Jensen, H. Holm, Noise reduction in the accomodation of ships by means of "constrained layer" damping.

S. Weyna, Vibro-acoustic tests of ship cabins ceilings.

M. J. A. de Regt, Transfer of structure-borne sound to ship's cabin.

K. Fukuzawa, C. Yasuda, Studies on structure-borne sound in ships.

J. W. Verheij, Airborne sound transmission via the cavity under a resiliently mounted ship diesel engine.

M. Ohlrich, Response of bedplates and the transmission of structureborn sound on ships.

A. Nilsson, Propeller induced hull plate vibrations.

E. Kozaczka, Investigation of the noise transmission into ship hull produced by a ship propeller.

U. P. Tyvand, B. Persson, Prediction of noise from a cavitating propeller.

A. Løvik, Scaling of propeller cavitation noise.

A. de Bruijn, Acoustic source strength of propeller cavitation.

H. P. Steenhoek, Resilient mounting of medium speed diesel engines in ships.

Poster-form papers

A. Rauch, Flexible approach to analysis of experimental model.

Section J — Aircraft and airport noise

Chairmen: J. B. Large, J. Šulc.

J. B. Large, M. E. House, The status of airport noise prediction.

K. Matschat, E.-A. Müller, How to estimate the differences between two aircraft noise indices without full knowledge of the leveltime-history.

Contributed papers

D. Commins, J.-P. Sirieys, Short term equivalent levels for the evaluation of aircraft noise.

G. Nishinomiya, F. Suzuki, F. Sasaki, Aircraft noise identification system by correlation technique.

D. E. Bishop, Current aircraft modelling concepts for developing aircraft noise exposure contours.

B. N. LOHANI, H. KARWARZ, W. PHOTHIPICHITR, Aircraft noise at Bangkok international airport.

G. H. VULKAN, Aircraft noise monitoring for local authorities.

Poster-form papers

T. ONO, M. OKUDA, I. ONO, N. HAYASHI, G. NISHINOMIYA, F. SUZUKI, F. SASAKI, Aircraft noise monitoring system with identification by correlation technique.

J. ŠULC, J. HOFR, Noise and pressure fluctuations on the fuselage of light propeller driven aircraft.

A. RUDIUK, Problem of repeatability of results during the helicopter noise measurements.

Section K — Environmental noise

Chairmen: Z. MAEKAWA, W. STRASZEWICZ.

Invited papers

A. KOMORN, A general method for objective description of noise environments.

P. M. da SILVA, Urban noise, noise prediction models.

M. AOKI, K. KONISHI, Practical technique on the *m*-sequence correlation method applied to measurements of long distant noise propagation.

K. KONISHI, M. AOKI, Measurement of the long distant noise propagation over ground.

Z. MAEKAWA, M. MORIMOTO, Measurement of the long distant noise propagation over sea.

R. SEZNEC, Integral equation technique for computing sound attenuation by wide barriers.

D. HAHAULT, G. CORSAIN, P. FILIPPI, Ground effect: a new theoretical approach and experimental results.

B. MERIEL, Noise in a French medium size town – an example of Blois.

D. C. STEVENSON, R. H. PALMER, Background noise level variations in a suburban environment.

H. ONO, S. SAITO, H. FUKUDA, K. MIZOI, K. NAKASHITA, Development of a vibratory pick-up type microphone and communication devices for daily use.

R. KÜRER, V. IRMER, Noise reduction by legislative incentives demonstrated by the lawn mower ordinance of the Federal Republic of Germany.

Poster-form papers

S. ISRAELSSON, C. LARSSON, Metereological data for determination of sound propagation in the atmospheric surface layer.

M. MUNTEANU, Estimation of noise annoyance effect with a new noise pollution level LNP index.

R. J. KUCHARSKI, M. BAREŃSKI, Projected evidence system for main sources of outside noise.

E. G. TERRY, Measurements in low noise environments.

J. C. TUKKER, Noise control alternatives for Zandvoort racing circuit.

J. ABŁAMOWICZ-POTAPOWICZ, Protection of sanatorium and hospital buildigs from noise made by free-standing generating sets.

Section L — Personal noise exposure

Chairman: E. ZWICKER.

Invited papers

S. NAMBA, S. KUWANO, An experimental study on the relation between long-term annoyance and instantaneous judgment of level-fluctuating sounds.

S. Kono, T. Sone, T. Nimura, A study on personal noise exposure in three cities different in population.
P. V. Brüel, Limits for intrasound and ultrasound in factories.

Contributed papers
M. Kumagai, T. Sone, T. Nimura, A study on the loudness of impact sound.
M. M. Hawkins, Human response to domestic appliance sound.
M. Vallet, J. M. Abramowitch, J. Lambert, Subjective effect of a roadside barrier: a case study at L'Hay Les Roses.
V. Velas, K. Přerovsky, Hygienic aspects of the control of noise in railway transport.
J. Starck, B. Godenhielm, K. Perkiö, Effect of helmet-liners on the attenuation of earmuffs.

Poster-form papers
T. Poulsen, Measurements on hearing protectors.
D. Trynkowska, R. Michalski, Attenuation and classification of ear protectors.
W. Sułkowski, A. Lipowczan, Impulse noise induced hearing loss in drop forge operators.

Section M — Noise control engineering in buildings

Chairmen: A. Illenyi, J. Sadowski.

Invited paper
D. N. Keast, Acoustic location of air-infiltration openings in buildings.
E. E. Unger, L. E. Wittig, A. Paolillo, Propagation of vibrations and noise from New York subway tunnels into nearby buildings.
S. Czarnecki, A. Rakowski, A. Różycki, J. Smurzyński, Reduction of piano noise in living apartments.

Special punel session

Education in noise control engineering.

Stefan Czarnecki (Warsaw)

XXVI Open Seminar on Acoustics

Oleśnica 17-25 September, 1979

XXVI Open Seminar on Acoustics was held at Oleśnica near Wrocław on September 17-21, 1979. The conference was sponsored by Committee on Acoustics of Polish Academy of Sciences, Wrocław Division of Polish Acoustical Society and Institute of Telecommunication and Acoustics of Wrocław Technical University. The chairman of the Organizing Committee was Prof. Zbigniew Żyszkowski. The conference was attended by 182 participants, including 8 foreign guests.

The sessions of the Seminar were held in three sections:

A. Speech, telecommunication, and physical acoustics, electroacoustics and signal processing.

B. Musical, psychological, building, and interior acoustics, and problems of noise, vibration and infrasound.

C. Quantum and molecular acoustics, hydroacoustics and ultrasound technique.

3 plenary papers concerning speech acoustics and 102 papers in sections were delivered. Moreover, there were 4 round table sessions on the problems of computer analysis of acoustic signals, methods of contemporary psychoacoustics, acoustooptics and holography, and the current problems of shaping the acoustic environment.

Andrzej Dobrucki (Wrocław)