ARCHIVES OF ACOUSTICS
4, 1, 3-10 (1979)

DYNAMIC APPROACH TO SOUND PITCH*

HALINA RYFFERT, ALICJA CZAJKOWSKA, URSZULA JORASZ, RUFIN MAKAREWICZ

The Chair of Acoustics, Adam Mickiewicz University (60-769 Poznań, ul. Matejki 48/49

Theoretical considerations have provided an analytical expression describing simultaneous changes of the intensity and pitch of the signal emitted by a moving source. Experimental investigations of the perceptibility of the sound-pitch changes accompanied simultaneously by appropriate intensity changes have confirmed mutual coupling of these two quantities in the process of the formation of a general sonic impression. A qualitatively new effect of automasking has been observed and shown to consist in an interaction of the two varying psychoacoustic parameters in the process of sound perception.

1. Introduction

The existence of the Doppler effect leads, from the psychoacoustic point of view, to the formation of complex sonic impression on an observer listening to the signal produced by a moving source. A simultaneous change of the momentary values of two physical parameters of the signal, i.e., its intensity and frequency leads to the formation of a sonic impression of variable intensity and pitch. In the case of a complex signal, a change of timbre occurs additionally.

The function for intensity changes and the corresponding frequency changes for two different velocities of a source of the sinusoidal tone of a frequency $f_0 = 1000$ Hz are presented in Fig. 1.

The source moves steadily along a straight line with respect to an immobile observer. The trajectory of the motion passes at a distance $d = 1$ m from the observer. The value $t = 0$ on the time axis is arbitrarily chosen and denotes the moment when the source passes the observer.

It follows from Fig. 1 that a continuous change of intensity (the right-hand side scale in dB) is accompanied by an almost stepwise change of frequency

* Supported by the Committee on Acoustics of the Polish Academy of Sciences.

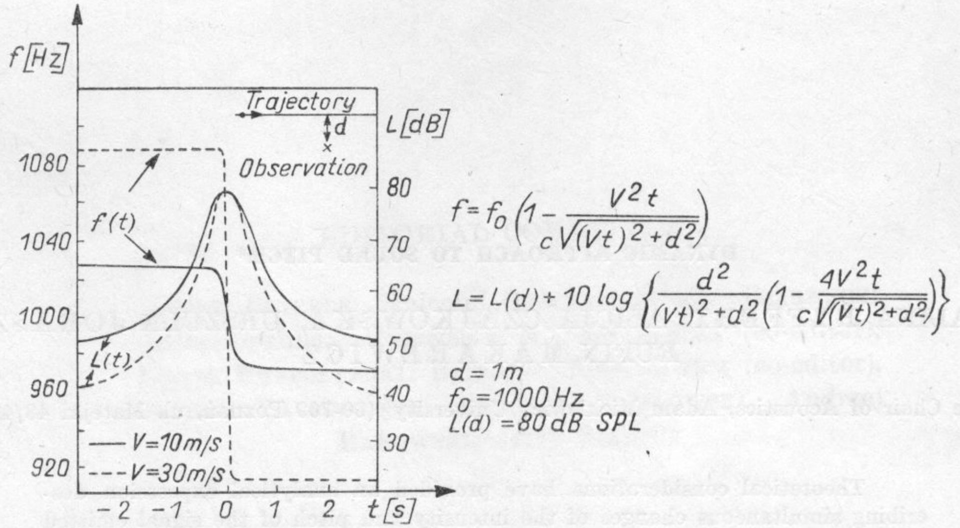

Fig. 1. The changes of intensity (the curves in the form of pulse) and the corresponding changes of frequency (the curves with a distinct fall) in the case of two different motion velocities of a sound source passing the observer

(the left-hand side in Hz). Both, the magnitude of the sudden frequency drop and the moment when it happens depend on the velocity of the source motion.

In our investigations we attempted to estimate the ability to perceive the pitch fall accompanied by continuous loudness changes due to intensity changes.

2. Measurement methods

The system designed for carrying out experimental investigations generated a signal of the form presented in Fig. 2. The duration of the frequency fall varied between 5 ms for $f = 2$ Hz to 10 ms for $f = 24$ Hz which was dictated by experimental facilities. It may be thus assumed that this duration was almost constant. No click was produced at the fall. The intensity pulse was formed photoelectrically. The signal had a form of the tone pulse with a predetermined envelope and a step-like change of carrier frequency. The signal intensity was controlled by a light beam masked by an appropriate diaphragm. The investigations were carried out for three different durations of intensity pulse, i.e., for $t_1 = 0.5$ s, $t_2 = 0.9$ s, and $t_3 = 1.3$ s. The pulse appeared against the background of a signal of a constant intensity lower by 20 dB than the maximum intensity of the signal pulse.

The duration of a single signal, including the time intervals with the appropriate background, preceding and following the intensity pulse was 6 s. The test presented to the observers in a single measurement series consisted

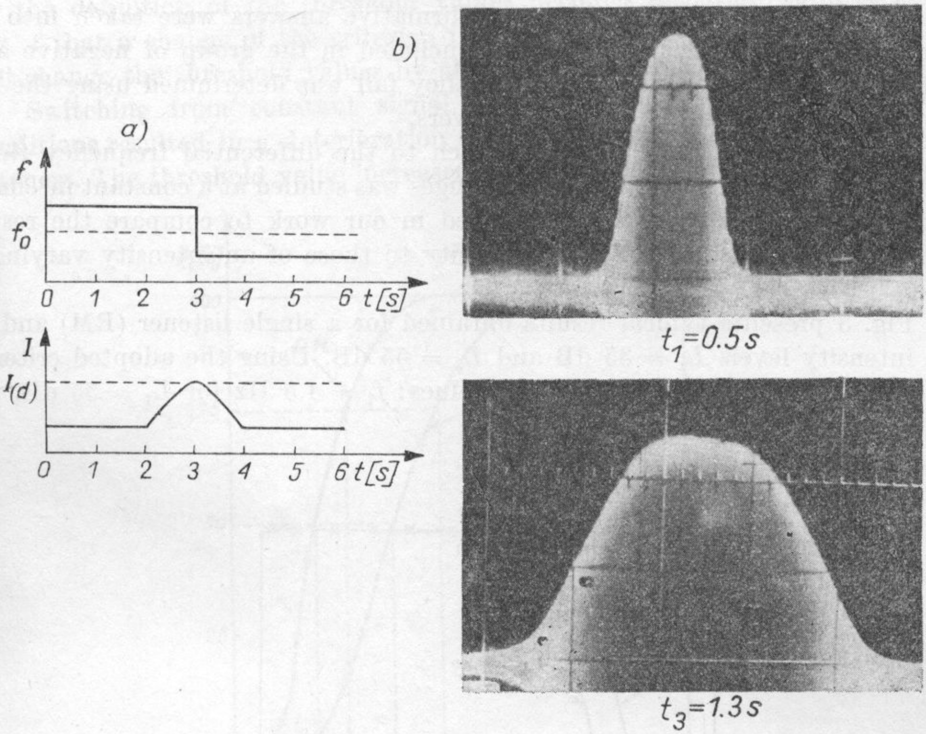

Fig. 2. (a) Schematic presentation of the course of a frequency fall accompanied by an appropriate intensity change. (b) The form of intensity pulses for two various pulse durations

of thirty nine 6-second signals. The interval between the signals, destined for an answer was 3 s. The course of intensity changes in a single measurement series was the same whereas the magnitude of the intensity fall was selected randomly in the range 0-24 Hz in 2 Hz steps. Individual series differed in the duration of intensity pulse and the value of the reference level with respect to the intensity. The dynamics of the process, expressed in dB remained unaltered.

The measurements were performed in a semi- anechoic room. The individual series of signals were tape-recorded and next reproduced from a GD 12/5 (5 W, 4 Ω) loudspeaker located at a height of the observer's head and at a distance of 1 m from him; 5 observers participated in the measurements (3 men and 2 women), 4 of them were musically trained. The observer's task was to evaluate each signal, i.e., to declare whether they had perceived a pitch or not.

3. Analysis of experimental results

The measurement results are presented in the form of diagrams. The values of frequency fall are plotted along the abscissa (the central frequency of each fall was $f_0 = 1000$ Hz) and the percentage of affirmative answers is plotted along the coordinate axis. Only the affirmative answers were taken into account. The hesitant answers [?] were included in the group of negative answers. The threshold value of the frequency fall was determined using the criterion of the 50 % of affirmative answers.

In the conventional, static approach to the differential frequency threshold, the perceptibility of frequency changes was studied at a constant intensity. Such investigations were also performed in our work to compare the results obtained in the case of constant intensity to those of an intensity varying in time.

Fig. 3 presents typical results obtained for a single listener (RM) and for two intensity levels $L_1 = 35$ dB and $L_2 = 55$ dB. Using the adopted criterion we obtained the following threshold values: $f_1 = 4.5$ Hz for $L_1 = 35$ dB, and $f_2 = 6.5$ Hz for $L_2 = 55$ dB.

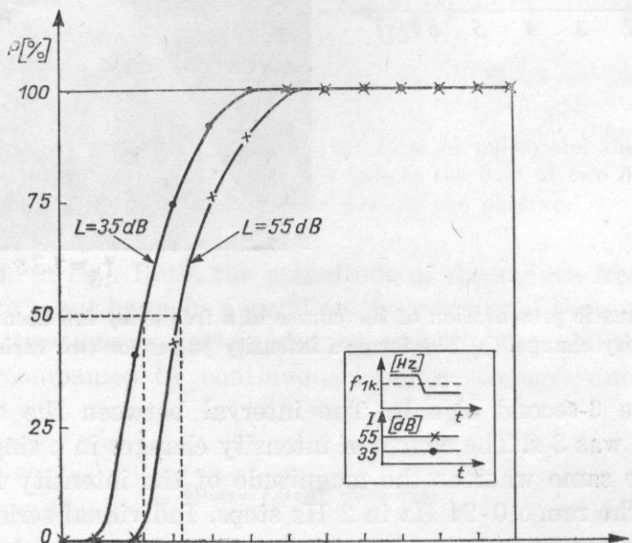

Fig. 3. The results of the preception measurements for a frequency fall of a constant-intensity signal for two various intensity values for a single observer (RM)

Thus, at a higher level of intensity, the threshold value of the perceivable frequency fall increased, i.e., this fall was more difficult to perceive.

The investigations have revealed significant individual differences between the results obtained for various observers. Fig. 4 presents typical results of the same investigations for two listeners (RM and HP) with the lowest and the highest threshold values (both persons were musically trained).

The differences occurred in the threshold value itself as well as in the scatter of answers revealing a certain weakness of perception. In view of a fairly large frequency range (exceeding 10 Hz in individual cases) the choice of the criterion for the definition of the threshold values becomes essential. It is seen from Fig. 4 that a change of the criterion from 50 % of affirmative answers to 75 % will change the threshold values by about 1.5 Hz (RM) and 2.5 Hz (HP).

Switching from constant signal intensity conditions to pulse-intensity conditions resulted in a deterioration of the frequency fall perceptibility for all listeners. The threshold value increased, the weak perception range broadened,

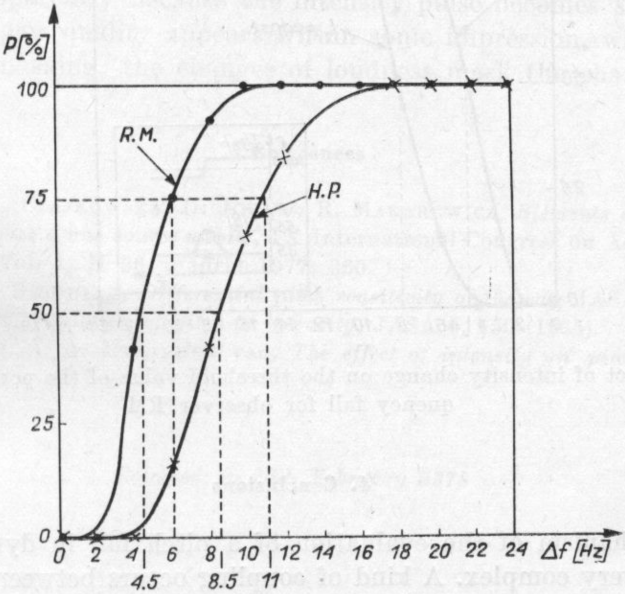

Fig. 4. Threshold values of frequency-fall perception at a constant intensity level for two observers: RM and HP

and some irregularities of psychometric curves began to appear. The increase of threshold values was different for different listeners and ranged between 1 Hz and 8 Hz. Fig. 5 presents a comparison of the results obtained for a single listener at a constant intensity level $L = 35$ dB and the results obtained in the case of an intensity pulse of a duration of 1.3 s. The increase of the threshold value amounted to 3.5 Hz for this observer (RM). The increase has even greater at higher intensity levels.

Not always did the results obtained in the conditions of intensity pulse yield such a regular curve and such a distinct deterioration of the threshold value. Particularly large discrepancies occurred in the case of an observer who was not musically trained.

It has been also found that the percentage of uncertain answers and the threshold value increase when the duration of intensity pulse decreases.

Frequently the intensity pulse itself gave the impression of a pitch fall even if the fall was absent (for instance, the observer with no musical training gave 40 % affirmative answers at $\Delta f = 0$ Hz).

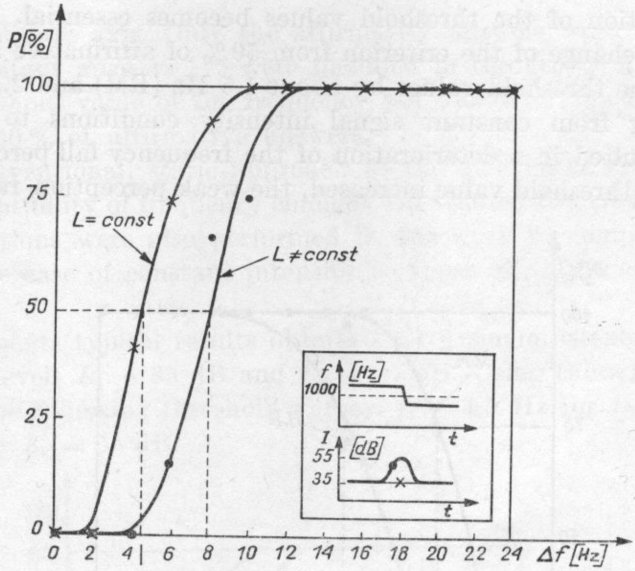

Fig. 5. The effect of intensity change on the threshold value of the perceptibility of a frequency fall for observer RM

4. Conclusions

The mechanism of the evaluation of a pitch fall in dynamic conditions seems to be very complex. A kind of coupling occurs between the impressions of loudness and pitch changes. The sonic impression is apparently dominated by the effect of loudness changes whereas the performed investigations required the observers to concentrate on the perception of a pitch fall. Thus it happened that loudness changes were perceived by the observers as pitch changes. Hence this type of investigation is very difficult to conduct with listeners with no musical training.

The problem of mutual relation between two psychoacoustical parameters: loudness and pitch was undertaken by Stevens [3] as early as in 1935. However, those investigations as well as subsequent attempts (cf. e.g. [4]) were limited to only static conditions. The observers compared loudness or the pitch of a tone of fixed parameters (intensity and frequency). The compensation method was used which balanced the difference in loudness by a change of the signal frequency. The results confirmed the mutual affect of the two psychoacoustic parameters on the formation of sonic impression. The investigations carried out in this work in dynamic conditions and thus concerned with the effect of loudness changes on the evaluation of pitch changes with the esential physical parameters of the signal continuously changing in time have confirmed the existence of a mutual coupling between the impression of loudness and pitch. In all dynamic conditions this coupling is very distinct and complex. It depends on the assumed theoretical model associating the changes of intensity with those of frequency. For instance, in the Doppler effect model, adopted in the present investigations, a decrease of intensity-pulse duration occurs simultaneously with the increase of the value of frequency fall. Thus it could be anticipated that the signal perceptibility should improve. Actually the experiment shows that the uncertainity of the observers as for the perception of a pitch fall increases, apparently because the intensity pulse becomes simultaneously steeper. Thus a new quality appears within sonic impression, which we could refer to as automasking: the changes of loudness mask the changes of pitch.

References

[1] H. Ryffert, A. Czajkowska, U. Jorasz, R. Makarewicz, *Eléments de perception des signaux acoustiques d'une source mobile*, IX International Congress on Acoustics, Contributed papers, Vol. 1, H 36, Madrid 1977, 360.

[2] G. Shower, R. Biddulph, *Differential pitch sensitivity of the ear*, JASA **3**, 275 (1931).

[3] S. S. Stevens, *The relation of pitch to intensity*, JASA **6**, 150 (1935).

[4] J. Verschuure, A. A. Meeteren van, *The effect of intensity on pitch*, Acustica **32**, 1, 33-44 (1975)

Received on 26th February 1978

ARCHIVES OF ACOUSTICS
4, 1, 11-26 (1979)

TRAFFIC NOISE MEASUREMENTS IN ANTWERP AND BRUSSELS AND THEIR RELATION WITH ANNOYANCE

H. MYNCKE, A. COPS, R. GAMBART, P. STEENACKERS

Laboratorium voor Akoestiek en Warmtegeleiding K. U. – Leuven, 3030 Heverlee – Belgium

This work is a part of a large scale traffic noise survey in two Belgian cities: Antwerp and Brussels.

A social survey (interviews of 1800 persons) was coupled with intensive acoustical measurements (40 places in Antwerp and 25 in Brussels, covering 1100 hours of measurements). Part I concerns the physical measurements, analysis and synthesis of the data. A prediction method based on these measurements is also included. Part II concerns the results of the enquiry about annoyance and its correlation with physical quantities.

PART I: PHYSICAL MEASUREMENTS

1. Measurement facilities

A technical team of 3 persons disposed of an autonomous measuring car equiped with two independant measuring systems composed out of:

– a microphone with special protection against humidity;

– a measuring amplifier;

– a level recorder giving continous registration of the sound level as a function of time;

– statistical analyzers with a resolution of 2 dB(A) and a sampling rate of 10 times per second;

– an automatic photo unit for the registration of the state of the statistical counters each 10 minutes;

At least one operator was always present during the measurements.

2. Measurements

Contin uous 24-hours measurements were done at 40 places in Antwerp while the total observation periods at 25 places in Brussels were reduced to 12 hours.

The microphone was placed at the edge of the kerbside at a height of 1.2 m.

The measurement sites covered the whole range from very quite places to very noisy and they presented mainly free flowing traffic conditions.

During the acoustical measurements the traffic intensity was automatically counted.

In order to study the effect of parameters which influence the sound level, series of additional measurements were done simultaneously, e.g. in respect with streetwidth, streetheight, ground cover, traffic lights etc.

In this context a study was also made on the reverberation phenomena in town streets, the results of which are published in Acustica [1].

For each measurement site data are presented as follows:

— exact location and description;

— statistical distribution histograms per hour;

— L_{eq}, L_{10}, L_{90} levels as a function of time (hourly values);

— For three periods namely: day (7-19h), evening (19-23h) and night (23-7h), mean values were calculated of the following quantities: L_1, L_5, L_{10}, L_{50}, L_{90}, L_{95}, L_{99}, L_{eq}, TNI, NPL, the standard deviation and L_{dn} over 24 hours.

3. Analysis of the measurement results

The most important information is given by the statistical noise distribution as a function of time. All these distributions are given in the final report.

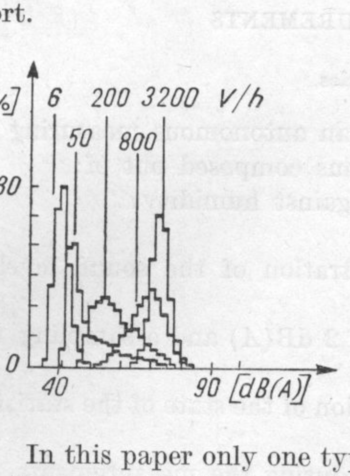

Fig. 1. Evaluation of the noise distribution with increasing traffic flow

In this paper only one typical noise distribution is given, e.g. the evaluation of the noise distribution with increasing traffic intensity in Fig. 1.

To have a good overall picture of the traffic noise the L_{eq}, the L_{10} and L_{90} sound levels are presented as a function of time. In Fig. 2 and 3 those values are given for a street with respectively low and dense traffic flow. The L_{eq}-values are nearly constant during the day period (7-19h) for streets with dense traffic flow.

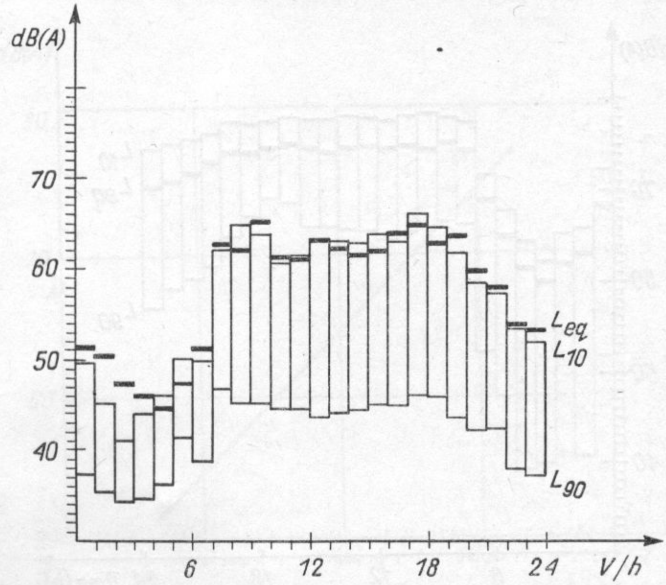

Fig. 2. Werkstraat – low traffic flow L_{eq} and the noise climate (L_{10} - L_{90}) as a function of time

For streets with relative low traffic flow there can be some greater variation. The mean L_{eq}-value over this period gives the best correlation with the enquiry about annoyance (cf. Part II).

For the evening and night periods the variation is much higher. This is also the case for the L_{10} and L_{90} values. It is less significant to speak about mean values for these both periods.

The L_{10}-values are comparable with L_{eq}-values in the case of dense traffic flow. They give values which are appr. 3 dB higher than the L_{eq}-values (see Fig. 3).

In streets with low traffic flow, the L_{10}-values even can give smaller levels than the L_{eq}-values. For normal (Gaussian) distributions this can be explained by the equation $L_{eq} = L_{50} + 0.115\sigma^2$. As traffic noise is not normally distributed, the relationship of L_{eq} with other percentiles of distribution can even be determined experimentally. L_{eq} will always be greater than L_{50}.

In order to make correlations with the enquiry about annoyance mean values of L_1, L_{10}, L_{50}, L_{90}, L_{eq}, TNI, NPL, vehicles/hour and L_{dn} were calculated during the three above mentioned periods [2].

4. Synthesis and prognosis

The L_{eq}-levels. L_{eq}-levels roughly give a 3 dB(A) increase per doubling of the amount of vehicles per hour. This is shown in Fig. 4, where the L_{eq}-levels per hour are given as a function of traffic intensity. All the measuring points fall nearly within the ± 5 dB(A) region.

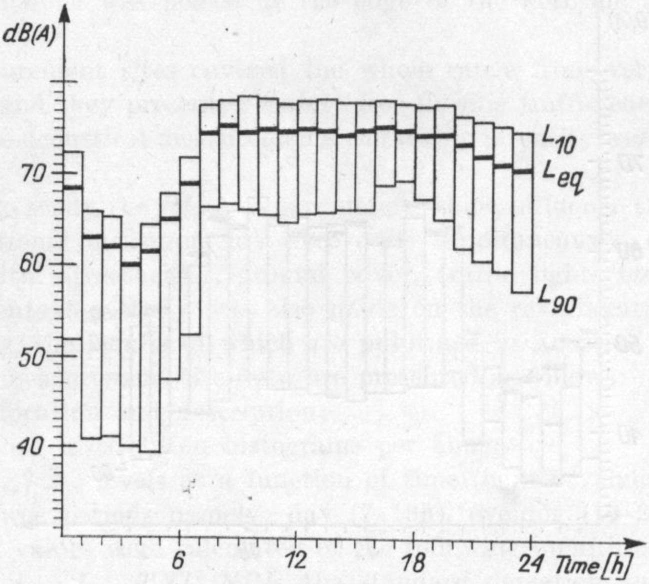

Fig. 3. Plantin en Moretusleki – dense traffic flow L_{eq} and the noise climate (L_{10}-L_{90}) a a function of time

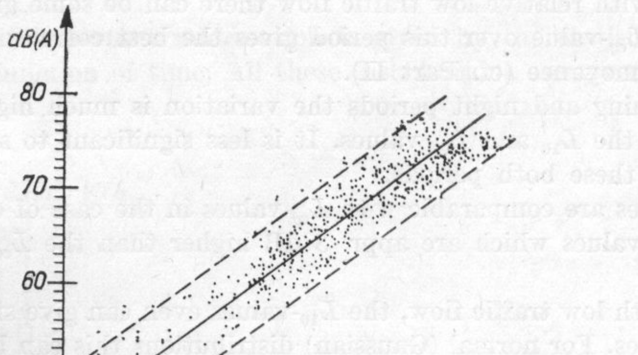

Fig. 4. L_{eq} levels per hour as a function of traffic intensities (Antwerp)

The measuring points are nearly completely within the ± 5 dB(A) region

The spread of these points around the theoretical straight line is caused by:

– errors of measurement due to foreign noise sources. The influence of these noises can be statistically smoothed out by taking mean values over larger observation periods. Fig. 4 turns then over to Fig. 5.

– a systematic influence of the ground cover. Fig. 6 shows mean lines for respectively stone cover and asphalt (difference appr. 2.4 dB(A)).

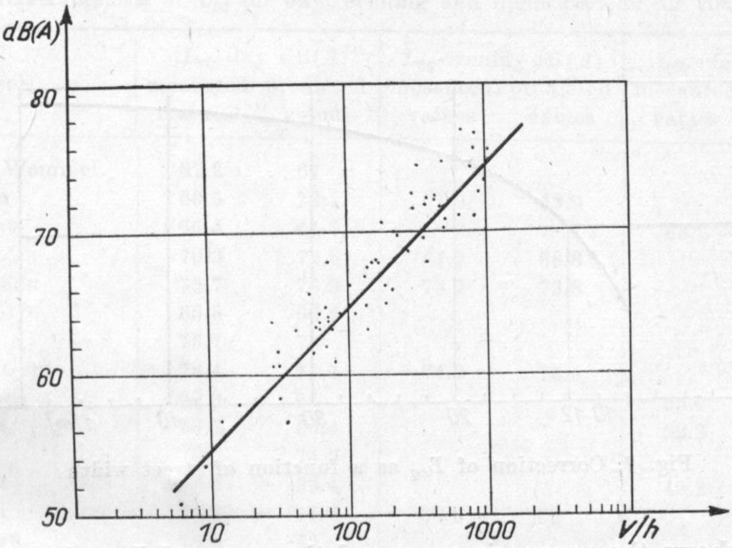

Fig. 5. Mean values for L_{eq} as a function of traffic intensities. Day – Evening – Night

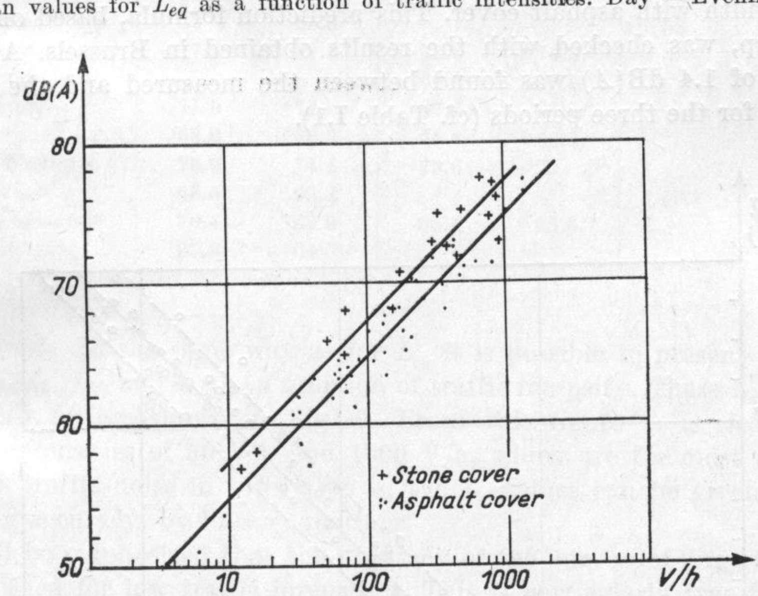

Fig. 6. Mean values for L_{eq} as a function of traffic intensities

— a systematic influence of the streetwidth. This influence was determined partly on a theoretical and partly on an experimental basis. Fig. 7 gives the proposed correction factor.

Summarised the following prediction formula can be used

$$L_{eq} = 10 \log I + 44.8 + C_1 + C_2,$$

where I is the traffic intensity (vehicles/hour). C_1 is the correction for ground cover, C_2 is the correction for streetwidth.

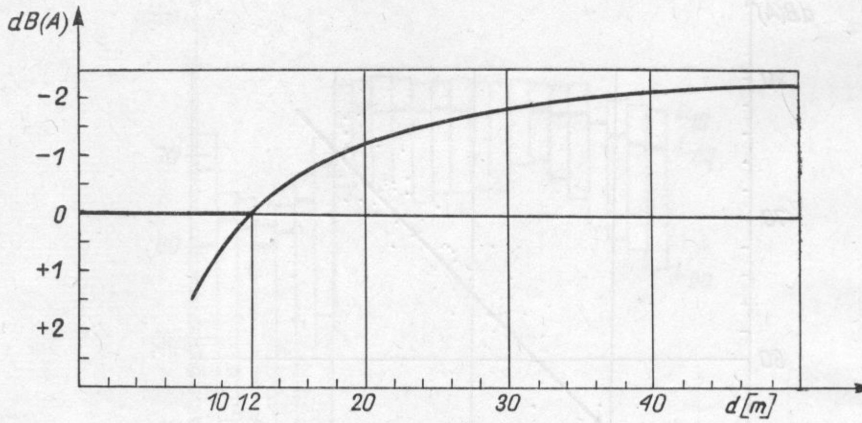

Fig. 7. Correction of L_{eq} as a function of street width

Fig. 8 shows the prognosis curve and the corrected L_{eq}-values for a street of 12 m width with asphalt cover. This prediction formula, based on the data of Antwerp, was checked with the results obtained in Brussels. A standard deviation of 1.4 dB(A) was found between the measured and the predicted L_{eq}-values for the three periods (cf. Table I.1).

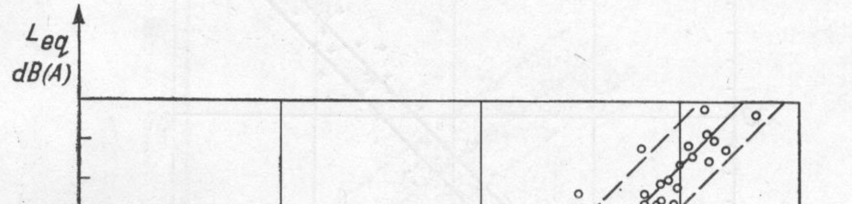

Fig. 8. Prognosis and measuring results for a 12 m street width an asphalt cover

Table I.1. Prognosis of L_{eq} for day, evening and night periods for Brussels

Brussels	L_{eq} day dB(A) measured values	L_{eq} day dB(A) predicted values	L_{eq} evening dB(A) measured values	L_{eq} evening dB(A) predicted values	L_{eq} night dB(A) measured values	L_{eq} night dB(A) predicted values
Steenweg op Wemmel	67.2	67				
C. Woestelaan	69.5	72.1	67.9	68.9		
Heidekenstraat	65.4	64.7			55.8	55.3
Vriheidslaan	70.3	72.6	67.7	68.8		
Keizer Karellaan	75.7	76.9	73.7	73.8		
Basilieklaan	65.6	66.2				
Piersstraat	75.1	74				
Steenweg op Gent	78.4	77.4	74.2	73.5		
Koekelberglaan	65.9	64			55.0	54.4
Begijnenstraat	66.2	66.4			52.2	53.5
Parklaan	72.3	74.6				
J. Robiestraat	62.5	59.4			49.2	48
P. Jansonlaan	71.6	71.6	68.6	69		
Clemenceaulaan	75.3	76				
Demolderlaan	70.3	69			58	59
P. Deshanellaan	69	68.8			58	55.8
Lambermontlaan	77.1	77.5	73.6	74.3		
Plaskylaan			69.2	68.5		
F. Guillaumelaan	71.9	72	67.9	67.2		
Slegerslaan	68.9	69.3	64.4	66.4		
Steenweg op Wemmel	76.9	74.4	73.6	72		
Hertogendalstraat	63.5	63.2			49.1	49.8
T. Van Der Elststraat	70.4	69.9	65.4	65.5		
Gem. Godshuisstraat	63.8	64				

The L_x-levels. In the same way as for L_{eq} it is possible to present a mean value for L_1, L_{10}, L_{50} and L_{eq} as a function of traffic intensity. These L_x-indices compared with L_{eq} are shown in Fig. 9. From this figure it is clear that, between the intensities of 50 V/h and 1000 V/h, which are the most realistic situations for traffic noise in town streets, the L_x-values can be given within the measuring accuracy, by linear equations.

It should be emphasized that the accuracy of the curves of Fig. 9 is severely diminished for low traffic intensities. This is particularly true for high x-indices.

PART II: ENQUIRY CONCERNING ANNOYANCE

In part I the physical aspects of the research program were explained. Another major aim of this research work consisted in determining the relation between objective characteristics of noise and the annoyance.

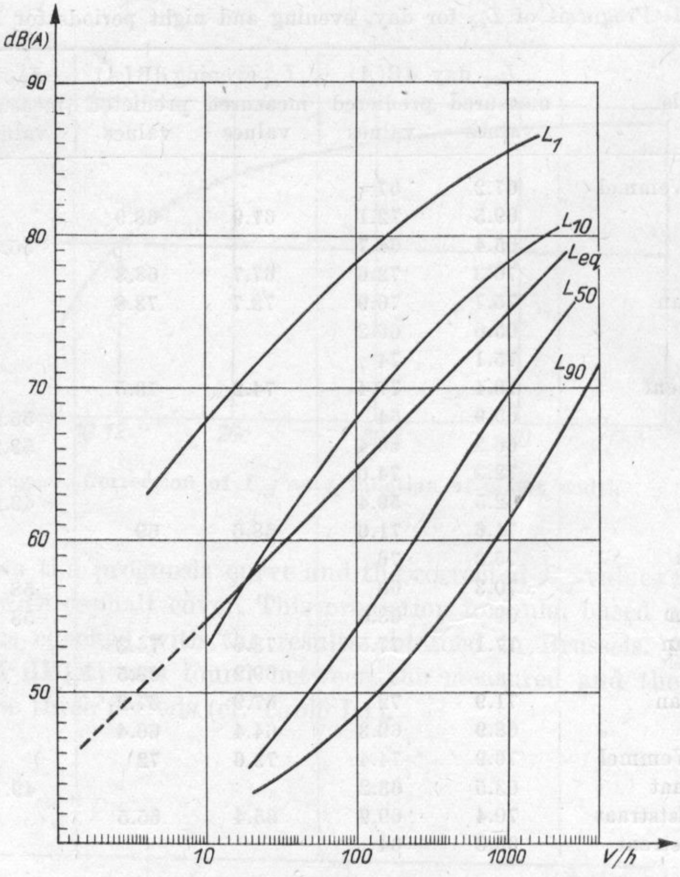

Fig. 9. Mean values for 40 streets

II.1. General information about the survey

A large-scale survey was carried out in two cities: Antwerp (1975) and Brussels (1976). In Table II.1 more details of these investigations are shown.

The subjects were invited to cooperate by letter. A second letter was mailed if the first one was not replied positively by at least 25 subjects. If this number

Table II.1. Extent of the surveys in Antwerp and Brussels

Town	Number of streets	Number of interrogated subjects
Antwerp	40	1279
Brussels	25	496
Totals	65	1775

was not yet reached after two written invitations, the collaborators ringed the bell of every house until they had interrogated 25 inhabitants.

The subjects who replied positively to our invitation, were interrogated at home by one of our collaborators. During this interview a questionnaire was presented to the subjects, who marked themselves the answers.

The questionnaire was composed of about 70 questions. The subjects gave an answer by putting a stripe on a full-linescale of 155 millimetres long. For example:

5. We would like to know your opinion on the traffic noise you hear during the day

Not disturbing at all ________________________________ Very disturbing

After each investigation (Antwerp, Brussels) a factor analysis was performed on the data. The results of this method enabled us to purify the questionnaire: questions with a loading lower than 0.500 were discarded for the following investigation, since they were of little use.

However, the factorial structure of the investigations was remarkably stable through the three investigations, and the high-loading items did not switch from one factor to another.

The results of the factor analysis further show that we can compute scores on five factors:

1. Disturbance of diurnal activities by traffic noise.
2. Nocturnal disturbance.
3. General statements on traffic noise.
4. Supposed physical (physiological) effects of traffic noise.
5. Satisfaction with the environment.

II.2. Some remarks upon the research method

Results of a survey can be influenced by the method. In fact, our survey was carried out with subjects who cooperated of their own free will, without any reward. In this circumstances it is possible that only highly annoyed inhabitants of a town make an appointment with our collaborators to express their displeasure.

However, there are different reasons why we can contradict this supposition:

(a) All letters to the inhabitants contained the following sentence: "Even if your quarter is quiet, or if you are not annoyed by traffic noise, your cooperation is very important, since we have to investigate quiet as well as noisy streets in order to obtained clear and useful results."

(b) In two streets the survey was carried out while the inhabitants were not previously notified or invited to cooperate: the collaborators ringed from door to door. In these two streets higher mean annoyance scores were computed than in streets with a comparable traffic noise level.

(c) The correlation between the percentage replied letters and the L_{eq} is not significant (see Table II.2).

The negative correlations in Brussels even suggest that more inhabitants of quiet streets are willing to cooperate.

Table II.2. Correlations between the percentage replied letters and the L_{eq}

% replied	Antwerp	Brussels
1st letter $- L_{eq}$ (day)	0.05	-0.12
1st letter $- L_{eq}$ (night)	0.02	-0.60
2nd letter $- L_{eq}$ (day)	0.02	-0.14
2nd letter $- L_{eq}$ (night)	0.02	-0.06

(d) There should be no correlation at all between the annoyance and noise in the case that only highly annoÿed people cooperate.

(e) The representativity of the sample is good.

We determined the relation between annoyance and noise by measuring the traffic noise outside, and by comparing the results of these measurements with the annoyance the inhabitants claimed to support. These people of course live inside their appartment. This is a kind of contradiction, and our method has been criticized for this reason: we should measure the traffic noise inside the appartment where we interview an inhabitant.

We could refute this argument by pointing out that by our knowledge, this method has not been used until now. But of course, this is not a valid refutation of the fundamental objections to the reliability of the method.

However, there is a second argument in favour of our method. In all the investigations we know, the correlation between the individual annoyance and the noise level is poor. Research workers usually explain this lack of correlation by pointing out that the reaction of an inhabitant to traffic noise is not only determined by the intensity. They state that physiological, psychological, sociological and environmental factors are at least as important as the intensity of the traffic noise. In Table II.3 we give a clear example of this lack of correlation between individual annoyance scores and the noise intensity (r = correlation coefficient).

Table II.3. Correlation between the annoyance scores on the first factor (disturbance of diurnal activities) and the noise indices (investigated town : Antwerp)

	Individual annoyance scores		Mean annoyance scores	
	r	r^2	$\bar{r}$	$\bar{r}^2$
L_1 day	0.40	0.16	0.83	0.69
L_{10} day	0.42	0.18	0.85	0.72
L_{50} day	0.42	0.18	0.86	0.74
L_{90} day	0.38	0.14	0.78	0.61
L_{eq} day	0.42	0.18	0.86	0.74

This table shows that our investigation in Antwerp is no exception to the rule that has been described by other research workers: the intensity of noise determines maximally 18 % of the variance of the individual annoyance. The mean annoyance on the contrary is determined for maximally 74 % by the same intensity.

This large difference between 74 % and 18 % enables us to suppose that we will be more lucky in computing the relation between individual annoyance and noise, if we pay attention to the causes of individual differences in annoyance, rather than by measuring the traffic noise in every appartment or dwelling. Indeed, when mean annoyance scores are computed, we minimize approximatively the influence of physiological, psychological, sociological and environmental factors.

II.3. The relation between noise and annoyance

In Table II.4 we can see that the correlation between the disturbance of diurnal activities and the diurnal traffic noise is good and highly significant. The other variables (nocturnal disturbance, general statements on traffic noise, supposed physiological effects, and satisfaction with the environment), correlate rather poorly.

Table II.4. Correlation between the mean factorscores and the noise indices (day)

		Antwerp					Brussels				
		FS1	FS2	FS3	FS4	FS5	FS1	FS2	FS3	FS4	FS5
L_1	$r =$	0.83	0.15	−0.04	0.32	0.11	0.82	0.55	−0.35	−0.15	0.52
	$p =$	0.001	0.184	0.415	0.021	0.241	0.001	0.003	0.047	0.236	0.004
L_{10}	$r =$	0.85	0.11	0.00	0.32	−0.20	0.86	0.36	−0.40	−0.18	0.51
	$p =$	0.001	0.250	0.496	0.023	0.115	0.001	0.043	0.027	0.195	0.005
L_{50}	$r =$	0.86	0.13	−0.04	0.38	−0.21	0.83	0.29	−0.25	−0.17	0.51
	$p =$	0.001	0.220	0.393	0.007	0.094	0.001	0.086	0.117	0.220	0.005
L_{90}	$r =$	0.78	0.11	−0.09	0.42	−0.14	0.78	0.27	−0.14	−0.13	0.47
	$p =$	0.001	0.241	0.292	0.003	0.189	0.001	0.101	0.259	0.266	0.010
L_{eq}	$r =$	0.86	0.14	−0.03	0.33	−0.18	0.86	0.41	−0.32	−0.15	0.56
	$p =$	0.001	0.199	0.419	0.019	0.139	0.001	0.023	0.062	0.239	0.002
TNI	$r =$	0.47	0.04	0.13	−0.02	−0.16	0.51	0.38	−0.53	−0.14	0.20
	$p =$	0.001	0.405	0.220	0.450	0.168	0.005	0.034	0.004	0.255	0.121
NPL	$r =$	0.70	0.12	0.05	0.16	−0.16	0.76	0.50	−0.48	−0.16	0.43
	$p =$	0.001	0.239	0.391	0.163	0.160	0.001	0.007	0.009	0.235	0.019
$\log_{10} V/h$	$r =$	0.83	0.04	−0.06	0.31	−0.18	0.85	0.18	−0.24	−0.19	0.44
	$p =$	0.001	0.417	0.365	0.031	0.146	0.001	0.209	0.135	0.200	0.018

FS1: Disturbance of diurnal activities by traffic noise
FS2: Nocturnal disturbance
FS3: General statements on traffic noise;
FS4: Supposed physical (physiological) effects of traffic noise
FS5: Satisfaction with the environment

When we compare the results in Table II.4 with those of Table II.5, where the correlation between the different annoyance variables with the nocturnal traffic noise is given, it is quite surprising that we cannot find large differences. In fact, the disturbance of diurnal activities correlates as well with the measurements of nocturnal traffic noise as with the measurements of diurnal traffic

Table II.5. Correlation between the mean factorscores and the noise indices (night)

		Antwerp					Brussels				
		FS1	FS2	FS3	FS4	FS5	FS1	FS2	FS3	FS4	FS5
L_1	$r =$	0.82	0.12	0.05	0.40	−0.22	0.50	0.46	−0.21	−0.05	0.54
	$p =$	0.001	0.229	0.392	0.007	0.093	0.069	0.091	0.284	0.449	0.055
L_{10}	$r =$	0.80	0.15	−0.08	0.41	−0.23	0.48	0.54	−0.01	0.02	0.61
	$p =$	0.001	0.187	0.317	0.005	0.084	0.079	0.053	0.485	0.478	0.029
L_{50}	$r =$	0.68	0.21	−0.17	0.29	−0.13	0.40	0.39	0.44	0.21	0.10
	$p =$	0.001	0.103	0.160	0.038	0.226	0.125	0.132	0.104	0.277	0.391
L_{90}	$r =$	0.22	−0.04	−0.10	0.16	0.02	0.50	0.01	0.41	0.12	−0.26
	$p =$	0.089	0.418	0.278	0.164	0.446	0.069	0.492	0.120	0.370	0.235
L_{eq}	$r =$	0.82	0.14	−0.02	0.41	−0.18	0.47	0.52	−0.16	−0.01	0.58
	$p =$	0.001	0.201	0.456	0.005	0.142	0.085	0.062	0.333	0.490	0.039
TNI	$r =$	0.80	0.17	−0.06	0.39	−0.26	0.25	0.52	−0.19	−0.03	0.71
	$p =$	0.001	0.147	0.373	0.007	0.059	0.240	0.06	0.300	0.465	0.011
NPL	$r =$	0.81	0.14	0.00	0.41	−0.25	0.35	0.56	−0.19	0.01	0.71
	$p =$	0.001	0.202	0.497	0.005	0.064	0.163	0.047	0.302	0.487	0.011
$\log_{10} V/h$	$r =$	0.82	0.07	−0.09	0.25	−0.13	0.66	0.37	0.35	0.38	−0.49
	$p =$	0.001	0.352	0.316	0.094	0.251	0.052	0.206	0.220	0.202	0.135

FS1: Disturbance of diurnal activities by traffic noise
FS2: Nocturnal disturbance
FS3: General statements on traffic noise
FS4: Supposed physical (physiological) effects of traffic noise
FS5: Satisfaction with the environment

noise, while the nocturnal disturbance correlates as poorly with the nocturnal traffic noise as with the diurnal traffic noise! In our investigation in Brussels the correlation between nocturnal disturbance and nocturnal traffic noise is better, but not convincing!

There are several reasons why nocturnal disturbance correlates poorly:

1. During the survey the research workers observed that some inhabitants of quiet streets declared themselves highly disturbed by the few passing cars at night, while in noisy streets some inhabitants declared that they were never disturbed by nocturnal traffic noise. Presumedly, this is a kind of "natural selection": people who cannot sleep in a noisy environment, will not move into a noisy quarter, or will not stay there for a long time.

2. In Antwerp and Brussels, the noise levels at night are proportionally lower to the diurnal noise levels. This means that noisy streets are also noisy at night, and quiet streets will also be quiet during the night period. This is very clear when we compute the correlation between the diurnal noise measu-

rements and the nocturnal measurements. The correlation between L_1 (Day) and L_1 (night) is 0.93; between L_{10} (Day) and L_{10} (night): 0.86. Reasoning in statistical terms: the correlation between diurnal noise levels (A) and nocturnal noise levels (B) is high.

Since the correlation between diurnal noise levels (A) and nocturnal disturbance (C) is low, the correlation between nocturnal noise levels (B) and nocturnal disturbance (C) will also be low: $ArB = +++$, $ArC = ---$, thus $BrC = ---$.

Fig. 10 shows the relation between the different noise indices and the variable (factor) 1: disturbance of diurnal activities. These figures

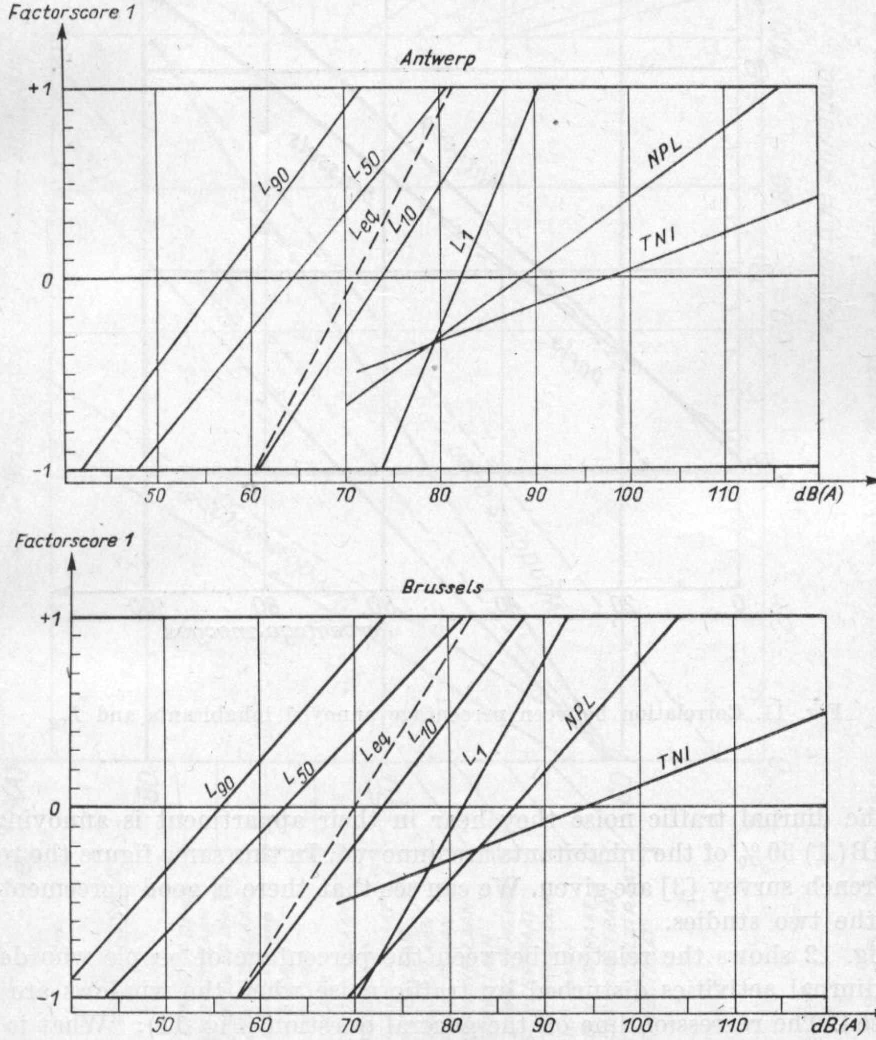

Fig. 10. Correlation between different noise indices and disturbance of diurnal activities (Faktorscore 1)

show clearly the different slopes and positions of L_1, L_{10}, L_{50} L_{eq}, NPL and TNI (Day period).

In every street where the survey was carried out, the percentage of inhabitants who declared themselves annoyed, was calculated. The following Fig. 11 shows the relation between the percentage of inhabitants who declare

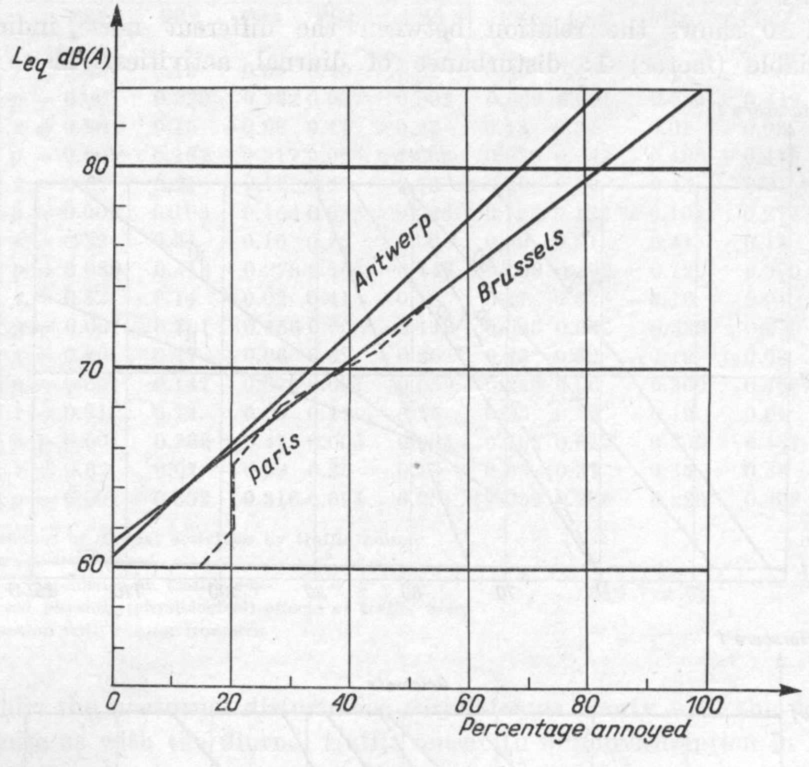

Fig. 11. Correlation between percentage annoyed inhabitants and L_{eq}

that the diurnal traffic noise they hear in their appartment is annoying. At ± 72 dB(A) 50 % of the inhabitants are annoyed. In this same figure the results of a French survey [3] are given. We can see that there is good agreement between the two studies.

Fig. 12 shows the relation between the percentage of people who declare their diurnal activities disturbed by traffic noise when the windows are open or closed. The regression line on the general question (Fig. 11): "What is your opinion on the traffic noise you hear during the day?" is also repeated in this Fig. 12.

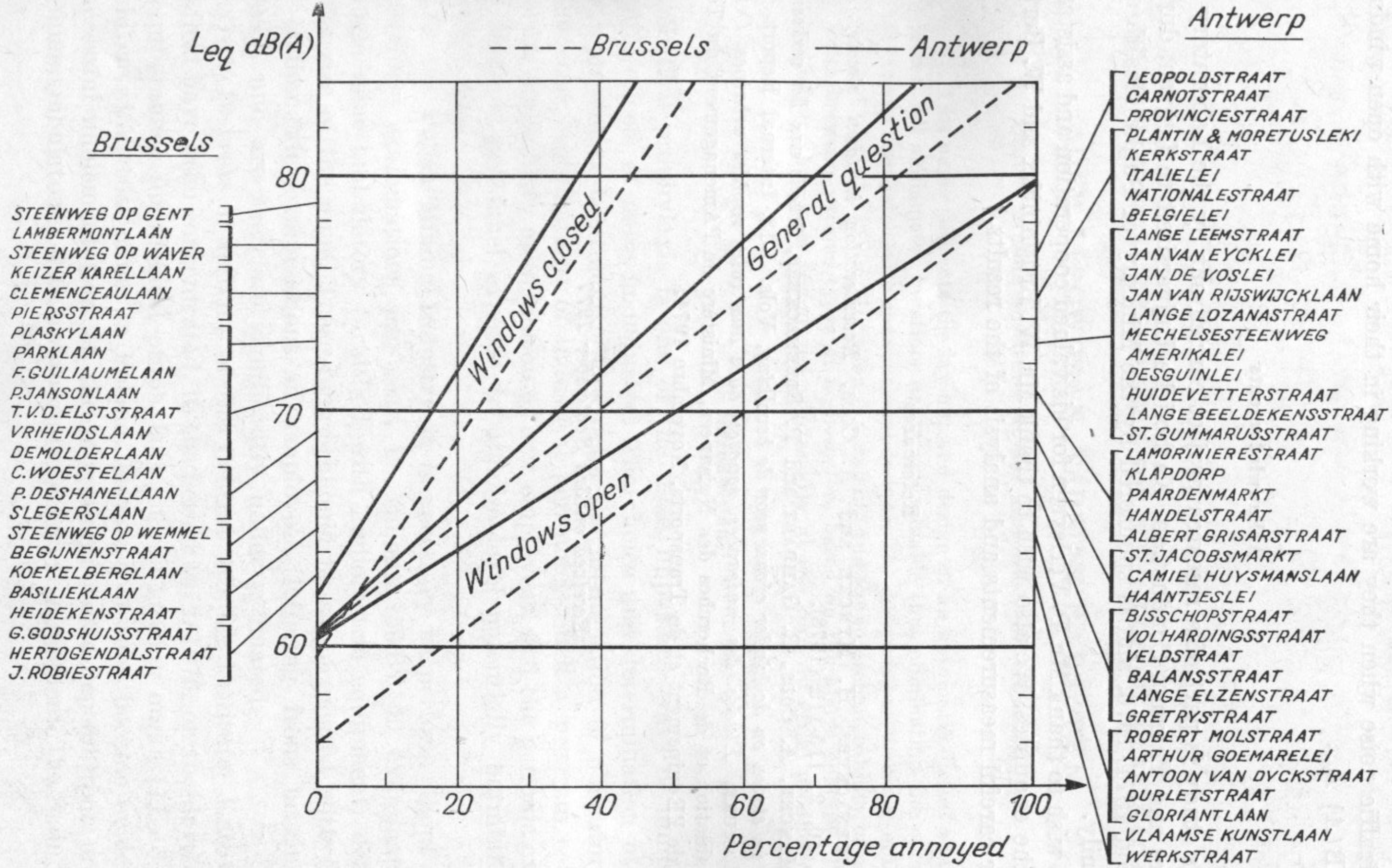

Fig. 12. Correlation between the percentage inhabitants disturbed when windows are open or closed, the general question and L_{eq}

We can see that approximatively 50 % of the inhabitants are disturbed by the traffic noise when they are working in their home with open windows at 70 dB(A) L_{eq}.

Acknowledgments

This study is part of a research project carried out by the Laboratory for Acoustics and Thermal Conduction, Katholieke Universiteit te Leuven, during the period 1974-1976, and is sponsored by the Ministry of Public Health and the Family.

We wish to thank Mr. P. Verleysen for his valued cooperation and assistance during the enquiries. We also wish to thank Mr. W. Bruyninckx and W. Bertels for their careful measurements and analysis of the results.

References

[1] P. Steenackers, H. Myncke and A. Cops, *Reverberation in town streets,* Acustica, 40, 2, 115-119 (1978).

[2] H. Myncke, A. Cops, R. Gambart and P. Steenackers, *Studie van het verkeerslawaai in steden en de hinder ervan voor de bevolking,* Vols. 1-13, Internal Report.

[3] C. Lamure, *Etudes des nuisances sonores dues aux voies rapides urbaines.* Centre d'Evaluation et de Recherches des Nuisances, Ministère de l'Amenagement du Territoire,de l'Equipement et de Transports. Novembre 1974.

Received on 3rd September 1977

ARCHIVES OF ACOUSTICS
4, 1, 27-38 (1979)

AUTOMATIC DIGITAL AUDIOMETER FOR MASS HEARING EXAMINATIONS*

ADAM LIPOWCZAN, TADEUSZ RABSZTYN

Central Mining Institute (40-951 Katowice, Pl. Gwarków 1)

This paper presents the lay-out and design of an automatic digital audiometer. All audiometer functions are programmable thus eliminating the operator errors associated with conventional designs. The design is based on a modular concept resulting in simplification in servicing and the interchangeability of modules between various types of audiometers based on the above concept. The audiometer has been tested and is presently at the stage of serial production.

1. Introduction

Medical prophylaxis is one of the most important ways of limiting the harmful effects of noise in industry. The choice and determination of appropriate preventive treatments relies upon the introduction of obligatory mass audiometric examinations of industrial personnel. Such a program comprises both the preliminary examination of new employees, and the periodic examination of these personnel exposed to a noise level potentially harmful to the hearing organ.

In the Polish Mining Industry it is necessary to perform about 150000 audiometric examinations per year, i.e., about 1500-3500 examinations in a single mine ambulatory furnished with audiometric equipment, depending on the size of the mine. These figures should be supplemented with those of the audiometric examinations of employees suffering from hearing-organ diseases, and are therefore significantly under estimated.

Over 10 years of activity of the Central Mining Institute, Katowice, in this field have been concentrated, in the period 1972-1976, on the development of a programme for medical prophylaxis of the hearing organ [11-13] which would cover all coal mines in Poland. However, it soon became evident that a successful implementation of the programme would depend upon providing the mine ambulatories with appropriate technical equipment, the main compo-

* Supported by the Ministry of Mining Research Programme No 124.

nents being an audiometer and a sound-proof chamber for audiometric examination [12].

In this paper we report on the development of the audiometric equipment.

2. Analysis of the existing status of audiometric equipment and the lay-out of audiometer design

Evaluation of existing audiometric equipment involved consideration of the following factors:

– the number of audiometric examinations anticipated per year to be made using a single instrument,

– the present state of furnishing of the Mining Industry Health Service with audiometric equipment,

– the availability of appropriate equipment in the country and abroad,

– the training level of specialized personnel for performing and interpretating the audiometric examinations,

– the reasons for faults in audiometric equipment and the feasibility of repair and periodic calibration required to maintain reproductibility, uniformity and comparability of the results of examinations and their recording. The analysis revealed that the existing state of this field was unsatisfactory [1, 10].

The above conclusion was the reason for initiating a wide research program in the Technical Acoustics Group of the Central Mining Institute, Katowice. The goals of the programme were derived from the results of the above analysis. It was decided to develop an audiometer for selective – screening mass examinations, satisfying the following requirements:

1. Total elimination of contact switches in the systems for sound-level and frequency switching. The fulfilment of this requirement would eliminate the most common source of audiometer faults.

2. Full automation of acoustic pressure level and frequency changes, provided by an appropriate programme controlling these changes.

3. Full conformation with the current international and national standards and requirements concerning audiometric equipment. This requirement is necessary in view of anticipated serial production and export.

4. Maximum use of indigenous components bearing in mind the technological possibilities of the expected producer.

In view of the above assumptions, it was necessary to reject the concept of a Békésy type automatic audiometer [6, 7, 9]. This audiometer satisfies conditions 1, 3 and 4 but the result of an examination is presented in the form of plot with a subject dependent fluctuation of the record about a threshold value. The determination of the essential effect of a number of factors on the fluctuation value [8, 9] requires considerable skill for proper interpretation of the results. This is unacceptable for the mass examination of a large population. Moreover the high cost of the recording system eliminates the Békésy type audio-meter from the range of considered design concepts.

3. Design considerations

Audiometric measurement in a threshold examination consists in determining the values of three variables:

A – an independent variable defining the type of signal generated by the audiometer.

This variable is described by two parameters: frequency f and acoustic pressure level p.

B – a dependent variable defining the reaction of the hearing organ to the signal described by variable A.

This variable is also defined by two parameters: the parameter describing the reaction of the central nervous system – B_1 and the so-called *effectoric reaction* – B_2 [5] describing excitation of effectors which signal the reaction of the hearing organ of the subject.

C – a mediating variable describing the psychic features of the subject, which trigger the reaction in agreement or in disagreement with the previously obtained instruction.

The instruction informs the subject of the necessity of immediate signalling of the moment of the first perceivable hearing impression of the tone emitted by the earphones. The purposeful or accidental (e.g. as a result of misunderstanding the instruction) refraining from signalling the hearing impression results in defining the reaction as being in disagreement with the instruction. It should be stressed that determination of the intermediating variable C is only possible when the subject is able to react correctly to the audiosignal, i.e., when his hearing organ reacts to an audiosignal with parameters described by variable A and he understands the meaning of the instruction received.

The procedure of a correct audiometric examination should enable all three variables to be controlled. Most of the so-called selective (screening) audiometers [3, 4] (e.g., Klaman Granert type MA-10, DDR), Medicor type Audiomin-2 (Hungary), Kamplex type DA-2) are based on the assumption that variable C does not influence the result of the examination. Consequently the control is limited to variables A and B which are supposed to be related to each other explicitly.

The authors assumed that when control of variable C is introduced to a selective audiometer, as is required by international standards, the measurement reliability will improve substantially and it will, moreover, be possible to make use of a number of the modules of the system which have been already developed when the design is further developed towards a clinical audiometer.

The result of an audiometric measurement is defined by the hearing state function

$$L_0 = \varphi(f), \qquad L_0 \in (L_{\min}\text{-}L_{\max}), \qquad f \in (f_d\text{-}f_g), \tag{1}$$

where L_0 is the threshold pressure level (in dB), L_{min} — the minimum acoustic pressure level produced by the audiometer (in dB), L_{max} — the maximum acoustic pressure level produced by the audiometer (in dB), f_d-f_g — the frequency range of the signal produced by the audiometer, $\varphi(f)$ — a function of the signal frequency (Hz).

Relation (1) can be obtained by means of two measurement procedures P_1 and P_2 defining the set of hearing threshold values, but retaining the scheme of determining the values of the three variables, described above.

Procedure 1 – P_1:

$$P_1 = \varphi(f_i, L_i), \qquad \text{where } L_i = L_0;\ L_0 \in (L_{\min}\text{-}L_{\max}). \tag{2}$$

Procedure 2 – P_2:

$$P_2 = \varphi(f_i, L_i), \qquad \text{where } f_i = f_0;\ f_0 \in (f_d\text{-}f_g). \tag{3}$$

The results of the measurements of threshold values of the acoustic pressure, obtained by using the above procedures may fall into one of two categories:

Category I: $P_1 = P_2$. The measurement results obtained from the procedures are identical. It means that each of the two procedures realized by the audiometer is correct.

Category II: $P_1 \neq P_2$. The measurement results are not identical. It thus makes a difference in this case as to which of the procedures will be applied.

The existing literature does not provide any example of the measurement of the category II. Therefore, in further considerations, it has been assumed that the audiometer will realize procedure P_1 as is commonly used in audiometers currently produced.

It should be pointed out here that the unequivocal choice of the procedure realized by the audiometer is of essential importance for the design of an automatic audiometer. In the case of an audiometer with manual control of the switches, the operator himself fixes the examination procedure and therefore there is no need to choose it in the course of designing the instrument.

The algorithm of the program realized by the audiometer has the form given by relation (2). In the instrument developed this algorithm is realized automatically thus eliminating the influence of operator on the course and accuracy of the measurement. The operator needs only to start the program and to control the intermediating variable C. The control consists in turning on a continuous or intermittent signal facilitating perception of the acoustic stimulus at a near – threshold level and starting the subprogram for repeating the examination several times at the same frequency for accurate determination of the value of the hearing threshold. The control of variable C has been thus included in the lay-out of the audiometer in an attempt to provide an easy and functional way of limiting the influence of the error introduced by the subject.

4. Description of the system

The audiometer is composed of two functionally interrelated circuits: a digital circuit and an analog circuit [14]. The digital circuit consists of:

– a control desk with a signal button to be pressed by the subject,

– a generator controlling the automatic circuits,

– a system for the automatic control of the acoustic pressure level and signal frequency,

– a system for displaying the level values and signal frequency.

The analog circuit consists of:

– acoustic generators

– a power amplifier,

– an attenuator for the acoustic pressure level output

– earphones.

The principle of co-operation of the blocks is presented in Figs. 1 and 2 is a photograph of the instrument which is serially produced by Śląskie Zakłady Elektronicznej Aparatury Medycznej (Silesian Factory of Electronic Medical Equipment). The manufacturer's code for the instrument is AUK 431.

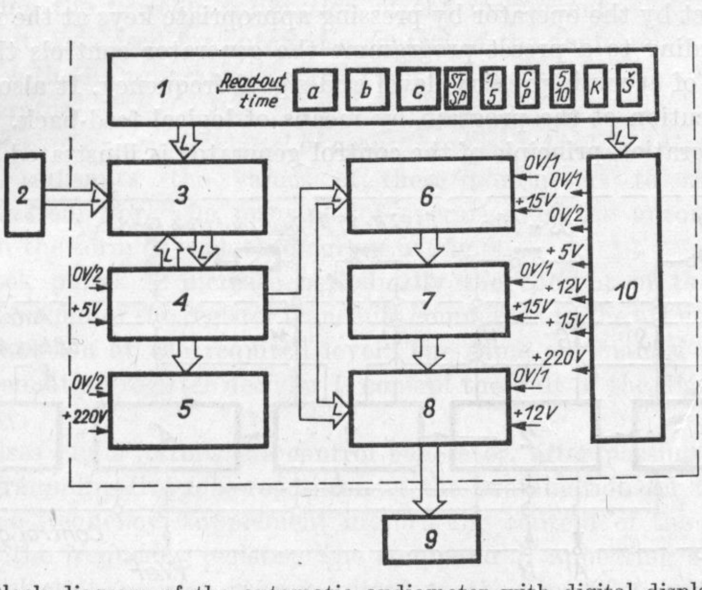

Fig. 1. Block diagram of the automatic audiometer with digital display

1 – Control desk: (a) left ear, (b) read-out, (c) right ear, 2 – Button, 3 – Control generator, 4 – Automatic system, 5 – Displays, 6 – Acoustic generators, 7 – Power amplifier, 8 – Acoustic attenuator, 9 – Earphones, 10 – Power supply

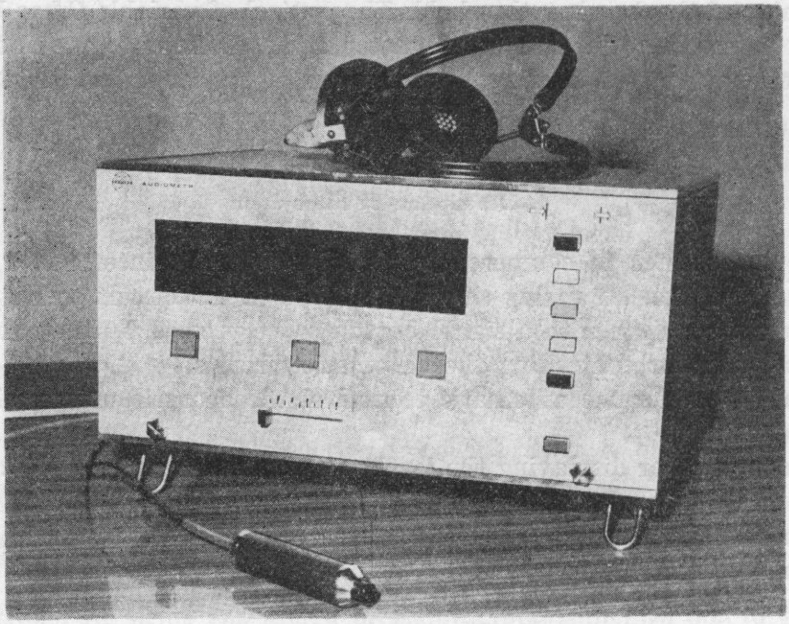

Fig. 2. Automatic audiometer with digital display manufactured by Śląskie Zakłady Elektronicznej Aparatury Medycznej w Zabrzu (Type AUK 431)

4.1. Control generator. The main block of this system is formed by a generator controlling the automatic systems. The functions executed by this generator are preset by the operator by pressing appropriate keys at the manipulation desk. According to a preset programme the generator controls the automatic adjustment of acoustic pressure level and signal frequency. It also controls the correct execution of the program by means of logical feed-back.

The operation principle of the control generator is illustrated by the block

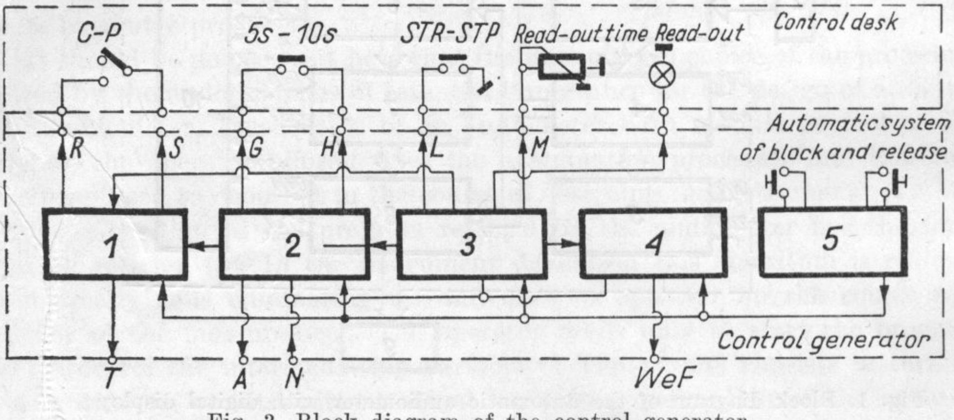

Fig. 3. Block diagram of the control generator

1 – Signal-level reduction system logic, 2 – Generator, 3 – Monostable multivibrator, 4 – Single-pulse generator, 5 – Automatic system block and control generator

diagram presented in Fig. 3. The rhythm of operation is fixed by a 1 Hz/0.5 Hz generator. At output A there appear pulses which change the state of the register of the acoustic pressure level divider. The repetition time of the pulses is controlled by the operator and is either 5 or 10 seconds. The logical systems of signal level reduction, operating synchronously with the generator, fulfil four functions:

– lowering the level of the signal in the earphones at the moment of changing the state of the level divider register,

– lowering the level of the signal in the earphones at the moment of changing the state of the frequency register,

– choosing the signal emitting system; continuous or intermittent,

– lowering the signal level in the earphones at the moment of subject reaction.

A monostable multivibrator controls the period reserved for the operator for registration of the hearing threshold of the examined person at a given frequency. The operator may control the read-out time within the range 0.5 - 50 s. The multivibrator starts at the moment of the reaction of the subject to the signal of a particular frequency. As a result the 1 Hz/0.5 Hz generator becomes blocked and the tuning out system is triggered.

After a preset operation time the multivibrator triggers a singlepulse generator. The pulses at the output of the generator serve for erasing the level attenuator register, and for increasing the value of frequency register. A possibility is provided for blocking the action of the control generator, which is utilized during the periodic calibration of the instrument.

The input N is used for automatic control of the conformity of the preset programme with the programme carried out by the automatic systems of the control generator.

4.2. System of automatic control of acoustic pressure level and signal frequency. This block transmits the values of these parameters to the generator and output attenuator. The principle of operation of the automatic unit is illustrated in the form of a block diagram in Fig. 4.

The clock pulses A increase periodically the content of the attenuator register. The output of the register transmits commands to the attenuator unit to set the attenuation at the required level. The same commands, suitably decoded in attenuation register decoder I, control the unit of the digital pressure-level display.

The pulses "in F" from the control generator, after passing through the system programming five-fold repetition of the examination for the same ear and the same frequency, supplement in turn the content of the ear-selection register and the frequency register. The command U appearing at the output of the ear selection register is transmitted to the amplifier unit and selects the ear to be examined at this moment. The output of the frequency register is supplied to attenuator register decoder I and frequency register decoder II.

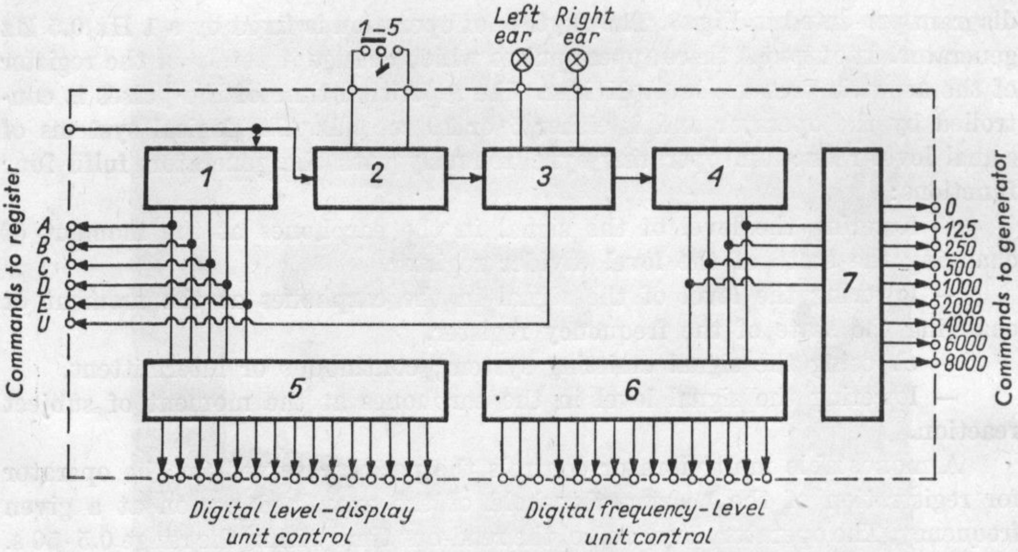

Fig. 4. Block diagram of the system for automatic control of the acoustic pressure level and the frequency

1 – Attenuator register, 2 – Five-fold repetition programme, 3 – Ear register, 4 – Frequency register, 5 – Frequency register decoder II, 6 – Attenuator register decoder I, 7 – Frequency register decoder II

The output of decoder I transmits commands to the generator to generate the signal of the required frequency. The output of decoder II produces pulses controlling the unit of digital display of the current frequency value.

4.3. Audio signal generators and power amplifier. The audio generators of the normalized frequencies required in audiometric examinations are built in the conventional lay-out of a three stage D.C. amplifier with a positive feed-back loop between the first and third stages. The amplifier contains a set of variable resistors for fixing the calibration of the output signal level and its frequency. The generator is protected against the influence of loading by subsequent units of the instrument and provided with circuits necessary for output voltage and frequency stabilization.

The power amplifier amplifies the signal from the generators to a value dictated by the input impedance of the attenuator. The circuit of the amplifier contains an optically – coupled isolator used for disconnecting the link between the amplifier stages and lowering the signal level in the earphones to a level set by standards [4]. The lowering of the level occurs at moments: of a change of signal frequency, of a change of acoustic pressure level, of receiving a signal from the subject that the test tone has been heard, and periodically testing with intermittent signal.

4.4. Output attenuator. The output of the power amplifier is connected to a step attenuator controlling the acoustic pressure level in 5 dB steps. The application of the step attenuator results both from standardization requirements [2, 4], and the adopted concept of digital presentation of the examination data. The operation principle of the attenuator is presented in Fig. 5. The power amplifier supplies a sinusoidal signal to the input denoted by "In T". The frequency decoder switches one of keys K21-K24 to open and directs

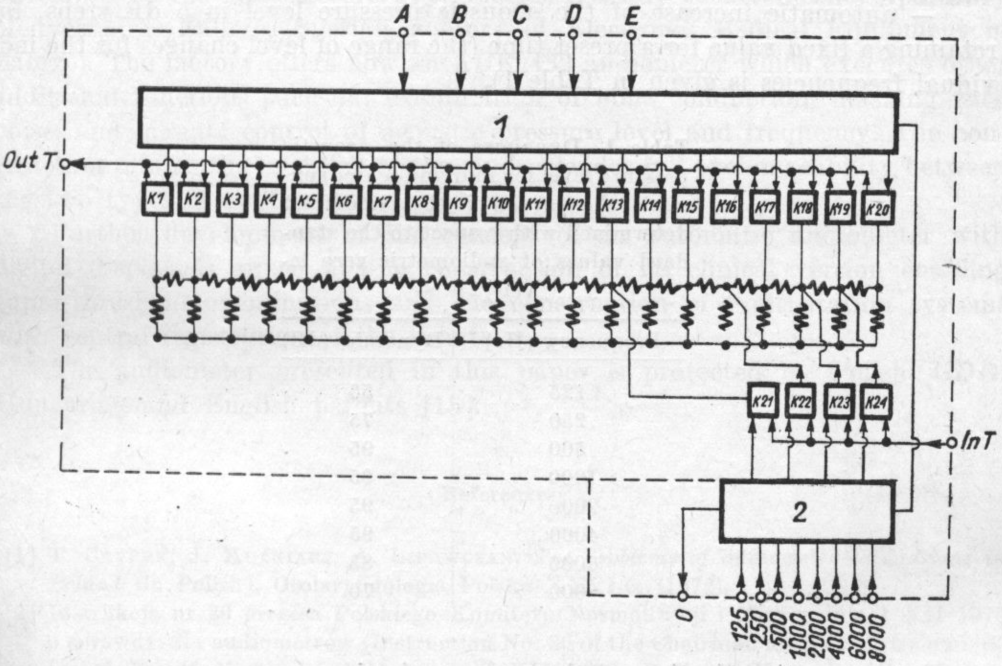

Fig. 5. Block diagram of the acoustic attenuator
1 – Level register decoder II, 2 – Frequency commands decoder

the sinusoidal signal to an appropriate point of the 5 dB resistance chain. The commands from level register II, after decoding in the decoder, open one of twenty output keys and result in the supply to the attenuator output (out "T") of an acoustic signal at a level determined by the signals at the input of the logic control A, B, C, D, E.

The logical state (0 or 1) of the voltage at output N controls the maximum value of the acoustic signal level supplied to the output attenuator for successive frequencies.

4.5. Display unit and power supply. The display unit is installed on the front plate of the instrument. It comprises two groups of digital display's (giving values of the acoustic pressure level and frequency) and three optical indicators which indicate the ear examined, i.e. left or right, and indicate a signal from the subject. The stabilized power supply built with a conventional lay – out provides the necessary voltages for supplying the functional units of the instrument.

5. Summary

The device described above performs the following functions:

— automatic selection of the frequency of the test tone for 8 normalized signal frequencies,

— automatic increase of the acoustic pressure level in 5 dB steps, but retaining a fixed value for a preset time (the range of level changes for the individual frequencies is given in Table 1),

Table 1. Dynamics of the acoustic pressure level for individual frequencies of the test tone in the audiometer determined with respect to the standard values of audiometric zero level [2]

Frequency (Hz)	Dynamics (dB)
125	55
250	75
500	95
1000	95
2000	95
4000	95
6000	85
8000	70

— automatic selection of the ear examined,

— automatic 5-fold repetition of the examination for any frequency after selecting the programme on the control desk,

— continuous control of the time of examination results read-out. After this time an automatic selection of one of the four above-described functions occurs,

— automatic and manual erasing of the performed functions and a return to the initial states.

The device is constructed in a modular form. This feature facilitates servicing by the customer. The service consists in replacing a defective module and subsequent repair of the module in the factory work shop. The design of the basic modules of the instrument permits their use in further development of the audiometer.

The block for acoustic pressure level and frequency calibration is separated from the system to enable precise periodic calibration of the audiometer. The audiometer is provided with high quality Telephonics TDH – 39 earphones which, together with establishing the audiometric zero in agreement with existing standards, make the instrument conform with all national and international requirements [2-4]. The operation of the audiometer is reduced to turning on the supply and registering the results at the moment of signalling by the subject. There is a future possibility of coupling the audiometer with a printer.

The audiometer based on the above-described principle has been further developed by the manufacturer, i.e., Śląskie Zakłady Elektronicznej Aparatury Medycznej w Zabrzu (Silesian Factory of Electronic Medical Equipment in Zabrze). The factory offers now an AUK 432 audiometer which executes other additional functions such as: examination of bone conduction, masking with noise, and manual control of acoustic pressure level and frequency. The construction retains the modular system and provides full exchangeability between the two types of audiometers.

Further development of the concept of an automatic audiometer with digital display is aimed at the construction of its clinical version enabling supra-threshold examinations and the construction of multi-station systems with central registration of the results.

The audiometer presented in this paper is protected by Polish, GDR, Hungarian and English patents [15].

References

[1] T. Ceypek, J. Kuźniarz, A. Lipowczan, *The problems of audiometer calibrations in Poland* [in Polish], Otolaryngologia Polska XXVI-5, (1972).

[2] Instrukcja nr 20 prezesa Polskiego Komitetu Normalizacji i Miar z dnia 1 XII 1975 o sprawdzaniu audiometrów (Instruction No. 20 of the Chairman of the Polish Committee of Standardization and Measures, I XII 1975), Dziennik Normalizacji i Miar 1, 9. 01. 1976.

[3] ISO Recommendation R 389, Standard Reference Zero for the Calibration of Pure-Tone Audiometers, Nov. (1964) and Addendum 1 to ISO Recommendation R 389-1964, Oct. 1970.

[4] IEC, *Pure tone audiometer for diagnostic purpose*, Publ. 177, Geneve (1965).

[5] W. Jakimowicz, *Neurologia kliniczna w zarysie* (*Introduction to Clinical Neurology*) PZWL Warsaw (1973), p. 12-19.

[6] A. Rakowski, A. Jaroszewski, T. Łętowski, *Békésy-type audiometer constructed using conventional laboratory equipment* [in Polish], Archiwum Akustyki, **4**, 3, 247-262 (1969).

[7] A. Jaroszewski, A. Rakowski, *Békésy-type audiometer with electronic control of signal level and its application in psychoacoustic measurements* [in Polish], Archiwum Akustyki, **11**, 3, 291-302 (1976).

[8] A. Jaroszewski, A. Rakowski, *Investigations of the initial phase of the process of threshold recording stabilization in automatic audiometry*, Archives of Acoustics.

[9] A. Jaroszewski, A. Rakowski, *Variability of the record of hearing threshold level after the initial stabilization phase in automatic audiometry*, Archives of Acoustics.

[10] A. Lipowczan, W. Beblo, H. Olszyczka, *Experience on legalization of acoustic and audiometric equipment* [in Polish], Proc of III-rd Conf. on Noise Prevention, Pol. Acad. Sci. Warszawa (1973).

[11] A. Lipowczan, *Noise prevention in mining industry* [in Polish], Przegląd Górniczy, 1 (1973).

[12] A. Lipowczan, L. Fajfrowski, T. Malinowski, R. Wagstyl, *Noise prevention system in the Polish coal mine industry*, Proc. of Noise Control Conf. 76, Pol. Acad. Sci. Warszawa 1976, 259-264.
[13] T. Malinowski, A. Lipowczan, *Harmful effects of noise and the methods of their determination in the light of the standing regulations*, [in Polish], Bezpieczeństwo Pracy w Górnictwie, 12, 44-48 (1971).
[14] T. Rabsztyn, A. Lipowczan, H. Olszyczka, *Automatic audiometer with digital display*, [in Polish], Proc. XX Seminar on Acoustics, Poznań 1973.
[15] T. Rabsztyn, A. Lipowczan, T. Malinowski, H. Olszyczka, *Audiometer*, Pol. Patent 86804, GDR, Pat. II 4220, U. K. Pat. 38 9401/74, Hungarian Pat. 601289.

Received on 25th Febrary 1978

ARCHIVES OF ACOUSTICS
4, 1, 39-54 (1979)

COMPUTER-AIDED RECOGNITION OF POLISH VOWELS IN CONTINUOUS SPEECH

WIKTOR JASSEM, DANUTA GEMBIAK

Acoustic Phonetics Research Unit of the Institute of Fundamental Technological Research, Polish Academy of Sciences (Poznań)

ANDRZEJ DYCZKOWSKI

Department of Computer Techniques, Mickiewicz University (Poznań)

Vowels pronounced by male voices in typical Polish sentences were the object of the recognition. Vowel formants as variable time functions were measured from the spectrograms. In the first experiment, the patterns for each phoneme in the form of two-element mean vectors and the appropriate covariance matrices were averaged over various combinations of 10 voices. In the second experiment the patterns were averaged separately for each of 10 persons. Quadratic and linear discriminant functions were used for the recognition. In general, the recognition scores in experiment I reached 75% and in experiment II – 90%. It is assumed that in the automatic recognition of Polish vowels in a computer-aided man-machine system and using two features, high scores may be obtained. They are improved by an adaptation of the system to the operator's voice.

1. Introduction

In order to define the method and technique of automatic speech recognition several decisions concerning a number of fundamental questions must be taken, of which the most crucial are the following[1]:

1. The application range of the system.
2. One or more recognition levels of the system.
3. The number of operators of the particular system.
4. The range of the dictionary and grammar of the language being recognized.
5. The range of computerization.

[1] The fundamental problems of the automatic recognition of speech were presented by Newell at al. [21].

6. The number of distinctive features and their types.

7. The mathematical identification-classification model.

The present study continues the preliminary research-work where the following assumptions concerning the above questions were made:

1. The system may be applied wherever it is sufficient to have the dictionary limited to several hundred entries uttered and transmitted in favourable conditions – especially with respect to the signal-to-noise ratio, the transmission characteristics and the reverberation conditions.

2. The recognition of each item is executed in the following steps: (a) the extraction of acoustic-phonetic parameters from the speech signal, (b) the normalization (or adaptation) that takes into account the type of voice, (c) therecognition of phonoids, (d) the recognition of words.

3. The number of operators (or types of voices) necessitating the adaptation is approximately 10.

4. The dictionary entries may be chosen freely but their number must be limited (compare above).

5. The recognizing system is of the hybrid type, i.e. peripheral analysis of the signal is executed in an analog system.

6. The extraction of acoustic-phonetic parameters and the determination of the distinctive features is carried out in the frequency domain (spectral characteristics of the signal) and at the same time the number of the features should be minimum whilst ensuring high correct-recognition scores at the final stage (words).

7. Statistical models of classification are applied in multi-dimensional spaces with the use of discriminant functions.

Although there is no apriori definition of the application of the system, the maximum reduction of costs is presupposed. This will in turn induce mass production of sufficiently versatile systems, equipped with either simple and cheap minicomputers or relatively inexpensive microprocessors.

For the time being it is not possible to state more definitely the cost of the system because of the introductory character of the present research-work and changing prices.

The basic element to be recognized in the system is the phonoid. In view of the great variety of existing and designed recognizing systems, there is no accurate and generally-accepted definition of a phonoid. It is assumed in this study that phonoid is a segmental element (i.e. one that is time-limited and defined by acoustic-phonetic parameters) subject in a particular system to the process of recognition at one of the lower levels of this process and, therefore, is an object in the sense of statistical-mathematical theory of pattern-recognition. The phonoid is a unit of the sequences constituting elements of the higher rank subject to the process of recognition (in the present case – the words). From the linguistic point of view a phonoid is represented by an allophone (see. eg. [9]) at the lower level, and by a phoneme at the higher level (ibid.). At the present stage of research it is admissible to ignore the distinction between these linguistically different levels. A phonoid may be a set of allophones of one or two phonemes (probably, never three or more). If a particular linguistic difference does not have a heavy distinctive load and necessitates many operations to ensure satisfactory discrimination (in the statistical sense), allophones of different phonemes may be treated as one phonoid[2].

2. The hypothesis

In a parametrically represented speech signal points in time which constitute boundaries between acoustic-phonetic segments can be defined – [7]. These segments are simply related to phonemes, idiophonemes and allophones (ibid.). Some of these segments also stand in particular perceptual-linguistic relations with the phones constituting stationary events. These are the segments representing the so called continuant (or liquid) sounds. Such phones may be pronounced in isolation, i.e. as separate sounds, with any duration between some 50 ms and several seconds. In continuous speech only some segments or parts of the segments may be treated as stationary events. However, there exist fairly simple relations between particular segments representing liquid sounds in continuous speech and their sustained correlates pronounced in isolation.

In the Polish language the syllabic vowels /i, ɨ, e, a, o, u/ belong to the continuant sounds. Formant frequencies are the acoustic-phonetic parameters describing them. It was indicated in [13], [14], [16] that stationary isolated Polish vowels pronounced by different male voices may be characterized by four formant frequencies allowing 100 %-correct classification and identification to be made with the adoption of appropriate probabilistic models. There is only a slight reduction of the scores after limiting the number of distinctive parameters to two, viz. F_1 and F_2. In running speech vowels do not constitute stationary events. Their formant frequencies in this case become variable time functions. These functions depend not only on the phonemic membership of the vowel but also on the proceeding and following phonetic context. Thus, the hypothesis adopted here can be formulated as follows: If the instantaneous values of the formant movements F_1 and F_2 for a particular vowel are measured at small time intervals and for each class (e.g.a phonoid, a phoneme, an allophone, an idiophoneme) a sufficient sampling of the sounds representing a given class is taken, then, on the basis of the set of data from this sample, the class may be represented in such a way as to obtain in a two-dimensional

[2] Should, for instance the difference between the allophones of the Polish phonemes /ɕ/ and /ʃ/ require the application of a considerable number of distinctive features, then this difference could be ignored, because in a lexicon of several hundred words it would never play a discriminating role. Only a few Polish words are differentiated purely by the opposition /ɕ/ : /ʃ/, e.g., proszę: prosię / proʃe(ŋ)/ : /proɕe(ŋ)/.

space (F_1, F_2), the identification area for a given class. A particular phone being the object of classification or identification at present will be represented as a trajectory in the (F_1, F_2) plane. It should be expected that at least an unambiguously definable part of this trajectory will be found in the identification area of the class which the phone to be recognized represents, so on the basis of this part the whole trajectory can unambiguously be assigned to one of the distinct classes.

3. Materials

The sequences of two, three or four phonemes most frequently occurring in Polish were given in [17]. For the purpose of the present study 6 short sentences satisfying the following conditions were composed: (a) in each sentence the most typical (i.e. the most frequent) triads (three-phoneme sequences) will be as numerous as possible; (b) the sentences will represent colloquial speech and constitute simple, casual utterances; (c) each sentence will be composed of various vowels in order to obtained an approximate balance of the total number of vowels in the materials. Trying heuristically to satisfy the above conditions the following sentences were formulated:

(1) Która godzina? /ˈktura goˈd͡ʑna/ 'what's the time?'

(2) Był pan tu już? /ˈbɨwpan ˈtujuʃ/ 'Have you been here before?'

(3) Możesz nic nie mówić? /ˈmoʒeʃ ˈɲit͡sɲe ˈmuvit͡ɕ/ 'Can you keep quiet?'

(4) Przyda ci się taki? /ˈpʃɨda ˈt͡ɕiɕe ˈtaci/ 'Can you use one like this?'

(5) Wszyscy ludzie tak robili?/ˈfʃɨst͡sɨ ˈlud͡ʑe ˈtak roˈbili/ 'Everyone did that?'

(6) Odpowiedział tylko: "Nie wiem" /otpoˈvjed͡ʑaw ˈtɨlko ˈɲevjem/ 'He only answered, 'I don't know'.

The total number of the individual phonemes represented in the above sentences is as follows: /i/7, /ɨ/5, /e/7, /a/7, /o/6 and /u/5.

It is a fact well known from the spectral analysis of speech that some speakers show more distinct vowel formants than others. There is a certain dependence (that has not been so far experimentally examined), among other things, upon the mean pitch of the voice. Therefore, in accordance with the assumptions discussed above, a greater number of voices than were actually used for the measurements were recorded and for all the voices introductory spectrographic analysis was made; the speakers whose spectrograms showed distinct formant movements, were chosen. This procedure may indicate the probable selection of the operators' voices for the future system realizing the method presented here. The number of rejected voices was however very small, the proportion being approximately 1 to 4. F_1 and F_2 values were read from the spectrograms, with an accuracy of 50 Hz, at intervals of $\Delta t = 20$ ms. The number of the bivariate measurements for each vocalic sound was always more than two, but never exceeded 10. Fig. 1 shows an illustrative wide-band spectrogram and the corresponding F_1 and F_2 formant movements, which in the entire material were calculated either from the wide or narrow-band spectrograms, or if necessary, from both. In Fig. 2 the measurements of the successive (F_1, F_2) values within the vowels occuring in the sentence from Fig. 1 are marked as points in the (F_1, F_2) plane. These points are connected and the resultant trajectory represents the individual phones. The borders between

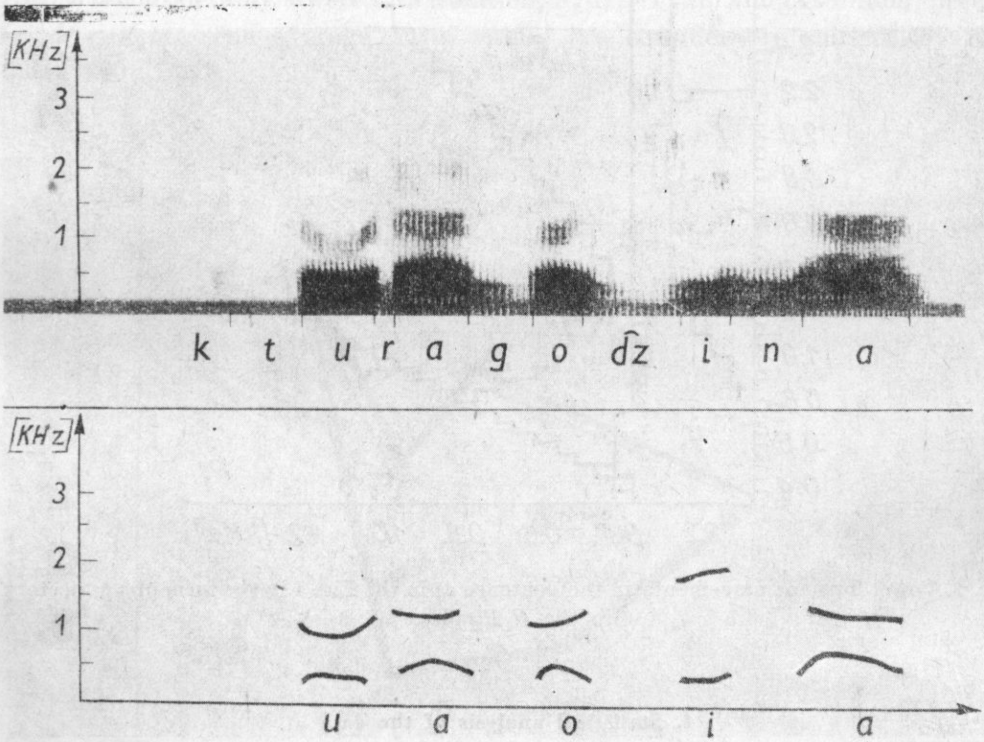

Fig. 1. Formants as time functions. A wide-band spectrogram and the formant movement in the phrase "ktura godzina" voice MN

the identification areas are marked in the (F_1, F_2) plane. These areas will be discussed in detail in the following sections.

In the present study the preliminary processing of the signal is of a semi-automatic nature because the data were obtained from visual measurements. A fully automatic system of speech recognition obviously presupposes the direct data input to the digital computer, the data being obtained either from the direct quantization of the signal or from an A/D converter after introductory processing of the signal in an analog system. At the present moment in the Acoustic Phonetics Research Unit definite progress has been made on the way to a complete automatization of the process of speech recognition (see, e.g., [19] and [20]).

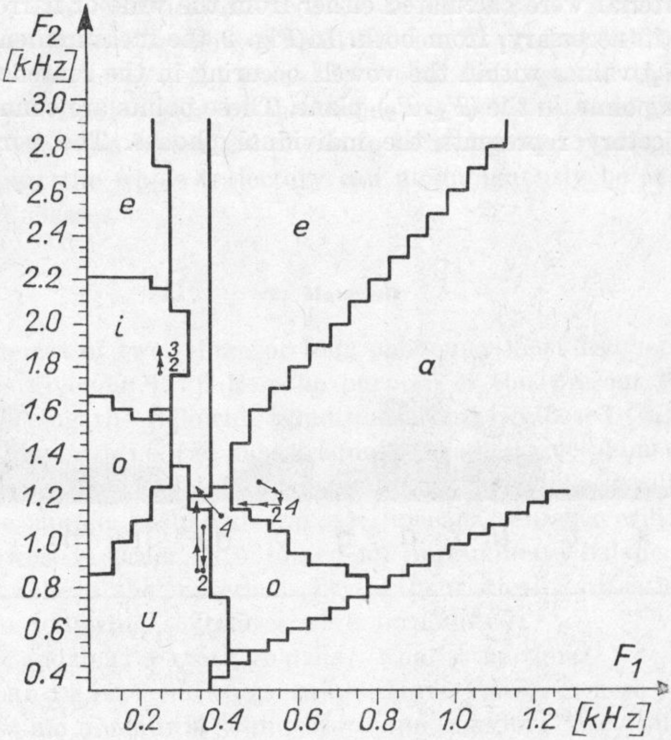

Fig. 2. Vowel-formant movements in the sentence as in the Fig. 1 in the form of a trajectory in the F_1F_2 plane

4. Statistical analysis of the data

11 speakers took part in experiment I. The data obtained from the measurements were processed in 11 variants. Each time the measurement obtained from 10 speakers constituted the design (learning) set for the individual phonemes, while the data obtained from the 11th voice constituted the test set. In each of the eleven variants the data from a different speaker were excluded from the design set. Each sample (design set) was described as a bivariate statistical distribution characterized by a mean vector and a covariance matrix (see, e.g. [5] and [24]). As each variant of experiment I differed from every other variant in the participation of one person (for 10 participants) the frequency values for a given formant of a given vowel are either identical or differ by some 1 to 2 %. The comparison of the formant values in continuous speech[3] with the data obtained for stationary vowels [13] shows the following differences:

[3] These values were presented in detail in [18].

(1) Average F_1 values are less differentiated for the vowels in continuous speech than for the isolated vowels and F_1 in continuous speech is higher for the closed vowels /i, ɨ, u/, and lower for the open vowels /e, a, o/.

(2) The average F_2 values are also closer. In running speech F_2 is lower for the vowels /i, ɨ/ and higher for the remaining vowels. The average value for /ɨ/ is considerably lowered. In isolation, $F_2/\dot{\text{ı}}/ > F_2/\text{e}/$ and in running speech – $F_2/\dot{\text{ı}}/ < F_2/\text{e}/$. The average F_2/u/ values are considerably shifted (by more than 400 Hz).

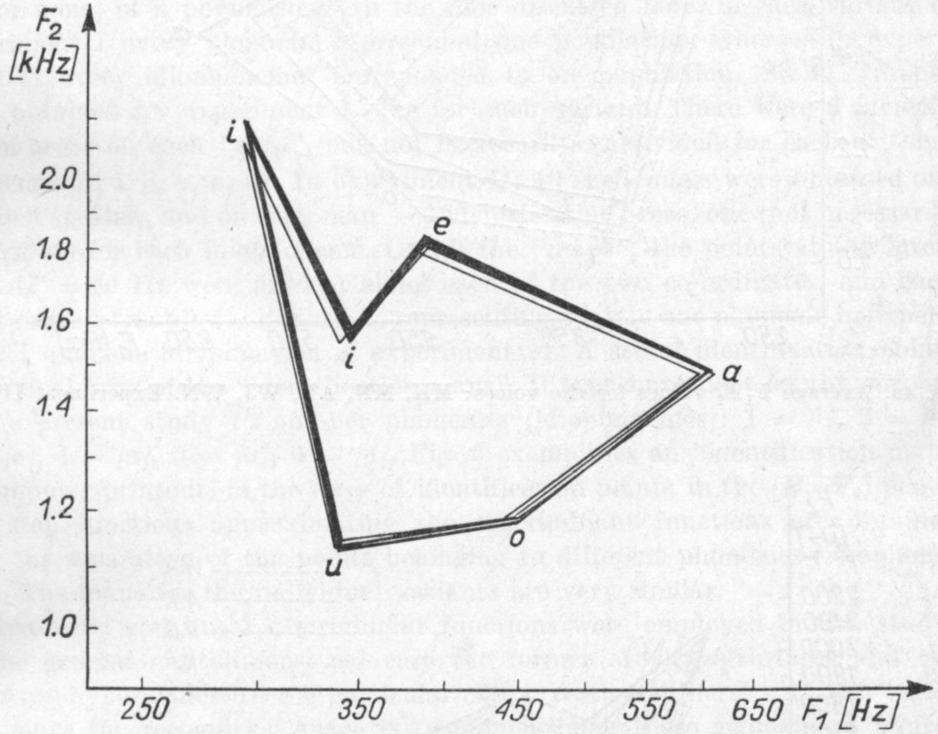

Fig. 3. Average F_1F_2 values. Experiment I

The covariance matrix is square and symmetric, so in the present case, where $s_{12} = s_{21}$, the three values s_{11}, s_{12} and s_{22} (cf. [18]) are necessary to construct it. For experiment II in which for each of the 10 voices, 5 repetitions of each sentence were analysed, the mean vectors and the covariance matrices were calculated from 4 replications constituting the design set; and the data from the fifth replication constituted the test set. Figs. 3 and 4a, b represent graphically the design data. The average F_1 and F_2 values are much more differentiated for the particular voices than for the variants of experiment I. On the other hand, the variations of individual idiophonemes in experiment II

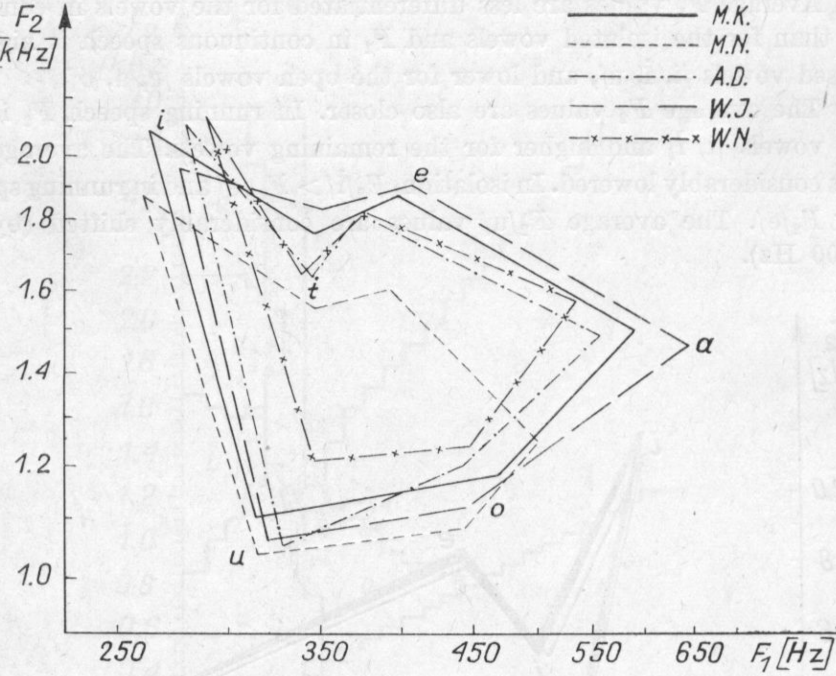

Fig. 4a. Average F_1F_2 values for the voices: MK, MN, AD, WJ, WN. Experiment II

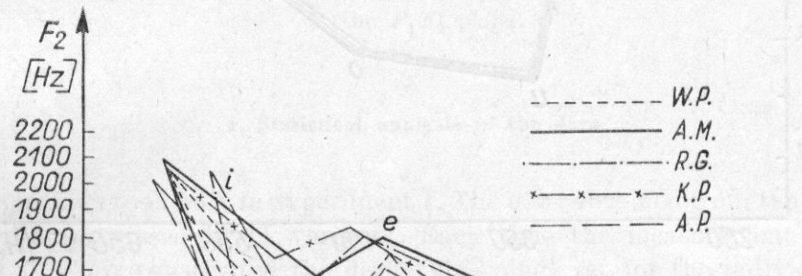

Fig. 4b. Average F_1F_2 values for the voices: WP, AM, RG, KP, AP. Experiment II

are distinctly smaller than those of the phonemes in experiment I. This suggests that the identification of vowels will give much better results in experiment II, than in experiment I.

5. Identification points

As indicated in [13], [14] and [15], in the plane with the coordinates F_1, F_2 in an orthogonal coordinate system, any point may be defined as being situated in exactly one identification area representing one (in the case of linear functions one or none) of n populations. In the case discussed here, in each variant of experiment I every phonome represented one population, whereas in experiment II every idiophoneme[4] coresponded to on population. So 11 "maps" were obtained for experiment I, one for each variant. There were 6 identification areas on each "map", one not necessarily unidivided for each of the 6 phonemes /i, ɨ, e, a, o, u/. In experiment II, 10 such maps were obtained one for each speaker, and on each map — 6 identification areas, one (not necessarily undivided) for each idiophoneme. On all the "maps", the points at the intervals $\Delta F = 50$ Hz were marked along each of the two co-ordinates, and each point on any "map" was defined as representing exactly one phoneme in experiment I and one idiophoneme in experiment II. A set of identification points is the final form of the "identification maps". It was convenient for the purpose of the present study to number phonemes (idiophonemes): 1 = /i/, 2 = /ɨ/, 3 = /e/, 4 = /a/, 5 = /o/, 6 = /u/. Fig. 5 exemplifies an "identification map" (a computer printout) in the form of identification points in the (F_1, F_2) plane. The step functions approximating the discriminant functions are obtained after the separation of the points belonging to different phonemes (idiophonemes). The maps for the individual variants are very similar.

Quadratic and linear discriminant functions were employed in this study. In the general multidimensional case the former are hypersurfaces and the latter are hyperplanes. In the particular case investigated here, with two variables, since the recognition space is twodimensional, these geometrical figures are reduced to second-order curves and straight lines respectively. The quadratic discriminant functions are expressed by the formula

$$v_{ij}(\boldsymbol{x}) = \boldsymbol{x}'(\boldsymbol{\Sigma}_j^{-1}-\boldsymbol{\Sigma}_i^{-1})\boldsymbol{x}+2(\boldsymbol{\mu}_i'\boldsymbol{\Sigma}_i^{-1}-\boldsymbol{\mu}_j'\boldsymbol{\Sigma}_j^{-1})\boldsymbol{x}+ \\ +\boldsymbol{\mu}_j'\boldsymbol{\Sigma}_j^{-1}\boldsymbol{\mu}_j-\boldsymbol{\mu}_i'\boldsymbol{\Sigma}_j^{-1}\boldsymbol{\mu}_i+\ln\frac{|\boldsymbol{\Sigma}_j|}{|\boldsymbol{\Sigma}_i|}+2\ln\frac{q_1}{q_j}, \quad (1)$$

whereas the linear functions are expressed by $\boldsymbol{b}'\boldsymbol{x}+c = 0$ in which

$$\boldsymbol{b}'\boldsymbol{x}+c \leqslant 0 \Rightarrow \boldsymbol{x}_0 \in \pi_1, \quad \boldsymbol{b}'\boldsymbol{x}+c > 0 \Rightarrow \boldsymbol{x}_0 \in \pi_2, \quad (2)$$

[4] An idiophoneme is a phonologically distinctive class of sounds consisting of sound elements pronounced by a particular voice (particular speaker) (see [9]).

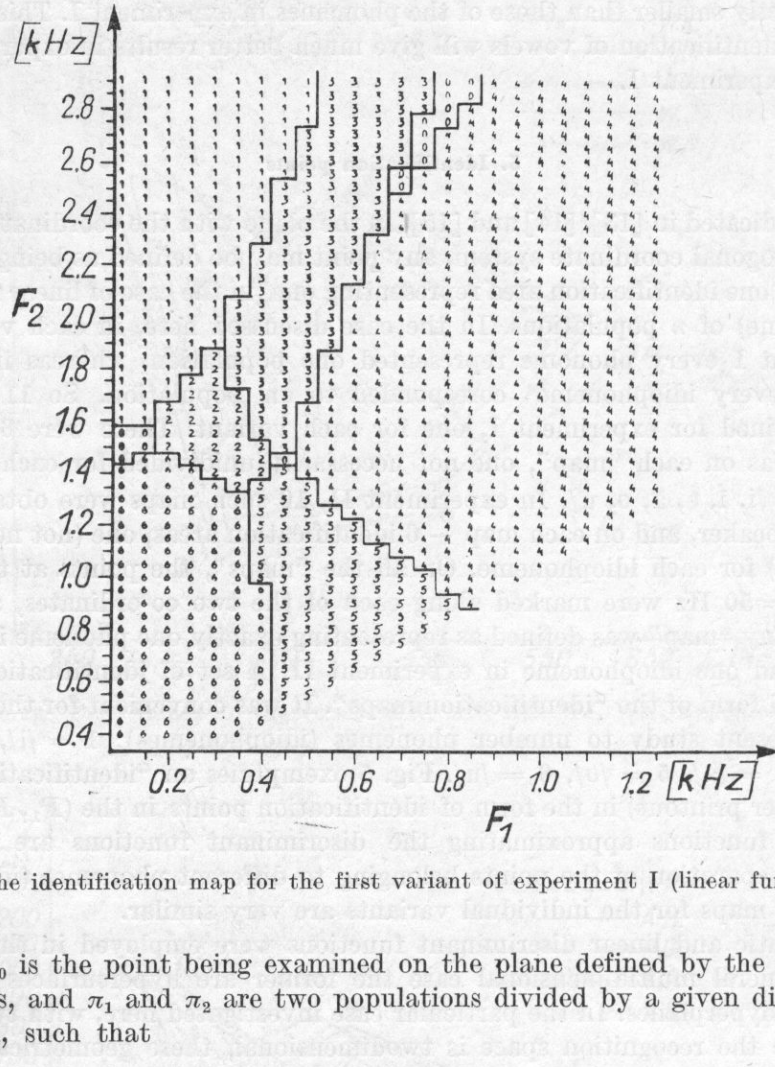

Fig. 5. The identification map for the first variant of experiment I (linear functions)

where $\boldsymbol{x}_0$ is the point being examined on the plane defined by the F_1 and F_2 variables, and π_1 and π_2 are two populations divided by a given discriminant function, such that

$$P(\pi_2\,|\,\pi_1) = P(\boldsymbol{b}'\boldsymbol{x}+c > 0) = P\left[\frac{\boldsymbol{b}'\boldsymbol{x}-\boldsymbol{b}'\boldsymbol{\mu}_1}{(\boldsymbol{b}'\Sigma_1\boldsymbol{b})^{1/2}} > \frac{-c-\boldsymbol{b}'\boldsymbol{\mu}_1}{(\boldsymbol{b}'\Sigma_1\boldsymbol{b})^{1/2}}\right] \tag{3}$$

$$= 1-\Phi-\frac{c+\boldsymbol{b}\boldsymbol{\mu}_1}{(\boldsymbol{b}'\Sigma_1\boldsymbol{b})^{1/2}} = 1-\Phi(t_1),$$

$$P(\pi_1\,|\,\pi_2) = P(\boldsymbol{b}'\boldsymbol{x}+c \leqslant 0) = P\left[\frac{\boldsymbol{b}'\boldsymbol{x}-\boldsymbol{b}'\boldsymbol{\mu}_2}{(\boldsymbol{b}'\Sigma_2\boldsymbol{b})^{1/2}} \geqslant \frac{-c-\boldsymbol{b}'\boldsymbol{\mu}_2}{(\boldsymbol{b}^{0}\Sigma_2\boldsymbol{b})^{1/2}}\right]$$

$$= 1-\Phi\left[\frac{c+\boldsymbol{b}'\boldsymbol{\mu}_2}{(\boldsymbol{b}'\Sigma_2\boldsymbol{b})^{1/2}}\right] = 1-\Phi(t_2),$$

where μ_1 and Σ_1 are the parameters (the mean vector and the covariance matrix) characterizing the population π_1, and μ_2 and Σ_2 are the parameters of the population π_2.

More details about the mathematical methods employed are presented in [14] and [18].

6. Two-stage segment→(idio)phonoid recognition

The measurements of the F_1 and F_2 values at each point in time may be represented as a point on the appriopriate identification map, and at the same time each bivariate measurement (each bivariate observation) can be classifield (identified) as representing exactly one phoneme in experiment I or exactly one idiophoneme in experiment II. Then a sequence of bivariate measurements will correspond to a sequence of classificatory decisions. Each F_1, F_2 bivariate measurement is treated as defining one 20 ms segment, and each individual vowel sound subject to the classification (identification) and represented by a sequence of bivariate measurements may then be represented as a sequence of decisions determining the successive segments. Because the formant movements within the particular vowel sounds are slow-varying continuous time functions, it may be expected that at least some of the subsequent segmental decisions will be identical (i.e. will assign segments to the same (idio)phonemes).

The method of recognizing the vowel sounds as the sequences of segments is as follows:

Each sequence of identical decisions within the vowel sound is defined as a partial sequence eg. *EEE*, *AA*, *OOO*, etc. Some of the vowel sounds were recognized as single partial sequences (e.g. one vowel sound was represented by a sequence of segmental decisions *IIIII*). These cases will be defined as unitary recognitions. Other vowel sounds were represented by two or more partial sequences. For instance, one vowel sound consisted of the sequence of decisions *EYYY*, in which two partial sequences occur, the first consisting of one element, and the second of three. Another vowel sound was recognized at the level of segmental decisions as *OEEEY*, i.e. as three partial sequences.

If an entire sequence of segmental decisions corresponding to a given vowel sound was composed of two or more partial sequences, then either one partial sequence was more numerous than the others. e.g. *EEAOOO*, or two (or possibly more) partial sequences were equally numerous, eg. *YYEEY*, *EAO* (three patrial sequences). In the first case – recognition according to the majority – the final decision was taken in favour of the longest sequence, in the second – recognition according to the order – the final decision was taken in favour of the last of the equally numerous partial sequences. So for the examples quoted above the decisions defining the asignment of a vowel sound to a particular phoneme or idiophoneme (in this study synonymous with a patricular phonoid or idiophonoid) are as follows: *IIIII→I*, *EEAOOO→O*, *YYEEY→E*, *EAO→O*.

The recognition of vowels in running speech according to the algorithm adopted here may be presented in the form of a flow diagram, as presented in Fig. 6. The computer Odra 1204 was used for the calculation of the statistical

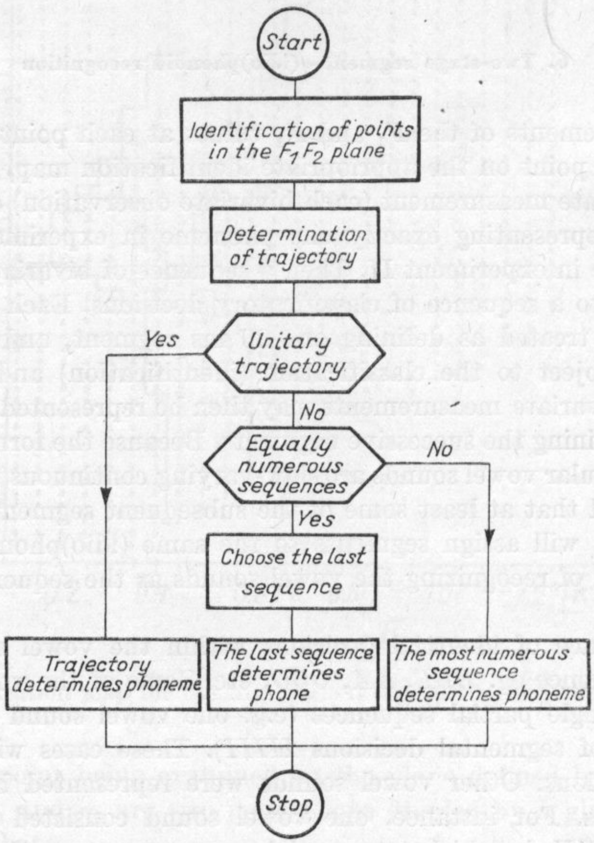

Fig. 6. The recognition of phonemes on the basis of 20 ms segments: a flow chart

parameters, the determination of the decision function, the construction of the identification maps and the execution of the decision algorithm. The programs for the calculations were presented in [13].

On the basis of the results the segment→(idio) phonoid recognition may be described as follows:

(1) Among the correct recognitions the unitary recognitions were considerably much more frequent than the recognitions according to the majority, whereas the recognitions according to the order were the least common.

(2) Among the wrong recognitions, the decisions taken according to the majority were the most numerous, the unitary decisions were slightly less

numerous, and the least frequent ones were the recognitions according to the order.

(3) The type of discriminant function seems to have no significant effect on the type of the classificatory decisions.

Among the recognitions according to the order there were more uncorrect decisions than correct ones. However, no conclusion as to the fallacy of this part of the method should be drawn from this fact for the following reasons:

(a) different types of sequences of various size were present among these recognitions and the decisions were based on various numbers of patrial sequences.

(b) according to the available material from both experiments, a slight increase in the number of the correct decisions in this type of sequences — comprising the equally numerous patrial sequences — may be acheived at the cost of considerably complicating the general form of the algorithm.

(c) the sequences of this type are very rare, and for the determination of the optimum algorithm which would take into consideration these cases, an examination of very extensive material would be necessary.

Adopting, for the present, the procedure described above, the general algorithm of the two-stage recognition on the bases of segment→(idio) phonoid, may be formulated as follows:

Choose the last of the equally numerous partial sequences all longer than the remaining sequences.

This algorithm comprises the three methods of recognizing the entire sequences representing vowel sounds as explained above. If an entire sequence is unitary, then the number of the equally numerous patrial sequences equals one. It is the only and at the same time the last patrial sequence and it is more numerous than the remaining ones, the size of which is 0. In the case of two or more patrial sequences one of which is the longest, the number of the remaining ones is 1 or more.

7. The results of the recognition

7.1. Experiment I. Table 1 presents the results of the recognition of the vowel sounds in each variant of experiment I using quadratic and linear functions. The figure representing the percentage of correctly recognized sounds of each variant are contained between 70 and 80% for the quadratic functions and 60-80% for the linear functions. The results of the classification in the design set (i.e. the recognition in training) differ only slightly form the results of one identification in the test set. This indicates that the choice of the discriminant functions was correct. The absence of a distinct difference in the overall results as between the methods based on the quadratic and the linear functions results from the fact that the segments of the conical curves separating the identification areas show very small curvatures.

Averaging over all the vowels and variants the percent recognition scores for each person are as shown in Table 2. Some voices (e.g. III and IX) give better results than the remaining ones. This reflects the fact that the formant movements of the vowels pronounced by these voices are more typical in the sense of the method adopted here.

The classification in the design set and the identification in the test set were much better in experiment I for the extreme vowels /i, a, u/ than for the remaining vowels[5].

7.2. Experiment II. As could be expected from the data Figs. 3 and 4a, b the results of the recognition of vowels in experiment II turned out to be much better than those of experiment I.

The data in Table 3 indicate that the differences between the voices are inconsiderable. The classification in the design set, as well as the identification in the test set resulted in scores of about 90 %. The best results were obtained for the idiophones /i, a, o/. The overall scores are considerably lowered by the results obtained for /e/, which are markedly lower than those for the remaining phonemes. It may be assumed that an allophonic division of the (idio) phoneme /e/ would yield better results. Two dependent phonoids would represent the phoneme /e/ in automatic recognition.

8. Conclusions

8.1. Experiment I simulates the situation in which a technical system recognizes the vowels within an utterance pronounced by a certain number (assumed to be sufficient) of other voices. Experiment II, on the other hand, simulates the situation where the system has either previously recognized the operator's voice (or at least the type of the operator's voice) or has been tuned to his voice.

If for certain purposes the recognition accuracy of vowels of about 75 % would be acceptable, then there is no need for tuning the system to the operator's voice or recognizing it by the system.

However, if accuracy of about 90 percent is required, then the tuning or recognition of the operator's voice is necessary.

8.2. If the spoken vowels are assigned to the proper idiophonemes, then some of them viz. those representing the phoneme /i/ will get almost one-hundred-percent correct recognition scores, whilst the others will get about 90 %.

Only /e/ and possibly /ɨ/ result in less satisfactory recognition, /e/ and possibly /ɨ/ could be recognized according to more idiophonemes than one

[5] See detailed data in [18].

Table 1. Correctness of recognition for individual variants of experiment II [%]

Variant	Quadratic functions		Linear functions	
	pattern set	test set	pattern set	test set
1	78	70	77	81
2	80	70	79	68
3	77	78	75	84
4	77	70	76	73
5	79	68	77	70
6	77	73	78	60
7	77	78	76	81
8	78	78	76	78
9	78	81	75	87
10	77	81	76	78
11	76	81	76	76
Average	78	75	77	76

Table 2. Correctness of classification for individual voices in experiment I [%]

Voice	Quadratic functions	Linear functions
I	79	81
II	69	67
III	80	84
IV	73	74
V	73	71
VI	74	62
VII	79	81
VIII	79	79
IX	81	86
X	82	79
XI	84	77

Table 3. Correctness of recognition for individual voices in experiment I [%]

Voice	Quadratic functions		Linera function	
	pattern set	test set	pattern set	test set
AD	95	95	93	92
WN	92	95	92	95
WJ	93	97	92	97
MŁ	92	97	92	95
MN	93	89	92	89
WP	80	81	83	81
AM	90	89	88	86
RG	92	92	93	92
KP	88	84	86	78
AP	81	92	78	86
Average	90	91	89	89

(probably not more than 2) should this turn out to be necessary. This would presumably result in overall recognition scores approaching one-hundred-percent correctness, at least for some selected voices.

8.3. It has been shown that the recognition of Polish vowels in continuous speech is possible with the use of only two classificatory features, namely the instantaneous values of F_1 and F_2 measured with an accuracy of 50 Hz, at 20 ms intervals. This method permits the identification to be executed in a hybrid analog-digital system processing a small number of data. Assuming the range for F_1 from 200 Hz to 1100 Hz, and 400-2700 Hz for F_2, and taking into consideration that by definition $F_2 > F_1$ and a part of the (F_1, F_2) plane is not used because of articulatory constraints (in real speech high F_2 values do not occur together with relatively high F_2 values), the number of the identification points may be limited to no more than about 500. The possibility of realizing the method adopted here in real time using hybrid systems is at present the object of further research.

References

[1] T. Caliński, A. Dyczkowski, Z. Kaczmarek, *Identification of observation using a dividing hyperplane* [in Polish], Rocznik Akademii Rolniczej w Poznaniu, *Algorytmy biometryczne i statystyczne*, No 4 (1975).

[2] G. I. Gjemel, *Recognition of speech signals* [in Russian], izd. Nauka, Moskwa 1971.

[3] M. Dierkacz, R. Gumieckij, Ł. Miszin, M. Owierczenko, M. Czaban, *The perception of speech in recognition models* [in Russian], izd. Iwowskowo Uniwersiteta, Lwów 1971.

[4] H. Dudley, S. Balashek, *Automatic recognition of phonetic patterns in speech*, JASA, **30**, 721-732 (1958).

[5] K. Hope, *Methods of multivariate analysis*, Univ. of London Press, London 1968.

[6] W. Jassem, *Vowel formant frequencies as cues to speaker discrimination*, *Speech Analysis and Synthesis*, Vol. I, PWN, Warszawa, 1968, 9-42.

[7] W. Jassem, *Phonetic-acoustic assumptions for the automatic recognition of phonemes* [in Polish], IFTR Reports 17/70, Warszawa (1970).

[8] W. Jassem, *Phonological segmental units in the speech signal, form and substance*, Akademik Verlag, Odense, 181-192 (1971).

[9] W. Jassem, *Speech and communication*, [in Polish], PWN, Warszawa 1974.

[10] W. Jassem, T. Caliński, Z. Kaczmarek, *Vowel formant frequencies as personal voice characteristics* [in Polish], IFTR Reports 5/70, Warszawa (1970).

[11] W. Jassem, T. Caliński, Z. Kaczmarek, *Investigation of vowel formant frequencies as personal voice characteristics by means of multivariate analysis of variance*, *Speech Analysis and Synthesis*, V. II, PWN, Warszawa 1970, 7-40.

[12] W. Jassem, L. Frąckowiak, *Vowel formant frequencies as a distinctive feature of speakers' voices* [in Polish], Biuletyn Polskiego Towarzystwa Językoznawczego, V. XXVI, Kraków, 67-99 (1968).

[13] W. Jassem, M. Krzyśko, A. Dyczkowski, *Classification and identification of Polish vowels on the basis of formant frequencies* [in Polish], IFTR Reports 14/72, Warszawa (1972).

[14] W. JASSEM, M. KRZYŚKO, A. DYCZKOWSKI, *Identification of isolated Polish vowels* [in Polish], Archiwum Akustyki, **9**, 3, Warszawa, 261-287 (1974).

[15] W. JASSEM, M. KRZYŚKO, A. DYCZKOWSKI, *Sequential identification of vowels* [in Polish], IFTR Reports 6/74, Warszawa (1974).

[16] W. JASSEM, M. KRZYŚKO, A. DYCZKOWSKI, *Verification of voices on the basis of vowel formant frequencies* [in Polish], Archiwum Akustyki, **9**, 1, Warszawa, 3-26 (1974).

[17] W. JASSEM, P. ŁOBACZ, *Phonotactic analysis of Polish text*, [in Polish], IFTR Reports 63/71, Warszawa (1971).

[18] W. JASSEM, D. SZYBISTA, A. DYCZKOWSKI, *Recognition of Polish vowels in typical sentences* [in Polish], IFTR Reports 43/75, Warszawa (1975).

[19] H. KUBZDELA, *Technical realization of the formant method for the recognition of Polish vowels* [in Polish], IFTR Reports 90/75, Warszawa (1975).

[20] K. MYTKOWSKI, *Analog function channel of the type KF-01 for the input and output of data in an "On – line" system to/from the storage of the minicomputer MOMIK-8B/100*, [in Polish], IFTR Reports 39/76, Warszawa (1976).

[21] A. NEWELL, J. BARNETT, J. W. FORGIE, C. GREEN, D. KLATT, J. C. R. LICKLIDER, J. MUNSEN, D. R. REDDY, W. A. WOODS, *Speech understanding systems, Final report of a study group*, North Holland, American Elsevier, 1973.

[22] G. E. PETERSON, M. L. BARNEY, *Control methods used in a study of the vowels*, JASA, **24**, 175-184 (1952).

[25] R. W. A. SCARR, *Zero-crossing as a means of obtaining spectral information in speech analysis*, IEEE Trans. on Audio Electroacoustics, AU-16, 247-255 (1968).

[24] M. M. TATSUOKA, *Multivariate analysis: techniques for educational and psychological research*, Wiley and Sons, New York 1971.

[25] N. G. ZAGORUJKO, *The methods of recognition and their application* [in Russian], izd. Sow. radio, Moskwa, 1972.

Received on 14th September 1977

ARCHIVES OF ACOUSTICS
4, 1, 57-64 (1979)

PROPAGATION OF ELASTIC WAVE IN SOLID LAYER-LIQUID SYSTEM*

ANNA GRABOWSKA

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

The paper presents the solution of the problem of a flat wave propagating without attenuation along a solid layer placed between a semispace filled with liquid and the vacuum.

The wave equation of scalar and vector potentials of displacement has been solved for this case. A characteristic equation accounting for boundary conditions has been derived. This equation has been solved numerically and it has been shown that in these conditions the wave can propagate at a velocity slightly smaller than the wave velocity in the liquid.

The distributions of stress, acoustic pressure, and displacement of the propagating wave have been determined numerically for a layer of a thickness $a = 0.075$ cm and $a = 0.010$ cm, contacting with water on one side, and for a frequency of 3×10^6 Hz. The type of wave is close to a surface wave.

1. Introduction

The problem of the propagation of an elastic wave in the solid layer-liquid system originated in the course of ultrasonic investigations of tumours. Using a probe containing a piezoelectric transducer and generating ultrasonic waves, a needle is driven through the patient's skin toward the tumour. The punctured organ is observed by an ultrasonic visualization system. The needle passes through a hole in the centre of the piezoelectric transducer which has the form of a plate.

In the course of these investigations it has been observed that the wave propagating along the needle placed already inside the body is accompanied by a wave which after reaching the end of the needle is reflected backwards and returns giving an image of the needle end on the oscilloscope screen. This effect makes it possible to locate precisely the puncture and to sample the tumour tissue instead of cutting the whole organ apart.

---

* The paper was written under problem MR.I.24.

This work is aimed at investigation of the effects accompanying the propagation of the above wave along the needle surrounded by the body tissue. We shall reduce this problem to the consideration of a wave propagating in a flat solid layer-liquid system. In addition we shall assume that the considered wave is a running, continuous sinusoidal wave.

The solution of this problem consists in a description of an elastic wave propagating in an infinite isotropic homogeneous solid layer in contact with an mimobile and infinitely deep liquid on one side and with the vacuum on the other.

2. Basic equations

The coordinate system is chosen as follows (cf. Fig. 1). The x-axis coincides with the upper edge of the layer and is parallel to the direction of the wave propagation. The z-axis is directed vertically upwards. The layer thickness

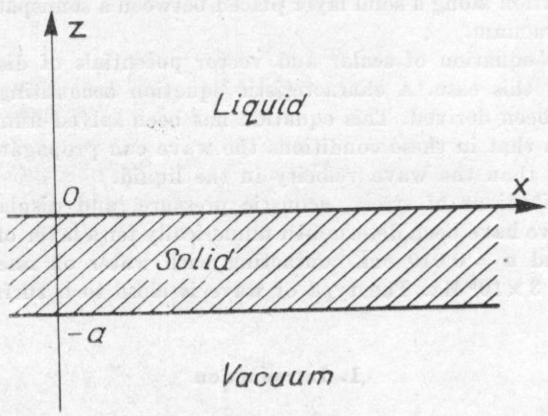

Fig. 1. The considered solid layer contacting liquid on one side

is a; the densities ϱ_s, λ and μ are the Lamé constants and the velocities of longitudinal and transverse waves in the material are c_d and c_t, respectively. The liquid denisty is ϱ_c and the longitudinal wave velocity in the liquid is c_0.

The displacement vector $\boldsymbol{u}$ can be presented in the form

$$\boldsymbol{u} = \boldsymbol{v} + \boldsymbol{w} \tag{1}$$

with conditions curl $\boldsymbol{v} = 0$, div $\boldsymbol{w} = 0$.

It follows from the vector analysis that such representation of a vector field is always possible. This is a representation of a vector in the form of a sum of the gradient of a scalar potential φ and the curl of a vector potential $\boldsymbol{\psi}\,(\psi_x, \psi_y, \psi_z)$:

$$\boldsymbol{u} = [u, v, w] = \operatorname{grad}\varphi + \operatorname{rot}\boldsymbol{\psi}. \tag{2}$$

In view of a two-dimensional character of the problem, the vector potential $\boldsymbol{\psi}$ contains only one component ψ_y.

Eq. (2) leads to the following expressions for the components of displacement u, described in the layer by potentials $\varphi_s(x, z, t)$, $\psi_y(x, z, t)$ and in the liquid by $\varphi_0(x, z, t)$:
– in the layer

$$u_s = \frac{\partial \varphi_s}{\partial x} - \frac{\partial \psi_y}{\partial z}, \quad w_s = \frac{\partial \varphi_s}{\partial z} - \frac{\partial \psi_y}{\partial x}; \tag{3a}$$

– in the liquid

$$u_0 = \frac{\partial \varphi_0}{\partial x}, \quad w_0 = \frac{\partial \varphi_0}{\partial z}, \tag{3b}$$

where u_s, w_s are components of displacement vector, parallel to the x-axis, w_s, w_0 – components of displacement vector normal to the x-axis.

Potentials φ_s, ψ_y, φ_0, satisfy the following wave equation:

$$\nabla^2 \varphi_s = \frac{1}{c_d^2} \frac{\partial^2 \varphi_s}{\partial t^2} \quad \text{in the layer},$$

$$\nabla^2 \psi_y = \frac{1}{c_t^2} \frac{\partial^2 \psi_y}{\partial t^2} \quad \text{in the layer} \tag{4}$$

$$\nabla^2 \varphi_0 = \frac{1}{c_0^2} \frac{\partial^2 \varphi_0}{\partial t^2} \quad \text{in the liquid}.$$

Eq. (4) is satisfied by any periodic function. By separating the variables with respect to eqs. (4) we obtain the following solution:

$$\varphi_s(x, z, t) = [A_1 \cos k_d z + A_2 \sin k_d z] e^{-jkx} e^{j\omega t},$$

$$\psi_y(x, z, t) = [B_1 \cos k_t z + B_2 \sin k_t z] e^{-jkx} e^{j\omega t}, \tag{5}$$

$$\varphi_0(x, z, t) = E e^{-jk_0 z} e^{-jkx} e^{j\omega t},$$

where

$$k_d^2 = \frac{\omega^2}{c_d^2} - k^2, \quad k_t^2 = \frac{\omega^2}{c_t^2} - k^2, \quad k_0^2 = \frac{\omega^2}{c_0^2} - k^2 \tag{5a}$$

and $c = \omega/k$. The potentials φ_s, ψ_y, φ_0 in (5) describe waves propagating along the x-axis with a phase velocity c and a wavelength λ related to the wave number by relation $k = 2\pi/\lambda$. The frequency f is given by relation $f = \omega/2\pi$.

The normal stress τ_{zz} and shear stress τ_{zx} in the solid can be expressed by potentials φ_s and ψ_y and by elastic constants as follows:

$$\tau_{zz} = \lambda \nabla^2 \varphi_s + 2\beta \left(\frac{\partial^2 \varphi_s}{\partial z^2} + \frac{\partial^2 \psi_y}{\partial x \partial z} \right),$$

$$\tau_{zx} = \eta \left(\frac{\partial^2 \psi_y}{\partial x^2} - \frac{\partial^2 \psi_y}{\partial z^2} + 2 \frac{\partial^2 \varphi_s}{\partial x \partial z} \right). \tag{6}$$

The acoustic pressure in liquid is given by relation

$$p = -\varrho_c \frac{\partial^2 \varphi_0}{\partial t^2}. \tag{7}$$

3. Boundary conditions and characteristic equation

The solution of the problem should satisfy appropriate boundary conditions posed by the requirement of continuity of stresses and displacements perpendicular to the surface of the layer. These conditions are as follows:

$$\begin{aligned} &\tau_{zz} = -p \ (z = 0), \quad \tau_{zx} = 0 \ (z = 0), \quad w_s = w_0 \ (z = 0), \\ &\tau_{zz} = 0 \ (z = -a), \quad \tau_{zx} = 0 \ (z = -a). \end{aligned} \tag{8}$$

In agreement with the convention used in acoustics [2] it has been assumed that positive tensile stress corresponds to negative pressures. Hence the negative sign appears at acoustic pressure in the first boundary condition (8).

Putting relations (5) into boundary conditions (8) and making use of (3a), (3b), (6) and (7) we obtain a set of five homogeneous equations with unknown coefficients A_1, A_2, B_1, B_2, E. The solution to this set requires the determinant W formed of the coefficients of this set to vanish. This determinant has the form

$$W = \begin{vmatrix} \omega^2\varrho_s - 2\mu k^2 & 0 & 0 & 2\mu jkk_t & -\varrho_c\omega^2 \\ 0 & 2jkk_d & k^2 - k_t^2 & 0 & 0 \\ 0 & k_d & -jk & 0 & jk_0 \\ (\omega^2\varrho_s - 2\mu k^2)\cos k_d a & (2\mu k^2 - \omega^2\varrho_s)\sin k_d a & 2\mu jkk_t \sin k_t a & 2\mu jkk_t \cos k_t a & 0 \\ 2jkk_d \sin k_d a & 2jkk_d \cos k_d a & (k^2 - k_t^2)\cos k_t a & (k_t - k^2)\sin k_t a & 0 \end{vmatrix}. \tag{9}$$

The characteristic equation is the very condition of vanishing of this determinant and has the form

$$\begin{aligned} &(\omega^2\varrho_s - 2\mu k^2)^2 (k^2 - k_t^2)^2 k_0 \sin k_t a \cdot \sin k_d a + \\ &\quad + 4\varrho_c \omega^2 \mu j (k^2 + k_t^2)^2 k^2 k_d^2 k_t \cos k_t a \cdot \sin k_d a + \\ &\quad + 16\mu^2 k^4 k_d^2 k_t^2 k_0 \sin k_d a \cdot \sin k_t a + \\ &\quad - (\omega^2\varrho_s - 2\mu k^2)\varrho_c \omega^2 k_d (k^4 - k_t^4) j \sin k_t a \cdot \cos k_d a - \\ &\quad - 8(\omega^2\varrho_s - 2\mu k^2)\mu (k^2 - k_t^2) k^2 k_d k_t k_0 + \\ &\quad + 8(\omega^2\varrho_s - 2\mu k^2)\mu (k^2 - k_t^2) k^2 k_d k_t k_0 \cos k_d a \cdot \cos k_t a = 0. \end{aligned} \tag{10}$$

If we put $\varrho_c = 0$ in characteristic equation (10) we obtain the characteristic equation for the layer in the vacuum as given by Ewinge and Jardetzky [1].

The characteristic equation (10) determines the relation between the phase velocity c of the wave and the wave number k. The relation between the phase velocity c and the angular frequency ω can be obtained after putting the relation $k = \omega/c$ in (10).

The solutions of (10) can be real or complex. The real values of k satisfying (10) correspond to wave propagation along the x-axis without attenuation. The complex values of k correspond to propagation with attenuation.

The characteristic equation (10) has been solved numerically for the following parameters.

(a) The layer is made of steel:

$$\varrho_s = 7.7 \text{ g/cm}^3, \quad \lambda = 1.07 \times 10^{12} \text{ g/cm}\cdot\text{s}^2,$$
$$\mu = 8.03 \times 10^{11} \text{ g/cm}\cdot\text{s}^2, \quad c_d = 5.9 \times 10^5 \text{ cm/s},$$
$$c_t = 3.23 \times 10^5 \text{ cm/s}.$$

The layer thickness was $a = 0.075$ cm in the first case and $a = 0.010$ cm in the second case. These thickness are equal to the thickness of the needle used for puncturing the body tissue.

(b) Liquid—water:

$$\varrho_c = 1 \text{ g/cm}^3, \quad c_0 = 1.48 \times 10^5 \text{ cm/s}.$$

(c) Frequency:

$$f = 3\text{MHz}.$$

Only the real values of k were taken into account when solving (10) since they correspond to the waves which are not attenuated in the x-direction. Under such an assumptions it has been obtained: for $a = 0.075$ cm, propagation constant $k = 127.4 \text{ cm}^{-1}$, wave numbers, $k_d = j\cdot 123 \text{ cm}^{-1}$, $k_t = j\cdot 113 \text{ cm}^{-1}$, $k_0 = -j\cdot 3.12 \text{ cm}^{-1}$. Then the phase velocity of the surface wave is

$$c = \omega/k = 1.47840 \times 10^5 \text{ cm/s},$$

i.e., it is slightly below the assumed velocity of wave in water. As it follows from (5a), k_0 is imaginary. Then, by virtue of (5) the first exponent of potential $\varphi_0(x, z, t)$ will be real. Thus the wave considered decays in liquid with the increase of depth z.

The distributions of normal and shear stresses, of the acoustic pressure, and of the components of displacement vector of the propagating wave have been calculated numerically for the above example with the aid of eqs. (6), (7), (3a), (3b) (for $x = t = 0$). These equations, after some elementary transformations, assume the form:

$$\tau_{zz} = \lambda[C_1\cos k_d a - C_2 \sin k_d a](-k^2 - k_d^2) +$$
$$+2\mu[-k_d^2(C_1\cos k_d a - C_2\sin k_d a) + (-D_1 \sin k_t a - D_2\cos k_t a)jkk_t], \tag{11}$$

$$\tau_{zx} = \mu[(D_1 \cos k_t a - D_2 \sin k_t a)(k_t^2 - k^2) + 2jkk_d(-C_1 \sin k_d a - C_2 \cos k_d a)], \tag{12}$$

$$
\begin{aligned}
p &= -\omega^2 \varrho_c E e^{-jk_0 a},\\
w_s &= (C_1 \sin k_d a + C_2 \cos k_d a) k_d - (D_1 \cos k_t a - D_2 \sin k_t a) jk,\\
u_s &= -(C_1 \cos k_d a - C_2 \sin k_d a) jk + (D_1 \sin k_t a + D_2 \cos k_t a) k_t, \qquad (13)\\
w_0 &= -jk_0 E e^{-jk_0 a}, \quad u_0 = -jkE e^{-jk_0 a}.
\end{aligned}
$$

The plots of the stresses and displacements are presented in Figs. 2 and 3.

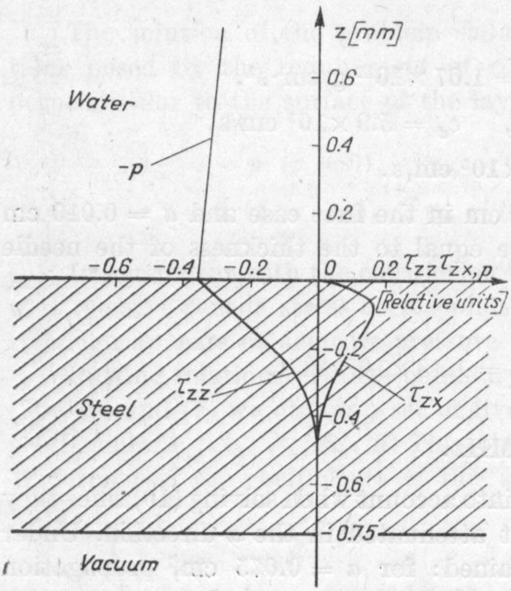

Fig. 2. Distribution of normal stress τ_{zz} and shear stress τ_{zx} in the steel layer and the distribution of acoustic pressure p in water; $f = 3\times 10^6$ Hz, $a = 0.075$ cm

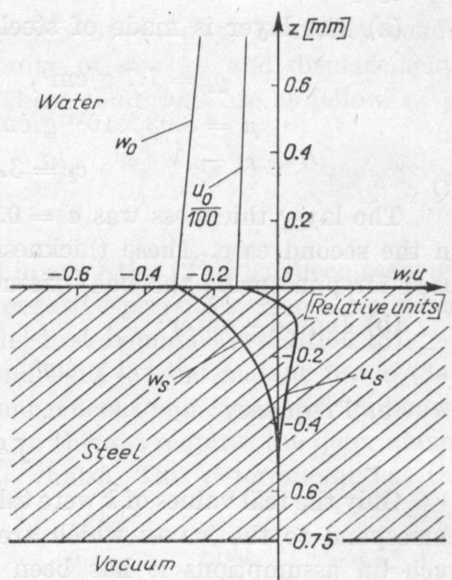

Fig. 3. Distribution of the components of displacement vector in the steel layer and in water for $a = 0.075$ cm

For $a = 0.010$ cm the propagation constant $k = 132\ \text{cm}^{-1}$, wave numbers $k_d = j\cdot 128\ \text{cm}^{-1}$, $k_t = j\cdot 118\ \text{cm}^{-1}$, $k_0 = -j\cdot 34.7\ \text{cm}^{-1}$. The velocity of surface wave is $c = 1.43\cdot 10^5$ cm/s. The distributions of acoustic pressure, stresses, and displacements are presented in Figs. 4 and 5.

4. Conclusions

The paper is concerned with the problem of propagation of an elastic, running wave along an infinite solid layer placed between the semispace filled with liquid and the vacuum.

By solving wave equations (4) with boundary conditions (8) the characteristic equation (10) has been obtained. Unlike in the case of the layer in the vacuum [3], this equation cannot be decomposed into two terms describing the symmetric and antisymmetric modes.

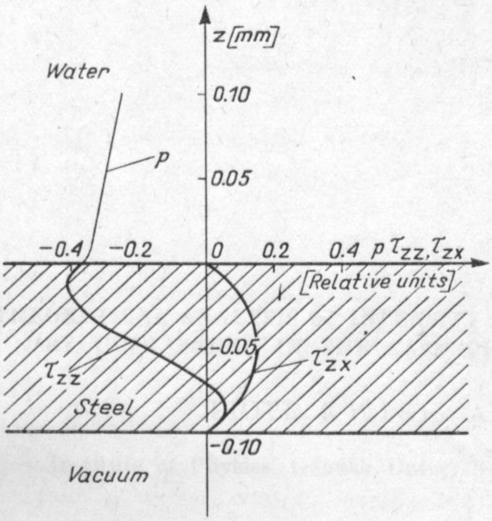

Fig. 4. Distribution of stresses τ_{zz} and τ_{zx} in the steel layer and the distribution of acoustic pressure p in water; $f = 3\times 10^6$ Hz, $a = 0.010$ cm

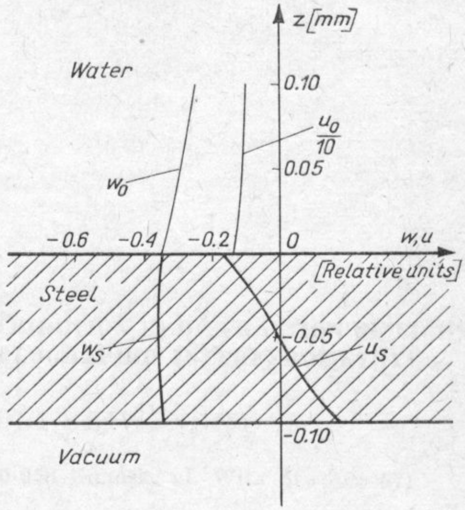

Fig. 5. Distribution of components w and u of displacement vector in the steel plate and in water for $a = 0.010$ cm

The characteristic equation (10) determines the phase velocity c of the wave as a function of the angular frequency ω. The phase velocity c may be real or complex and in a general case depends on the frequency.

In the case considered in this paper it has been assumed that the wave in question should be unattenuated in the direction of its propagation. For steel layers of thicknesses of 0.075 cm and 0.01 cm, equal to the thicknesses of needles used for puncturing the body tissue, and for the assumed wave velocity in water $c_0 = 1.48\times 10^5$ cm/s it has been found that for a frequency of 3×10^6 Hz the velocities are $c = 1.47840\times 10^5$ cm/s and $c = 1.42810\times 10^5$ cm/s for the thick and the thin layer, respectively. These velocities are only slightly lower than c_0. The wave decays exponentially with the increase of the penetration depth in water. The penetration depth of this wave in water is much larger than its penetration depth in the layer. The wave is conducted without attenuation along the solid-liquid interface; its character resembles that of a surface wave.

References

[1] W. Ewing, W. Jardetzky, *Elastic waves in layered media*, McGraw-Hill Book Company, INC, 1957, 281-288, 293-299.
[2] I. Malecki, *Teoria fal i układów akustycznych* (*The theory of waves and acoustic systems*), PWN, Warszawa, 1964, 79-82.
[3] M. Redwood, *Mechanical wave-guides*, Pergamon Press, Oxford 1960, 119-121.

Received on 15th November 1977

ARCHIVES OF ACOUSTICS
4, 1, 65-72 (1979)

THEORETICAL ANALYSIS OF INTENSITY DISTRIBUTION IN DIFFRACTION OF LIGHT BY ULTRASOUND IN THE CASE OF HOLOGRAPHIC INTERFEROMETRY*

PIOTR KWIEK, ANTONI ŚLIWIŃSKI

Institute of Physics, Gdańsk University (80-950 Gdańsk, ul. Wita Stwosza 57)

The theoretical background for optical holographic interferometry used in ultrasonic field investigations is presented. The double-exposure pulse technique, the time averaged technique, and also the stroboscopic real time technique are considered. Formulae for distributions of the light intensity in ultrasonic field holograms are derived for the techniques considered.

1. Introduction

Theoretical considerations of the light intensity distribution in optical holographic interferometry used for ultrasonic field examination have been performed by E. I. Kheifets [1] and P. Kwiek [2]. In this paper these results will be presented in a discussion of three methods which have been theoretically described and verified experimentally in our laboratory. These methods are:

1. pulse holographic interferometry (an ultrasonic wave as a phase object which moves with the high velocity, requires very short times of exposure for hologram recording) – usually the double-exposure technique is used.
2. holographic interferometry averaged in time.
3. stroboscopic interferometry in real time.

Let us consider the theoretical relations between the light intensity distribution and the ultrasonic wave intensity distribution for the three methods, successively.

2. Double-exposure pulse interferometry

In this case the main assumption is that the duration of the light pulse used for the exposure must be very much shorter than a period of the ultrasonic wave (for example 1/10 of a period).

* The work was carried under research programme MR.I.24.

Let us consider that such a pulse of light presents a parallel beam Σ (Fig. 1) for the construction of the ultrasonic wave hologram. The beam Σ is divided into two beams: Σ_p being the object beam and Σ_0 being the reference beam. The beam Σ_p is illuminating the ultrasonic wave which represents the object

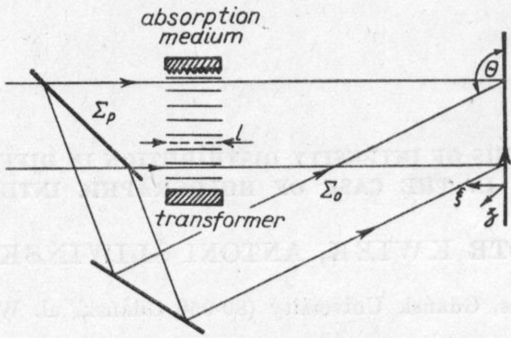

Fig. 1. Creation of an optical hologram of an ultrasonic wave

of observation. The object and the reference beams meet each other at a plane H where they interfere, creating the hologram. The beam Σ_0 makes an angle Θ with the plane H which corresponds to the plate on which the hologram can be recorded. The system of coordinates is chosen in such a way that η and ζ are coplanar to H.

Firstly, let us write the light intensity distribution for the case when the ultrasonic wave is absent from the path of the beam Σ_p, i.e. for the so called "empty" hologram. Denoting by $a_1(\eta, \zeta) = a_0 e^{-ik\zeta}$ the complex amplitude of the light wave of the beam Σ_p at the plane H, and by $F(\eta, \zeta) = F_0 e^{-ik\theta\eta}$ the amplitude of the reference beam Σ_0 one can write the expression for the illumination at the plane (η, ζ) as

$$E_1 = (a_1 + F)(a_1^* + F^*) = |a_1|^2 + |F|^2 + a_1^* F + a_1 F^*, \tag{1}$$

where * stands for the conjugate complex, a_0, F_0 are the real amplitudes of light for the object and reference beams, respectively, and k is the wave number of the light.

The hologram is recorded on the plate during an exposure time T. The energy recorded on the hologram is thus as follows:

$$W_1 = E_1 T = T|a_1|^2 + T|F|^2 + Ta_1^* F + Ta_1 F^*. \tag{2}$$

Now, let us introduce (switch on) the ultrasonic wave which changes the object beam. The thickness of the ultrasonic layer is L and the angular frequency Ω. The progressive ultrasonic wave changes the refractive index, of the medium Δn in proportion ot the acoustic pressure: $\Delta n \sim p$. In consequence the phase of the light passing through the ultrasonic wave will be changed by an amount $\Delta\psi$, and

$$\Delta\psi = \frac{2\pi L \Delta n}{\lambda} = \frac{2\pi L \delta n}{\lambda} \sin(K\eta - \Omega t) \cdot e^{-\alpha\eta}. \tag{3}$$

Thus the ultrasonic wave is treated as a phase diffraction grating (the Raman-Nath approximation). In this approximation the curvature of the light beam resulting form the gradient of the refractive index is neglected (taking this gradient into account one can consider the so-called amplitude diffraction grating as in the Lucas-Biquard theory).

In the presence of the ultrasonic wave the complex amplitude of the light beam Σ_p can be written as

$$a_2 = a_0 e^{-i[k\xi + \Delta\psi]} = a_0 \exp\left\{-i\left[k\xi + \frac{2\pi\delta n L}{\lambda}\sin(K\eta - \Omega t)e^{-\alpha\eta}\right]\right\}, \tag{4}$$

where K is the wave number of the ultrasonic wave,

α is the absorption coefficient of ultrasound in the medium,

$\delta\eta$ is the amplitude of changes in refractive index,

λ is the wavelength of the light.

When $T \ll 2\pi/\Omega$ (single pulse or stroboscopic exposure), expression (4) may be rewritten as

$$a_2 = a_0 \exp\left\{-i\left[k\xi + \frac{2\pi\delta\eta L}{\lambda}\sin K\eta e^{-\alpha\eta}\right]\right\}. \tag{5}$$

In this case the energy recorded at the plane H during the exposure time T will be equal to

$$W_2 = E_2 T = T|a_2|^2 + T|F|^2 + T a_2^* F + T a_2 F^*. \tag{6}$$

The total energy after two exposures ("empty" and "filled in") is

$$W = W_1 + W_2. \tag{7}$$

Let us assume that the recording process on the holographic plate uses the linear part of the amplitude transmission characteristic t_W of the plate (Fig. 2), i.e.,

$$t_W = t_0 - \beta(W - W_0), \tag{8}$$

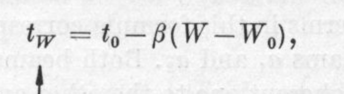

Fig. 2. Amplitude transmission characteristic of a holographic plate

where W_0 is the background energy corresponding to the amplitude transmission t_0. We can put

$$W_0 = T(|a_1|^2+|a_2|^2), \tag{9}$$

where β is the slope of the linear part of the transmission characteristic.

In the double-exposure process the transmission function of the hologram is the following:

$$t_W = t_0 - \beta T[2|F|^2 + F^*(a_1+a_2) + F(a_1^*+a_2^*)]. \tag{10}$$

a_2 in this expression contains all the information about the changes in refractive index of the medium which are caused by the acoustic pressure. Thus it can be seen that this information is involved in the transmission function t_W and we can speak about a hologram of the ultrasonic wave.

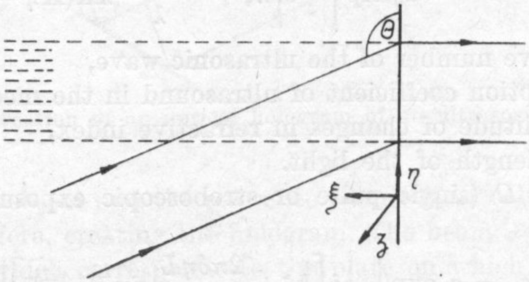

Fig. 3. Reconstruction of an optical hologram of the ultrasonic wave

Let us consider the reconstruction of this hologram, from which we shall obtain a visualization of the ultrasonic wave. The ultrasonic wave hologram of the transmission function t_W is placed in the parallel reconstruction light beam (Fig. 3) which is identical, to within to a constant factor, with the reference beam in the construction process. The output beam from the hologram can be written as:

$$bt_W = t_0 b - 2\beta' b|F|^2 + \beta' bF^*(a_1+a_2) + \beta' bF(a_1^*+a_2^*), \tag{11}$$

where $\beta' = \beta T$ and $b = b_0 e^{-ik\theta\eta}$ is the complex amplitude of the light of the reconstruction beam at the hologram plane, and θ is the angle of incidence.

The third and fourth terms in this formula correspond to the reconstruction of images obtained with beams a_1 and a_2. Both beams are reconstructed simultaneously. The beams are coherent one to the other and produce an interference image. The distribution of light intensity of that image, to within a constant factor, can be written as

$$I \sim (a_1+a_2)(a_1^*+a_2^*). \tag{12}$$

After substitution of a_1 and a_2 from formulae (1) and (5) one obtains

$$I = \left(a_0 e^{-ik\xi} + a_0 \exp\left\{-i\left[k\xi + \frac{2\pi\delta n L}{\lambda}\sin(K\eta)e^{-\alpha\eta}\right]\right\}\right) \times \\ \times \left\{a_0 e^{ik\xi} + a_0 \exp\left\{+i\left[k\xi + \frac{2\pi\delta n L}{\lambda}\sin(K\eta)e^{-\alpha\eta}\right]\right\}\right\}. \tag{13}$$

After some calculations and transformations using the known relations

$$\cos z = \frac{e^{iz}+e^{-iz}}{2} \quad \text{and} \quad 1+\cos\varphi = 2\cos^2\frac{\varphi}{2}$$

formula (13) appears in the form

$$I = 4a_0^2\cos^2\left[\frac{\pi\delta n L}{\lambda}\sin(K\eta)e^{-\alpha\eta}\right]. \tag{14}$$

This distribution of light intensity in the reconstructed image from the hologram corresponds to the ultrasonic field distribution. It can be used for determination of the phase and amplitude distributions, the velocity and the attenuation of the ultrasonic wave because all of these quantities are responsible for the light distribution obtained. The formula (14) gives the basis for experimental measurements of the ultrasonic field distribution with interferometric methods. [8].

3. Holographic interferometry averaged in time

Now, we shall consider the case when the exposure time $T \gg 2\pi/\Omega$ (the exposure time is much greater than a period of the ultrasonic wave). Due to the long exposure time the light intensity distribution is averaged in time.

Assuming the same geometry as in the Fig. 1 we can write the expression for the (single-exposure) hologram as

$$E = (a+F)(a^*+F^*) = |a|^2+|F|^2+aF^*+a^*F, \tag{15}$$

where F is the complex amplitude of the reference beam, a is the complex amplitude of the object beam passing through the ultrasonic wave ($a = a_2$ from the previous case and is given by formula (4)).

The energy accumulated in the hologram during the exposure can be written as

$$W = \int_0^T E\,dt = \int_0^T |a|^2\,dt + \int_0^T |F|^2\,dt + \int_0^T aF^*\,dt + \int_0^T Fa^*\,dt. \tag{16}$$

If the recording process corresponds to the linear part of the hologram transmission characteristic (Fig. 2), then the images which arise during the reconstruction of the hologram would have light intensity distributions proportional to the individual terms of formula (16). Including the first two terms in the constant W_0 (describing the background) we shall obtain as a result of the reconstruction two images: the real one (determined by the third term) and the imaginary one (determined by the fourth term). We shall deal with the real image, only.

We can write (substituting for F and a) the expression for the light intensity distribution recorded in the hologram:

$$I \sim \int_0^T F^* a dt = \int_0^T F_0^{ik\theta\eta} a_0 \exp\left\{-i\left(k\xi + \frac{2\pi\delta n L}{\lambda}\sin(K\eta - \Omega t)e^{-a\eta}\right)\right\}dt$$

$$= F_0 a_0 e^{i(k\theta\eta - k\xi)} \int_0^T \exp\left\{-i\frac{2\pi\delta n L}{\lambda}\sin(K\eta - \Omega t)e^{-a\eta}\right\}dt. \quad (17)$$

Developing the expression under the integral (of type $e^{iA}\sin\varphi$) into series of Bessel functions [5], one obtains

$$\int_0^T F^* a dt = F_0 a_0 e^{i(k\theta\eta - k\xi)}\left\{\int_0^T J_0\left(\frac{2\pi\delta n L}{\lambda}e^{-a\eta}\right)dt + \right.$$

$$+\int_0^T 2\sum_{n=1}^{\infty}\left[J_{2n}\left(\frac{2\pi\delta n L}{\lambda}e^{-a\eta}\right)\cos 2n(K\eta - \Omega t) + \right.$$

$$\left.\left. + iJ_{2n+1}\left(\frac{2\pi\delta n L}{\lambda}e^{-a\eta}\right)\sin(2n-1)(K\eta - \Omega t)\right]\right\}. \quad (18)$$

For $T \gg 2\pi/\Omega$ the second term in { } is equal to 0.

Illuminating the hologram during the reconstruction with the reference beam we obtain the amplitude distribution of the object beam as follows

$$A \sim F\int_0^T F^* a dt = F_0^2 a_0 e^{ik\theta\eta} T J_0\left(\frac{2\pi\delta n L}{\lambda}e^{-a\eta}\right)$$

$$= BJ_0\left(\frac{2\pi\,\delta n\,L}{\lambda}e^{-a\eta}\right)e^{ik\theta\eta} \quad (19)$$

and the intensity distribution

$$I \sim J_0^2\left(\frac{2\pi\,\delta n\,L}{\lambda}e^{-a\eta}\right). \quad (20)$$

Formula (20) does not give the possibility of phase and velocity determination, however, it is possible to determine attenuation of the ultrasonic wave a measuring the change in light intensity along the direction of propagation η.

4. Stroboscopic holographic interferometry in real time

The double-exposure pulse holography described in the first section allows comparison of only two states of the moving object being examined. However, it is possible to observe continuous changes of movement using only one hologram in the procesure called real time interferometry. The procedure depends on the interference of the object beam approaching the hologram with the beam which is previously recorded on the hologram, and which is thus simultaneously reconstructed. In our case the beam of the image of the "empty" medium (without the ultrasonic wave) recorded on the hologram interferes with the beam actually passing through the medium with the ultrasonic beam. On the base of the continuous superposition of the single hologram with the momentary image of the ultrasonic wave one can examine dynamical states of the ultrasonic field.

If in that case of real time interferometry a stroboscopic illumination is used with durations pulse much shorter than the period of the ultrasonic wave, then the light intensity distribution which arises in the interference image is described by the same formulae as for the case of double-exposure holographic interferometry.

Some experimental illustrations of the realization of the three methods described above has been presented in the paper of P. Kwiek et al. [8].

5. Conclusions

Using holographic interferometry one can visualise the space distribution of the ultrasonic wave as the distribution of the intensity of light diffracted on the ultrasonic hologram.

The theory presented here allows a quantitative comparison between measured and calculated distributions.

In the double-exposure and real time stroboscopic methods, examination of the wave fronts and phase relations is possible. In the time averaged method one can only determine the distribution of the acoustic energy density and the coefficient of attenuation of the ultrasonic wave.

References

[1] E. I. Kheifets, *Kolitchestviennyi analiz zvukovykh poley pri visualisatsii ikh holografitcheskim metodom*, Akust. Zhurn. **19**, 1973, 434.

[2] P. Kwiek, *Theoretical background of the time averaged hologram interferometry applied to ultrasonic field observations* [in Polish], Proc. of the 24-th Open Seminar on Acoustics, Gdańsk-Władysławowo, 1977, 365-367; and *Application of double exposure hologram interferometry to the investigation of the ultrasonic field distribution in liquids* [in Polish], Proc. of the 24-th Open Seminar on Acoustics, Gdańsk-Władysławowo, 1977, 368-371.

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

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

[5] V. I. Smirnov, *Kurs vysšey matematiki*, 5$^{\text{th}}$ ed., t. III, ch. II, M-L., 1952, 519-534.

[6] R. J. Collier, *Optical holography*, Ac. Press, N. York – London, 1972.

[7] A. E. Ennos, *Hologram interferometry*, Adv. in Quantum Electr., N. York – London, 1970.

[8] P. Kwiek, A. Śliwiński, I. Wojciechowska, *An arrangement for ultrasound field measurements using optical holography*. Ultrasonics, **17** (1979), No 3.

Received on 18th February 1978

ARCHIVES OF ACOUSTICS
4, 1, 73-86 (1979)

COMPLEXITY INVESTIGATION IN SOLUTIONS OF $ZnCl_2$ AND LiCl IN WATER USING AN ACOUSTIC METHOD*

WŁODZIMIERZ BOCH, JACEK GOC

Chair of Acoustics, A. Mickiewicz University (Poznań)

The formation of complex ions brings about in water solutions a destruction of the water structure, thus leading to the formation of a new structure in the solution which is considerably less cohesive than the original water structure.

The investigations performed with zinc chloride and lithium chloride solutions in water indicate that those solutions in which the ratio of zinc atoms to chlorine atoms is near to 1 : 4, i.e. near to the ratio for the formation of $ZnCl_4^{-2}$, have maximum adiabatic compressibility, minimum hydration number and minimum viscosity. Measurements of the absorption coefficient of ultrasonic waves have shown that the process of relaxation of the formation and disintegration of complex ions with a relaxation time of 10^{-8} s exerts a predominating influence on the sound absorption.

1. Introduction

A complicated interaction between ions and molecules in the solvent leads to the formation, in electrolyte solutions, of strictly defined strucural forms. In some cases these forms are precisely defined but sometimes their definition is obtained only by investigating the solution with many methods. We have to deal with such a situation in the case of water solutions of zinc chloride.

The aim of this paper is to present an acoustic method for investigating the complexity of water solutions of zinc chloride.

Zinc ions Zn^{2+} are comparatively small and according to the data of PAULING [1] their crystallographic radius is equal to $0.74 \cdot 10^{-8}$ m and the charge surface density is 3.67. Chlorine ions Cl^- are considerably larger and have a crystallographic radius of $1.81 \cdot 10^{-8}$ m and a charge denisty equal to 0.31.

In order to obtain a surplus of ions in zinc chloride solution lithium chloride is (preferably) introduced into the solution. Lithium ions Li^+ have a crystallographic radius equal to $0.60 \cdot 10^{-8}$ m and a charge surface density of 2.78.

*The paper was written under problem MR.I.24.

The first proposal that zinc may have a co-ordination figure of 4 was made by OSWALD and JAGGI [2]. On the basis of x-ray analysis they found that in anhydrous crystals of $ZnCl_2$, zinc co-ordinates four chloride ions in tetrahedral structures.

Supported by the statements of HARRIS and PARTON [4] that in concentrated water solutions of $ZnCl_2$, the transfer number of zinc ions Zn^{2+} is negative, ROBINSON and STOKES [3] have come to the conclusion that zinc occurs in solution mainly in the form of the $ZnCl_4^{2-}$ complex which, being more weakly hydrated, can have a higher mobility than its anti-ion Zn^{2+}.

In investigating the Raman spectrum DELWAULLE et al. [6] have observed in a water solution of $ZnCl_2$ the occurrence of two weak bands at 275 cm^{-1} which they, after HIBBEN [5], ascribe to the molecule $ZnCl_2$, and 307 cm^{-1}. The addition of a surplus of Cl^- ions causes the transformation of this band into a narrower line with a maximum at 282 cm^{-1} which the authors ascribe to the $ZnCl_4^{2-}$complex. Delwaulle concludes that such complexes also exist in solutions which do not have a surplus of chlorine ions.

KĘCKI [7-10] has investigated the Raman spectrum in methanol solutions of $ZnCl_2$, in which a surplus of chlorine ions was obtained by adding lithium chloride.

In a methanol solution of $ZnCl_2$ the asymmetry of the contour indicates, in addition to $ZnCl_2$, the presence of the $ZnCl_4^{2-}$ complexes. With the molecular ratio of $ZnCl_2$ to LiCl being equal to 1 : 2 and 1 : 4 these contours become symmetrical at a maximum of 282 cm^{-1} which corresponds to the $ZnCl_4^{2-}$ complex.

The occurrence of $ZnCl_4^{2-}$ complexes in water solutions should be also apparent in the behaviour of characteristics as adiabatic compressibility, hydration number, viscosity etc. Such data can be obtained by the use of an acoustic method applicable for the investigation of the solutions under consideration.

2. Method of investigation

The zinc chloride and lithium chloride used for the preparation of solutions were (partly for analysis) dried at a temperature of 120°C for 10 hours. Doubly distilled water was used as a solvent.

The solutions were prepared in such a way that in 6 moles of water 0,6 moles of electrolyte were dissolved, thus the ratio of salt water to water moles was constant and amounted to 1 : 10. Both zinc chloride and lithium chloride were included in 0.6 mole of salt, their relative amounts changing in such a maner that

$$n_1 + n_2 = 0.6, \tag{1}$$

where n_1 is the number of moles of LiCl, and n_2 is the number of moles of $ZnCl_2$. The mutual concentration k expressed as the percentage of LiCl in the total amount of electrolyte moles (0.6) can be written in the form

$$k = \frac{n_1}{n_1 + n_2} 100\%. \tag{2}$$

In the following text the mutual concentration k will define the percentage of LiCl in the total amount of electrolyte moles.

Seventeen solutions with mutual concentrations: 10.0; 20.0; 30.0; 40.0; 50.0; 60.0; 62.5; 65.0; 66.6; 67.5; 70.0; 72.5; 75.0; 80.0; 90.0; 100.0 % LiCl were prepared for the investigation. For better understanding it can be said that at a temperature of 25° C a solution with a mutual concentration $k = 100$ % LiCl has a concentration of 4.97 mole/l and this amounts to a concentration by weight of 19 %, while a solution with $k = 0$ % LiCl has a concentration of 4.58 mole/l (this would be a solution of zinc chloride alone) and this amounts to a concentration by weight of 43 %.

Ultrasonic waves propagate in a liquid medium at a phase velocity defined by the expression

$$c = \sqrt{\frac{1}{\varrho \beta_s}}, \tag{3}$$

where c denotes the propagation velocity of the ultrasonic waves, ϱ is the density of medium, and $\beta_s = -1/v((\partial v/\partial p)_s)$ is adiabatic compressibility.

Knowledge of the propagation velocity of ultrasound and of the solution density permits evaluation of the coefficient of adiabatic compressibility.

$$\beta_s = \frac{1}{\varrho c^2}. \tag{4}$$

This is the only method for simply evaluating the quantity β_s.

Measurements of the propagation velocity of ultrasonic waves at a frequency of 1 MHz were made with the aid of a pulse-phase interferometer type U-13 which permitted measurement of the phase velocity with a relative error (precision) less than 10^{-3}.

The solution density was determined by the specific gravity bottle method with a relative error not exceeding $2 \cdot 10^{-5}$.

If the adiabatic compressibilities of the solution and the pure solvent are denoted by β_s and β_{s_0} respectively and the incompressible part of the solvent in the solution [11] by X, one can write

$$V\beta_s = n_1 V_0' \beta_{s_0} (1 - X), \tag{5}$$

where n_1 denotes the total number of moles of the solvent in the solution, V_0 is the mean molar volume of pure solvent, and V is the volume of solution. As a result it is possible to evaluate the hydration number from the expression

$$Z = \frac{n_1}{n_2}\left(1 - \frac{V\beta_s}{n_1 V_0' \beta_{s_0}}\right), \tag{6}$$

where n_2 is the number of moles of the dissolved electrolyte.

While propagating in a medium the ultrasonic wave loses some of its energy in dissipative processess. The displacement of a particle of the medium due to the propagating ultrasonic wave is described by the expression

$$u = Ae^{-\alpha x}\, e^{i(\omega t - kx)}, \tag{7}$$

where α — denotes the absorption coefficient of the ultrasonic wave.

Stokes and Kirchhoff, from consideration of the viscosity and thermal conductivity losses have derived a formula for the absorption coefficient of ultrasonic waves [12]

$$\alpha = \frac{2\pi^2 f^2}{\varrho c^3}\left(\frac{4}{3}\eta + \frac{C_p/C_v - 1}{C_p}K\right), \tag{8}$$

where η denotes the coefficient of dynamic viscosity, C_p is the specific heat at constant pressure, C_v is the specific heat at constant volume, ϱ is the density of medium, f is the frequency of the ultrasonic wave, c is the velocity of propagation of the ultrasonic wave, and K is the coefficient of thermal conductivity.

It follows from this formula that

$$\frac{\alpha}{f^2} = \text{const}. \tag{9}$$

If molecular processes induced by the ultrasonic wave occur in a medium, then the dependence of α/f^2 on the frequency takes another form, as becomes apparent in the graph of the dependence of $\alpha\lambda$ on the frequency.

A point of inflexion on the curve α/f^2 and a maximum on the curve $\alpha\lambda$ can be used to determine the frequency of a relaxation process.
Hence, it is easy to determine the relaxation time

$$\tau = \frac{1}{2\pi fr}. \tag{10}$$

The absorption coefficients of ultrasonic waves at frequencies of 9.8; 13.3; 16.6; 20.0; 30.0; 42.0; 50.0 MHz were measured by means of a high-frequency ultrasonic unit US 4/6 with a maximum relative error in the determination of the absorption coefficient of 9 % at 9.8 MHz and 1 % at 50 MHz.

The dynamic viscosity of the solutions examined was determined from measurements made using a Höppler viscosimeter type BH-2 with a relative error not exceeeding 0.15 %. The konwledge of the viscosity coefficient permitted comparison of the calculated classical coefficient of the absorption of ultrasonic waves with the values actually measured. In these investigations the measuring vessels were thermostaticaly controlled; the temperature was measured with the aid of a thermistor type OTP-11 with a nominal resistance of 100 Ω, operating in a Wheatsone bridge circuit. The temperature was measured to an accuracy of 0.1 deg. It is also worth noting that all the measurements were performed in perfectly sealed measuring vessels, thus preventing a change in the concentration of the examined solutions as a result of water evaporating.

3. Results of the investigations

The measurements of the velocity of propagation of ultrasonic waves in the seventeen solutions examined were performed at five temperatures: 25, 30, 35, 40 and 45° C. Fig. 1 shows the dependence of the sound velocity on the temperature.

Concurrently with the measurements of the sound velocity the density of the liquids was also measured. The measured values of the density are shown as a function of the temperature in Fig. 2.

Knowledge of the velocity of propagation of ultrasonic waves and of the density of the liquids permits calculation of the coefficient of adiabatic compressibility. The results of these calculations are presented in Fig. 3.

On the basis of the results obtained for the coefficients of adiabatic compressibility of the solutions, the hydration numbers were evaluated from formula (6). Fig. 4 shows the relationship between the hydration number and the mutual concentration of water solutions of $ZnCl_2$ and LiCl.

The dependence of the measured values of the coefficient of dynamic viscosity of the liquids on the mutual concentration is shown in Fig. 5.

The classical theory of Stokes and Kirchhoff on the loss of energy of an ultrasonic wave as a result of viscosity and thermal conductivity in a medium makes it possible to calculate the absorption coefficient of ultrasonic waves according to the equation (8). The results of these calculations in the form of the dependence of α/f^2 on the frequency for various mutual concentrations of solutions are shown in Fig. 6.

The measured values of the absorption coefficient of ultrasonic waves in the solutions examined differ from the values calculated according to classical theory. Figs. 7 and 8 present the results of measurements of the absorption coefficient of sound in the form of the dependencies of α/f^2 and $\alpha\lambda$ on the frequency of ultrasonic waves.

4. Analysis and discussion of the results of the investigations

From previous investigations of solutions of zinc chloride it followed that zinc chloride exists in water and methanol solutions mainly in the form of $ZnCl_4^{2-}$. If one introduces a surplus of chloride ions into the water solution of zinc chloride by dissolving lithium chloride, then by changing the amount of lithium chloride it is possible to change the ratio of zinc atoms to chlorine atoms in the solution. The method of preparing the solution selected for these

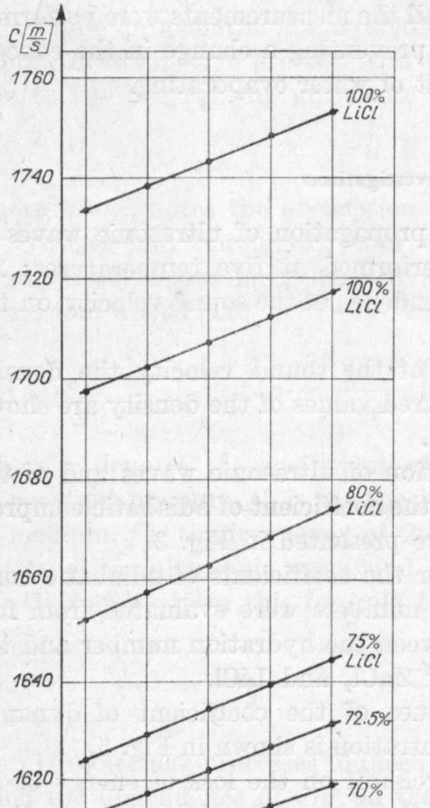

Fig. 1. The dependence of the velocity of propagation of ultrasonic waves on the temperature in water solutions of $ZnCl_2$ and LiCl, for various concentrations

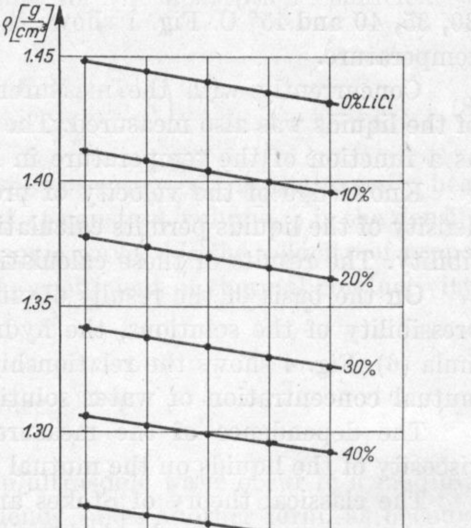

Fig. 2. The dependence of the density of water solutions of $ZnCl_2$ and LiCl on the temperature, for various mutual concentrations

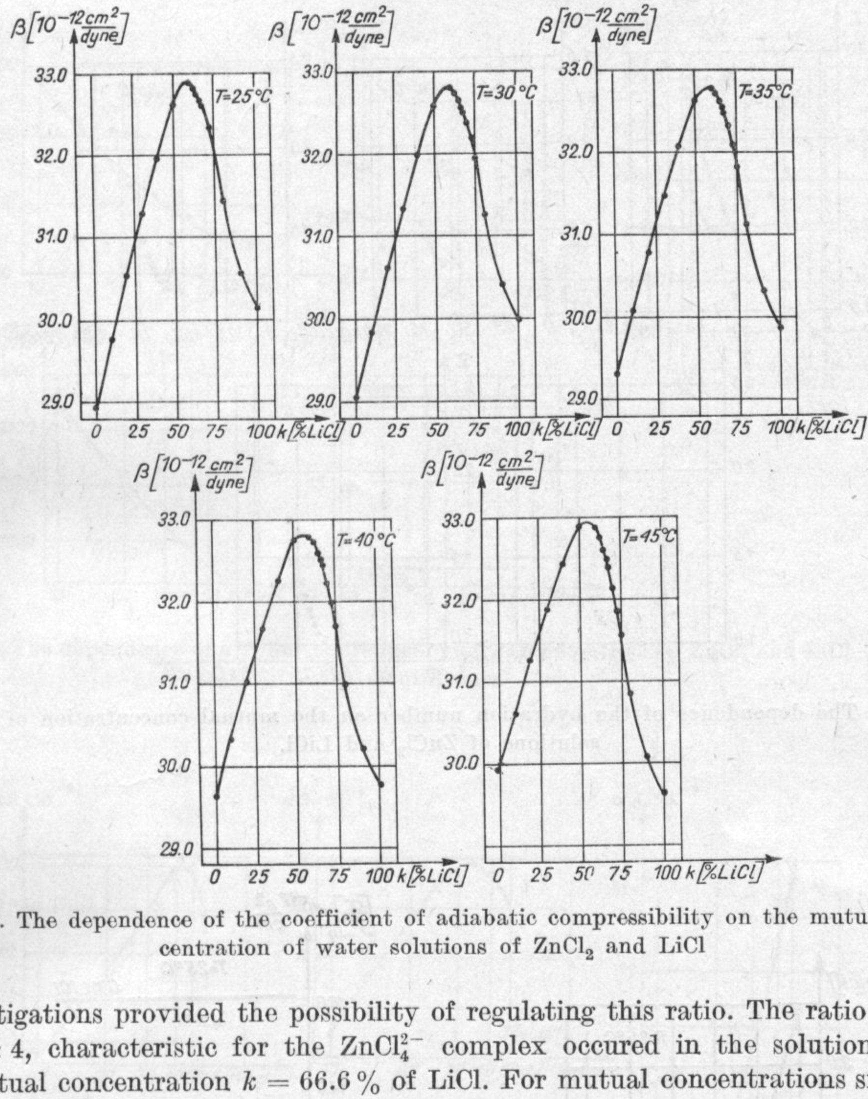

Fig. 3. The dependence of the coefficient of adiabatic compressibility on the mutual concentration of water solutions of $ZnCl_2$ and LiCl

investigations provided the possibility of regulating this ratio. The ratio equal to 1 : 4, characteristic for the $ZnCl_4^{2-}$ complex occured in the solution with a mutual concentration $k = 66.6\,\%$ of LiCl. For mutual concentrations smaller than this value, the ratio of zinc atoms to chlorine atoms was smaller than 1 : 4; for higher concentrations the ratio was greater.

The dependence of the velocity of propagation of ultrasonic waves in water on the temperature is parabolic, being described by the expression given by Willard [14],

$$c = 1557 - 0.0245(74 - t)^2 \text{ m/s}, \tag{15}$$

with a maximum at a temperature of 74°C. Dissolving electrolyte in the water causes a shift in the position of the maximum of the sound velocity towards smaller temperatures.

It follows from Fig. 1 that a solution of pure zinc chloride with a mutual

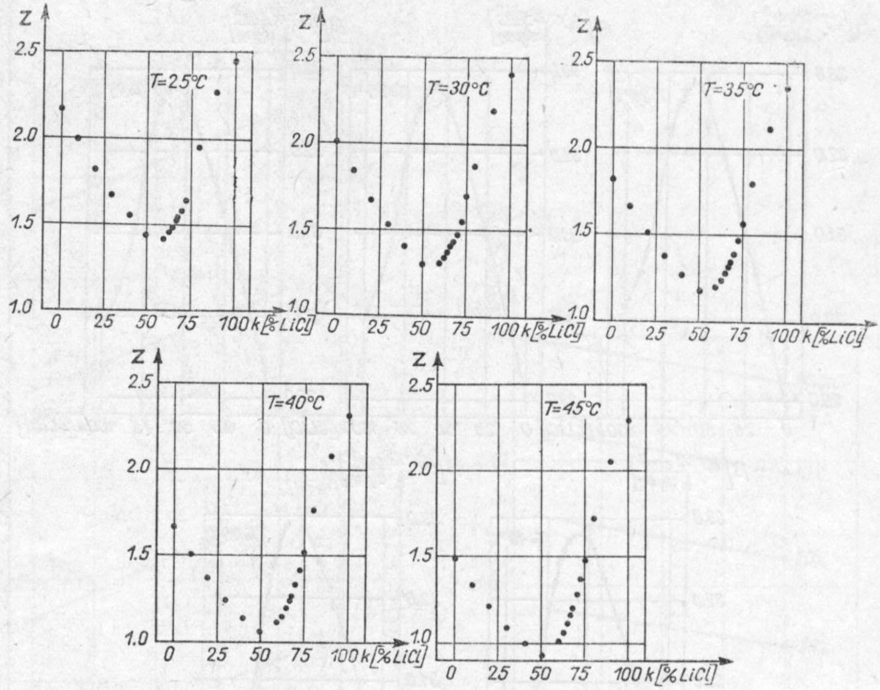

Fig. 4. The dependence of the hydration number on the mutual concentration of water solutions of $ZnCl_2$ and LiCl.

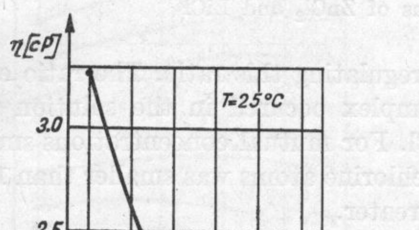

Fig. 5. The dependence of the coefficient of dynamic viscosity on the mutual concentration of water solutions of $ZnCl_2$ and LiCl

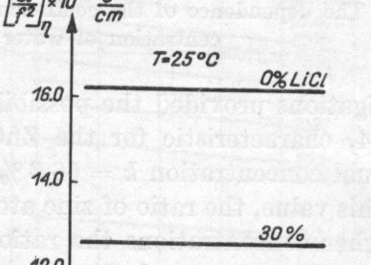

Fig. 6. The dependence of α/f^2 on the frequency in water solutions of $ZnCl_2$ and LiCl. Calculated according to the Stokes-Kirchhoff theory

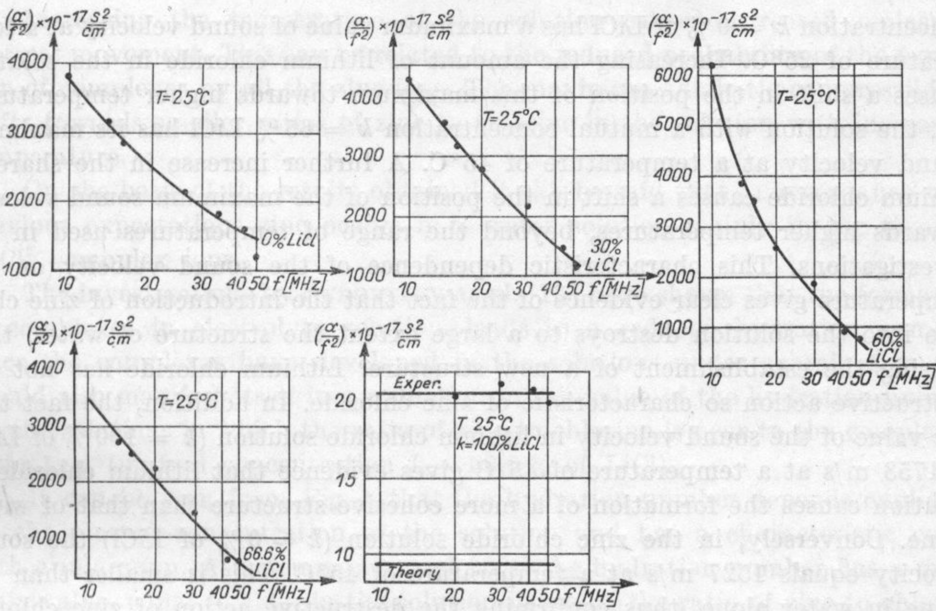

Fig. 7. The dependence of α/f^2 on the frequency in water solutions of $ZnCl_2$ and LiCl. Experimental data

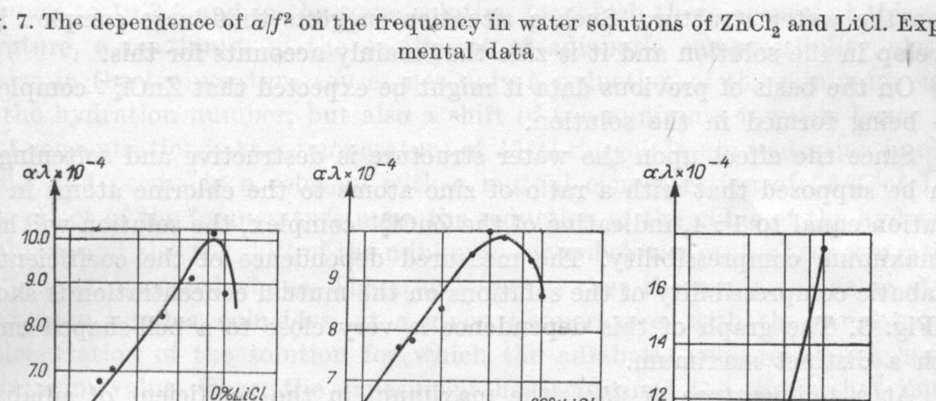

Fig. 8. The dependence of $\alpha\lambda$ on the frequency in water solutions of $ZnCl_2$ and LiCl. Experimental data

concentration $k = 0\,\%$ of LiCl has a maximum value of sound velocity at a temperature of 25°C. Increasing the amount of lithium chloride in the solution causes a shift in the position of this maximum towards higher temperatures, e.g. the solution with a mutual concentration $k = 65\,\%$ LiCl has its maximum sound velocity at a temperature of 45°C. A further increase in the share of lithium chloride causes a shift in the position of the maximum sound velocity towards higher temperatures, beyond the range of temperatures used in the investigations. This characteristic dependence of the sound velocity on the temperature gives clear evidence of the fact that the introduction of zinc chloride into the solution destroys to a large extent the structure of water, thus causing the establishment of a new structure. Lithium chloride has not the destructive action so characteristic of zinc chloride. In addition, the fact that the value of the sound velocity in lithium chloride solution ($k = 100\,\%$ of LiCl) is 1753 m/s at a temperature of 45°C gives evidence that lithium chloride in solution causes the formation of a more cohesive structure than that of water alone. Conversely, in the zinc chloride solution ($k = 0\,\%$ of LiCl) the sound velocity equals 1527 m/s at a temperature of 45°C. This is smaller than the value in water alone, thus confirming the destructive action of zinc chloride upon the water structure. Such a situation suggests that large complex ions develop in the solution and it is zinc that mainly accounts for this.

On the basis of previous data it might be expected that $ZnCl_4^{2-}$ complexes are being formed in the solution.

Since the effect upon the water structure is destructive and softening, it can be supposed that with a ratio of zinc atoms to the chlorine atoms in the solution equal to 1 : 4, indicative of the $ZnCl_4^{2-}$ complex, the solution will have a maximum compressibility. The measured dependence of the coefficient of adiabatic compressibility of the solutions on the mutual concentration is shown in Fig. 3. The graph of this dependence is very close to a bell-shaped curve with a distinct maximum.

At a temperature of 25°C the maximum in the coefficient of adiabatic compressibility occurs for a mutual concentration of the solution at which there are 3.5 chlorine atoms per zinc atom. With incresaing temperature, the maximum shifts towards lower mutual concentrations, so that at a temperature of 45°C it occurs at a mutual concentration corresponding to a zinc-chloride ratio equal to 1 : 3.25. On this basis it can be concluded that substantial quantities of $ZnCl_4^{2-}$ complexes develop in the solutions under consideration, and that these complexes exert a strong influence upon the adiabatic compressibility of the solution. If one takes into consideration the fact that the probability of all the zinc ions in the solution forming complex ions is very low, it is not surprising that experimentally we find a maximum in the adiabatic compressibility in solutions in which the ratio of the number of zinc atoms to the number of chlorine atoms is some-what smaller than the complexity ratio 1 : 4. Some of the ions will be bonded in ion pairs of the form $Zn^{2+}\cdot ZnCl_4^{2-}$.

Increasing the temperature of the solution causes increased molecular thermal movement. This can be related to the reduced probability of the formation of complexes by all the zinc ions. The maximum adiabatic compressibility shifts towards smaller ratios of zinc to chlorine in the solution with increasing temperature.

On the basis of the results obtained it can be said that in accordance with previous expectations zinc occurs in a water solution mainly in the form of $ZnCl_4^{2-}$ complex ions.

The investigations by SUROVTZEV at al. [15] have shown that the formation of complexes in electrolyte solutions leads to a reduced hydration number. Once the complexes have developed in the solutions under examination this should also manifest itself in a reduction in the value of the hydration number for the solutions in which the ratio of zinc to chlorine is near to the complesity ratio 1 : 4 (mutual concentration $k = 66.6\,\%$ of LiCl).

It can be seen from Fig. 4 that the hydration number depends explicitly on the mutual concentration of the solution and has a characteristic curve with a minimum. At a temperature of 25°C the hydration number has a minimum value, equal to 1.42, in the solution in which the ratio of zinc to chlorine is equal to 1 : 3.5 and in the same solution for which there occurs, at this temperature, a maximum in the coefficient of adiabatic compressibility. An increase in the temperature causes not only a reduction of the minimum value of the hydration number, but also a shift of the minimum towards lower mutual concentrations. At a temperature of 45°C the minimum hydration number is 0,9 and occurs in a solution with a mutual concentration of 52 % of LiCl. The effect of the temperature upon the reduction of the value of the hydration number, and also the shift of the minimum towards lower mutual concentrations have been discussed above. The crucial fact is that the minimum value of the hydration number coincides, at a given temperature, with the same mutual concentration of the solution for which the adiabatic compressibility has its maximum value. Thus, the experiment has confirmed the thesis that $ZnCl_4^{-2}$ complexes are formed in the solution.

The formation of these complexes, i.e. of large forms which react electrically and mechanically, in the solution should also have its effect upon the mechanical properties of the solutions, including the dynamic viscosity of the solutions. In a solution in which there exists a maximum number of complex ions the dynamic viscosity should possess a minimum value because of the reduction of liquid cohesion by the complexes. The dependence of the dynamic viscosity of the solutions on the mutual concentration is shown in Fig. 8. The minimum dynamic viscosity (1.887 cP) is exhibited by the solution with a mutual concentration of $k = 66.6\,\%$ of LiCl, and this corresponds to the complexity ratio of zinc to chlorine. This is another experimental fact which supports the conception that zinc in water solutions occurs chiefly in the form of $ZnCl_4^{2-}$ complexes.

It seems to be very interesting that the viscosity coefficient of the zinc chloride solution ($k = 0\%$ of LiCl) is 3.280 cP while the viscosity coefficient of the lithium chloride solution ($k = 100\%$ of LiCl) has a considerably smaller value equal to 2.045 cP. In addition, the increase in the number of chlorine ions in the solution causes a reduction of the viscosity of the solution, as a result of the formation of an ever greater number of $ZnCl_4^{2-}$ complexes. It seems to support the hypothesis stated by Kęcki [10] that zinc chloride also exists in the solution in the form of $ZnCl_2$"polymers", the occurrence of which involves a high viscosity for the solution.

The measurements made of the absorption coefficient of ultrasonic waves in the solutions examined also provide much valuable information. The quantity α/f^2 has a value independent of the frequency of the ultrasonic waves only for the lithium chloride solution, $k = 100\%$ of LiCl, although the experimentally determined value is more than twice as high as the value calculated from the Stokes-Kirchhoff formula. For all the solutions containing zinc chloride the dependence of α/f^2 on the frequency has a character which gives evidence of the effect of the molecular relaxation process on she sound absorption in the liquids examined. This effect can be seen distinctly when one considers the dependence of the sound absorption times the wavelength $\alpha\lambda$ and its relationship to the frequency of the ultrasonic waves. The graphs of this dependence represent relaxation sound absorption in the solutions examined. On the basis of both dependencies it is possible to estimate the time of the relaxation process: which attains its maximum value of $1 \cdot 10^{-8}$ s for the solution with a mutual concentration of $k = 60\%$ od LiCl (Zn : Cl = 1 : 3.5). It seems most warranted to associate the relaxation process discovered with the formation and disintegration of complex $ZnCl_4^{2-}$ ions. This suggestion seems to be confirmed by the fact that the longest time for the relaxation process is found for solutions with a mutual concentration of $k = 60\%$ of LiCl in which optimal conditions are created for the formation of complex ions, and this in turn effects a high stability for the structure of the solution.

5. Summing up and conclusions

The acoustic method used in the investigations seems to be a good method for the verification of the previous hypotheses and suggestions as to the form in which zinc chloride occurs in water solutions.

The results obtained show clearly that zinc chloride occurs in water solutions chiefly in the form of complex $ZnCl_4^{2-}$ ions. The solutions in which the ratio of the number of zinc ions to the number of chlorine ions is some-what smaller than the ratio required for the formation of the complex $ZnCl_4^{2-}$, are characterized by a maximum adiabatic compressibility, a minimum hydration number, and a minimum dynamic viscosity; while the relaxation process of the formation and disintegration of the complexes has a decisive effect on the absorption of ultrasonic waves in the solutions. The time of this relaxation process is estimated on the basis of the analysis of the sound absorption, at 10^{-8} s.

The results of the investigations carried out do not indicate that $ZnCl_3^-$ complexes are formed in water solutions of zinc chloride. Should such complexes exist in large quantities, then this would have influenced the adiabatic compressibility of the solutions and the values of the hydration number in those solutions in which the ratio of the number of zinc atoms to the number of chlorine atoms would be equal to or somewhat smaller than the ratio 1 : 3

The proven fact of the diminishing dynamic viscosity of the solution with an increasing number of chlorine atoms in the solution gives evidence for the validity of the hypothesis stated by Kęcki that zinc chloride also occurs in the solution in the form of $(ZnCl_2)_n$"polymers". The increase in the amount of chlorine in the solution must cause an increase in the number of complex $ZnCl_4^{2-}$ ions, thus reducing the chain numbers of the "polymers" and thus causing the viscosity to decrease.

References

[1] E. Józefowicz, *Inorganic chemistry* [in Polish] PWN, Warszawa 1959, 137-153.
[2] H. R. Oswald, H. Jaggi, Helv. Chim. Acta **43**, 72 (1960).
[3] R. A. Robinson, R. H. Stokes, *Electrolyte solutions*, Butterworths Sciences Public, London 1959.
[4] A. C. Harris, H. N., Parton, Trans. Faraday Soc., **36**, 1139 (1940).
[5] J. H. Hibben, Proc. Nat. Acad. Sci. USA, **18**, 532 (1932). J. Chem. Phys. **5**, 710 (1937). Phys. Rev. **51**, 593 (1937); *The Raman effect and its chemical applications*, Reinhold Publ. Corp, New York 1939.
[6] M. L. Delvaulle, Bull. Soc. Chim. France, 1294 (1955).
[7] Z. Kęcki, *A dissertation qualifying for assistant-professorship*, Warsaw University 1961.
[8] Z. Kęcki, Spectrochim. Acta, **18**, 1165 (1962).
[9] Z. Kęcki, J. Mańkowski, Chemical Annales **36**, 345 (1962).
[10] *Spectral investigations of the structure of electrolyte solutions*, PWN, Collective work, Warszawa 1969, 210-239.
[11] A. Pasynski, Zhurn. Fiz. Chim. **20**, 981 (1946).
[12] I. G. Michajłow, W. A. Sołowiew, I. P. Syrnikow, *Osnowy moliekuliatnoj akustiki* Izd. Nauka, Moskwa 163-186.
[13] K. F. Herzfeld, I. A. Litovitz, *Absorption and dispersion of ultrasonic waves*, Academic Press New York, London 1959.
[14] G. W. Willard, JASA **19**, 235 (1947).
[15] W. I. Surowcew, W. M. Szewczenko, E. J. Gorienbiejn, W. S. Rajewkij, Zhurn. Fiz. Chim. **46**, 2530 (1972).

Received on 8th March 1978