ARCHIVES OF ACOUSTICS
6, 1, 3-12 (1981)

THE ACOUSTICS OF THE EARLY-CHRISTIAN MONUMENTS OF THESSALONIKI

EMMANUEL G. TZEKAKIS

School of Technology, University of Thessaloniki, Greece

1. Introduction

This paper is the last of a series of four papers concerning the acoustics of historic monuments and churches of Thessaloniki. All monuments dealt with in these series have been built inside the very center of the old walled part of the town.

Although Thessaloniki has suffered many destructions, a number of fourteen buildings with some interest for the room acoustics, are to be found today witnessing her over two thousand year old history. Of these buildings five come from the early-Christian era, 4th to 7th century, and nine from the Byzantine era, 11th to 14th century.

The present paper reports on data accumulated on the first group of five early-Christian monuments, while data on the other group is at the moment already in print [4].

After the descriptions of the monuments and the measurement results, a discussion follows, comparing the main characteristics of both groups.

2. Descriptions

The early-Christian monuments described below are St. Demetrius, St. Sophia, Acheiropoietos, Hosios David and St. George. All of them except Hosios David have exceptionally large volumes and share many characteristics. St. Demetrius is an early-Christian basilica built during the first quarter of the 5th century. The building was destroyed by fire and rebuilt once during the 7th century. A little after 1917 a new fire destroyed the building as well as the largest part of the town and so today's form is the result of large scale reconstructions completed around 1950.

The basilica of St. Demetrius is a large rectangular building measuring 57 ×33 m² with a height at the center axis of about 18 m. It is divided by four colonnades into five aisles. There is an upper storey above each of the four side aisles and the narthex. The nave and each aisle have seperate tilted roofs.

The volume of the church is 22100 m³, while the total internal surface is 6500 m². This gives a mean-free-path value of 13.60 m. The floor and a large percentage of the wall surfaces are covered with marble slabs. The rest of the walls is covered with mosaic and fresco paintings. The roof is made of reinforced concrete, imitating the old destroyed wooden construction, without an underceiling. The plans of the church are presented in Fig. 1.

St. Sophia is a scaled down imitation of the famous St. Sophia of Constantinople built during the 7th century. It has a transitional form between domed basilica and domed cruciform church. Also destroyed by fire in 1890, the building was reconstructed between 1907 and 1910.

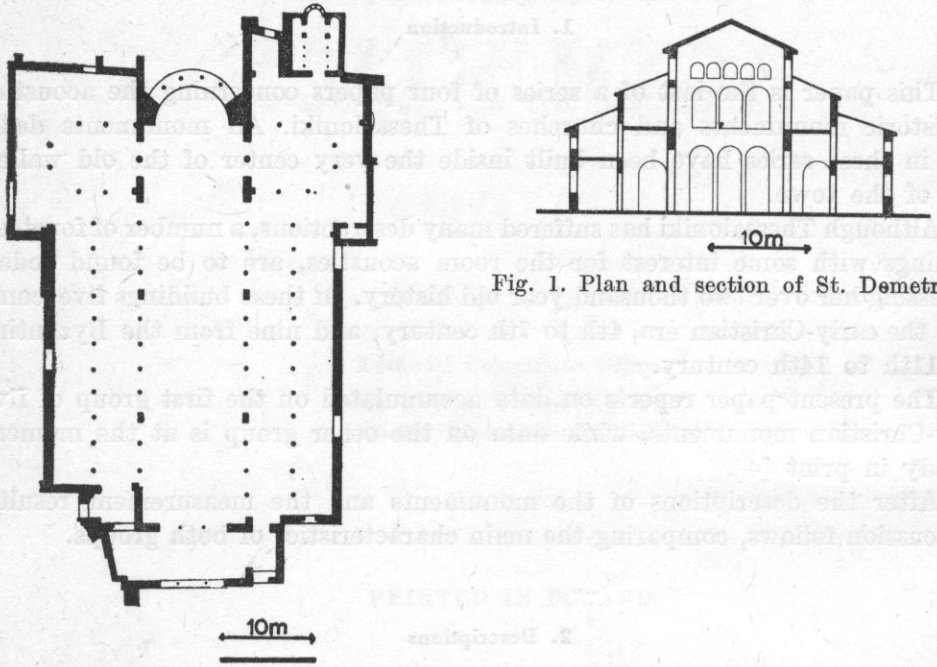

Fig. 1. Plan and section of St. Demetrius

The interior of the church is divided into three aisles by two rows of columns alternating with piers. The church of St. Sophia is square as in the domed cruciform type. In the center of the nave is a dome resting on four barrel vaults of which the western is wider.

The barrel vaults of the dome rest on thick piers. The aisles are roofed at a lower level by uninterrupted vaults and the narthex by smaller domes. Above the narthex and both aisles there are galleries.

The church has a volume of 15250 m³. A large percentage of the floor and wall surfaces is covered with marble, while the largest part of the roof and the rest of the walls are covered with mosaic and fresco paintings.

The plans of the church are presented in Fig. 2.

Acheiropoietos is another early-Christian basilica, built during the second quarter of the 5th century, which was never destroyed or altered. The building is large and rectangular with dimensions comparable to those of St. Demetrius. It is divided by two colonnades into three aisles. Above each of the two side aisles there is an upper-storey.

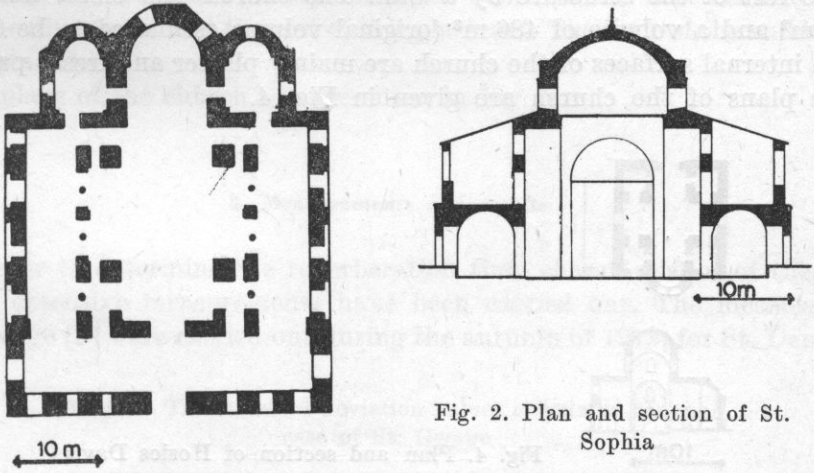

Fig. 2. Plan and section of St. Sophia

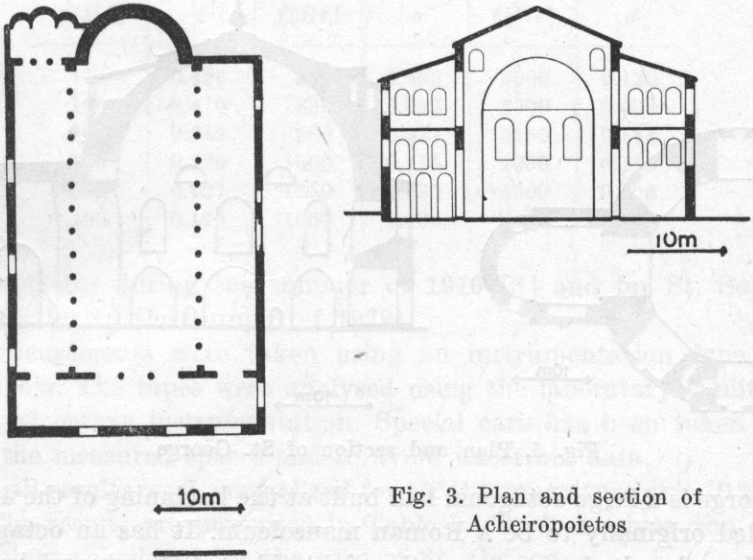

Fig. 3. Plan and section of Acheiropoietos

The volume of the monument is 19.250 m³ and the total internal surface is 5900 m², which gives a mean-free-path value of 1305 m. The floor area is

made of large marble blocks, and the wall surfaces are mainly of plaster. The roof of the church is a wooden tilted construction without an underceiling.

The plans of the church are presented in Fig. 3.

Hosios David is a small domed cruciform church, belonging to the old monastery of Latomou, built at the end of the 5th century. The original form of the church was partly changed during the Turkish age. The dome of the church was covered by a wooden underceiling which has existed until today. The entrance was transfered from the west to the south side, and the western part of the church, part of which has disappeared altogether, has been seperated from the rest of the structure by a wall. The church has small dimensions 11×8.5 m^2 and a volume of 486 m^3 (original volume estimated to be 650 m^3).

The internal surfaces of the church are mainly plaster and fresco paintings.

The plans of the church are given in Fig. 4.

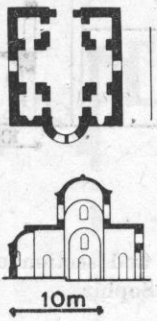

Fig. 4. Plan and section of Hosios Davis

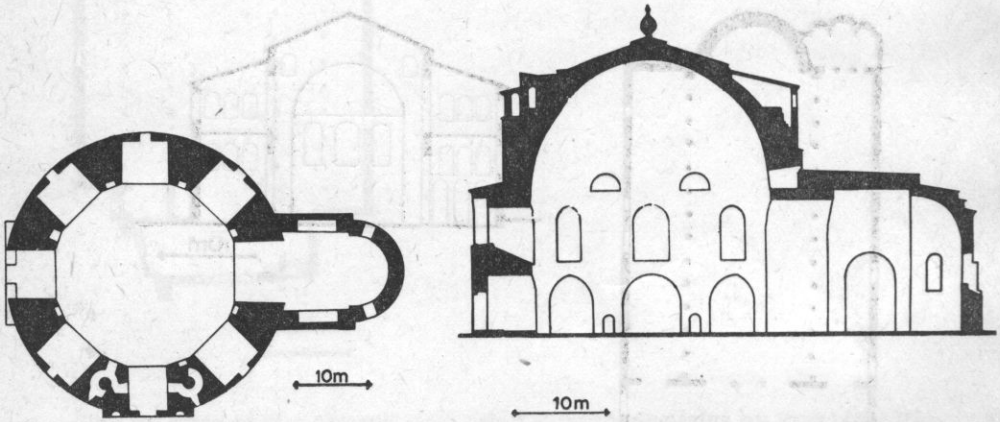

Fig. 5. Plan and section of St. George

St. George is a large octagonal hall built at the beginning of the 4th century and intended originally to be a Roman mausoleum. It has an octagonal form and large barrelvaulted recesses at the eight sides of the ground level. Above them there are eight similar but smaller recesses with windows. A little higher at the base of the dome there is a ring of even smaller recesses.

The central space has a diameter of about 24 m and a total height of about 26.50 m although the original floor level was more than 2 m lower than it is today. The eight recesses that surround the central space have an average opening of 6.20 m and a height of 6 m each. The hall has undergone two Byzantine transformations, the second one during the 10th century, which transformed it into a Byzantine church. The eastern recess was made a sanctuary with a width of 8 m, a height of 13.50 m and a depth of 20 m.

The hall's total volume today is 15340 m^3 out of which 10710 m^3 belongs to the central space. The largest part of the surfaces of the hall is made of glazed brick, interrupted by strips of small stone blocks. The surface of the dome is mostly mosaic and the rest of it is plastered. The floor of the hall is made of terracota tiles.

The plans of the church are presented in Fig. 5.

3. Measurements and results

In order to determine the reverberation time characteristics of these monuments, extensive measurements have been carried out. The measurements for St. George [2] were carried out during the autumn of 1973, for St. Demetrius

Table 1. The standard deviation values calculated for the case of St. George

f[Hz]	σ	f[Hz]	σ	f[Hz]	σ
125	0.426	500	0.203	2000	0.120
160	0.410	630	0.093	2500	0.098
200	0.343	800	0.137	3150	0.144
250	0.220	1000	0.115	4000	0.110
315	0.227	1250	0.180	5000	0.106
400	0.176	1600	0.166	6300	0.110

and Acheiropoietos during the summer of 1976 [3] and for St. Sophia and Hosios David during the autumn of 1978.

The measurements were taken using an instrumentation tape recorder and pistol shots. The tapes were analysed using the laboratory facilities, with standard third octave instrumentation. Special care has been taken to cover thoroughly the measured spaces and to avoid uncertain data.

Finally all results were normalized for 15°C temperature and 70 % relative humidity. The results are presented in Table 2. Table 1 gives the standard deviation values calculated for the case of St. George.

Examples of decay curves at various frequencies in two points are given in Fig. 6.

Table 2. Analytical measurement results of the five monuments

Church	Frequency [Hz]																	
	125	160	200	250	315	400	500	630	800	1000	1250	1600	2000	2500	3150	4000	5000	6300
St. Demetrius	6.50	7.00	6.90	7.20	6.00	5.90	5.90	5.70	5.30	4.90	4.50	4.00	3.40	2.80	2.30	1.85	1.50	1.10
St. Sophia	3.25	3.25	4.10	4.10	4.10	3.90	3.90	3.90	3.50	3.35	3.10	2.90	2.65	2.30	2.05	1.75	1.45	—
Acheiropoietos	5.10	4.80	4.40	3.90	3.80	3.70	3.70	3.70	3.50	3.10	2.80	2.60	2.30	2.00	1.75	1.50	1.25	—
Hosios David	0.82	0.73	0.64	0.66	0.65	0.64	0.63	0.65	0.62	0.59	0.56	0.58	0.54	0.53	—	—	—	—
St. George	5.30	5.15	5.05	5.00	4.90	4.90	4.90	4.95	4.75	4.60	4.40	4.25	4.10	4.00	3.80	3.35	2.70	2.20

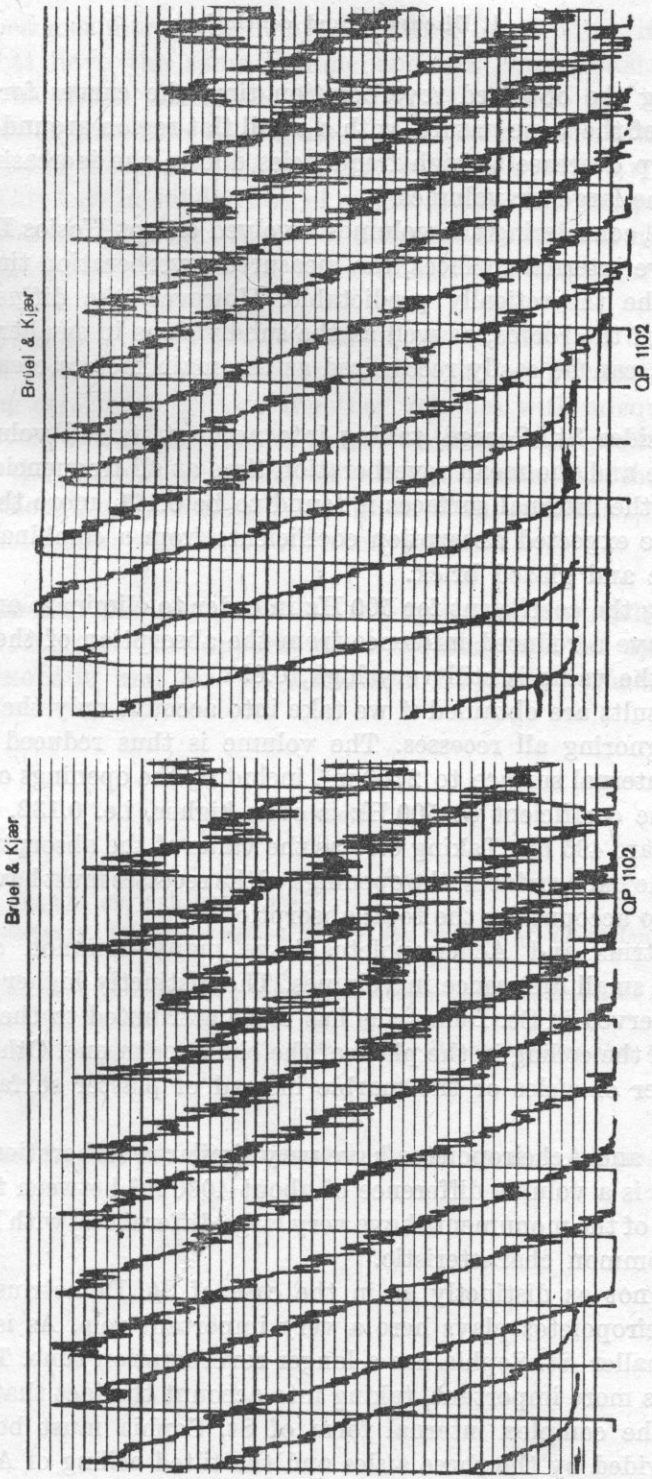

Fig. 6. Decay curves for St. George

4. Discussion and conclusions

Considering the obtained reverberation time the curves for all churches are essentially of the same nature, with a small flat region around mid-frequencies and a sharp decrease at high frequencies due to the increasing absorptive properties of the large air volumes.

Further on, considering the volumes involved except Hosios David and the totally reflective internal surfaces, the measured reverberation times are about one third of the theoretically predictable. However, the diffusion provided by the shapes of the rooms, as well as the subdivisions in smaller, acoustically coupled spaces, can be easily recognized as the main factors creating this discrepancy.

If we consider St. George, taking into account its full volume, its total internal surface and the mean reverberation time at all frequencies as 4.15 sec, the mean $\bar{a}$ of the internal surfaces is found to be 0.111, more than two times higher than the expected absorption coefficients from a combination of stone, marble, mosaic and glazed brick.

Calculating the coefficient for 500 Hz in order to eliminate any air absorption, and to have a reduced influence from the absorption of the very limited glass surface, the value is still very high, 0.094.

Similar results are obtained if we take into account only the central space of the room ignoring all recesses. The volume is thus reduced to 10710 m^3 and the total internal surface to 2700 m^2, including the openings of the recesses. In this case the coefficient at 500 Hz is even higher, i.e. 0.132. The openings of the recesses are 495 m^2. Taking 0.05 as the value of the absorption coefficient at 500 Hz for the hard surfaces, the openings of the recesses must have a coefficient equal to 0.50 to account for the total absorption.

St. Demetrius and Acheiropoietos have many common characteristics in form, and a small difference in volumes. The distinctly higher reverberation time value observed in St. Demetrius has to be attributed to the new concrete construction of the ceiling in the place of the old wooden one. Other differences, e.g. the number of aisles or the marble instead of plaster surfaces play only a minor role.

St. Sophia and Acheiropoietos have very similar reverberation time curves, although there is a volume difference of about 4000 m^3 between them.

The forms of the monuments have very large differences, with large diffusion as the only common characteristic.

Although not as distinctly as in the case of St. Demetrius, the wooden ceiling of Acheiropoietos plays here a very important role. As is shown in the results the smaller St. Sophia has a longer reverberation time. The role of the ceiling becomes more important, taking into account the fact that the diffusion provided by the complex internal form of St. Sophia must be much larger than that provided by the three aisles and the tilted ceiling of Acheiropoietos.

This fact becomes more evident by comparing St. Sophia with St. George. Both monuments have the same volume and are constructed from similar materials with no noted difference. However, St. George has a one second longer reverberation time than St. Sophia, a fact that must be attributed to the very large volume of the central space of St. George, a situation not to be found in St. Sophia. This lack in subdivision of spaces is the only recognizable factor. It seems also probable that the absence of diffusing elements at higher frequencies in St. George is the cause of the noted linearity of the reverberation time curve.

In spite of the differences between them, these 1500-year old monuments all point to the fact, which is also supported by earlier publications [1], that rooms consisting of smaller coupled together volumes with adequate diffusing elements have very low reverberation times, which are not predictable in calculation using the known reverberation time formulae. Another interesting point is that although not acoustically designed all these rooms have reverberation time characteristics well suited for their purposes.

Aknowledgments. The author wishes to express his gratitude to Professor Dr. J. TRIANTAFILLIDIS who supports the laboratory and his appreciation to Dr. G. PAPANIKOLAOU and Mr. M. LAZARIDIS for their valuable assistance.

References

[1] R. S. SHANKLAND, H. K. SHANKLAND, *Acoustics of St. Peter's and patriarchal basilicas in Rome*, JASA, **50**, 389–396 (1971).

[2] E. G. TZEKAKIS, *The reverberation time of the rotunda of Thessaloniki*, JASA, **57**, 1207–1209 (1975).

[3] E. G. TZEKAKIS, *Reverberation times of two early-Christian basilicas of Thessaloniki*, Proccedings 9th I.C.A., Madrid 1977 C 16.

[4] E. G. TZEKAKIS, *Data on the acoustics of the Byzantine churches of Thessaloniki*, Acustica, **43**, 275–279 (1979).

Received on May 25, 1979; revised version on 15 september 1980.

ARCHIVES OF ACOUSTICS
6, 1, 13-22 (1981)

PROBLEMS IN RECORDINGS FOR LISTENING EVALUATION OF THE QUALITY OF VIOLINS

ANDRZEJ SASIN, ANTONI KARUŻAS, TOMASZ ŁĘTOWSKI

Sound Recording Department, Chopin Academy of Music (00-368 Warszawa, ul. Okólnik 2)

This paper considers the possibility of performing a comparative evaluation of violin sound quality, based on sound recordings. The investigations dealt with the effect of the microphone techniques – mono, stereo *AB* and stereo *XY* – and the acoustics of the room, on the usefulness of the recordings made in given conditions for the evaluation of the sound quality of violins. The paper presents the investigation results and a comparison of these with the results of the "live" sound quality assessment performed by expert violinists.

1. Introduction

The ultimate test of the quality of any musical instrument is an evaluation made on the basis of auditory sensations. Investigations aimed at such an auditory evaluation can have the following forms;

(a) evaluation by an instrument player who can directly get to know an instrument under test, i.e. direct evaluation,

(b) evaluation of the sound quality of an instrument made on the basis of sound recordings, i.e. indirect evaluation.

In the previous investigations of the sound quality of musical instruments the so called listening tests containing music recorded for the purpose of underlining the characteristics of instruments investigated, have been widely used. It seems that for the evaluation based on such tests to be considered valid, it is necessary to pay more attention to the methods of sound recording being used. The aim of the present paper is to answer the question whether such a recording is possible that could be an index of quality of a given instrument, and possibly whether a set of principles could be established, which could be used in making recordings for the listening tests of the sound quality of instru-

ments. The investigation was carried out using the violin as the specimen instrument. The violin was selected because

(a) there is a great variety of relatively easily available instruments,

(b) there is a large interest in the investigations of the quality of the violin,

(c) the instrument is portable.

The investigations were carried out on three instruments. The sound qualities of all the instruments were previously evaluated by the direct method by three outstanding soloists employed at the Stringed Instruments Department of the Chopin Academy of Music in Warsaw. The results of this evaluation can be presented in the form of the following rank series:

1. Violin no. 1, unknown violin-maker, an instrument of very good quality;
2. Violin no. 2, made by NEUNERT, an instrument of good quality;
3. Violin no. 3, a mass made specimen, of very bad quality.

The experts at the same time pointed to the considerable difference between violins nos. 1 and 2 and violin no. 3.

The results of the experts' evaluation were unknown both to the workers carrying out the experiment and the listeners until the final results of the investigations were obtained.

2. Procedure

The first stage of the investigation consisted in a listening evaluation of selected instruments on the basis of sound recordings made in different ways. Established as a result of various discussions and tests, the sound material contained:

1. diatonic gamuts of one octave performed on each string;
2. three chords; *G* major, *D* major, *C* major, one following the other,
3. a part of Ysaye's Sonata;
4. a part of J. S. Bach's Chaconna.

All the musical pieces were recorded on the *mf* – *f* dynamic level.

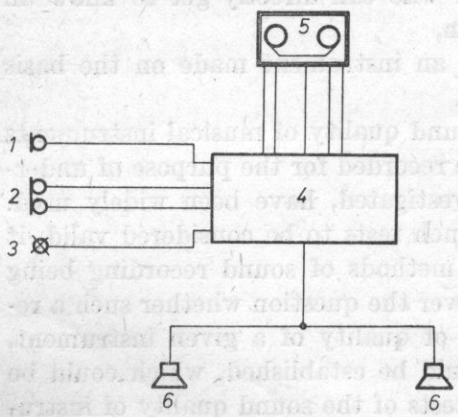

Fig. 1. A block diagram of the electroacoustic system used in the recordings

1 – a mono *U* – 67 Neumann microphone, *2* – two *U* – 67 microphones in the system *AB*, *3* – a stereo *SM* – 69 Neumann microphone in the system *XY*, *4* – a Siemens control panel, *5* – an eight-track Studer tape recorder, *6* – *ZG* – 60 *C* loudspeaker columns

The material mentioned above was recorded by three microphone techniques, i.e. mono, XY, and AB. A comparison of sound recordings obtained by these techniques was expected to provide an answer to the question whether the technique used does not matter in terms of sound quality evaluation, or otherwise whether the differences due to the different principles of microphone techniques are so essential with regard to this type of investigation that one of them should be preferred.

A block diagram of the system used for the implementation and monitoring of the recordings is shown in Fig. 1.

By using the multitrack technique a single performance of a musical piece performed on one instrument could be recorded at the same time by three independent recording systems. Thus employing one person only who reproduced the musical piece successively on three instruments, the minimum influence of external factors of the recording was assured.

All the recordings were made at studio $S-1$ and the concert hall of the Chopin Academy of Music in Warsaw. These rooms are distinctly different in terms of volume and reverberation time. Their parameters are the following:

– at studio $S-1$ the volume is 1080 m^3 and the reverberation time is 1.1s,

– at the concert hall the volume is 6000 m^3 and the reverberation time is 1.8s.

In both cases the reverberation time is given as a mean for the frequency range from 500 to 2000 Hz.

Thus the investigations also included the problem of the influence of the acoustics of the room both on the recording and on the sound quality of the instrument.

The final recordings were preceded by a selection of the position of the player in a given interior and of the optimum arrangement of microphones in each of the system used. In each case such a position of microphones was assumed as optimum that to a highest degree assured the condition of the similarity of a recording to the original sound of the violin recorded. This choice was made on the basis of a number of preliminary experiments.

The music material was recorded in one session for each of the rooms. In order to evaluate the quality of a given violin on the basis of the material recorded, listening tests were carried out. In all the tests the method of pairs comparison was employed [3].

The durations of the individual stimuli were the following:

1. diatonical gamuts – 7s (on one string);
2. chords – 4s;
3. a part of Isaye's sonata – 8s.

The test set consisted of 36 trials of the $A-B$ type arranged randomly and balanced in terms of succession. One-second intervals between stimuli in a trial and two and half second intervals between trials were used. The intervals between trials were intended to give the expert enough time to put his

Table 1

No.	mono						*XY*						*AB*					
	violin						violin						violin					
	I		II		III		I		II		III		I		II		III	
	S−1	*SK*	*S*−1	*SK*	*S*−1	*SK*	*S*−1	*SK*	*S*−1	*SK*	*S*−1	*SK*	*S*−1	*SK*	*S*−1	*SK*	*S*−1	*SK*
1	19	20	14	10	3	6	12	18	16	12	8	6	16	16	14	13	6	7
2	19	18	10	12	7	6	15	11	8	13	13	12	7	14	10	13	19	9
3	19	19	11	13	6	4	19	17	12	11	5	8	17	18	13	14	6	4
4	19	20	12	11	5	5	19	17	12	13	5	6	16	15	13	13	7	8
5	16	23	14	11	6	2	17	15	8	13	11	8	13	13	12	13	11	5
6	17	23	13	11	6	2	14	18	12	13	10	5	14	17	12	16	10	3
7	19	20	10	12	7	4	13	18	15	15	8	5	14	19	13	13	9	4
8	16	20	13	12	7	4	12	17	10	10	14	9	15	18	10	13	1	5
9	17	19	14	11	5	6	10	16	15	13	11	7	17	15	15	16	4	5
10	18	20	12	11	6	5	16	21	13	8	7	7	15	20	14	12	7	4
11	16	20	13	11	7	5	13	20	12	10	11	6	14	18	10	12	12	6
12	16	18	12	13	8	5	12	18	10	11	14	7	15	13	10	20	11	3
13	15	18	11	12	10	6	13	12	12	14	11	10	14	15	12	12	10	9
14	15	19	13	13	8	4	13	14	12	14	11	8	13	12	16	15	7	9
15	16	18	12	13	8	5	15	15	11	9	10	12	14	12	10	15	12	9
16	11	17	11	15	14	4	16	10	15	14	5	12	14	15	14	16	8	5
17	11	18	17	10	8	8	15	15	14	14	7	7	12	13	15	13	9	10
18	11	16	12	12	13	8	12	14	13	13	11	9	16	18	9	14	11	4
19	14	22	12	8	10	6	11	16	16	13	9	7	13	18	16	14	7	4
20	16	20	15	8	5	8	10	11	16	16	10	9	13	16	13	14	10	6
	320	388	251	229	149	103	277	313	252	249	191	160	282	315	251	281	177	119

S−1 – studio *S*−1, *SK* – concert hall of Chopin Academy of Music, I, II, III – violins.

mark in the estimation sheet. The whole test consisted of three sets of trials, each containing the same sound material recorded in one of the three selected microphone systems. Each set was intended to give an evaluation of a violin on the basis of recordings made in a given microphone system. The total duration of one set did not exceed 18′. The tests were performed on the basis of a 20-member expert group consisting of students from higher grades in the departments of sound engineering, of instruments (violins section) and of theory and composition (sections of theory and conducting). The listening test were carried out at the concert hall and at studio $S-1$ of the Chopin Academy of Music, both cases involving the whole expert group. The loudness level was 80 phons. 20′ rest intervals were introduced between the test sets. In order to give a clear distinction between instruments investigated, a preferential evaluation: worse-better, was used in the test [3]. The results of the experiment are shown in Table 1. It presents numbers of points given to a particular instrument by each expert, according to the instruction preceding the test. The evaluation marking was the following: choosing an instrument in a given pair as better — 1 point, lack of choice — 0 points. The table gives the sum totals of points assigned to the individual instruments in the comparison made by experts.

The points given to the particular violins and shown in Table 1, were the basis for a statistical analysis of the results, which used

(a) the Kolomogorov test — the investigation of the normalcy of the distribution of the results obtained [1].

(b) the t-Student test — the investigation of the statistical significance of differences [2].

The results of the Kolomogorov test permitted a hypothesis that the distribution is normal for the estimates of the whole listener group to be taken. The above investigation was performed in order to detect the possible cases of multimodal distributions or a rectangular distribution, which would suggest that it is impossible to obtain an "objectivized" qualitative judgment of the sound quality of violins under investigation.

The results from the t-Student test obtained for the individual pairs of instruments investigated with the three microphone systems and for the two rooms in which the recordings were made, are given in Table 2.

According to the rule frequently used in mathematical statistics, the following significance levels were taken for the hypotheses assumed [3]:

1. a highly significant result — the probability of the validity of the null hypothesis, $p \leqslant 0.1\,\%$;

2. a significant result — the probability of the validity of the null hypothesis within the limits from 0.1 to 1.0 %;

3. a probably significant result — the probability of the validity of the null hypothesis within the limits from 1.0-5.0 %;

4. an insignificant result — the probability of the validity of the null hypothesis $p \geqslant 5\,\%$.

The null hypothesis was the statement that the investigation method used does not permit to distinguish qualitatively between instruments under investigation.

Table 2

Violin		Technique		
		mono	XY	AB
I–II	concert hall	11.362	2.232	2.321
	studio $S-1$	4.232	1.277	2.867
I–III	concert hall	20.225	7.248	10.332
	studio $S-1$	7.423	4.011	4.164
II–III	concert hall	9.703	6.429	9.460
	studio $S-1$	6.158	2.939	2.783

The values of the t-Student test, which evidence a given significance level of differences, as read from the statistical tables for the t-Student distribution with $S = N-1$ degrees of freedom, where N is the number of listeners ($N = 20$), are the following [4]:

(1) $t \geqslant 3.883$ – highly significant;
(2) $t \geqslant 2.861$ – significant;
(3) $t \geqslant 1.729$ – probably significant;
(4) $t \geqslant 1.729$ – insignificant.

The values of the t-Student test obtained from the experiment are shown in Table 2. A comparison of the empirical values with the critical t-Student values taken from the tables shows a real difference between instruments investigated. Only once, i.e. the difference in evaluation between violin I and violin II, based on the recording by the XY technique at studio $S-1$, it appeared to be insignificant. A comparison of the empirical values of the t-Student test obtained for the recordings made at studio $S-1$ and at the concert hall, using the same microphone system, shows a significant difference between the results obtained. This suggests an effect of the acoustics of the room on the sound quality in the recording.

Since the values of the t-Student test obtained as a result of the above experiment suggest a distinction between instruments in each of the three systems used, the investigators determined to carry out another test in order to answer the question which of the sound recordings made in the three microphone systems was closest in terms of sound quality to the "live" performance (the assessment of the fidelity of a recording). A piece of J. S. Bach's Chaconna which was recorded previously was used as sound material. The test was based on the $A-W-B$ method with a presentation of the standard W ("live") performance between test stimuli. Each sound sample lasted 9 seconds. In view of the "live" W presentation it was difficult to take a precise duration of the intervals between the stimuli $A-W-B$ in the test. The real intervals between

the stimuli lasted about 2 seconds, while the interval between the test trials lasted 3 seconds. The whole test was divided into three trial sets, each consisting of six randomly arranged trials (three trials of the $A-W-B$ type and three trials of the $B-W-A$ type). The total duration of one set did not exceed 6′. The aim of each task set was to select such a microphone system that would give in its recording the most faithful representation of the quality of the individual instruments investigated. The test was carried out at the concert hall of the Chopin Academy of Music for a twelve-member expert group. The performer was situated half-way between loudspeakers.

The principle of point giving was the same as in the first test: choosing a given microphone system — 1 point, lack of choice — 0 points. Table 3 shows the sum total of points assigned to a given microphone system by each expert in each trial set. The principle of point summation was the same as in Table 1. The data in this table were then the basis for statistical processing of the results in the same way as the statistical elaboration of the results of the previous experiment.

Table 3

No.	Violin I			Violin II			Violin III		
	mono	*XY*	*AB*	mono	*XY*	*AB*	mono	*XY*	*AB*
1	4	0	2	3	1	2	4	1	1
2	4	0	2	3	0	3	2	3	1
3	3	3	0	3	1	2	1	1	4
4	3	1	2	3	1	2	1	2	3
5	3	2	1	2	3	1	4	1	1
6	3	1	2	2	1	3	3	2	1
7	3	0	3	2	3	1	1	2	3
8	2	3	1	2	2	2	2	1	3
9	2	2	2	2	2	2	3	2	1
10	2	2	2	2	1	3	2	1	3
11	2	0	4	2	2	2	0	2	4
12	1	2	3	1	2	3	0	2	4
	32	16	24	27	19	26	23	20	29

Under the assumption of significance levels according to the same principle as in the first test, there are now four critical values of the t-Student distribution read from the statistical tables for $S = N-1$ degrees of freedom (where $N = 24$, i.e. 12 listeners times two successions of presentation);

(1) $t \geqslant 3.496$ — highly significant;
(2) $t \geqslant 2.500$ — significant;
(3) $t \geqslant 1.714$ — probably significant;
(4) $t \geqslant 1.714$ — insignificant.

The empirical values of the t-Student test obtained for the individual pairs of microphone systems in the case of each instrument are given in Table 4.

Table 4

Technique	Violin		
	I	II	III
mono – XY	1.796	1.623	0.495
mono – AB	1.430	0.266	0.640
AB – XY	1.100	1.277	1.650

The results obtained show that the fidelity of a recording (similarity ot the original) was, according to the listeners, highest:

(a) in the case of violin I when the mono system was used;

(b) in the case of violin II, similar when the mono and the stereo AB systems were used;

(c) in the case of violin III when the stereo AB system was used.

In the case of all the three violins the representations in the stereo XY technique were found to be least faithful.

3. Discussion and conclusions

The results obtained in both experiments and a statistical analysis of them lead to the following conclusions and remarks.

The results of the evaluation of the sound quality of selected violions, made using the listening test, coincided with the results of direct evaluation made by experts (the members of the Department of Bow Instruments). The recordings made in each of the microphone systems used permitted a good estimation of differences between the instruments investigated. The credibility of the estimates made by the expert group, that is visible from a statistical processing of the results, was very high. A comparison of the values of the t-Student test obtained for the qualitative evaluation of recordings made in the particular microphone systems (Table 2) suggests the conclusion that difference between instruments was most conspicuous in the case of a recording made using a mono microphone, and successively in a recording made using the stereo techniques AB and XY. A comparison of the values of the t-Student test obtained for recordings made at studio $S-1$ and at the concert hall seems to indicate that each microphone system responded to a higher degree to the differences between instruments at the concert hall, i.e. a room with better acoustic conditions. An exception here is the t-Student test value for the comparison of violin I and violin II, based on recordings in the AB technique.

Since the aim of the investigations was not to solve the problem of the effect of the acoustics of a room, but only to show whether this effect occurs at all in terms of evaluation of an instrument recorded, it is difficult to draw speci-

fic conclusions on the basis of the results obtained. It seems useful, therefore, to perform investigations of the problem in nearest future, which may shed more light on it. This demand also applies to other variable conditions of the listening technique and to microphone types used in recording.

It follows from the results obtained from the second test that it was most troublesome for the listener group to select that microphone system which would most faithfully represent an instrument. The present results permit, however, the following conclusions to be drawn:

1. in the case of good instruments the mono technique assures highest fidelity;

2. in the case of poor instruments the stereo AB technique assures highest fidelity;

3. the stereo XY technique appears to be least useful in the case of auditory evaluation of the sound quality of the violin on the basis of microphone tests.

The present results also permit a conclusion that the better the violin is, the greater differences occur between recordings obtained by means of different microphone techniques. This leads to a further hypothesis that the poorer an instrument is, the less effect the recording technique has on the representation of the sound quality of an object investigated. This hypothesis requires, however, further and wider experimental evidence.

In summary it can be stated that it is useful to perform subjective investigation of the sound quality of the violin using the comparison method on the basis of sound recordings. Coincidence of estimates made by expert violinists in the so called direct evaluation with estimates made by a listener group by way of indirect evaluation indicates that a recording can sufficiently well evidence the quality of a given instrument. Since the most frequent aim of auditory quality evaluation is selecting the best of a group of objects tested, the present results are in favor of the recordings made in the mono system.

References

[1] A. GÓRALSKI, *Methods of description and statistical conclusion in psychology* (in Polish), PWN, Warszawa 1974.

[2] J. P. GUILDFORD, *Psychometric methods*, McGraw Hill Book Co., New York 1954.

[3] T. ŁĘTOWSKI, *Auditory assessment of electroacoustic devices* (in Polish), Reports of Research and Development Centre of Polish Radio and Television (1976).

[4] *Statistical tables* (in Polish) W. SADOWSKI, ed. PWN, Warszawa 1957.

Received on April 23, 1980.

ARCHIVES OF ACOUSTICS
6, 1, 23-44 (1981)

AVERAGING THE FREQUENCY OF THE LARYNX TONE IN THE CORRELATION METHOD OF ITS ESTIMATION USING THE TECHNIQUE OF LINEAR PREDICTION

ANDRZEJ DZIURNIKOWSKI

PESEL (02-106 Warszawa, ul. Pawińskiego 17/21)

This paper presents some problems in the effects of averaging the results obtained from the analysis of a speech signal, related to the use of the correlation method of the estimation of the frequency of the larynx tone. The present considerations are concerned with the analysis of a speech signal using the algorithm of linear prediction. The thesis that the results obtained from averaging depend on the parameters of the analysis assumed and the character of a signal is proposed and the reasons for this phenomenon discussed. This dependence must not be neglected in choosing the methods of speech signal analysis in real investigations.

1. Introduction

The autocorrelation analysis of a speech signal is one of the earliest methods of the estimation of the frequency of the larynx tone (the fundamental frequency). Since, in general, a speech signal is an implementation of a certain stochastic process which for adequately long durations in its selected classes, can be regarded as a stationary signal, it is necessary to take into consideration during the analysis of it the dependencies which arise from this fact and conditions for the estimators of an autocorrelation function of this type of signals.

This paper presents numerical methods of the estimation of the frequency of the larynx tone by means of the autocorrelation analysis using the technique of linear prediction. The direct considerations and examples are based on the analysis of the autocorrelation sequence of the error signal of linear prediction.

It is proposed that the precision of the results obtained using the estimation method assumed depends directly both on the parameters of the analysis assumed and the character of the signal itself in a predetermined interval of the analysis. The averaging of the estimated values of the periods of the larynx tone is an effect of these dependencies.

The present paper discusses this problem and presents a method for its analysis, based on a measure of the deviation of the estimated values of the periods of the larynx tone from their real values, as proposed by the author, and the smoothing coefficient defined on the basis of this measure.

The existence of these factors in the methods for the determination of the value of the periods of the larynx tone presented here must be considered in the determination of the aim of investigation and also calls in question its use in the spectral analysis of a speech signal, synchronised by the larynx tone.

2. Estimation of the autocorrelation function of a discrete stochastic process

All stochastic processes are characterised by giving a n-dimensional probability distribution of random variables for $n \to +\infty$. A human speech signal is one of these processes, experimental investigations of which are in view of practical constraints performed on a signal of finite duration. It is, therefore, impossible to determine precisely all the parameters of the probability distribution on the basis of experimental data. In this situation, functionals defined by selected implementations of random processes obtained from investigation of signals are assumed as the parameters of the probability distribution and are called the estimators of the parameters of processes analysed. In the correlation methods for the investigation of random signals, which are most frequently used to determine the regularity of the structure of a speech signal and for determination of its successive periods, called the larynx tone, the autocorrelation function of the process (or its transform) is assumed as the parameter, which is defined as

$$R_x(t_1, t_2) = E[X_0(t_1) X_0(t_2)], \tag{1}$$

where $X_0(t) = X(t) - E[X(t)]$. For the stationary processes this function depends only on $\tau = t_2 - t_1$. In numerous cases in view of the fact that the variance does not depend on the time t, the normalised autocorrelation function is assumed

$$R_x^u(\tau) = \frac{R_x(\tau)}{\sigma_{x^2}}. \tag{2}$$

Estimation of the parameters of the probability distribution of stochastic processes, based on the estimators, must always satisfy the following requirements: the estimators should be compatible, unweighted, "best" from the point of view of a criterion assumed (e.g. effective where effectiveness is the ratio of the minimum optimum variance to the variance of an estimator assumed) [12]. There is a number of estimators for determination of the autocorrelation function $R(\tau)$ of a stationary process, based on its implementation. The estima-

tor defined by the formula

$$P_x(\tau) = \int_0^{T-\tau} a(t, \tau) X_0(t) X_0(t+\tau)\, dt \tag{3}$$

is most frequently used, where $a(t, \tau)$ is the weight function whose selection affects in a significant way the quality of the estimator. In the case when the mean value of a random process is known, the quantity $P_x(\tau)$ defined by formula (3) is an unweighted estimator of the function $R_x(\tau)$. Determination of the optimum weight function a_{opt} in view of the minimum variance of the estimator $P_x(\tau)$ is here troublesome and requires knowledge of the autocorrelation function $R_x(\tau)$. Therefore, this estimator is quite often determined on the basis of the integral mean as

$$P_x(\tau) = \frac{1}{T-\tau} \int_0^{T-\tau} X_0(t) X_0(t+\tau)\, dt \tag{4}$$

which gives a relatively low estimation error [12]. The situation is different when the mean value of a random process is unknown and it is estimated on the basis of the implementation of the process itself. In this case the error of the weight of the estimator cannot be avoided. In order to avoid the weight of this estimator it would be necessary to introduce additionally a coefficient of the general form $1/1-F[R_x(\tau)]$ [12]. Since the function $R_x(\tau)$ is unknown at the time of estimation, therefore in order to avoid the weight first the weighted autocorrelation function is quite often determined and then iterative operations are used to avoid the weight [12]. It is a very complicated process, therefore, many authors consciously or not assume in their investigations estimated values weighted by error. The estimator of the autocorrelation function of random signal is essentially the estimator of the function and not a parameter and also a random function. Therefore, its covariance should be determined in the estimation of the error of such estimator. Given in [12], the formula for the estimation of the covariance of the estimator permits a conclusion that the covariance between the values of the estimator decreases with increasing duration of the process. Therefore, the final values of the estimator in a sequence are insignificant. In view of the above fact and also the relevant relations for the variability of the estimator [12] it is possible to determine in practice the duration τ_{max}, for which it is worthwhile to calculate the estimator [4, 5].

In practice under the pressure of requirements resulting from the assumptions of an experiment the analysis of this type is not frequently carried out despite the fact that only in some cases this approach would have been justified by the requirement of the experiment e.g. when the absolute values of the estimator are not essential, but the character of its variation or the position of its extrema etc. As for a continuous stationary process, for the stationary

sequence $\{X(n)\}$ the estimator defined by the formula

$$P_j = \frac{1}{N-j} \sum_{n=1}^{N-j} x(n)x(n+j) \tag{5}$$

is determined, to which all the remarks concerning the estimation of the autocorrelation function of a random process apply.

In the further part of this paper for the sake of uniformity of notation, the values of the autocorrelation function or its estimators will expressed by $\varrho(j)$ (the coefficients of the autocorrelation function).

3. Numerical methods for the estimation of the frequency of the larynx tone by the autocorrelation analysis using the technique of linear prediction

Only those signals will considered that are represented by the sample sequences $\{x(n)\}$ of the signal, i.e. signals in the form of numerical data, which can be easily processed digitally using computer.

One of the basic methods for the estimation of the frequency of the larynx tone used in the literature is the autocorrelation analysis of the sample sequence of a speech signal, $\{x(n)\}$ carried out on the basis of the relation

$$\varrho_x(j) = \sum_{n=-\infty}^{+\infty} x(n)x(n+j). \tag{6}$$

In practice the autocorrelation coefficients are calculated for the cut-off signal, i.e. a signal for which $x(n) = 0$ for $n < 0$ and $n > N-1$. This signifies that

$$\varrho_x(-j) = 0 \quad \text{for } |j| \geqslant N, \tag{7}$$

and

$$\varrho_x(-j) = \varrho_x - j = \sum_{n=0}^{N-1-j} x(n)x(n+j) \quad \text{for } j = 0, 1, \ldots, N-1.$$

There is a number of algorithms for calculating the autocorrelation coefficients, for example: the algorithm based on the calculation of the Fast Fourrier Transform (FFT) [9, 11] or the algorithm which implements the so called continuous correlation [9]. However, in both these and other methods, the results obtained, which are the basis for the estimation of the frequency of the larynx tone, are, in addition to the estimation error, weighted by the higher harmonic frequencies of F_0 (the effect of formants is significant), particularly the first one, that occur in a speech signal. One of the methods for the estimation of the larynx tone freguency by the autocorrelation technique in a signal from which the effect of formants was eliminated is the one proposed by ITAKURA and SAITO [7], a method

for the estimation of the frequency of the larynx tone, using the technique of linear prediction. The algorithm for calculating the estimated frequencies f_0, as described in [4] is based on the calculation in the discrete time domain of the successive values of the error signal $\{e(n)\}$ (cf. Fig. 1) defined as

$$e(n) = \sum_{i=0}^{M} a_i x(n-i) = x(n) + \sum_{i=1}^{M} a_i x(n-i), \tag{8}$$

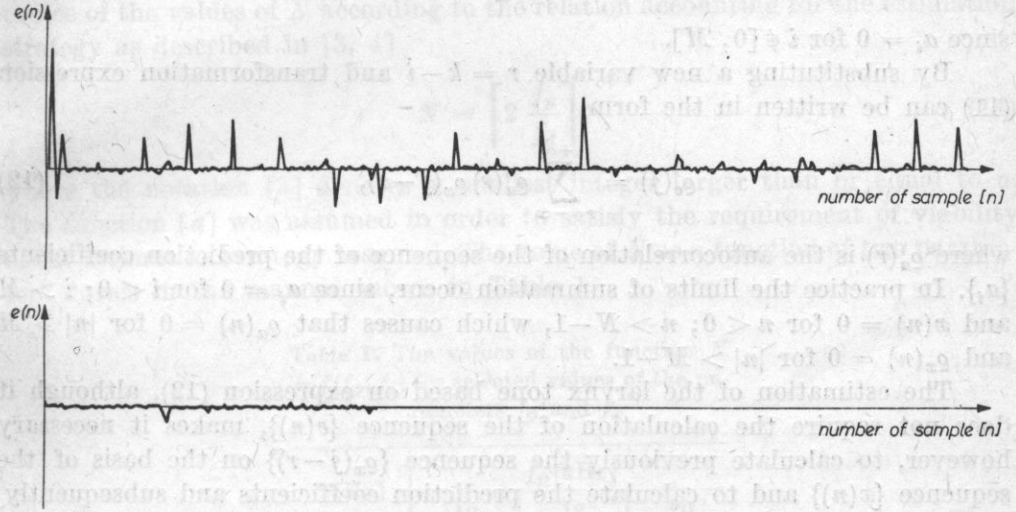

Fig. 1. An example of the error signal $\{e(n)\}$ in the word "pokoju"

where a_i are the coefficients of the inverse filter $A(z)$ of order M, defined in the z-domain as

$$A(z) = \sum_{i=0}^{M} a_i z^{-1}, \qquad a_0 = 1. \tag{9}$$

Given in [4], the analysis of the duration of the period of the larynx tone and the estimation of its frequency is based on the autocorrelation analysis of the sequence $\{e(n)\}$

$$\varrho_e(j) = \sum_{n=-\infty}^{+\infty} e(n)e(n+j). \tag{10}$$

There are a few specific variants of calculating the autocorrelation sequence $\{\varrho_e(j)\}$ using the relations between the coefficients of linear prediction and the real values of a signal, based on a direct implementation of the procedure described by formula (10) or using the autocorrelation function $\{\varrho_x(j)\}$ of the real data $\{x(n)\}$. The second variant uses the relation obtained from transformation

of formula (10) by insertion in it of the expression described by formula (8)

$$\varrho_e(j) = \sum_{n=-\infty}^{+\infty} \sum_{i=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} a_i a_k x(n-i) x(n+j-k). \tag{11}$$

This insertion involved the fact that

$$e(n) = \sum_{i=-\infty}^{+\infty} a_i x(n-i)$$

since $a_i = 0$ for $i \notin [0, M]$.

By substituting a new variable $r = k-i$ and transformation expression (11) can be written in the form

$$\varrho_e(j) = \sum_{r=-\infty}^{+\infty} \varrho_a(r)\, \varrho_x(j-r), \tag{12}$$

where $\varrho_a(r)$ is the autocorrelation of the sequence of the prediction coefficients $\{a_i\}$. In practice the limits of summation occur, since $a_i = 0$ for $i < 0$; $i > M$ and $x(n) = 0$ for $n < 0$; $n > N-1$, which causes that $\varrho_a(n) = 0$ for $|n| > M$ and $\varrho_x(n) = 0$ for $|n| > M-1$.

The estimation of the larynx tone based on expression (12), although it does not require the calculation of the sequence $\{e(n)\}$, makes it necessary however, to calculate previously the sequence $\{\varrho_x(j-r)\}$ on the basis of the sequence $\{x(n)\}$ and to calculate the prediction coefficients and subsequently on their basis the sequence $\{\varrho_a(r)\}$. When using one of the many available techniques of the calculation of the error signal given in [8] it is, however, more convenient to analyse on the basis of relation (10) in the limited summation range

$$\varrho_e(j) = \sum_{n=0}^{N-1-j} e(n) e(n+j), \tag{13}$$

where the sequence $\{\varrho_e(j)\}$ is in practice calculated for $j \leqslant N/2$. This method, which essentially consists in simple calculation of the autocorrelation coefficients, based on the values of the error signal of linear prediction will be the basis of the further considerations.

4. The effect of the method for estimation of the larynx frequency on the precision of results obtained

It follows from (13) that if one takes the sequence $\{\varrho_e(j)\}$ as the direct basis for tracing the duration of the period of the larynx tone by determining the values of $\max_j \{\varrho_e(j)\}$ in a predetermined interval Q depending on f_p (the sampling frequency) and a predetermined period of the analysis of the frequency

of the larynx tone, then the selection of the values of N in formula (13) essentially affects the averaging of the values of $\{f_{om}\}$, which are the successive values of the estimated frequency of the larynx tone.

It is interesting to consider how the values of N are chosen in relation to the frequency f_p and a predetermined analytical range of the larynx tone limited by the lower boundary f_d and the upper boundary f_g and how this affects the averaging of the estimated values of $\{f_{om}\}$, $m = 1, 2 \ldots$, in relation to their real values. In practice only the lower boundary essentially affects the choice of the values of N according to the relation accounting for the estimation strategy as described in [3, 4]

$$N = \left\lceil 2 \frac{f_p}{f_d} \right\rceil, \tag{14}$$

where the notation $\lceil a \rceil$ denotes a smallest integer larger than or equal to a. The function $\lceil a \rceil$ was assumed in order to satisfy the requirement of viability of the estimation strategy assumed. The value of N as a function of two parameters varies in the manner shown in Table 1.

Table 1. The values of the function $N = f(f_p, f_d)$ for selected values of the parameters f_p and f_d

f_d [Hz]	f_p [kHz]		
	10	12	20
20	1000	1200	2000
40	500	600	1000
50	400	480	800
60	334	400	667
80	250	300	500
100	200	240	400

While the lower predetermined frequency f_d affects the determination of the lower boundary of the interval Q, in which $\max_j \{\varrho_e(j)\}$ is traced, the upper frequency f_g is the lower limit of the interval Q; $j \in [L(f_g), N/2]$. The interval thus defined is essential for the effect of averaging the estimated values of f_{om}. This effect can to a lesser or greater degree occur in relation to the real period of the larynx tone.

If T_m denotes the real m-th period of the larynx tone then it can be stated that the longer the period T_m the lesser averaging effect occurs. The shorter the period T_m the greater the effect is, depending in direct proportion on the value of N. The variable N is a function of the permissible value of $T_d = \max_m \{T_m\}$, defined by formula (14). Therefore, the higher frequency of the

larynx tone is examined and at the same time the lower frequency f_d is taken, the greater the effect of averaging the frequency of the larynx tone will be in its estimation based on the autocorrelation method and the technique of linear prediction.

The above problem can be presented in the following way. If T_h denotes the hypothetical period of the larynx tone defined by the number of samples of the signal sampled at the frequency f_p and α denotes the filling coefficient of interval $[0, N]$ of the form

$$\alpha = \frac{N}{T_h} \tag{15}$$

then the function $\alpha(T_h)$ which takes values from the interval $\left[2, \frac{Nf_g}{f_p}\right]$ behaves as in Fig. 2.

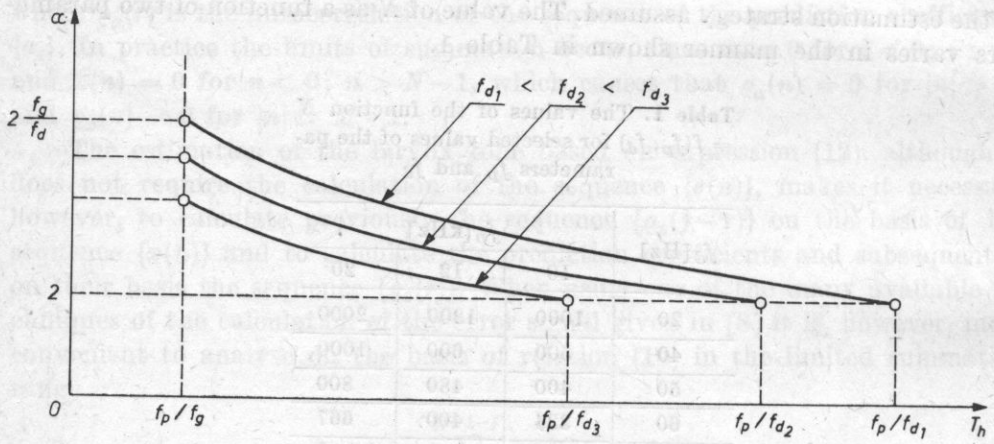

Fig. 2. The curve of the function $\alpha(T_h)$ in relation to the parameter f_d

It can be seen that depending on the frequency f_d assumed and in particular on the possible gap $f_r = f_g - f_d$, the filling coefficient varies taking a maximum value of

$$\alpha_{\max} = 2\frac{f_g}{f_p}. \tag{16}$$

When the function describing the signal is periodic, the coefficient α is insignificant in the calculation of the autocorrelation coefficients, since it is invariable for each successive period of a signal with the duration N. It is significant, however, in the autocorrelation analysis of quasiperiodic signals, e.g. a voiced speech signal. In this case $\alpha \neq$ const. and depending on the real values

of the period of the larynx tone T_m can be expressed by the formula*

$$\alpha_q = \sum_{l=1}^{P}\left[I\left(N-\sum_{m=0}^{l}T_m\right)+\frac{\left(N-\sum_{m=0}^{l-1}T_m\right)I\left(N-\sum_{m=0}^{l-1}T_m\right)}{T_l}I\left(\sum_{m=0}^{l}T_m-N\right)\right], \tag{17}$$

where

$$T_0=0, \quad p=\lceil \alpha_{\max}\rceil, \quad T_{p+1}=+\infty, \quad I(x)=\begin{cases}0 & \text{for } x<0,\\ 1 & \text{for } x\geqslant 0.\end{cases}$$

The coefficient α_q is a function of the successive periods of the larynx tone; $\alpha_q = f(T_0, T_1, T_2, T_3, \ldots, T_{\lceil \alpha_q\rceil})$ and takes the real positive values $\alpha_q \leqslant \alpha_{\max}$.

Formula (17) applies to the real conditions of the autocorrelation analysis of a signal of finite duration. The filling coefficient of a quasiperiodic signal, α_q, significantly affects the averaging of the estimated of the larynx tone, since it reflects the share of a given number of the variable periods in the calculation of the autocorrelation coefficients. There is still another aspect related to the behaviour of the autocorrelation function determined by the sequence of its coefficients, which is connected with a speech signal, and accordingly the error signal, being in most general terms a stochastic signal. When one considers only the voiced speech signal, which is quasiperiodic and therefore interesting for the estimation of the fundamental frequency f_{om}, then for adequately long signals it is often possible to assume that a signal is ergodic. Under this assumption for such signals the autocorrelation function behaves in a specific manner and is then defined by the relation

$$\varrho_x(j) = \lim_{N\to+\infty}\frac{1}{2N}\sum_{n=-N}^{n=N}X(n)X(n+j). \tag{18}$$

The autocorrelation function of an ergodic stochastic process is nonlinear and takes values from a limited range, which results from the fact that

$$\lim_{j\to 0}\varrho_x(j) = \varrho_x(0) = \lim_{N\to+\infty}\frac{1}{2N}\sum_{n=-N}^{n=N}X^2(n) \tag{19}$$

and

$$\lim_{j\to\pm\infty}\varrho_x(j) = \lim_{j\to\pm\infty}\sum_{X(n)}\sum_{X(n+1)}X(n)X(n+j)\varphi\left(X(n), X(n+j), j\right) = E^2[X(n)], \tag{20}$$

* Formula (17) is valid for the analysis of a quasiperiodic signal synchronised by the larynx tone as described in [4].

where φ is the probability density function, while $E[X(n)]$ is the first moment of the random variable $X(n)$* and $\varrho_x(0) = \max q_x(j)$. The possible assumption is here used, that for $j \pm \infty$ interaction of the random variables $X(n)$ and $X(n+j)$, represented by the values $x(n)$ and $x(n+j)$, disappears and in the boundary case they can be treated as statistically independent.

Similar relations apply to the error signal obtained using the technique of linear prediction. The foregoing considerations dealt with the autocorrelation signal analysis based on relation (6). In practice the analysis is carried out on a signal of finite duration using relations (7) and (13). This fact naturally causes a more rapid decrease in the value of the autocorrelation function. It can be expressed in approximation by the relation

$$\varrho'(j) = \beta(j)\,\varrho_x(j), \tag{21}$$

where $\beta(j) = \dfrac{N-1-j}{(N-1)}$; $j \in [L(f_g), N/2]$ and functions as the damping coefficient of the autocorrelation $\varrho_x(j)$ (see Fig. 3).

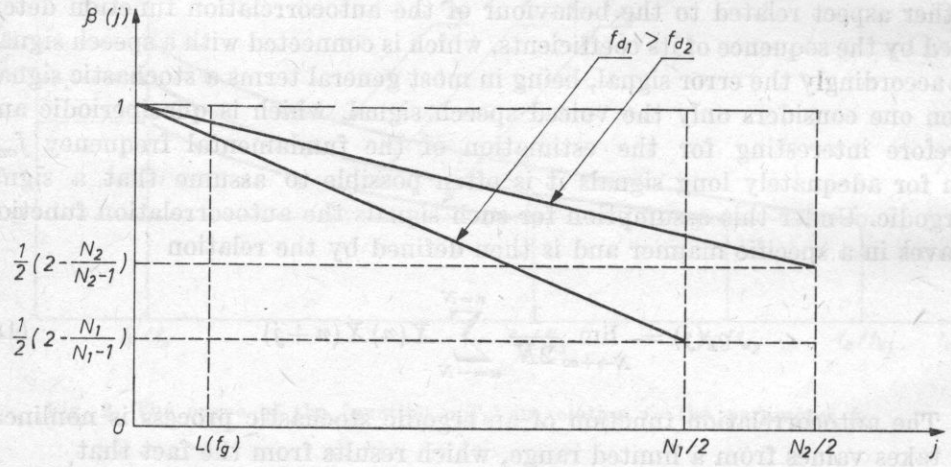

Fig. 3. The dependence of the linear "damping coefficient" of autocorrelation, $\beta(j)$ on the length of the calculation interval

The relation discussed above, which is connected with the random character of the signal, is distinctly reflected in Figs. 4-8. In practice, relation (21) only expresses a trend in the behaviour of the autocorrelation function, which can be disturbed by the quasiperiodicity of the signal. This phenomenon is shown in Figs. 4-8. A more exact consideration of how the signal itself affects the be-

* In the case of quasiperiodicity of the ergodic process

$$E[X(n)] = E[X(n+j)] = \ldots = E[X].$$

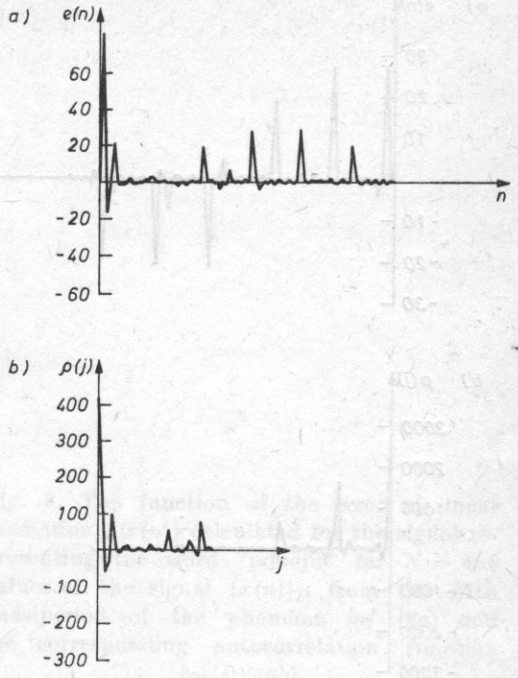

Fig. 4. The function of the error of linear prediction, $\{e(n)\}_D$ calculated for the signal $\{X(n)\}_D$ representing the word "pokoju", for the first $N = 480$ values of the signal $\{X(n)\}_D$ from the beginning of the phoneme /k/ (4*a*) and the corresponding autocorrelation function $\{\varrho_e(j)\}$ (4*b*)

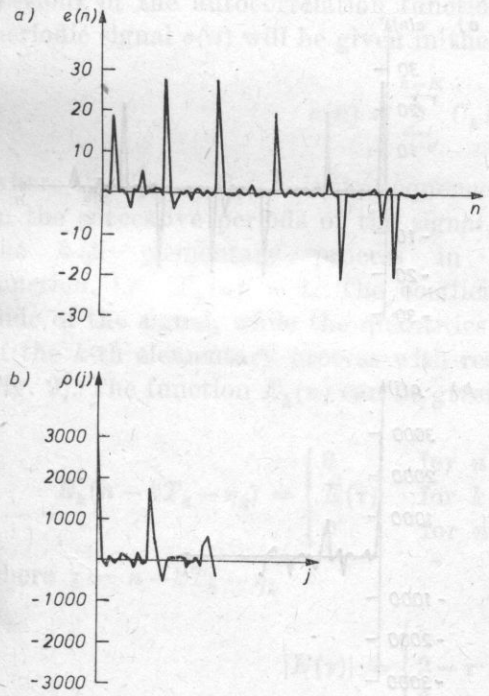

Fig. 5. The function of the error of linear prediction, $\{e(n)\}$ calculated for the signal representing the word "pokoju" for $N = 480$ values of the signal $\{x(n)\}_D$ from the second quasiperiod of the signal of the phoneme /o/ (5*a*) and the corresponding autocorrelation function $\{\varrho_e(j)\}$ (5*b*)

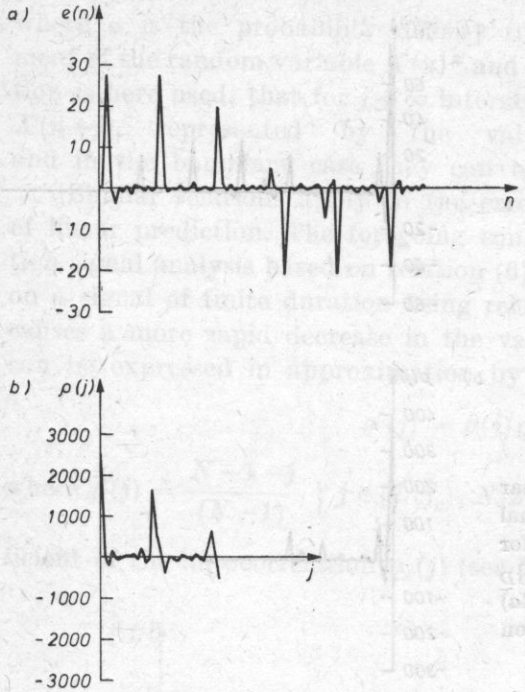

Fig. 6. The function of the error of linear prediction, $\{e(n)\}$ calculated for the signal representing the word "pokoju" for $N = 480$ values of the signal $\{x(n)\}_D$ from the third quasiperiod of the signal of the phoneme /o/ (6*a*) and the corresponding autocorrelation function $\{\varrho_e(j)\}$ (6*b*)

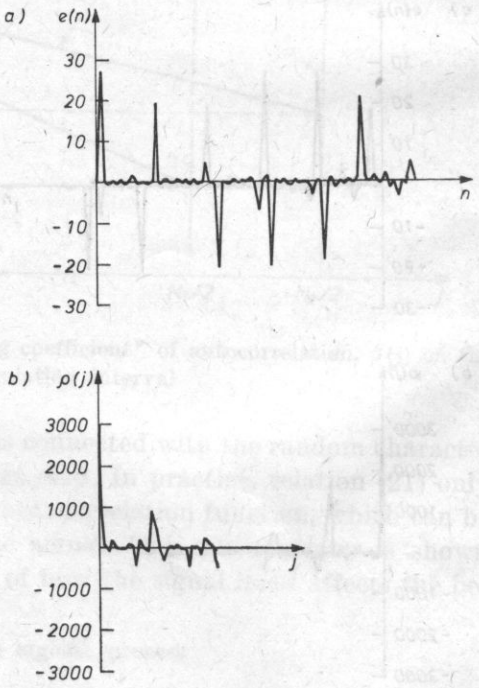

Fig. 7. The function of the error of linear prediction, $\{e(n)\}$ calculated for the signal representing the word "pokoju" for $N = 480$ values of the signal $\{x(n)\}_D$ from the fourth quasiperiod of the phoneme /o/ (7*a*) and the corresponding autocorrelation function $\{\varrho_e(j)\}$ (7*b*)

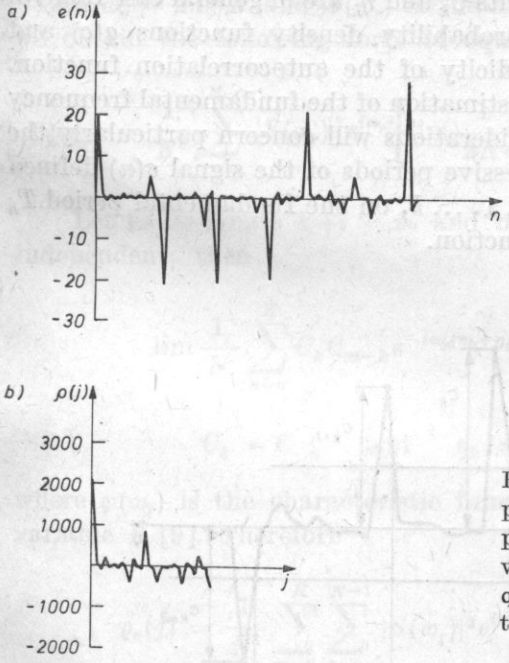

Fig. 8. The function of the error of linear prediction, $\{e(n)\}$ calculated for the signal representing the word "pokoju" for $N = 480$ values of the signal $\{x(n)\}_D$ from the fifth quasiperiod of the phoneme /o/ (8*a*) and the corresponding autocorrelation function $\{\varrho_e(j)\}$ (8*b*)

haviour of the autocorrelation function now follows. To that end the quasi-periodic signal $e(n)$ will be given in the form

$$e(n) = \sum_{k=0}^{k=K} C_k E_k(n - kT_h - \eta_k), \tag{22}$$

where $C_0, C_1, \ldots, C_k, \ldots$ is the sequence of the coefficients of amplitude changes in the successive periods of the signal, while $E_k(n)$ is the function describing the k-th elementary process in the period and is a normalised function, i.e. $|E_k(n)| = 1$. The coefficients C_k characterise changes in amplitude of the signal, while the quantities η_k describe changes in the initial phase of the k-th elementary process with respect to the hypothetical period T_h (see Fig. 9). The function $E_k(n)$ can be given in approximation in the following way

$$E_k(n - kT_h - \eta_k) = \begin{cases} 0 & \text{for } n < kT_h + \eta_k, \\ E(\tau) & \text{for } kT_h + \eta_k \leqslant n \leqslant (k+1)T_h + \eta_{k+1}, \\ 0 & \text{for } n > (k+1)T_h + \eta_{k+1}, \end{cases} \tag{23}$$

where $\tau = n - kT_h - \eta_k$

$$|E(\tau)| = \begin{cases} \tau & \text{for } 0 \leqslant \tau \leqslant 1, \\ 2 - \tau & \text{for } 1 < \tau \leqslant 2, \\ 0 & \text{for } \tau > 2. \end{cases}$$

It can be assumed that the coefficients C_k and η_k are in general case random variables characterised by unknown probability density functions $g(c)$ and $g(\eta)$, and essentially disturb the periodicity of the autocorrelation function. Since the interest here is first of all the estimation of the fundamental frequency of signals of this type, the present considerations will concern particularly the influence of change in duration of successive periods of the signal $e(n)$ defined by the sequence of values of η_k, $k = 0, 1, \ldots k$, on the fundamental period T_h estimated using the autocorrelation function.

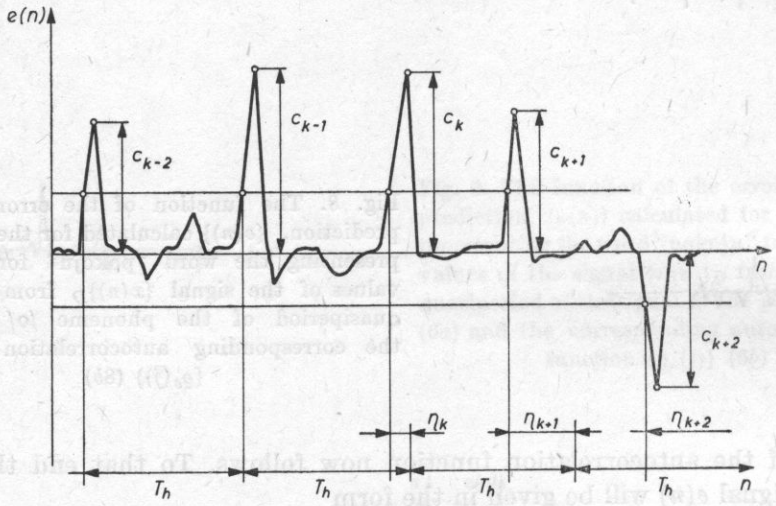

Fig. 9. A random sequence of pulses of a quasiperiodic signal

We insert into formula (18) the signal described by formula (22)

$$\varrho_e(j) = \lim_{N\to\infty} \frac{1}{2N} \sum_{n=-N}^{n=N} e(n)e(n+j)$$

$$= \lim_{N\to\infty} \frac{1}{2N} \sum_{n=-N}^{n=N} \sum_{k=0}^{k=K} C_k E_k(n - kT_h - \eta_k) \sum_{l=0}^{l=K} C_l E_l(n - lT_h - \eta_l + j)$$

$$= \lim_{N\to\infty} \frac{1}{2N} \sum_{k=0}^{k=K} \sum_{l=0}^{l=K} C_k C_l \sum_{n=-N}^{n=+N} E(n - kT_h - \eta_k) E(n - lT_h - \eta_l + j). \quad (24)$$

Substituting the relevant Fourier transforms [10] for the expression $E(n)$, with consideration of delay

$$E(n - kT_h - \eta_k) = \frac{1}{N} \sum_{r=0}^{N-1} S(\omega_r) e^{i\omega_r n} e^{-i\omega_r kT_h} e^{-i\omega_r \eta_k},$$

where $\omega_r = 2\pi r/N$ and $S(\omega_r)$ is the amplitude spectrum of the process $E(n)$, we obtain the following form of equation (24):

$$\varrho_e(j) = \frac{1}{N^2}\sum_{r=0}^{N-1}|S(\omega_r)|^2 e^{i\omega_r j}\left[\lim\frac{1}{2N}\sum_{k=0}^{K}\sum_{l=0}^{K}C_k C_l e^{-i\omega_r(\eta_k+\eta_l)}e^{-i\omega_r(k+l)T_h}\right]. \quad (25)$$

Let us designate $k+l = m$ and note that if C_k and C_l and η_k and η_l are independent, then

$$\lim\frac{1}{N}\sum_{k=0}^{K}C_k C_{m-k}e^{-i\omega_r(\eta_k+\eta_i)} = \begin{cases}\bar{C}^2 & \text{for } m = 0,\\ \bar{C}^2[\varphi(\omega_r)]^2 & \text{for } m \neq 0;\end{cases}$$

$$C_k = C_{-k} \quad \text{and} \quad \eta_k = \eta_{-k}, \quad \varphi(\omega_r) = E[e^{j\eta\omega_r}],$$

where $\varphi(\omega_r)$ is the characteristic function of the distribution of the random variable η [9]. Therefore

$$\varrho_e(j) = \frac{1}{N^2}\sum_{m=0}^{K}\sum_{r=0}^{N-1}|S(\omega_r)|^2 e^{i\omega_r j}e^{-i\omega_r mT_h}\left\{\lim\frac{1}{2N}\sum_{k=0}^{K}C_k C_l e^{-i\omega_r(\eta_k+\eta_l)}\right\}$$

$$= \frac{\bar{C}^2\varrho_0(j)}{2N^2} - \frac{\bar{C}^2}{2N^2}\sum_{m=1}^{K}\sum_{r=0}^{N-1}|S(\omega_r)|^2|\varphi(\omega_r)|^2 e^{-i\omega_r(mT_h-j)}, \quad (26)$$

where $\varrho_0(j)$ is the autocorrelation function of the periodic signal $E(n)$.

It should be noted that relation (26) is valid for $N\to+\infty$, i.e. at the same time $K\to+\infty$, which is not satisfied for investigations in practice, and also under the assumption that $E[e] = 0$. More exact consideration of these problems can be found in paper [1]. The above theoretical considerations lead to the conclusion that the autocorrelation function of the signal described by formula (22) is a complicated probability density function of the random variable η and C and therefore it is difficult to draw conclusions on the behaviour of values of the autocorrelation function without knowledge of these distributions.

Irrespective of the form assumed for the distributions of these variables, the results of analyses based on relation (22) do not have general character, since the shape of these distributions is strongly dependent on the individual characteristics of the speaker. It is, however, an essential result of these considerations that they confirmed strong dependence of the autocorrelation function on the length of the analytical interval and on the real duration of successive periods. The determination of the degree of averaging the frequency of the larynx tone, consisting in tracing $\max_j\{\varrho(j)\}$ in a predetermined analytical interval is therefore dependent on the values

of N and T_h, the latter being a hypothetical duration of the larynx tone period, chosen as the first period within the interval of the signal processed.

It seems justified to introduce a measure which would be a simple relation describing the degree of deviation of the estimated mean duration of the period of the larynx tone, from the real length of the first period within the predetermined l-th interval of the analysis. This measure can be expressed in the form

$$\delta_1^l = 1 - \frac{T_s^l}{T_{rz}^l}, \qquad T_{rz}^l = T_h^l = T_1^l. \tag{27}$$

The averaged estimated length can in approximation be described as the arithmetic mean

$$T_s^l = \frac{\sum_{i=1}^{a_k} T_i^l}{a_k^l} \cong \frac{N}{a_k^l}, \qquad a_k^l = [a_q^l]. \tag{82}$$

If we hypothetically assume the periodicity of the signal in the investigated interval N with a predetermined period equal to a chosen value T_i, $i = 1, 2, \ldots, a_k$, then we can determine the filling of this interval by relevant periods of the signal as $a_i = [N/T_i]$. Thus assuming that $T_1^l = T_{rz}^l \cong N/a_1^l$ we obtain the approximate value δ_1^l of the quantity δ_{1p}^l as

$$\delta_{1p}^l \cong 1 - \frac{a_1^l}{a_k^l}. \tag{29}$$

It should be noted here that expression (29) can be used to determine the approximate value δ_1 defined by formula (27) only for the relatively long analytical intervals with respect to the periods of the larynx tone under analysis (i.e. for high values of a_k). Fig. 10 shows schematically the examples of values taken by the function a_q depending on the distribution of the real successive periods of the larynx tone for a predetermined length of the analytical period N.

It follows from the foregoing considerations that in the autocorrelation analysis the estimated periodicity of the signal deviates in succesive steps of the analysis from the real values of periods of the quasiperiodic signal and depends in most general case on the value of a_k and also on the real period $T_1 = T_{rz}$ itself. This difference can be defined in approximation by the deviation measure δ_1 whose sign shows the direction of this deviation.

The deviation measure δ_1 does not define the measure of averaging irrespective of how this measure will be defined. There is, however, a relation between these two measures, since the averaging measure is a function of the variables (N, T_i^l, δ_1^l).

If we take as the averaging measure a smoothing coefficient of the sequence of the real durations of the period of the larynx tone in the form of a normalised sum of differences between the successive real values and estimated period

durations for their predetermined number L in the form

$$W^L = \frac{1}{L}\sum_{i=1}^{L} |(T^l_{rz} - T^l_s)/T_{rz}| \cong \frac{1}{L}\sum_{i=1}^{L} |\delta^i_{1p}| \tag{30}$$

then the thus defined smoothing coefficient is the arithmetic mean of successive deviation measures. The greater absolute values the coefficient W^L takes the more significant the smoothing of the real values of the sequence $\{T^i_{rz}\}$ becomes. Some conclusions can be drawn, therefore, on the averaging of these values in their estimation by the method of linear prediction, based on determination in a predetermined analytical interval of the averaged values of successive periods of the larynx tone [4]. The value of the coefficient W^L via the absolute values of the deviation measure δ^l_{1p} depends on the ratio a^l_1/a^l_k which to some extent reflects the quasiperiodicity of the signal investigated.

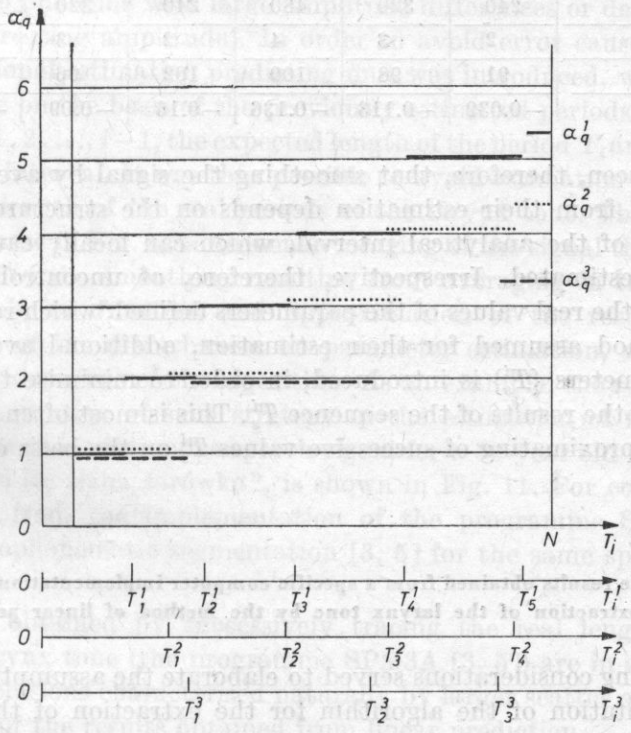

Fig. 10. An example of the values of the function depending on the length and distribution of quasiperiods of a signal

The longer the analytical interval defined by the value of N and the shorter the quasiperiods T_i contained in it, the greater value a_k takes, which affects determination of the averaged duration of the estimated period T_s. At the same time, however, the greater the measure δ_1 can become in the case of local steplike change in the real value of the larynx tone. In turn, for low values of N (short

analytical interval) and long quasiperiods $T_i a_k$ decreases, which causes large averaging errors to occur in the estimation of value of T_s and subsequently significant changes in values of δ_1. This can be demonstrated by the estimation of the period T_s for the two successive sequences of the quasiperiods T_i; the sequence $C_1 = \{88, 95, 103, 107, 111, 119\}$ and the sequence $C_2 = \{89, 115, 85, 108, 112\}$. The results obtained are shown in Table 2.

Table 2. Variation in the deviation of the estimated mean duration of the period of the larynx tone from the real duration of the first period, depending on the length of the analytical period N, for two chosen sequences of values of T_i

Coefficients	Sequence					
	C_1			C_2		
N	240	320	480	240	320	480
a_k	2	3	4	2	3	4
T_s	91	98	100	102	96	99
δ_1	−0.039	−0.113	−0.136	−0.16	−0.09	−0.12

It can be seen, therefore, that smoothing the signal by averaging its real values obtained from their estimation depends on the structure of the signal and the length of the analytical interval, which can locally cause large error of the values estimated. Irrespective, therefore, of uncontrolled smoothing or averaging of the real values of the parameters defined, which results from the analytical method assumed for their estimation, additional averaging of the estimated parameters $\{T_s^i\}$ is introduced, in order to minimise the local errors and smooth out the results of the sequence T_s^i. This is most often done by linear or nonlinear approximating of successive values T_s^i on the basis of the previous values [4].

5. Discussion of the results obtained from a specific computer implementation of the algorithm for the extraction of the larynx tone by the method of linear prediction

The foregoing considerations served to elaborate the assumptions for a computer implementation of the algorithm for the extraction of the larynx tone in a continuous speech signal. This algorithm was written in the Fortran language and used for statistical investigations aimed at elaboration of a method for speaker identification.

This algorithm described in [4] assumed the following implementation conditions. The lower frequency limiting the analytical range of the larynx tone frequency was taken as $f_d = 50$ Hz, which at the sampling frequency used in these investigations defined the length of the analytical segment $N = 480$ samples

according to formula (14) and Table 1, while the upper limiting frequency of the larynx tone was $f_g \cong 333$ Hz.

Thus the coefficient a_k was contained in the interval from 2 to 14 and in the investigation of male voices whose mean frequency F_0 oscillated about a frequency of ~ 120 Hz, this coefficient was close to 5 (Figs. 4-8). The averaging range in a signal of relatively low disturbance in its periodicity should be thus selected that it could be possible to estimate correctly successive values of F_{oi} and at the same time smooth out their variations. Moreover, it may happen during the estimation of the larynx tone by the autocorrelation technique that disturbances in the structure of the signal $\{X(n)\}$ and in $\{e(n)\}$ were so significant as to cause (as in the case in Fig. 4b) "local disturbances" in the work of the unit tracing $\max_j\{\varrho_e(j)\}$. This also happened (in addition to the events shown in Fig. 4) when the analytical segment contained transitions from phoneme to phoneme with large amplitude differences or decaying signals (in terms of decreasing amplitude). In order to avoid error caused by this influence an additional estimation predicting unit was introduced, whose function was to determine on the basis of the previously estimated periods of the larynx tone, $T_{i-n}, n = 1, 2, \ldots, i-1$, the expected length of the period T_i and accordingly of the successive interval. It was thus possible to avoid estimation error equal to half or multiple length of the real periods of the larynx tone. This caused, however, as in the case in Fig. 4, additional averaging of the signal through smoothing in addition to the smoothing resulting from averaging in the calculation of the autocorrelation function and approximation of the results, reducing partly the effect of the disturbances on parameter estimation, which did not occur in the previous steps of the algorithm. As an example, the result obtained from the implementation of such strategy of the estimation of the frequency of the larynx tone in a continuous speech signal for a 30-year old man who said "W pokoju paliła się słaba żarówka", is shown in Fig. 11. For comparison the results obtained from the implementation of the programme SPM3A based on primary microphonematic segmentation [3, 5] for the same speech signal is shown in Fig. 12.

The results obtained by successively tracing the real lengths of single periods of the larynx tone (the programme SPM3A [3, 5]) are in accordance to the previous conclusions characterised naturally by larger scatter of their values than in the case of the results obtained from linear prediction.

The results obtained by linear prediction reflect changes in the mean values of the results of primary segmentation calculated for a time constant of the order of scores of milliseconds, which fully confirms the foregoing conclusions on smoothing and averaging results obtained from the autocorrelation analysis used in the estimation of the larynx tone by the method of linear prediction. The differences at the beginning and at the end of the curves in Figs. 11 and 12 result from relatively high disturbances of the structure of the signal $\{X(n)\}$ and also

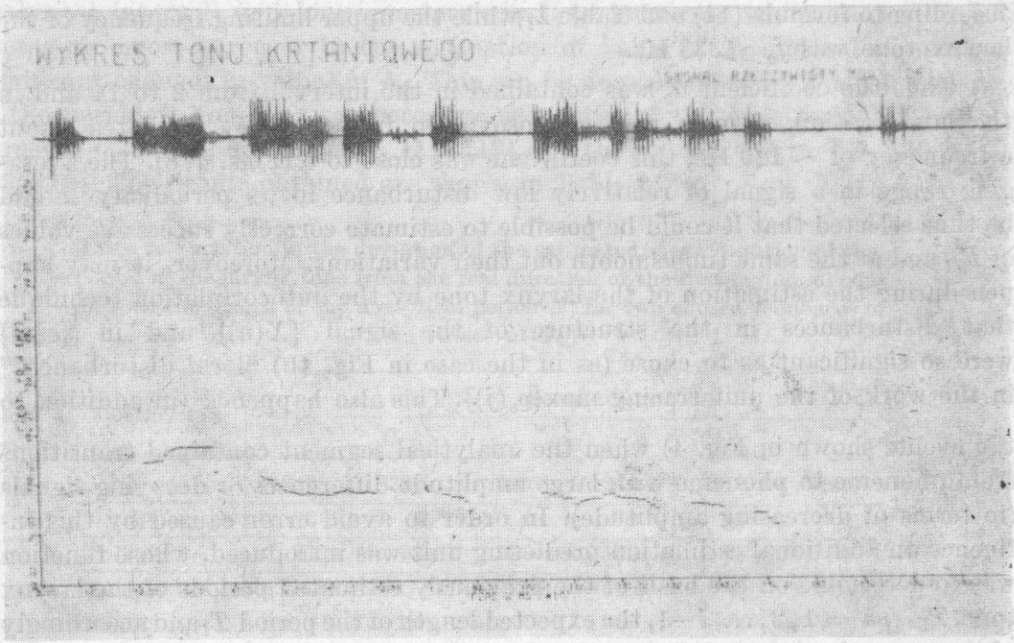

Fig. 11. The estimated pitch contour obtained by the method of linear prediction for the signal representing the sentence "w pokoju paliła się słaba żarówka"

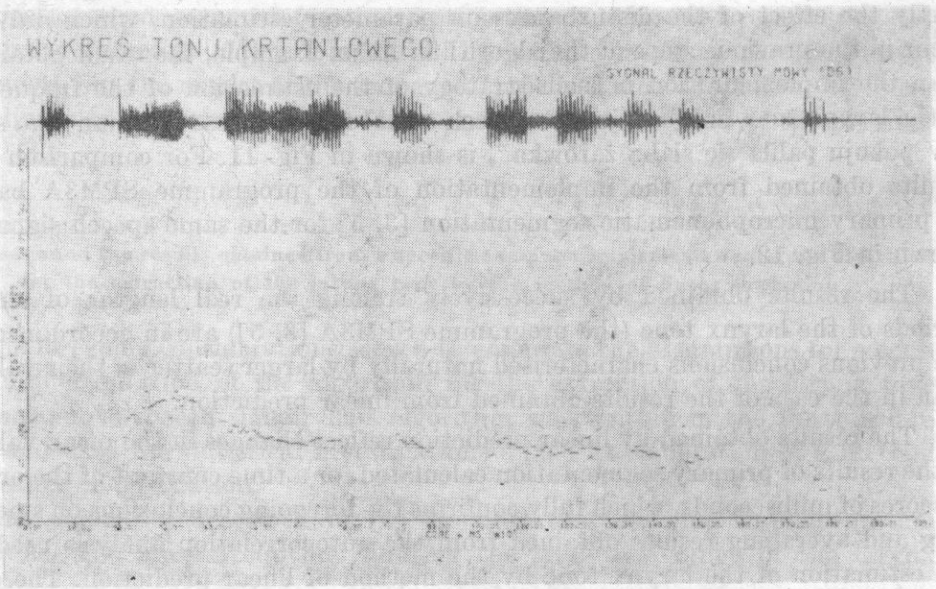

Fig. 12. The estimated pitch contour obtained by primary segmentation of the speech signal representing the sentence "w pokoju paliła się słaba żarówka"

from the assumed threshold values and the analytical intervals of F_0 (cf. [5]) and can be avoided by changing these values or their adequate fitting to a specific signal.

6. Conclusions

It can be stated on the basis of the results obtained by the two methods for the estimation of the parameter F_0 that both are viable from the point of view of their conditions and in agreement with theoretical predictions. These methods can be used for further analysis of human speech for different purposes. While the results obtained from primary segmentation serve for spectral analysis of the signal synchronised by the larynx tone in a system for speech recognition [2], the results obtained from the implementation of the programme based on the method of linear prediction are useless for this type of analysis. They can be used, however, in statistical investigation of the parameter F_0 for speaker identification or verification [6]. At present the author is performing extensive investigations of selected statistical parameters of the distributions of the frequency F_0 obtained by using both methods described above in order to determine the effect of the extraction methods used on the precision of identification and verification of speakers, depending on the classification algorithms employed. Preliminary investigations showed that there are parameters which show smaller scatter when the model of linear prediction is used than in the case when the method of primary segmentation is used, which is essential for the classification of characteristics. This confirms the conclusion that the choice of analytical method essentially affects the results of the investigations of speech signal and accordingly different directions of investigation (e.g. speech recognition or speaker identification require different analytical methods). In the present case the estimation of the frequency of the larynx tone by the method of linear prediction for spectral analysis of a speech signal synchronised by the larynx tone is not suitable despite advantages of the method itself (e.g. easy implementation of the digital algorithm for linear prediction).

References

[1] H. H. Beisner, *Spectrum of the lagged product in crosscorrelation*, JASA, **62**, 4, 916–921 (1977).

[2] L. Bolc, A. Dziurnikowski, K. Jaszczak, G. Kiełczewski, *Sistiema ponimania rieczi SUSY, Materiały konferencyjne: Trudy rasszíriennowo zasiedania rabocziej grupy II* KHBBT *po mietodam rozpoznawania, klasyfikacji i poiska informacji*, IOiK PAN, Warszawa 1976, 219–257.

[3] A. Dziurnikowski, *Primary segmentation of speech sound signals in the* SUSY *system*, Reports of JJUW, **52** (1976).

[4] A. DZIURNIKOWSKI, *Automatic determination of the frequency behaviour of the larynx tone by the method of linear prediction*, Archives of Acoustics, **5**, 1, 31–46 (1980).

[5] A. DZIURNIKOWSKI, *Microphonemes as fundamental segments of a speech wave; Primary segmentation – automatic searching for microphonemes*, Papers of the IJCAI – 75, 476–482, Tbilisi 1975.

[6] A. DZIURNIKOWSKI, *Automatic speaker recognition, problems and methods* (in Polish), Problemy kryminalistyki, **143**, 54–68 (1980).

[7] F. I. ITAKURA, S. SAITO, *Analysis-synthesis telephony based upon the maximum likelihood method*, Proc. 6th Int. Congress on Acoustics, C17–20, Tokyo 1968.

[8] J. D. MARKEL, A. H. GRAY, *Linear prediction of speech*, Springer–Verlag, Berlin, Heidelberg, New York, 1976.

[9] R. K. OTNES, L. ENOCHSON, *Digital time series analysis* (in Polish), WNT, Warszawa 1978.

[10] B. W. PAWŁOW, *Diagnostic investigation in technology* (in Polish), WNT, Warszawa 1967, 142–150.

[11] G. SANDE, *On the alternative method for calculating covariance functions*, Princeton Computer Memorandum, Princeton, New York 1965.

[12] S. J. WILENKIN, *Statistical methods for investigation of automatic control system* (in Polish), WNT, Warszawa 1969, 13–13, 61–62.

Received on August 1, 1979; revised version on August 7, 1980.

ARCHIVES OF ACOUSTICS
6, 1, 45-56 (1981)

DETECTABILITY OF SMALL BLOOD VESSELS AND FLAT BOUNDARIES OF SOFT TISSUES IN THE ULTRASONIC PULSE ECHO METHOD

L. FILIPCZYŃSKI

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

The detectability of a small blood vessel of a radius 0.1 mm by the pulse echo ultrasonic method using a frequency of 2.5 MHz was estimated. It was assumed that the vessel was surrounded by a homogeneous soft tissue (i.e. causing no reflection) and was in the near region of the far field radiated in a continuous manner by a plane transducer of a diameter of 2 cm. The soft tissue and the walls of the vessel were assumed to have the same elastic properties as those of a liquid.

The measurements were carried out in a plane polar coordinate system, where the incident wave, the reflected wave and the wave penetrating into the vessel were expressed in terms of Bessel and Hankel functions. The boundary conditions were assumed in the form of the equality of the acoustic pressures and the normal components of acoustic velocities on each side of the surface of the vessel. Thence the magnitude of the reflected wave was determined.

The losses of the signal due to the reflection of the wave, its divergence and absorption, are shown in the form of a graph from which it can be seen that the signal from the vessel considered is essentially detectible, although it lies near the noise level, and is critically dependent on the distance from the transducer due to attenuation in the tissues penetrated.

The detectability of the plane boundaries of soft tissues was also determined, indicating that at a distance of 20 cm from the transducer a difference in the characteristic acoustic impedance of the tissues of 0.2% is sufficient to give a detectable echo.

1. Introduction

Ultrasonography (pulse ultrasonography) gives valuable information on the human tissues investigated, permitting the detection and visualization of the boundaries of the tissues, both normal and pathological, and also visuali-

zation of the texture of these tissues. This is possible as a result of specular reflections arising from the boundaries of the tissues, and from the echoes caused by diffuse reflections, when the size of the anatomical structures is small compared to the wavelength. The present paper attempts to estimate quantitatively the detectability, using an ultrasonographic method, of a blood vessel within human soft tissues. The diameter of the vessel is assumed to be several times smaller than the ultrasonic wavelength so that the reflection of the wave is diffuse in character.

At the same time the minimum difference in the (characteristic) acoustic impedance of the tissue that is necessary to detect plane boundaries with a normally incident wave, is determined.

2. The assumptions of the analysis

It is assumed that a blood vessel with a radius $a = 0.1$ mm is in the initial ultrasonic region of the far field generated by a plane piezoelectric transducer (transmit-receive transducer) with a diameter of 2 cm and a frequency $f = 2.5$ MHz (wavelength $\lambda = 0.63$ mm).

Under the conditions assumed the boundary between the near and far fields is 16 cm. At a distance $r_0 = 20$ cm from the transducer one can assume that the acoustic pressure in this region of the field is approximately equal to the acoustic pressure averaged over the whole area of the ultrasonic beam near the transducer. The blood vessel is on the axis of the ultrasonic beam and the axis of the vessel is perpendicular to the axis of the beam. It is assumed that the ultrasonic wave falling onto the vessel is a plane, homogeneous wave, although this is essentially not satisfied in practice.

The calculations were carried out for a continuous wave despite the use of the pulse echo method in ultrasonography. However, since the diameter of the vessel is small compared to the wavelength and since the impedances of the soft tissue and blood hardly differ, the transient time becomes very short. In this case the assumption of a steady state introduces no serious error.

It was also assumed that the elastic properties of the tissues are the same as those of the liquid and also that the surrounding tissue and the walls of the vessels are sufficiently homogeneous for no reflections to occur inside them.

3. The reflection of a plane wave from a blood vessel

The calculations were made for a plane polar coordinate system where r is the radius, and θ is the azimuth. The axis of the coordinate system coincides with the axis of the blood vessel. The incident plane wave can be re-

presented here as a series of cylindrical functions [5]

$$\varphi_i = \varphi_M\left[J_0(k_t r) + 2\sum_1^\infty (-j)^m J_m(k_t r)\cos(m\theta)\right] e^{j\omega t}, \tag{1}$$

where φ_i is the acoustic potential of the incident wave, φ_M is the potential amplitude, J_M is a Bessel function of order m, m is an integer, $k_t = \omega/c_t$ is the wave number, $\omega = 2\pi f$, f is the frequency, and c_t is the wave velocity in the tissue.

The acoustic potentials of the reflected wave and of the wave penetrating into the vessel are of the forms respectively

$$\varphi_r = \sum_0^\infty A_m H_m^{(2)}(k_t r)\cos(m\theta)\, e^{j\omega t}, \tag{2}$$

$$\varphi_t = \sum_0^\infty B_m J_m(k_b r)\cos(m\theta)\, e^{j\omega t}, \tag{3}$$

where $H_m^{(2)}$ is a Hankel function of the second kind of order m, $k_b = \omega/c_b$ is the wave number, and c_b is the wave velocity in blood.

The functions $J_m(kr)$ and $H_m^{(2)}(kr)$ are solutions of the wave equation in the cylindrical coordinate system for the variable r. The function $H_m^{(2)}$ represents the wave propagating in the direction of increasing radius r. However, this function is not specified for r tending to zero. The wave penetrating into the blood vessel was therefore described by the function J_m. The constants A_m and B_m are determined from two boundary conditions, which must be satisfied on the perimeter of the vessel $r = a$. These conditions consist in the equality of the pressures p and the normal components of acoustic velocity, v_r, which are connected with the acoustic potential the by relations

$$p = \varrho \frac{\partial \varphi}{\partial t}, \tag{4}$$

$$v_r = -\operatorname{grad}_r \varphi, \tag{5}$$

where ϱ is the density of the medium.

All the tangential stresses were assumed to be equal to zero. It was thus assumed that the elastic properties of the blood vessel and the surrounding tissue are the same as those of a liquid.

This is certainly an approximation, since tangential stresses, e.g., those maintaining the cylindrical shape of the vessel, do occur in the walls of the vessel.

The acoustic pressure and velocity in the surrounding medium can be obtained from expressions (1) and (2), and for the internal medium from expres-

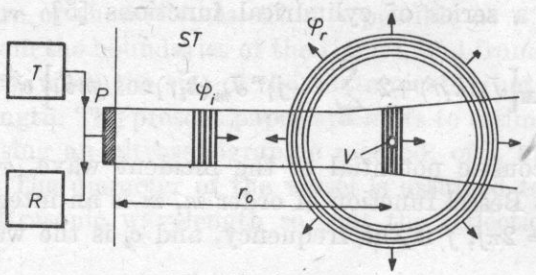

Fig. 1. Diffuse reflection of an ultrasonic pulse from a small blood vessel

T – transmitter, R – receiver, P – ultrasonic probe with a piezoelectric transducer, v – small blood vessel, φ_i – incident wave, φ_r – reflected wave (diffuse), ST – soft tissue, r_0 – distance between the piezoelectric transducer and the blood vessel

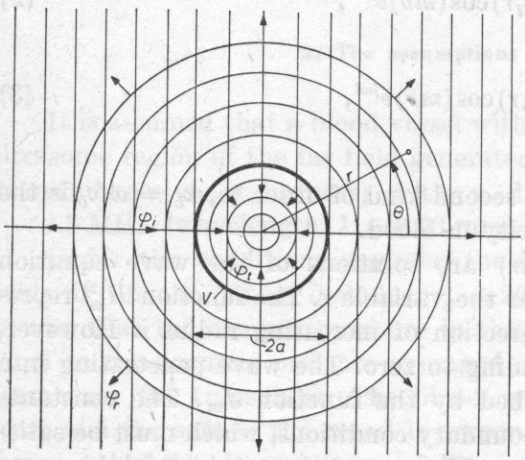

Fig. 2. A blood vessel (V) with incident (φ_i), reflected (φ_r) and penetrating (φ_t) waves

r, θ – cylindrical coordinates

sion (3). Bearing in mind the boundary conditions we obtained from equations (1)-(3) (for $r = a$)

$$\varphi_M\left[J_0(k_t a)+2\sum_1^{\infty}(-j)^m J_m(k_t a)\cos(m\theta)+\sum_0^{\infty}A_m H_m^{(2)}(k_t a)\cos(m\theta)\right.$$

$$=\frac{\varrho_b}{\varrho_t}\sum_0^{\infty}B_m J_m(k_b a)\cos(m\theta); \qquad (6)$$

$$\varphi_M\left\{-k_t J_1(k_t a)+2\sum_1^{\infty}(-j)^m\left[\frac{m}{a}J_m(k_t a)-k_t J_{m+1}(k_t a)\right]\cos(m\theta)\right\}+$$

$$+\sum_0^{\infty}A_m\left[\frac{m}{a}H_m^{(2)}(k_t a)-k_t H_{m+1}^{(2)}(k_t a)\right]\cos(m\theta)$$

$$=\sum_0^{\infty}B_m\left[\frac{m}{a}J_m(k_b-a)k_b J_{m+1}(k_b a)\right]\cos(m\theta). \qquad (7)$$

Multiplying the above relations by $\cos(n\theta)$ $(n = 0, 1, 2, \ldots)$ and integrating with respect to θ from 0 to π we eliminate all the terms for which $n \neq m$. Only those terms remain, for which $n = m$. In addition the dependence on the angle θ disappears. Thus we can successively determine from each of relations (6) and (7), the values of the constants A_m and B_m for each m [5].

Introducing into the calculations the values corresponding to the muscle tissue of the uterus and blood [3], i.e. $\varrho_t/\varrho_b = 1$ and $c_t = 1.63$ km/s, $c_b = 1.57$ km/s we obtain the following values for the constants

$$A_0 = 0.0495\, e^{-j93^\circ}, \quad A_1 = 0.0136\, e^{j180^\circ}, \quad A_2 = 0.0003\, e^{j90^\circ}, \; A_3 = 0. \tag{7a}$$

The constants A_m with higher indices quickly decrease in magnitude so they can be neglected. In order to determine the reflected wave, the expression approximating a Hankel function for large values of the parameters x [4] can be used,

$$H_m^{(2)}(x) \underset{x\to\infty}{=} \sqrt{\frac{2}{\pi x}}\, e^{-j[x-(2m+1)\pi/4]}. \tag{8}$$

Finally, neglecting the time factor, we obtain the acoustic potential of the reflected wave (2) in the form

$$\varphi_r = \varphi_M(A_0 e^{j\pi/4} + A_1 e^{j3\pi/4}\cos\theta + A_2 e^{j5\pi/4}\cos 2\theta + A_3 e^{j7\pi/4}\cos 3\theta)\sqrt{\frac{2}{\pi k_t r}}\, e^{-jk_t r}$$

$$= \varphi_M R D e^{-jk_t r}, \tag{9}$$

where R denotes the value of the expression in the brackets, and D is the square root expression in equation (9).

4. The ratio of the power of the signal received to the power of the signal transmitted

The power of the transmitted signal N_T can be expressed in terms of the acoustic potential in the following way

$$N_T = S\frac{|p|^2}{2\varrho_t c_t} = S\frac{\omega^2 \varrho_t}{2c_t}|\varphi_M|^2, \tag{10}$$

where S is the surface area of the piezoelectric transducer.

A similar expression can be used for the power of the signal received, N_R, by considering the blood vessel to be in the far field so that the wave falling onto the plane piezoelectric transducer is approximately a plane wave. Then the ratio of the power of the received and transmitted waves is, with consideration of (9) and (7a), equal to

$$\frac{N_R}{N_T} = \frac{|\varphi_r|^2}{|\varphi_M|^2} = R^2 D^2 = 0.051^2 \cdot 0.018^2. \tag{11}$$

Using the above formula for numerical calculations r was replaced by the distance of the vessel from the transducer $r_0 = 20$ cm, and the cosine functions were approximated-1 in formula (9), since the angle at which the piezoelectric transducer is "seen" is only 3°. R and D are the losses of the signal occurring respectively due to the reflection and the divergence of the reflected wave.

Formula (11) is only valid for a plane wave. When the vessel v is small (Fig. 3) the reflected wave is a cylindrical wave according to expression (9).

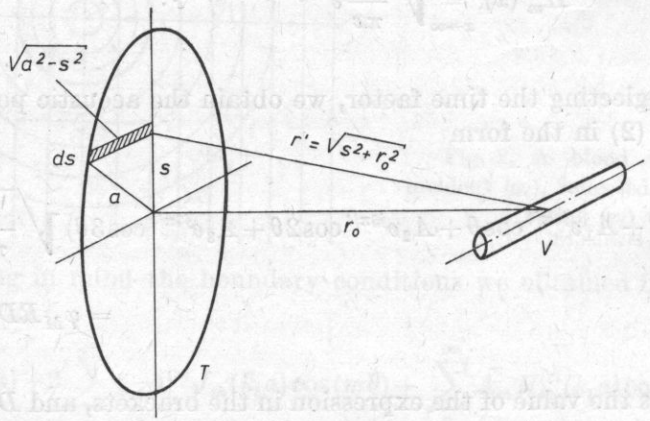

Fig. 3. Geometrical relationships in a cylindrical wave reflected from a vessel
P – a plane piezoelectric transducer, v – a blood vessel, r_0 – the distance from the vessel to the transducer, a – the radius of the transducer, s – the current coordinate

This wave propagating along the path r' falls onto a planar transducer T with a different phase than the wave propagating on the path r_0. The phase difference between the waves is

$$\frac{2\pi}{\lambda}(r' - r_0) = \frac{2\pi}{\lambda}(\sqrt{s^2 - r_0^2} - r_0).$$

As a result of this phenomenon a decrease of the elementary cylindrical waves reflected from the vessel v and falling onto the piezoelectric transducer T, occurs. Thus an electric signal, which is proportional to the value of the acoustic

pressure averaged over the whole surface area of the transducer, occurs on the electrical terminals. From such averaging we obtain a coefficient of proportionality

$$N = \sqrt{\frac{4}{\pi a^2}} \int_0^a e^{-j\frac{2\pi}{\lambda}\sqrt{s^2+r_0}-r_0} \sqrt{a^2-s^2}\,ds. \tag{11a}$$

It is responsible for the lack of parallelism between the equiphasal surface of the reflected wave and the surface area of the transducer. For the conditions assumed here with $r_0 = 20$ cm it takes the value $N = 0.92$.

Thus finally formula (11) can be expanded by the introduction of the factor N and represented on a logarithmic scale (with the numerical data, respectively),

$$\frac{N_R}{N_T} \doteqdot R\,[\mathrm{dB}] + D\,[\mathrm{dB}] + N\,[\mathrm{dB}] = -29\mathrm{dB} - 35\mathrm{dB} - 1\mathrm{dB} = 65\mathrm{dB}. \tag{12}$$

5. The electrical parameters of the apparatus

When one assumes $U_o = 10\ \mu$V as a typical voltage sensitivity of the ultrasonograph receiver, and the output voltage of the transmitter as $U_n = 250$ V, then the ratio of these two voltages is equal to $W = U_n/U_o = 2.5 \cdot 10^7 \doteqdot 148$ dB. Considering the losses due to the transducing in the transmission and reception of the electric and ultrasonic pulses (including the diffraction losses) to be equal to $T = -15$ dB, then the ratio of the amplitude of the minimum detectable echo to the amplitude of the pulse radiated corresponds to $W - T = 133$ dB.

Since in the present case this value is higher them the value of (12) expressed in dB, the blood vessel would be detectable. However, it is necessary to account for the attenuation of the tissue.

6. Detectability of the vessel

A detailed survey of the literature on the ultrasonic properties of tissues and organs shows over 140 papers containing data on attenuation [3]. One should, however, state a general lack of data obtained from measurements "in vivo". The present author therefore decided to use the "in vivo" measurements obtained in a typical situation for ultrasonic visualization in obstetrics [1]. Attenuation in the penetrated tissues of the abdomen was found on average to be 1.8 dB/cm at a frequency of 2.5 MHz. In this case the signal loss on the path of 40 cm (from the transducer to the blood vessel and back) would be $A = -72$ dB.

Graph 4 (left) shows the levels of the signals, starting from the signal transmitted, to the signal received by the ultrasonic pulse device (echoscope) with consideration of all potential sources of signal loss. It can be seen that it is not possible to detect a blood vessel with a radius of 0.1 mm when it is at a distance $r_0 = 20$ cm. The signal received has then a level 4 dB lower than the sensitivity level of the apparatus. However, when one decrease the distance of the vessel to the value $r_0 = 10$ cm, the attenuation losses in tissues are only $A = -36$ dB, and the losses due to the divergence of the reflected beam would decrease to a value of $D = -32$ dB. The losses due to the lack of parallelism between the surface of the equiphasal reflected wave with respect to the surface area of the transducer would, however, increase to the value of $N = -6$ dB. It follows from graph 4 (right) that the blood vessel would be detected, and the signal would be 30 dB above the noise level.

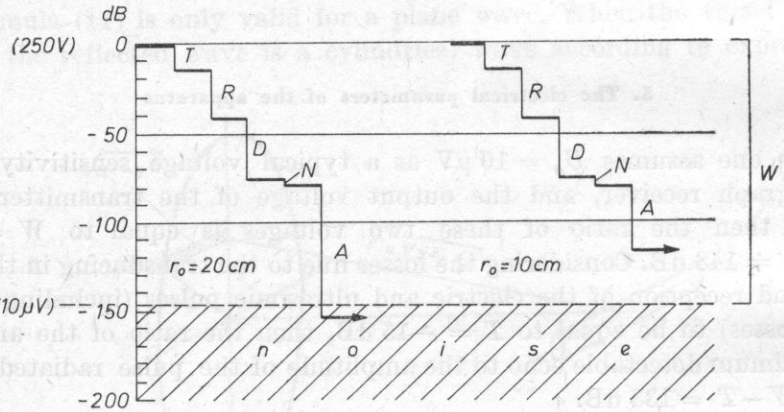

Fig. 4. The signal losses in the detection of a blood vessel with a radius $a = 0.1$ mm at different distances r_0 from the transducer

T – the electroacoustic transducing losses (of transmission and reception together), R – the losses in the reflection from the blood vessel, D – the losses in the divergence of the reflected wave (diffuse wave), A – the losses in the attenuation of the wave in tissue, N – the losses due to the lack of parallelism between the equiphasal wave reflected and the surface of the transducer, W – the ratio of the output voltage of the transmitter to the voltage sensitivity of the receiver

The fact that in reality the tissues penetrated are heterogeneous should, however, be considered. A number of echos occur due to the scattering of the ultrasonic wave from small structures, such as muscle fibres (diameter 10-150 μm, length 1-20 mm), or other small blood vessels (arterioles and capillaries of a diameter of 8-200 μm). The signal reflected from the single blood vessel considered may then be undetectable amongst the other signals with a random spatial distribution and similar amplitude.

7. Detectability of a plane boundary in soft tissues

Assuming normal incidence for an ultrasonic wave impinging on the plane boundary between two soft tissues of densities (ϱ) and wave velocities (c) that are only slightly different from each other so that their characteristic acoustic impedances are equal to ϱc and $\varrho c' = \varrho c + \Delta\varrho c$, one can determine the ratio of the potential of the reflected wave to the potential of the incident wave from the formula

$$\frac{\varphi_r}{\varphi_M} = \frac{\varrho c' - \varrho c}{\varrho c' + \varrho c} \cong \frac{\Delta\varrho c}{2\varrho c}. \tag{13}$$

Assuming accordingly the distance of the boundary of the tissues from the transducer to be $r_0 = 20$ cm, and thus the attenuation losses in the tissue to be $A = -72$ dB, one can determine the minimum detectable value of the ratio of the potentials

$$\frac{\varphi_r}{\varphi_M} \doteq (W - T - A) \quad [\text{dB}] = -61\ \text{dB}. \tag{14}$$

From a comparison of (13) and (14) we finally obtain a minimum detectable change in the acoustic impedance of the tissue, i.e. $\Delta\varrho c/\varrho c = 1.8 \cdot 10^{-3}$. At a distance $r_0 = 10$ cm, however, when the attenuation losses in the tissue are only $A = -36$ dB, this quantity is $\Delta\varrho c/\varrho c = 2.8 \cdot 10^{-5}$.

8. Conclusions

This analysis shows that even very small blood vessels of a radius 0.1 mm give potentially detectable signals at a frequency of 2.5 MHz. The level of these signals is close to the noise level of the apparatus and depends critically on the distance of the blood vessel from the transducer because of attenuation in the tissues penetrated.

The present calculations have an approximate character in view of a number of simplifying assumptions, which were necessary for the analysis to be performed.

Detection of blood vessels of larger radii is not difficult. Fig. 5 shows a longitudinal ultrasonogram of a pregnant woman with a visible placenta including a number of blood vessels of very small diameters [2]. Fig. 6 shows an ultrasonogram of the abdominal aorta with a large aneurism. Both ultrasonograms were obtained at a frequency of 2.5 MHz.

The detectability of plane boundaries of soft tissues situated perpendicular to the direction of the falling waves is very large. At a distance 20 cm from the transducer, differences in the acoustic impedances of even 0.2 % are potentially detectable. At shorter distances this detectability rapidly increases due to the smaller attenuation losses of the wave in the tissues penetrated.

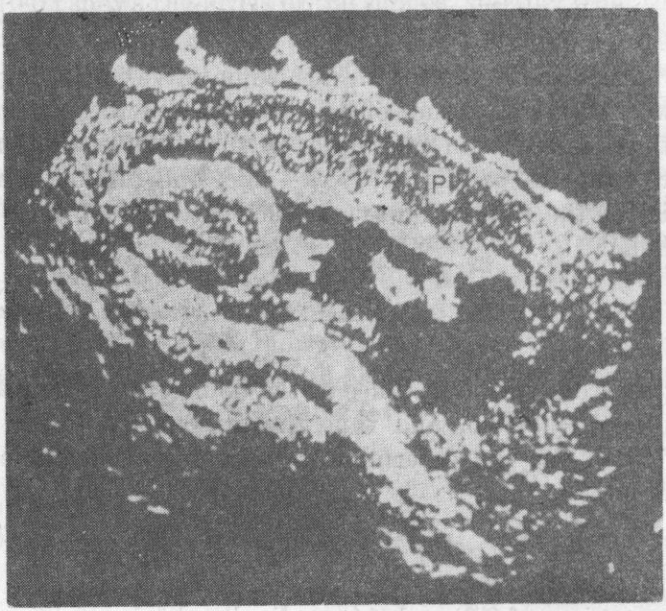

Fig. 5. A longitudinal ultrasonogram of a pregnant woman with a visible placenta *P*, obtained with a Polish USG – 10 apparatus [2]

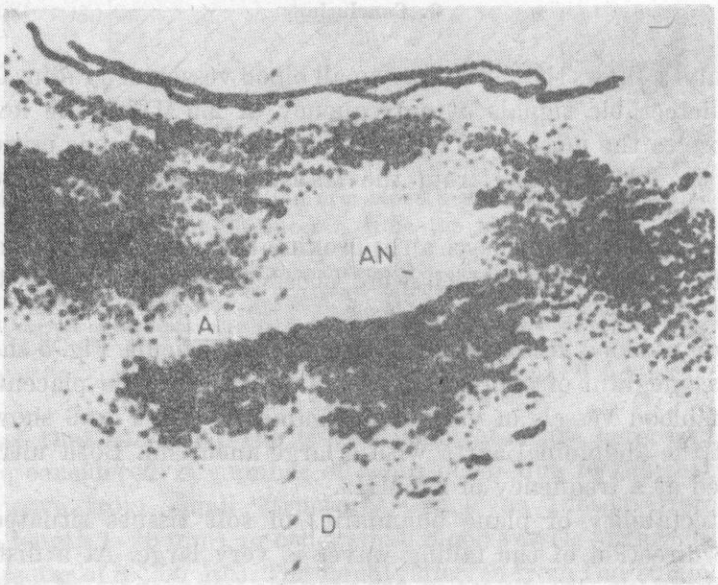

Fig. 6. An ultrasonogram of the abdominal aorta *A*, obtained with the USG-10 apparatus [2]
AN – aneurism, *D* – 1 cm distance markers

References

[1] J. ETIENNE, L. FILIPCZYŃSKI, A. FIREK, J. GRONIOWSKI, J. KRETOWICZ, G. ŁYPACEWICZ, J. SAŁKOWSKI, *Intensity determination of ultrasonic focused beams used in ultrasonography in the case of gravid uterus*, Ultrasound in Medicine and Biology, **2**, 119–122 (1967); and Archiwum Akustyki, **11**, 1, 35–43 (1976) (in Polish).

[2] L. FILIPCZYŃSKI, L. ROSZKOWSKI (ed.), *Ultrasonic diagnostic in obstetrics and gynaecology* (in Polish), PZWL, Warszawa 1977.

[3] S. A. GOSS, R. L. JOHNSTON, F. DUNN, *Comprehensive compilation of empirical ultrasonic properties of mamalian tissues*, JASA, **64**, 2, 423–457 (1978).

[4] W. MAGNUS, F. OBERHETTINGER, *Formeln und Sätze für die speziellen Funktionen der mathematischen Physik*, Springer, Berlin.

[5] P. M. MORSE, *Vibration and sound*, McGraw Hill, New York 1948.

Received on January 15, 1980.

ARCHIVES OF ACOUSTICS
6, 1, 57-62 (1981)

PROPAGATION OF ACOUSTIC WAVE ALONG A HOLLOW CYLINDER IMMERSED IN A LIQUID

ANNA GRABOWSKA

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

The problem of the propagation of an nonabsorbed, continuous, progressive and axially-symmetric acoustic wave along an infinite homogeneous and isotropic cylinder filled with air and immersed in an ideal liquid has been considered. The wave equations of displacement potentials have been solved. The characteristic equation has been derived for the preset boundary conditions and solved numerically for the selected data characteristic for the conditions of the biopsy performed in an ultrasonic field. It has been shown that a wave guided along a needle immersed in a liquid can propagate with the velocity only slightly smaller than the wave velocity of the surrounding liquid. The distributions of displacement, stresses and acoustic pressure of the propagating wave have been determined.

1. Introduction

The problem of the acoustic wave propagation along a hollow cylinder immersed in a liquid is an attempt to describe the wave phenomena occurring in the puncture of organs with a needle, as used in ultrasonic medical diagnostics.

A biological structure to be investigated is localized with an ultrasonic beam and the needle is subsequently entered into the body toward the biological structure through the hole in the piezoelectric transducer. It has been found that a wave propagates along the needle after it has been entered into the body. The wave after reaching the end of the needle returns, giving the image of the needle end on the oscilloscope screen.

The aim of this paper is to describe the phenomena accompanying the propagation of this wave. In the first approximation the needle was considered a solid layer [2]. It is now defined as an infinite elastic, isotropic and homogeneous hollow cylinder with the Lamé constants λ, μ and the density ϱ and

surrounded by an ideal liquid. The external and internal radii of the cylinder are a and b respectively, and the velocities of longitudinal and transverse waves in the material are c_d and c_t respectively. The liquid density is ϱ_0 and the longi-

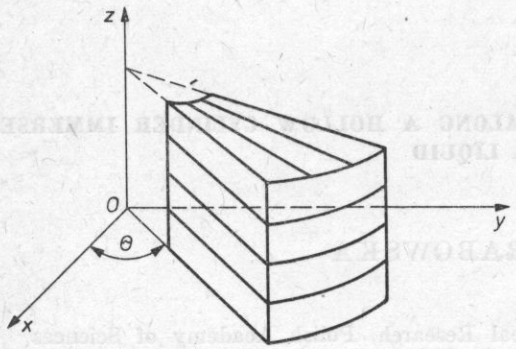

Fig. 1. A sector of circular section of the hollow cylinder

tudinal wave velocity in the liquid is c_0. The air inside the cylinder is regarded as the vacuum. This paper considers the case when a running, continuous and axially – symmetric wave propagates along the z-axis which coincides with the axis of the hollow cylinder.

2. Basic equations

In the cylindrical coordinate system displacements, strains and stresses are independent of the angle θ in the case of an axially – symmetric deformation with respect to the z-axis. The displacements in the hollow cylinder have the form

$$u_r = u_r(r, z, t), \qquad u_\theta = 0, \qquad u_z = u_z(r, z, t). \tag{1}$$

For harmonic vibrations the solutions of the wave equations (of displacement potentials):

$$\nabla^2 \Phi = \frac{1}{c_d^2} \frac{\partial^2 \Phi}{\partial t^2}, \qquad \nabla^2 \Psi = \frac{1}{c_t^2} \frac{\partial^2 \Psi}{\partial t^2} \tag{2}$$

were assumed in the form [3]:

$$\begin{aligned} \Phi &= [A_1 J_0(k_d r) + A_2 Y_0(k_d r)] e^{i\omega t - ikz}, \\ \Psi &= [B_1 J_0(k_t r) + B_2 Y_0(k_t r)] e^{i\omega t - ikz}, \end{aligned} \tag{3}$$

where

$$k_d^2 = \frac{\omega^2}{c_d^2} - k^2, \qquad k_t^2 = \frac{\omega^2}{c_t^2} - k^2, \tag{4}$$

$\omega = 2\pi f$, f – frequency, k – propagation constant, A_1, A_2, B_1, B_2 – constants, J_n and Y_n are the Bessel and Neuman functions respectively. Displacements and stresses are expressed in the cylindrical coordinate system as follows:

$$u_r = \frac{\partial \Phi}{\partial r} - \frac{\partial^2 \Psi}{\partial r \partial z}, \qquad \tau_{rr} = \lambda \left(\frac{\partial u_r}{\partial r} + \frac{u_r}{r} + \frac{\partial u_z}{\partial z} \right) + 2\mu \frac{\partial u_r}{\partial r},$$
$$u_z = \frac{\partial \Phi}{\partial z} - \frac{\partial^2 \Psi}{\partial^2 r} - \frac{1}{r} \frac{\partial \Psi}{\partial r}, \qquad \tau_{rz} = \mu \left(\frac{\partial u_r}{\partial z} + \frac{\partial u_z}{\partial r} \right). \tag{5}$$

For the liquid surrounding the needle we assume the following solution of the wave equation (of displacement potential)

$$\nabla^2 \Phi_0 = \frac{1}{c_0^2} \frac{\partial^2 \Phi_0}{\partial^2 t} \tag{6}$$

in the form

$$\Phi_0 = A^0 H_0^{(2)}(k_0 r) e^{i\omega t - ikz}, \qquad k_0^2 = \frac{\omega^2}{c_0^2} - k^2, \tag{7}$$

where $H_0^{(2)}(k_0 r) = J_0(k_0 r) - i Y_0(k_0 r)$. Generally the Hankel function of the second kind represents a wave travelling in the direction of increasing r. For the imaginary and negative wave number k_0 it results from the asymptotic representation

$$H_0^{(2)}(k_0 r) \xrightarrow[r \to \infty]{} \sqrt{\frac{2}{k_0 r \pi}} \exp\left[-i \left(k_0 r - \frac{\pi}{4} \right) \right] \tag{8}$$

that the wave is not radiated but decays exponentially with increasing r; thus it is guided along the needle. Our present analysis considers only this case. The displacements u_r^0, u_z^0 and acoustic pressure in liquid are described by the relations

$$u_r^0 = \frac{\partial \Phi_0}{\partial r}, \qquad u_z^0 = \frac{\partial \Phi_0}{\partial z}, \qquad p = -\varrho_0 \frac{\partial^2 \Phi_0}{\partial t^2}. \tag{9}$$

3. Boundary conditions and charakteristic equation

The following boundary conditions should be satisfied on the surfaces of the hollow cylinder

$$\tau_{rr} = -p, \quad \tau_{rz} = 0, \quad u_r = u_r^0 \quad \text{for } r = a,$$
$$\tau_{rr} = 0, \quad \tau_{rz} = 0, \quad \text{for } r = b. \tag{10}$$

After inserting relations (5) and (9) into (10) the system of five uniform equations with the unknowns A_1, A_2, B_1, B_2, A^0 is obtained

$$A_1\left[(\omega^2\varrho-2\mu k^2)J_0(k_d a)-\frac{2\mu k_d}{a}J_1(k_d a)\right]+$$

$$+A_2\left[(\omega^2\varrho-2\mu k^2)\,Y_0(k_d a)-\frac{2\mu k_d}{a}\,Y_1(k_d a)\right]+B_1\frac{2\mu i k k_t}{a}\,[J_1(k_t a)-k_t a J_0(k_t a)+$$

$$+B_2\frac{2\mu i k k_t}{a}\,[Y_1(k_t a)-k_t a Y_0(k_t a)]-\varrho_0\omega^2 A^0 H_0^{(2)}(k_0 a)=0,$$

$$2A_1 k_d i k\mu J_1(k_d a)+2A_2 k_d i k\mu Y_1(k_d a)+B_1 k_t(k^2-k_t^2)\mu J_1(k_t a)+$$

$$+B_2 k_t\mu(k^2-k_t^2)\,Y_1(k_t a)=0, \qquad (11)$$

$$-A_1 k_d J_1(k_d a)-A_2 k_d\,Y_1(k_d a)+B_1 k_t k i J_1(k_t a)+B_2 k_t k i Y_1(k_t a)+$$

$$+A^0 k_0 H_1^{(2)}(k_0 a)=0,$$

$$A_1\left[(\omega^2\varrho-2\mu k^2)J_0(k_d b)-\frac{2\mu k_d}{b}J_1(k_d b)\right]+A_2\left[(\omega^2\varrho-2\mu k^2)\,Y_0(k_d b)-\right.$$

$$\left.-\frac{2\mu k_d}{b}\,Y_1(k_d b)\right]+B_1\frac{2\mu i k k_t}{b}\,[J_1(k_t b)-k_t b J_0(k_t b)]+B_2\,\frac{2\mu i k k_t}{b}\,[Y_1(k_1 b)-$$

$$-k_t b\,Y_0(k_t b)]=0,$$

$$2A_1 k_d k i\mu J_1(k_d b)+2A k_d k i\mu Y_1(k_d b)+B_1 k_t(k^2-k_t^2)\mu J_1(k_t b)+$$

$$+B_2 k_t\mu(k^2-k_t^2)\,Y_1(k_t b)=0.$$

By eliminating the constants A_1, A_2, B_1, B_2 and A^0 from the system of equations (11) one obtains the characteristic equation

$$|a_{ij}|=0 \qquad (i,j=1,\ldots,5), \qquad (12)$$

where a_{ij} are coefficients of the constants A_1, A_2, B_2 and A^0. The sought propagation constant k occurs explicitely in (12) and also in k_d, k_t, k_0 (4), (7) and in the arguments of the Bessel and Neuman functions. In general, those complex values of k can occur in solutions (12), for which Im (k) defines attenuation of the wave along the z-axis, while imaginary values of k correspond to the unpropagating wave modes. Only the real propagation constants k were considered in the solution of the characteristic equation, which results from the assumption made that no attenuation occurs in the solid and the liquid.

The characteristic equation was solved numerically for the following data: $f=3$ MHz, $a=0.75$ mm, $b=0.5$ mm, a steel needle of the density $\varrho=7.7$ g/cm^2, $c_d=5.9$ km/s, $c_t=3.23$ km/s, $\lambda=1.07\cdot10^{12}$ g/(cms^2), $\mu=8.03\cdot10^{11}$ g/(cms^2). Water of the density $\varrho_0=1$ g/cm^3 was taken as a liquid, and $c_0=1.48$ km/s.

The previous experiments [4] showed that the velocity of this wave is close to the wave velocity in the surrounding liquid. Therefore such c was sougth that $c<c_0$, since for $c>c_0$ the wave would not be guided along the z-axis (compare (8) and (7)). The velocity of the wave propagating along the hollow cylinder, $c=1.44$ km/s, was obtained. The distributions of displace-

ments, stresses and acoustic pressure in the needle and the liquid, calculated from (5) and (9) are shown in Figs. 2 and 3.

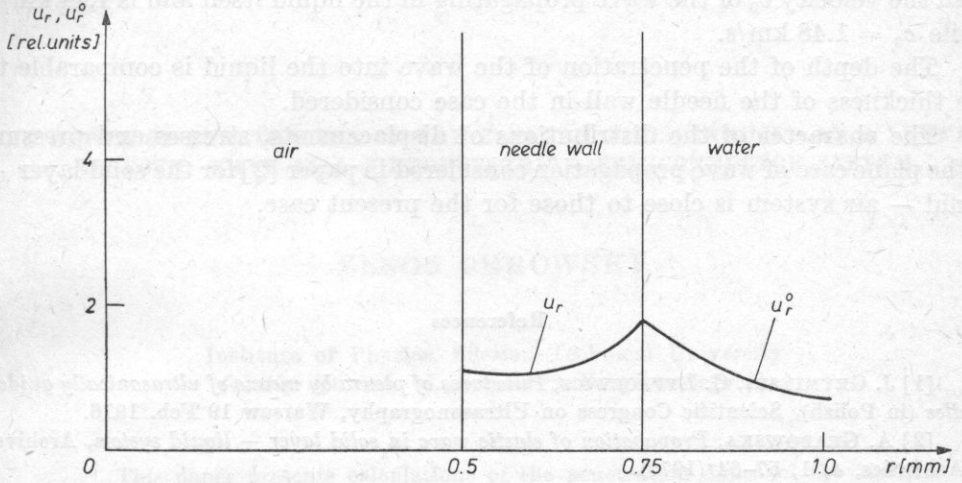

Fig. 2. The distribution of displacement in the needle and the liquid

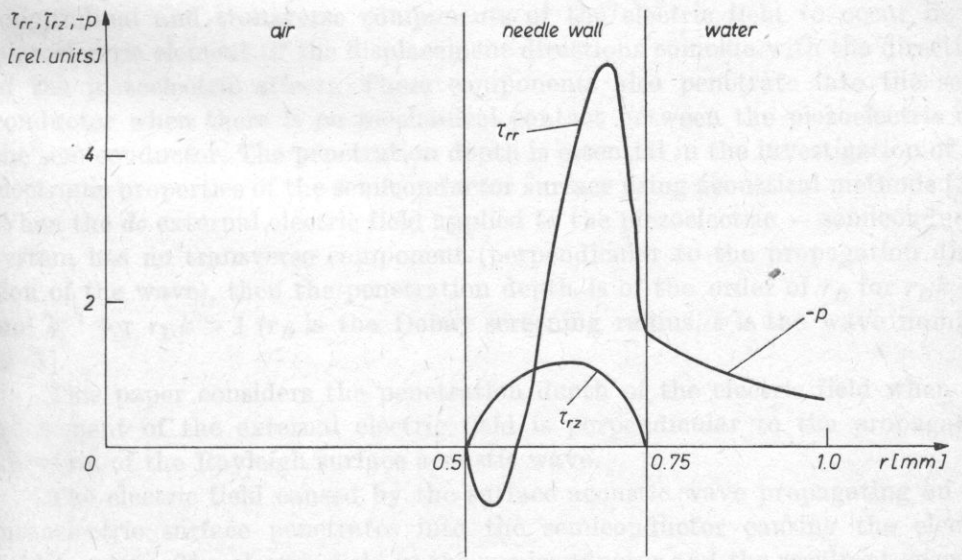

Fig. 3. The distributions of stresses in the needle and water

4. Conclusions

It has been shown that a wave can propagate along the needle immersed in a liquid. The phase velocity c of the progressive wave is then slightly smaller than the velocity c_0 of the wave propagating in the liquid itself and is 1.44 km/s, while $c_0 = 1.48$ km/s.

The depth of the penetration of the wave into the liquid is comparable to the thickness of the needle wall in the case considered.

The character of the distributions of displacements, stresses and pressure in the plane case of wave propagation considered in paper [2] for the solid layer – liquid – air system is close to those for the present case.

References

[1] J. GRYMIŃSKI, G. ŁYPACEWICZ, *Punctures of pleura by means of ultrasonically guided needles* (in Polish), Scientific Congress on Ultrasonography, Warsaw 19 Feb. 1976.

[2] A. GRABOWSKA, *Propagation of elastic wave in solid layer – liquid system*, Archives of Acoustics, **4**, 1, 57–64 (1979).

[3] M. REDWOOD, *Mechanical waveguides*, Pergamon Press, Oxford 1960.

[4] L. FILIPCZYŃSKI, *Ultrasonic wave propagation along the surface of a rod immersed in a liquid*, Archives of Acoustics, **4**, 3, 271–286 (1979).

Received on December 20, 1979.

ARCHIVES OF ACOUSTICS
6, 1, 63-68 (1981)

THE DEPTH OF PENETRATION OF THE ELECTRIC FIELD INDUCED BY A SURFACE ACOUSTIC WAVE IN A PIEZOELECTRIC — SEMICONDUCTOR SYSTEM

ZENON CEROWSKI

Institute of Physics, Silesian Technical University
(44-100 Gliwice, ul. Bolesława Krzywoustego 2)

This paper presents calculations of the penetration depth of the electric field accompanying an acoustic wave for penetration into a semiconductor layer over the plane of propagation when a *dc* electric field is applied to the piezoelectric – semiconductor system perpendicular to the propagation direction of the wave.

In a piezoelectric – semiconductor system the wave propagating on the surface of the piezoelectric, which has two displacement components, causes longitudinal and transverse components of the electric field to occur in the piezoelectric element (if the displacement directions coincide with the directions of the piezoelectric effect). These components also penetrate into the semiconductor when there is no mechanical contact between the piezoelectric and the semiconductor. The penetration depth is essential in the investigation of the electronic properties of the semiconductor surface using acoustical methods [17]. When the *dc* external electric field applied to the piezoelectric – semiconductor system has no transverse component (perpendicular to the propagation direction of the wave), then the penetration depth is of the order of r_D for $r_D k < 1$ and k^{-1} for $r_D k > 1$ (r_D is the Debay screening radius, k is the wave number) [1-3].

This paper considers the penetration depth of the electric field when the component of the external electric field is perpendicular to the propagation direction of the Rayleigh surface acoustic wave.

The electric field caused by the surface acoustic wave propagating on the piezoelectric surface penetrates into the semiconductor causing the electric field to occur. The electric field in the semiconductor and the resultant currents are described by Poisson's equations, the current equations and the equations

of the continuity of current

$$\varepsilon_0 \varepsilon \Delta\varphi = qn, \tag{1}$$

$$\boldsymbol{j} = \mu q(n_0+n)(\boldsymbol{E}+\boldsymbol{E}^0)+qD\nabla n, \tag{2}$$

$$\nabla \boldsymbol{j} = qn_t, \tag{3}$$

where n_0 is the density of carriers in the conduction band without the acoustic wave, n is the change in the density of carriers in the conduction band when the acoustic wave propagates, φ is the electric potential in the semiconductor, $\boldsymbol{j}$ is the current density in the semiconductor, $\boldsymbol{E} = \boldsymbol{E}_1+\boldsymbol{E}_3$ is the strength of the electric field in the semiconductor due to the propagation of the wave, $E_1 = -\partial\varphi/\partial x_1$ is the component of the field in the direction of propagation of the wave, $E_3 = -\partial\varphi/\partial x_3$ is the component perpendicular to the direction of propagation of the wave, $\boldsymbol{E}^0 = \boldsymbol{E}_1^0+\boldsymbol{E}_3^0$ is the strength of the external *dc* electric field, E_0^1 is the component along the direction of propagation, ε, ε_0 are the dielectric constants in a vacuum and in the semiconductor respectively, q is the charge of the carriers, μ is the mobility of the carriers, D is the diffusion coefficient, and E_3^0 is the component perpendicular to the direction of propagation.

From equations (1)-(3) and considering that n changes in the following way

$$e^{[i(kx_1-\omega t)-kax_3]}, \tag{4}$$

where x_1 is the direction of propagation of wave, x_3 is the direction perpendicular to the direction of propagation of the wave, ω is the angular frequency of the wave, k is the wave number, a is the coefficient of the depth of penetration into the semiconductor, the following equation for a can be obtained

$$a^2 \pm \frac{\mu k E_3^0}{\omega}\frac{\omega_D}{\omega}a - \left(1+\frac{1}{r_0^2k^2}\right)+i\gamma\frac{\omega_D}{\omega} = 0, \tag{5}$$

where $\omega_D = \omega^2/k^2D$ is the diffusion frequency, $\omega_C = \sigma_0/\varepsilon_0\varepsilon = \mu q n_0/\varepsilon_0\varepsilon$ is conduction relaxation frequency, $r_D k = \sqrt{\omega^2/\omega_D\omega_C}$, and $\gamma = 1+\mu k E_1^0/\omega$ is the drift parameter.

It follows from equation (5) that

$$a = a_1 + i a_2 \tag{6}$$

is a complex number. Accordingly equation (5) can take the form

$$a_1^2 - a_2^2 \pm \frac{\mu k E_3^0}{\omega}\frac{\omega_D}{\omega}a_1 - \left(1+\frac{1}{r_D^2k^2}\right) = 0, \tag{7a}$$

$$2a_1a_2 \pm \frac{\mu k E_3^0}{\omega}\frac{\omega_D}{\omega}a+\gamma\frac{\omega_D}{\omega} = 0. \tag{7b}$$

The real part of the penetration coefficient a_1 plays an important role in the

depth of penetration of the field. Solving equation (7*a*) and neglecting a_2 (i.e. without considering the oscillatory character of the decay of the field) one obtains

$$a_1 = \frac{\pm \frac{\mu k E_3^0}{\omega} \frac{\omega_D}{\omega} + \sqrt{\left(\frac{\mu k E_3^0}{\omega} \frac{\omega_D}{\omega}\right)^2 + 4\left(1 + \frac{1}{r_D^2 k^2}\right)}}{2}. \tag{8}$$

The variation of the value of a_1 as a function of $r_D k$ for different E_3^0 for $\gamma = 0$ and, as a result, for $a_2 = 0$ is shown in Fig. 1.

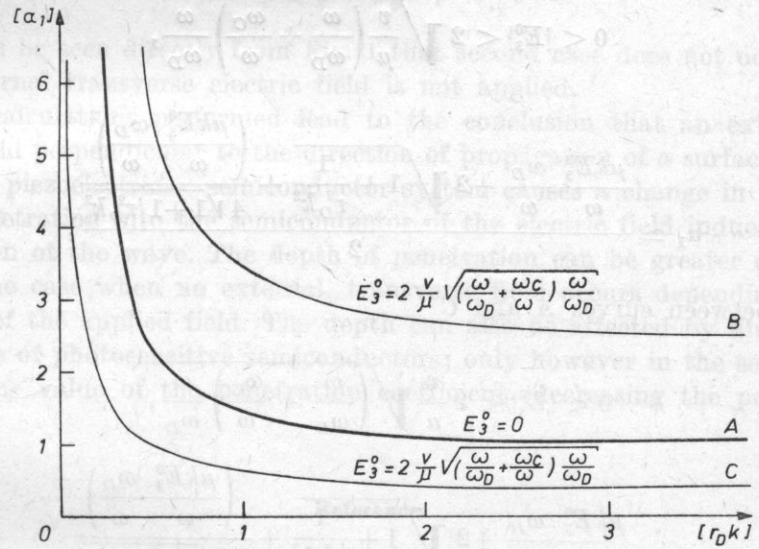

Fig. 1. The variation in the value of the coefficient of the depth of penetration as a function of the Debay screening radius $r_D k$ for different values of the *dc* external transverse field

Curve *A* shows the variation in the penetration coefficient as a function of $r_D k$ for $E_3^0 = 0$. It can be seen that the penetration coefficient is in this case always larger than unity (for $r_D k \ll 1$, $a_1 \simeq 1/r_D k$, i.e. the penetration depth is of the order of r_D; for $r_D k \gg 1$, $a_1 \simeq 1$, the penetration depth is of the order of k^{-1}) as was mentioned above.

The area above curve *A* represents the values of a_1 as a function of $r_D k$ for different values of a transverse field which brings the charge carriers to the surface on which the wave propagates, while the area below curve *A* is appropriate when the field is in the opposite direction.

It can be seen from Fig. 1 that the penetration coefficient for a field which brings the charge carriers to the surface on which the wave is propagating, is always larger than unity, while for a field in the opposite direction it can be

larger or smaller than unity, e.g. for $r_D k > 1$ it is practically always smaller than unity.

The approximate value of a_1 in the individual areas is, according to (8),

1 – above curve B: $|E_3^0| > 2\frac{v}{\mu}\sqrt{\left(\frac{\omega}{\omega_D}+\frac{\omega_C}{\omega}\right)\frac{\omega}{\omega_D}}$, v is the velocity of propagation of the surface wave, $a_1 = \frac{\mu k E_3^0}{\omega}\frac{\omega_D}{\omega}$ i.e. a_1 is directly proportional to E_3^0;

2 – between curves A and B

$$0 < |E_3^0| < 2\sqrt{\frac{v}{\mu}\left(\frac{\omega}{\omega_D}+\frac{\omega_C}{\omega}\right)\frac{\omega}{\omega_D}},$$

$$a_1 \simeq \frac{\frac{\mu k E_3^0}{\omega}\frac{\omega_D}{\omega} + 2\sqrt{1+\frac{1}{r_D^2 k^2}} + \frac{\left(\frac{\mu k E_3^0}{\omega}\frac{\omega_D}{\omega}\right)^2}{4\sqrt{1+1/r_D^2 k^2}}}{2},$$

3 – between curves A and C

$$0 < |E_3^0| < 2\frac{v}{\mu}\sqrt{\left(\frac{\omega}{\omega_D}+\frac{\omega_C}{\omega}\right)\frac{\omega}{\omega_D}},$$

$$a_1 \simeq \frac{\frac{\mu k E_3^0}{\omega}\frac{\omega_D}{\omega} + 2\sqrt{1+\frac{1}{r_D^2 k^2}} + \frac{\left(\frac{\mu k E_3^0}{\omega}\frac{\omega_D}{\omega}\right)}{4\sqrt{1+1/r_D^2 k^2}}}{2},$$

4 – below curve C

$$|E_3^0| > \frac{v}{\mu}\sqrt{\left(\frac{\omega}{\omega_D}+\frac{\omega_C}{\omega}\right)\frac{\omega}{\omega_D}}, \qquad a_1 \simeq \frac{1+1/r_D^2 k^2}{\frac{\mu k E_3^0}{\omega}\frac{\omega_D}{\omega}},$$

i.e. a_1 is inversely proportional to E_3^0.

Solving (1) with consideration of (4) leads to

$$\varphi = (C_1 e^{-kx_3} + C_2 e^{-akx_3}) e^{i(kx_1 - \omega t)}, \tag{9}$$

where C_1, C_2 are constants which can be determined from the boundary conditions. Thus the value of the penetration coefficient has a strong influence on the value and distribution of the potential within the semiconductor. The value

of the potential in turn effects the variation in the propagation velocity and the attenuation coefficient of the surface acoustic wave in a piezoelectric-semiconductor system [4].

If

(1) $\alpha_1 \ll 1$, then the second term in expression (9) is considerably less significant than the first and therefore

$$\varphi \sim e^{-kx_3} e^{i(kx_1 - \omega t)},$$

(2) $\alpha_1 \gg 1$, in this case the greater influence on the value of φ in equation (9) is exerted by the second term, i.e.

$$\varphi \sim e^{-akx_3} e^{i(kx_1 - \omega t)}.$$

It can be seen directly from Fig. 1 that second case does not occur at all if an external transverse electric field is not applied.

The calculations performed lead to the conclusion that an external, *dc* electric field perpendicular to the direction of propagation of a surface acoustic wave in a piezoelectric – semiconductor system causes a change in the depth of the penetration into the semiconductor of the electric field induced by the propagation of the wave. The depth of penetration can be greater or smaller than in the case when no external, transverse field occurs depending on the direction of the applied field. The depth can also be affected by illumination in the case of photosensitive semiconductors; only however in the sense of increasing the value of the penetration coefficient (decreasing the penetration depth).

References

[1] A. Opilski, *On the possibility of an investigation of semi-conductor surface properties using ultrasonic surface waves*, Archives of Acoustics, **1**, 1, 29–32 (1976).

[2] Ju. W. Gulajew, B. I. Pustowojt, *Usilenie powierchnostnych woln w poluprowodnikach*, ŻETF, 47, 2251 (1964).

[3] A. M. Kmita, A. W. Miedwied, *Akustoelektriczeskij efekt w sloistoj strukturie piezoelektrik – poluprowodnik*, Fiz. Twied. Tieła, **14**, 9, 2646 (1972).

[4] A. Opilski, *The influence of the surface states on the propagation of ultra- and hypersonic surface waves in semiconductors* (in Polish), Zeszyty Naukowe Politechniki Śląskiej, Seria Matematyka-Fizyka, **17** (1975).

Received on August 1, 1979; revised version on June 6, 1980.

ARCHIVES OF ACOUSTICS
6, 1, 69-78 (1981)

THE EFFECT OF $FeCl_3$ ON CHANGES IN THE RHEOLOGICAL PROPERTIES OF GLYCEROL DETERMINED BY AN ACOUSTIC METHOD

MAREK WACIŃSKI

Institute of Chemistry, Wrocław University
(50-383 Wrocław, ul. F. Joliot-Curie 14)

RYSZARD PŁOWIEC

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

The results of the shear mechanical impedance measurements are presented for solutions of $FeCl_3$ in glycerol at frequencies 30 and 500 MHz and over a temperature range from 218-303 K. These results complemented the previous investigations [4] performed at a frequency of 0.5 MHz. A distinct effect of $FeCl_3$ on the range of the viscoelastic relaxation of glycerol and on its molecular structure was found.

1. Introduction

An increasing interest in the rheological properties of liquid has been observed in the recent years, since it is expected that they are essential in the investigations of the liquid state and the principle of intermolecular forces.

In the case of solutions it can be expected that their response to shear stress, particularly over the range of viscoelastic relaxation can reflect the effect of the solute on the structure of the solvent and the molecular rearrangement. In terms of knowledge it is most interesting to note the case when an electrolyte is the solute and alcohol is the solvent, since alcohol can form with electrolyte ions not only specific solvate complexes, but also more developed and arranged molecular regions (clusters).

The above investigations were initiated with the measurements of the shear impedance in solutions of electrolytes in glycerol taken at a frequency of 0.5 MHz

[4]. The present paper complements the results of measurements in solutions of $FeCl_3$ in glycerol with the investigation of the shear impedance performed at frequencies 30 and 500 MHz over a temperature range from 218 to 303 K. The results of this investigation made it possible to determine the curve of the viscoelastic relaxation at high frequencies. It was found to be in good agreement with the theoretical conclusions in [4].

2. Experimental part

2.1 Preparation of samples

Anhydrous ferric chloride (manufactured by British Drug Houses, cz.d.a.) with a content of up to 2 % of ferrous chloride and glycerol (manufactured by POCh, Gliwice, b.f.) dehydrated by boiling under decreased pressure were used in the investigations. The water content in the glycerol determined by the Fischer method was 0.3 %. The solutions of $FeCl_3$ were obtained by solving a specific amount of ferric chloride in glycerol and its concentration was determined by weight [1]. The calculated values of the concentration are given in Table 1.

Table 1. Temperature dependence of the density and viscosity of glycerol with $FeCl_3$

Concentration of $FeCl_3$ in moles in 1 kg^3 of glycerol	Density $[kgm^{-3}]$	Viscosity $[Nsm^{-2}]$
0	$1.43605 \times 10^3 - 6.012 \times 10^{-1} T$	$-3.6344 - 9.3908 \times 10^7 T^{-3}$
0.06715	$1.4284 \times 10^3 - 5.460 \times 10^{-1} T$	$-3.4155 - 9.2425 \times 10^7 T^{-3}$
0.1069	$1.4404 \times 10^3 - 5.702 \times 10^{-1} T$	$-3.3328 - 9.1404 \times 10^7 T^{-3}$
0.1693	$1.4438 \times 10^3 - 5.614 \times 10^{-1} T$	$-3.2573 - 9.1768 \times 10^7 T^{-3}$
0.3058	$1.4707 \times 10^3 - 5.9190 \times 10^{-1} T$	$-3.3343 - 9.9207 \times 10^7 T^{-3}$
0.4015	$1.4904 \times 10^3 - 6.1571 \times 10^{-1} T$	$-3.6218 - 1.1342 \times 10^8 T^{-3}$

2.2. The measurements of density and viscosity

The density of the solutions was determined using a specific gravity bottle in a temperature range from 323 to 243 K with a precision of ± 0.05 K. For lower temperatures the density was extrapolated from the linear equation $\varrho = A + BT$, where A and B are constants and T is the temperature.

The static viscosity η was determined using a Höpler viscometer and capillary viscometers over a temperature range from 323 to 253 K. For lower temperatures the values of the density were extrapolated using the equation proposed by MEISNER [2] $\log \eta = c + D/T^3$, where c and D are constants.

The measured values of the viscosity and density of the solutions investigated are given in Table 1.

2.3. The measurements of the shear impedance

The measurements of the shear impedance of the solutions investigated and of its variation over the temperature range from 218-303 K at a frequency of 500 MHz were taken using a measuring system prepared at the Department of Physical Acoustics, Institute of Fundamental Research, Polish Academy of Sciences [3]. The measured values of the shear resistance are shown in Table 2.

Table 2. The real component of the shear impedance of solutions of $FeCl_3$ in glycerol at a frequency $f = 500$ MHz (in $[Nsm]^{-3}x10^{-5}$)

Temperature [K]	Glycerol	Glycerol – $FeCl_3$ $m = 0.4015$	Glycerol – $FeCl_3$ $m = 0.3058$	Glycerol – $FeCl_3$ $m = 0.1693$	Glycerol – $FeCl_3$ $m = 0.1069$	Glycerol – $FeCl_3$ $m = 0.06715$
218.15	22.2	22.65	22.55	22.35	22.3	22.25
223.15	21.8	22.3	22.15	22.05	22.0	21.95
228.15	21.4	21.95	21.85	21.75	21.65	21.65
233.15	21.0	21.65	21.45	21.4	21.35	21.3
238.15	20.6	21.05	20.75	20.7	20.7	20.7
243.15	20.2	20.4	20.0	19.7	19.9	20.0
248.15		19.8	19.2	18.9	19.2	19.3
253.15	19.4	19.1	18.5	17.8	18.3	18.25
263.15	17.85	17.6	16.9	16.6	16.3	16.3
273.15	15.4	15.8	15.0	14.5	14.3	14.2
283.15	12.9	14.0	12.95	12.3	12.2	12.0
293.15	10.05	12.1	11.1	10.5	10.3	10.0
303.15	7.85	10.1	9.35	8.85	8.68	7.85

3. The presentation of results

The literature proposes presentation of the measurement results of the viscoelastic relaxation range based on the Maxwell model or on the $B-E-L$ model.

3.1. Presentation based on the Maxwell model

Assuming a continuous Gaussian distribution of relaxation times the behaviour of standardized values of the shear impedance can be determined from

the formulae

$$\frac{R}{(\varrho G_\infty)^{1/2}} = \sqrt{\frac{1}{2}\int_0^\infty \frac{g_{(x)}\omega^2\tau_0^2x^2}{1+\omega^2\tau_2^0x^2}\,dx\left\{1+\left[1+\left(\frac{\int_0^\infty \frac{g_{(x)}x}{1+\omega^2\tau_0^2x^2}\,dx}{\int_0^\infty \frac{g_{(x)}\omega^2\tau_0^2x^2}{1+\omega^2\tau_0^2x^2}\,dx}\right)^2\right]^{1/2}\right\}^{1/2}}, \tag{1}$$

$$\frac{X}{(\varrho G_\infty)^{1/2}} = \sqrt{\frac{1}{2}\int_0^\infty \frac{g_{(x)}\omega^2\tau_0^2x^2}{1+\omega^2\tau_0^2x^2}\,dx\left\{\left[1+\left(\frac{\int_0^\infty \frac{g_{(x)}x}{1+\omega^2\tau_0^2x^2}\,dx}{\int_0^\infty \frac{g_{(x)}\omega^2\tau_0^2x^2}{1+\omega^2\tau_0^2x^2}\,dx}\right)^2\right]^{1/2}-1\right\}^{1/2}}, \tag{2}$$

where

$$g_{(x)} = (b/\pi^{1/2}x)\exp-[bln x]^2, \quad 0 < \tau_s < \infty, \tag{3}$$

R and X are the components of the shear impedance (shear mechanical resistance and reactance), ϱ is the density of the medium, G_∞ – is the limiting shear modulus of the liquid, ω is the angular frequency. $g_{(x)}$ represents the distribution of relaxation times, $g_{(x)}dx$ represents the part of the distribution of relaxation times over the range between x and $x+dx$ for the following parameters of the distribution width for b: $b = 0.4$ for glycerol and $b =$ from 0.33 to 0.3 for solutions of $FeCl_3$ in glycerol [4].

3.2. Presentation based on the *B—E—L* model

The $B-E-L$ model consists of two parallel acoustic impedances for the solid Z_S and for the Newtonian liquid Z_N

$$\frac{1}{Z^*} = \frac{1}{Z_N} + \frac{1}{Z_S}. \tag{4}$$

By transforming (4) one can derive the formula for the shear compliance J^* as a function of the characteristic constants of the medium investigated

$$\frac{J^*}{J_\infty} = \frac{1}{G_\infty} + \frac{1}{j\omega\eta} + 2k\left(\frac{1}{j\omega\eta G_\infty}\right)^\beta. \tag{5}$$

The components of the acoustic shear impedance of the liquid, R and X, determined from formula (5) are

$$\frac{R}{(\varrho G_\infty)^{1/2}} = \frac{(\omega\eta/2G_\infty)^{1/2}[1+(2\omega\eta/G_\infty)^{1/2}]}{[1+(\omega\eta/2G_\infty)^{1/2}]^2+\omega\eta/2G_\infty}, \tag{6}$$

$$\frac{X}{(\varrho G_\infty)^{1/2}} = \frac{(\omega\eta/2G_\infty)^{1/2}}{[1+(\omega\eta/2G_\infty)^{1/2}]^2+\omega\eta/2G_\infty}. \tag{7}$$

Formula (5) sufficiently well describes the results of the measurements taken of simple liquids for $k = 1$ and $\beta = 1/2$. In the case of complex liquids or their mixtures this formula is modified by the coefficients k and β which fit the curves to the measurement results [4]. Using formula (5) and the coefficients $k = 1.56$, $\beta = 0.4$ the effect of electrolytes on the viscoelastic relaxation curves of glycerol (Fig. 6) was described.

4. Discussion of the results and conclusions

The measurement results are shown in Figs. 1-5. They show that $FeCl_3$ widens distinctly the range of the viscoelastic relaxation time of glycerol. The changes observed in the distribution of the relaxation times induced by the presence of electrolyte reflect its effect on the cooperativeness of molecular rearrangements in the solvent [6]. Investigations of the structural and viscoelastic relaxation in associated liquids (polyhydroxide alcohols, hexachlorodiphenyl) [2, 6, 7], showed that an explanation of the changes in molecular moduli (i.e. bulk modulus and shear modulus) requires a wide spectrum of relaxation times to be assumed.

The width of the range of the viscoelastic relaxation of solutions of $FeCl_3$ in glycerol increases with increasing concentration of electrolyte (Fig. 6). This did not occur in other solutions [6]. It seems probable that a wider spectrum of the viscoelastic relaxation times is in the case of solutions of $FeCl_3$ in glycerol connected with formation of stable solvate complexes involving orbitals $3d$ of the ion Fe^{+3}.

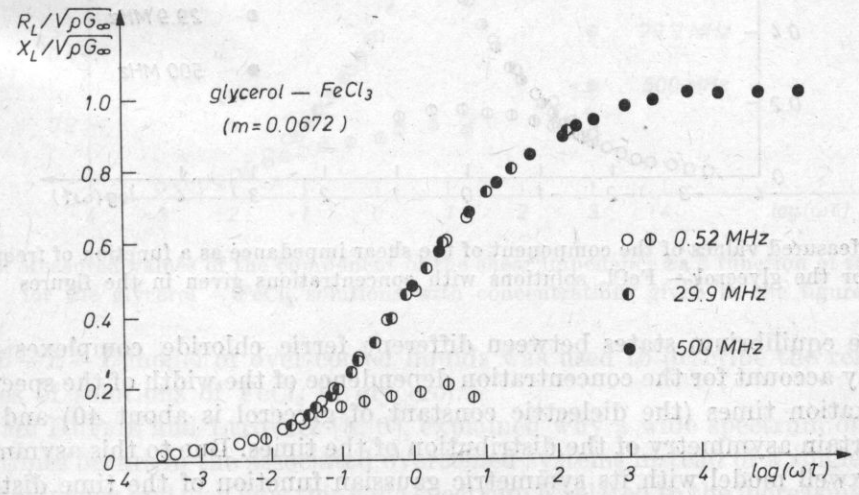

Fig. 1. Measured values of the component of the shear impedance as a function of frequency for the glycerol – $FeCl_3$ solutions with concentrations given in the figures

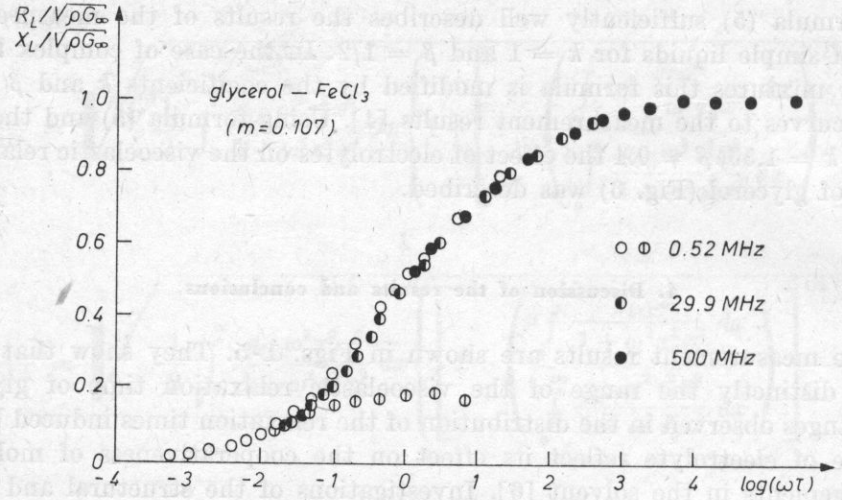

Fig. 2. Measured values of the component of the shear impedance as a function of frequency for the glycerol — $FeCl_3$ solutions with concentrations given in the figures

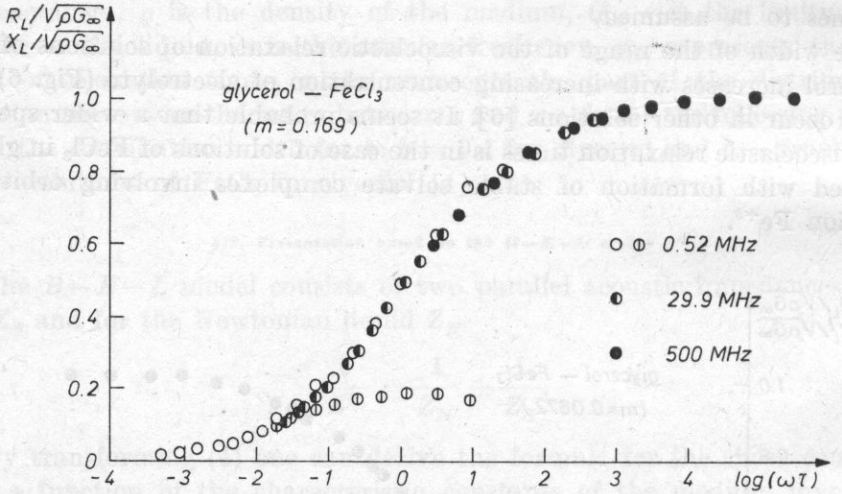

Fig. 3. Measured values of the component of the shear impedance as a function of frequency for the glycerol — $FeCl_3$ solutions with concentrations given in the figures

The equilibrium states between different ferric chloride complexes (III) probably account for the concentration dependence of the width of the spectrum of relaxation times (the dielectric constant of glycerol is about 40) and also for a certain asymmetry of the distribution of the times. Due to this asymmetry the Maxwell model with its symmetric gaussian function of the time distribution describes insufficiently the behaviour of the reduced relaxation curves, particularly in the lower temperature region (higher frequencies). Therefore

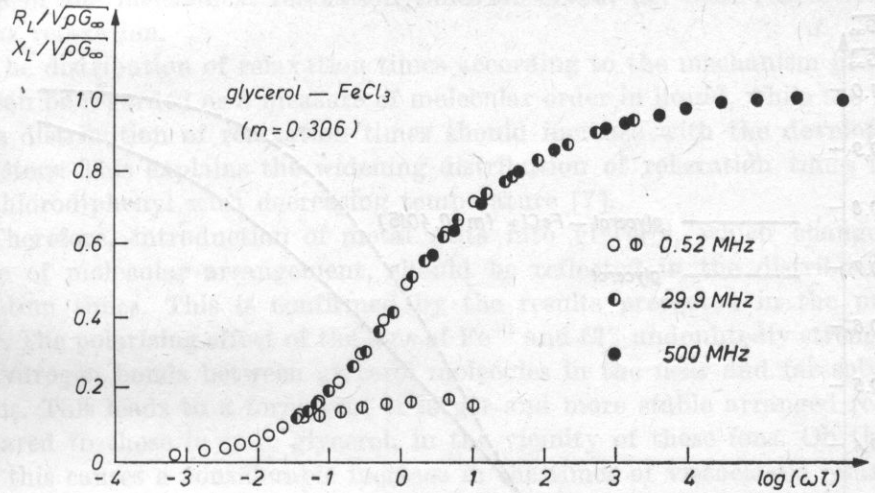

Fig. 4. Measured values of the component of the shear impedance as a function of frequency for the glycerol — $FeCl_3$ solutions with concentrations given in the figures

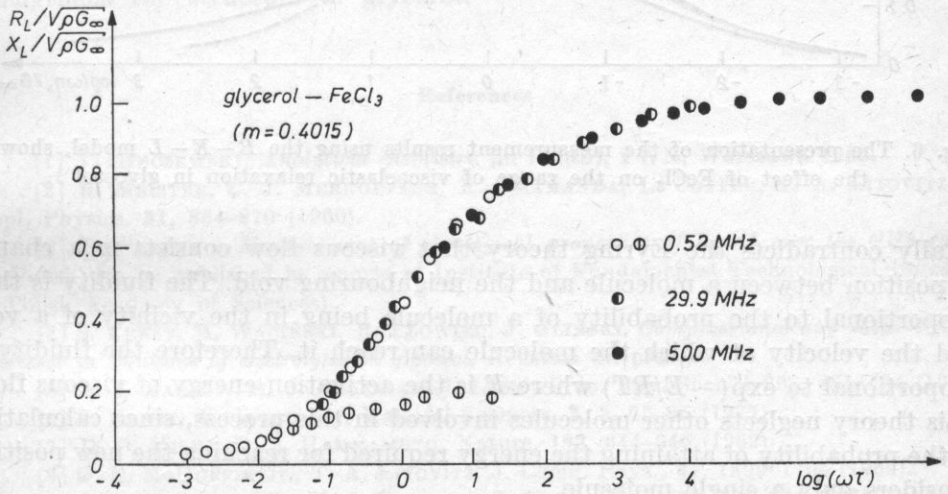

Fig. 5. Measured values of the component of the shear impedance as a function of frequency for the glycerol — $FeCl_3$ solutions with concentrations given in the figures

the $B-E-L$ model of over-cooled liquids was used to describe the relaxation curves of solutions of $FeCl_3$ in glycerol.

Mc DUFFIE and LITOVITZ [8, 9], explained why a wide spectrum of relaxation times occurs in the associated overcooled systems instead of a single relaxation frequency and also how this wide spectrum is related to the increasing temperature. They found that at temperatures of up to about 100 K above the glassy state temperature over-cooled liquids show a temperature dependence that

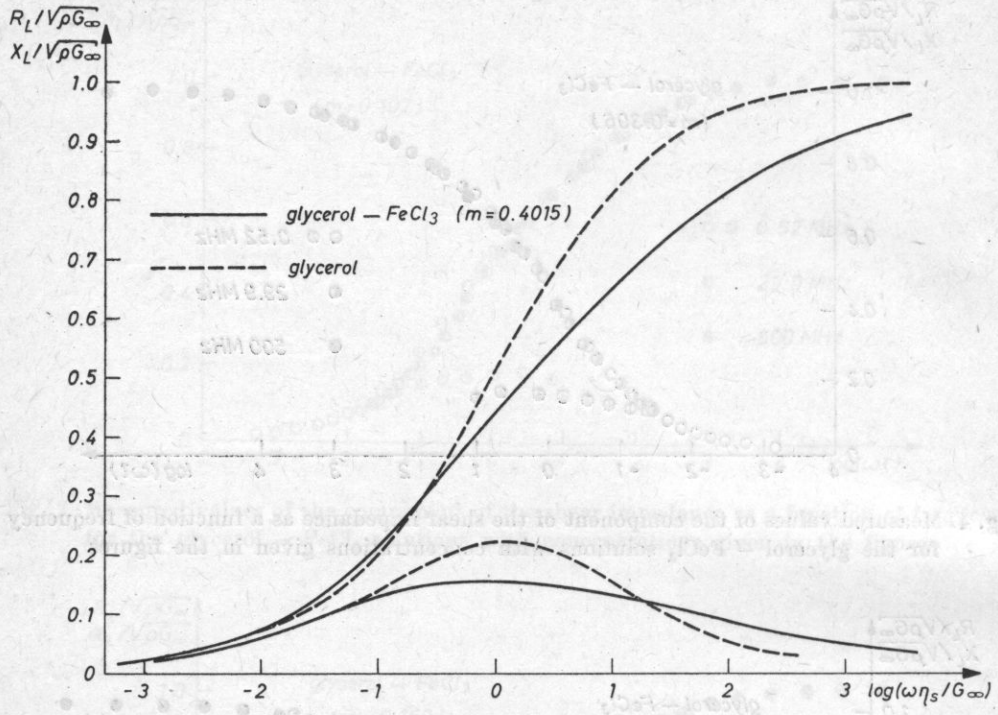

Fig. 6. The presentation of the measurement results using the $B-E-L$ model, showing the effect of $FeCl_3$ on the range of viscoelastic relaxation in glycerol

totally contradicts the Eyring theory that viscous flow consists in a change of position between a molecule and the neighbouring void. The fluidity is thus proportional to the probability of a molecule being in the vicinity of a void and the velocity at which the molecule can reach it. Therefore the fluidity is proportional to $\exp(-E/RT)$ where E is the activation energy of viscous flow. This theory neglects other molecules involved in this process, since calculation of the probability of attaining the energy required for reaching the new position considers only a single molecule.

The model of McDuffie and Litovitz is based on the following assumptions:

1. In liquid there are small arranged regions which disintegrate and reintegrate continually. (This is confirmed by investigations of diffraction of X-rays and neutrons.)

2. The disintegration of such a structure is cooperative in character, i.e. when a molecule changes its position, other molecules "cooperate" with it in order to provide the necessary space. Thus the degree of arrangement changes nonexponentially as a function of time.

3. This cooperative nonexponential behaviour is the cause of the distri-

bution of the mechanical relaxation times observed, i.e. bulk relaxation and viscous relaxation.

The distribution of relaxation times according to the mechanism proposed here can be regarded as a measure of molecular order in liquid, while the width of the distribution of relaxation times should increase with the development of clusters. This explains the widening distribution of relaxation times in the hexachlorodiphenyl with decreasing temperature [7].

Therefore, introduction of metal salts into glycerol, which changes the degree of molecular arrangement, should be reflected in the distribution of relaxation times. This is confirmed by the results presented in the present paper. The polarizing effect of the ions of Fe^{+3} and Cl^- undoubtedly strengthens the hydrogen bonds between glycerol molecules in the near and far solvation regions. This leads to a formation of larger and more stable arranged regions, compared to those in pure glycerol, in the vicinity of these ions. On the one hand this causes a considerable increase in the times of viscoelastic relaxation and on the other hand a considerable widening of the distribution of these times as a result of increasing cooperativeness of molecular groups. It can be stated, therefore, in terms usually used for aqueous electrolyte solutions, that $FeCl_3$ strengthens the structure of glycerol.

References

[1] J. Minczewski, *Analytical chemistry* (in Polish), PWN, Warszawa 1965.

[2] R. Meister, C. J. Merhoeffer, R. Sciamanda, L. Cotter, T. A. Litovitz, J. Appl. Physics, **31**, 854–870 (1960).

[3] R. Płowiec, *Measurement of rheological properties of liquid over the GHz range* (in Polish) (to be published in reports of Institute of Fundamental Technological Research of Polish Academy of Sciences).

[4] S. Ernst, M. Waciński, R. Płowiec, J. Gliński, *Compressional and shear viscous processes in solutions of electrolytes in glycerol*, Acustica (in press).

[5] A. J. Barlow, A. J. Erginsav, J. Lamb, Proc. Roy. Soc., A 298, 481–494 (1967).

[6] S. Ernst, M. Waciński, Material Sciences, **3**, 3, 75–85 (1977).

[7] D. O. Miles, D. S. Hammamoto, Nature, **193**, 644–646 (1962).

[8] G. E. McDuffie Jr., T. A. Litovitz, J. Chem. Phys., **37**, 1699–1705 (1962).

[9] G. E. McDuffie Jr., T. A. Litovitz, J. Chem. Phys., **39**, 729–734 (1963).

Received on May 26, 1980.

ARCHIVES OF ACOUSTICS
6, 1, 79-84 (1981)

A PROGRAMMED PARAMETRIC SYNTHESIZER

RYSZARD PATRYN

Institute of Applied Linguistics, Warsaw University
(00-311 Warszawa, ul. Browarna 8/10)

The synthesizer described generates signals with a duration of the pitch period, which follow the waveform of the respective segments of a natural speech signal. Parameters of each segment are determined on the basis of digital data using a computer. Some of these data are used to generate the constituents of a signal in digital to analogue convertors and some to shape the waveform obtained. The paper also gives the range of the available parameters of the sounds generated and also some examples of programmes.

1. Introduction

The main purpose of the construction of the synthesizer was to obtain a system, which would make it possible to shape a wide range of synthetic speech signals, so that while using it the analytical results, which define phonetically and acoustically the identifiable properties of speech sounds, could be verified. The principle on which it was based was that of a representation of the elementary waveform of a natural speech signal, without reference to the vocal tract transmission.

The parameters of the waveform generated by the synthesizer are determined by digital data from a computer. These data serve to generate a signal in the time domain corresponding to the pitch period.

From the technical point of view the synthesizer system is a specialized output computer device.

2. The performance of the synthesizer and its design

The synthesizer system consists of two independent tracts: one to generate sounds of harmonic (formant) structure, the other to generate signals with the character of noise. To generate the signal of the formant structure, four inde-

pendent systems of sound generation corresponding to the successive formants: $F1$, $F2$, $F3$, $F4$, which were later summed, were used.

2.1. The principle of formant signal generation

The generation of the formant signal with a waveform defined by the digital input data occurs in a digital to analogue converter system of a special type in which, in addition to the generation of a sinusoidal signal, it is modulated in amplitude at the same time, according to the analogue signal supplied by the envelope shaping system. This system is also a form of a digital to analogue converter. In both digital to analogue converters the signal generation consists in decoding the counter, to which the signal from the pulse generator is supplied. Subsequently each decoded value (in the range of 0 to 15) is reduced to a relevant analogue value for which, when the waveform corresponding to a specific formant is generated, the series of analogue values forms a sinusoidal variation, and for which, when the envelope signal is generated, it can have an essentially arbitrary shape, defined by the system. Fig. 1 shows schematically the relationships between the input signals and the resultant formant signal.

2.2. Generation of signals with the character of noise

The generator of white noise, whose signal is controlled simultaneously by 12 selective amplifiers, is used to generate a signal with a predetermined

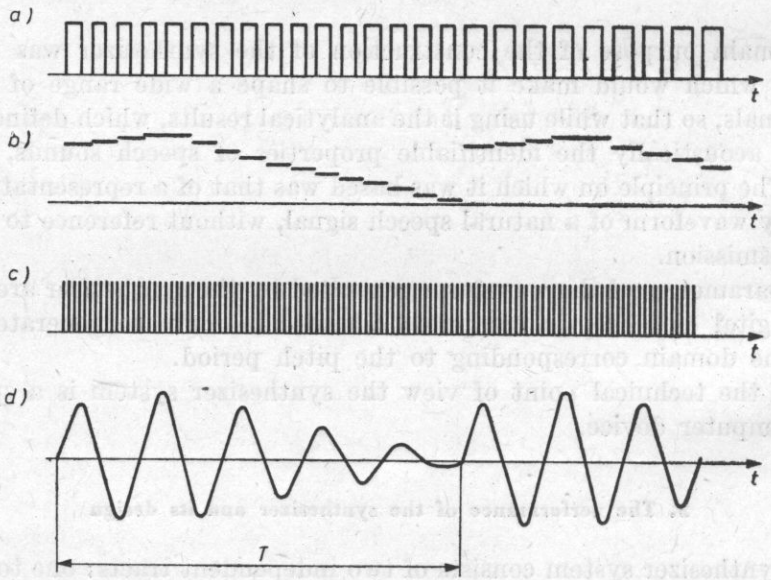

Fig. 1. An example of forming a formant signal: a) a signal from a pulse generator; b) an envelope shaping signal; c) a signal from a pulse generator F_n; d) a formant signal

spectral characteristic. The amplifiers are controlled by a digital signal, and the output signal is the sum of the outputs of the amplifiers.

2.3. The performance of the synthesizer and its block diagram

A block diagram of the synthesizer is shown in Fig. 2. The main units of the system are the following: a control unit, 8 input registers, 8 operative registers,

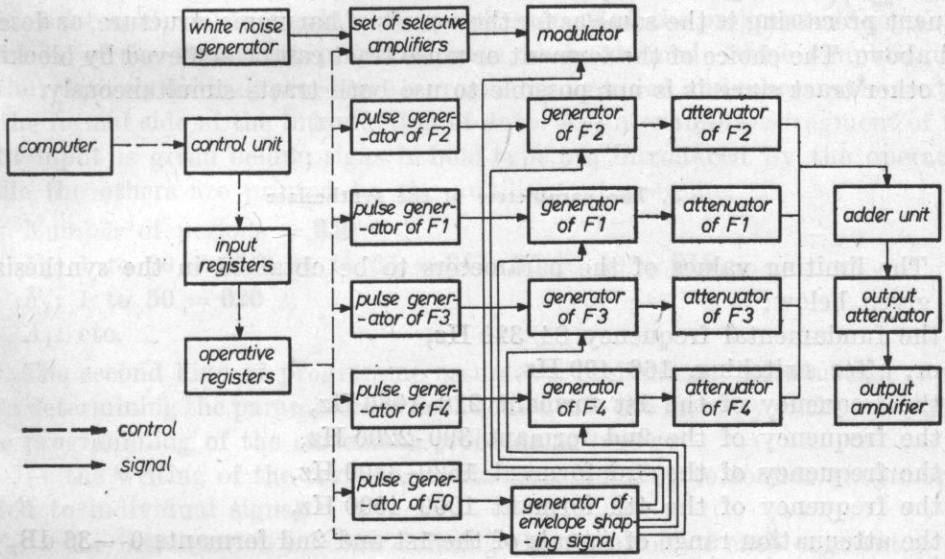

Fig. 2. A block diagram of the synthesizer system

a white noise generator, a pulse generator in the pitch circuit, a digital to analogue converter system for envelope shaping, 4 pulse generators in the formant circuits, 4 digital to analogue converters for modulating the formant circuits, 12 selective amplifiers, a modulator in the noise circuit, 4 attenuators of the formant signals, an adder, an output attenuator, an output amplifier.

The synthesizer functions in the following way: a series of 8 commands of 8 bits, defining the parameters of the signal, is sent from the computer. The end of each segment (period) is followed by a transcription of the contents of the input registers to the operative registers whose state defines the parameters of the signal for the ongoing segment whose duration is determined by 16 pulses from the F_0 pulse generator. In the case of the generation of sounds with a harmonic structure, the signals, each of which has an independent frequency, level and envelope shape, are generated by formant generators. The formant signals are amplitude modulated in the digitally controlled attenuators, and subsequently added in the adder system. The output signal is attenuated in the

output attenuator, which is also digitally controlled. This attenuator can be used to shape the envelopes of the segments which have a duration longer than the pitch period. An amplifier was used at the output of the system in order to achieve an output signal level of the order of a fraction of a volt.

A modulator, which is controlled from the envelope shaping system with a signal of the fundamental frequency is included in the noise tract. Without the modulating signal the signal from the outputs of the selective amplifiers does not change at the output of the modulator. The shaped noise signal is supplied to the input of the attenuator of the second formant and its subsequent processing is the same as for the sounds of harmonic structure, as described above. The choice of the formant or noise tract can be achieved by blocking the other tract since it is not possible to use both tracts simultaneously.

3. The parameters of the synthesizer

The limiting values of the parameters to be obtained in the synthesizer are given below:

- the fundamental frequency 84-320 Hz,
  or, after switching, 166-490 Hz,
- the frequency of the 1st formant 210-1080 Hz,
- the frequency of the 2nd formant 390-2200 Hz,
- the frequency of the 3rd formant 1620-3300 Hz,
- the frequency of the 4th formant 1900-4000 Hz,
- the attenuation range of signals of the 1st and 2nd formants 0- —36 dB,
- the attenuation range of signals of the 3rd and 4th formants 0- —26 dB,
- the attenuation range of the output signal 0- —23 dB,
- the number of the envelope shapes available without switching 7,
- the available levels at the outputs of the selective amplifiers 0, —8, —16 dB, no signal,
- the frequency range transmitted by the selective amplification unit 1-5 kHz,
- the relative width of the selective amplification bandwidth about 0.14.

The synthesizer was built on 4 boards of 300 ×300 dimensions and TTL technique was used in the digital systems.

4. Controlling the performance of the synthesizer

The performance of the synthesizer depends on the computer programme which defines parameters for each segment of period T_0. In order to define the parameters of the signal the following should be given

- the mode of structure: harmonic-noise,

– the fundamental frequency F_0,
– the frequencies of the four formants,
– the level of the four formants,
– the shape of the envelope of the waveform of the formant within the pitch period,
– the attenuation level of the resultant signal.

When data is introduced into the memory of a computer it has the form of three numerical values and is presented in a coded form.

In the use of the synthesizer two kinds of controlling programmes have been so far used. In the first kind the parameters of all the programmed segments (periods) of the statement prepared have to be defined. Some simplification of the relatively difficult procedure is programming-aided and concerns the control of the formal side of the introduction of data. As an example, a fragment of the data input is given below; signs in bold type are introduced by the operator, while the others are printed by the auxiliary programme.

Number of periods = **050**
F_0: 1 to 20 = **015**, 21 to 35 = **017**, 35 to 50 = **016**
F_1: 1 to 50 = **020**
A_1: etc.

The second kind of programme assumes a preliminary introduction of the data determining the parameters of each sound into the memory of the computer. The programming of the statement proceeds as below:
– the writing of the text (segments corresponding to sounds are subordinated to individual signs),
– the introduction of data defining the duration of each segment,
– the introduction of data defining the intonation curve,
– the definition the value of the fundamental frequency.

The introduction of the data is also assisted by the programme; the duration of the sound segments is given in a coded form as a single digit number each time after the given sign has been printed by the programme. The intonation is defined on the basis of the data introduced as a number of musical half-tones preceded by a sign (+ or –), or possibly 0 after a bracket printed after the sign of the segment and the number of periods (16 ms) defining the duration of the segment. The closing bracket is interpreted by the programme as a lack of change until the end of a given segment.

An example of programming: signs in bold type are introduced by the operator

janek text
j5a7n5e5 4 k3 duration
j06(**+1 +2**)a12(**+1 +1**)n05()e06(**–200** –2)04()*k*03() intonation
sf0 = **0.14.**

In view of the limited range of the use of the synthesizer to date it has not been necessary to use other controlling programmes which are easier to use if a more detailed definition of data is required than in the example presented above. The possibility of influencing the applying of a signal generated for short time intervals is considered by the users of the synthesizer as a characteristic meeting the initial assumptions and a useful application of the synthesizer.

Received on January 29, 1980.

C H R O N I C L E

1ST SPRING SCHOOL OF ACOUSTO-OPTICS AND ITS APPLICATIONS
Gdańsk-Wieżyca 26-30 May 1980

The School was organized by the Institute of Physics of the University of Gdańsk and the Sections of Quantum and Molecular Acoustics and Sonochemistry of Polish Acoustical Society and supported by the Institute of Fundamental Technological Research of the Polish Academy of Sciences. The Honorary Committee consisted of Prof. Dr. Ignacy MALECKI, Prof. Dr. Halina RYFFERT, Prof. Dr. Janusz SOKOŁOWSKI. The Organizing Committee consisted of Prof. Dr. Antoni ŚLIWIŃSKI, Prof. Dr. Aleksander OPILSKI, Prof. Dr. Jerzy RANACHOWSKI, Dr. Anna MARKIEWICZ and Maria BORYSEWICZ, M. Sc., Dr. Iwona WOJCIECHOWSKA, Dr. Marek KOSMAL, Dr. Piotr KWIEK and Dr. Bogumił LINDE.

The School had an international character both in terms of experts invited (9 from abroad and 7 from Poland) and participants (5 from abroad and 51 from Poland). 72 scientists took part in the school. In addition to 22 lectures there were also 15 poster form presentations.

The aim of the School was to survey the developments and enlighten the participants in the field of the physical problems of interaction between light and sound (mainly over the ultra- and hypersonic ranges) in liquids and solids. The acousto-optic phenomena have been known to physicists for several scores of years but it is in the recent decade that this field has seen veritable development. A good many original applications in the acousto-optic processors have appeared, e.g. ultrasonic deflectors of light beams, modulators, filters and so on, which are used in the integrated optics,ultrasonic visualization (also holography), signal analysis, spectroscopy etc.

The yearly developments in acoustooptics indicate that this is an integral and promising field, which made the idea of an international meeting of experts i.e. the School, which would have a workshop character, so much worthwhile.

The School appeared to be very useful both to the mind of the participants and organizers, therefore they welcomed the idea of such Schools every second year.

The School consisted of:

General papers

A. BREAZEALE, *Bragg imaging of finite amplitude ultrasonic waves.*

R. MERTENS, *Some recent developments in the theory of diffraction of light by an ultrasonic wave*: 1. *The wave equation — its establishment — its aproximation.* 2. *Methods of solution.* 3. *Diffraction of laser light by ultrasonics.*

R. STEPHENS, *The photoacoustic effect.*

A. ZAREMBOWITCH, *Bragg diffraction of light by ultrasonic waves, a specific tool for solid state investigations.*

J. K. Zieniuk, J. Litniewski, *On the influence of coherence and wave-length on ultrasonic images.*
J. C. Somer, *Application of an acousto-optic device in an optical deconvolver for blurred ultrasound – diagnostic images.*
A. Opilski, *Acousto-optical methods of solid state investigations.*
Z. Kleszczewski, *Nonlinear acousto-optic interaction.*
M. Szustakowski, *Acousto-optical devices of signal processing.*
R. Reibold, *Double exposure holography: its application to intensity measurements of arbitrarily shaped ultrasonic fields.*
M. Szustakowski, *Acousto-optic devices for laser beam control.*
I. Malecki, J. Ranachowski, *Propagation of ultrasonic waves in nonhomogeneous media as in piezoelectric ceramics used in acousto-optics.*
A. Śliwiński, *Recent results on the experimental verification of the Leroy's theory of diffraction of light by two adjacent ultrasonic beams.*
A. Śliwiński, *Optical holography and acousto-optics.*
I. Gabrielli, *Spatial and temporal light modulation by ultrasound: theory and experiments.*
A. Alippi, SAW – *acousto-optics.*
W. Pajewski, *Piezoelectric and elasto-optic properties of crystals.*
A. Defebvre, *Theoretical and experimental study of* SAW *propagation velocity in layered media.*
L. Pimonow, *Intercellular information and ultrasonics.*
D. Watmough, *An investigation by telemicroscopy and electron microscopy of the biological effects of ultrasound with a view to damaging malignant tumours.*

Poster form papers

H. J. Hein, *The importance of the threshold contrast for acoustical imaging.*
P. Kwiek, A. Markiewicz, A. Śliwiński, *An optical holography with a reference beam of the cosinusoidally modulated amplitude used for investigation of ultrasonic fields.*
I. Wojciechowska, *Application of optical holography to determination of amplitude distribution throughout ultrasonic transducers.*
P. Kwiek, A. Markiewicz, A. Śliwiński, *Experimental verification of light diffraction by two ultrasonic beams.*
M. Kosmol, *Diffraction of high intensity laser beam by ultrasonic wave.*
J. Berdowski, *Diffraction of laser light by* SAW.
B. Świetlicki, *Calculations of acousto-optic interactions in the* $LiNbO_3$ *planar diffusion waveguide.*
H. Kusek, *Acoustical signals recordered under strain.*
P. Hauptmann, R. Sauberlich, S. Wartewig, *Acoustic relaxation spectroscopy on polymer solutions.*

In addition to numerous discussions accompanying all the lectures and poster form sessions there also was a separate round table discussion, which showed a considerable topicality of the issues presented at the meeting and also mapped out the prospective directions of acoustooptic research, including the acousto-optic interaction in solids, Bragg's diffraction, light diffraction by two ultrasonic beams, applications of optic holography in ultrasonic visualization, application of acoustooptic processors in telecommunication signal processing etc.

The proceedings of the School were published by the end of 1980.

Antoni Śliwiński (Gdańsk)