ARCHIVES OF ACOUSTICS
2, 4, 235 (1977)

A SIMULATIVE MODEL OF THE VOCAL TRACT INCLUDING THE EFFECT OF NASALIZATION*)

JANUSZ KACPROWSKI

Department of Cybernetic Acoustics, Institute of Fundamental Technological Research, Polish Academy of Sciences (Warszawa)

The subject and the aim of this work is the development of an articulatory model of the speech organ, adapted to computer simulation, for investigation of the influence of nasalization due to cleft palate upon the formant structure of originally non-nasal speech sounds, e.g. oral vowels, under realistic operating conditions from the acoustical, physiological and clinical points of view. This was achieved by reproducing the anatomy of the vocal tract under different articulatory conditions including all the possible individual variants, which depend on personal voice characteristics in both normal physiological states and in various pathological conditions. An essential part of the model was the taking into account of the losses in the vocal tract, the radiation impedance of the mouth and/or nose orifice, and the introduction of an additional parameter, which expresses in a quantitative manner the degree of nasalization, depending on the extent of the cleft palate. The model is intended for clinical applications in computer-aided acoustic diagnostics in phoniatry.

1. Introduction

The subject of an earlier work of the present author and his colleagues [12] was a graphical analysis of the influence of cleft palate (med.: *palatoschisis molle*) upon the transmission characteristics of the vocal tract, which determine the spectral structure of the oral vowels. Experimental verification was performed on an analogue model simulating, in an extremely simplified manner, the geometrical configuration of the vocal tract during the articulation of the nautral vowel Н. The model had a scale of 5 : 1 dimensionally and, consequently, 1 : 5 in frequency. Good agreement was obtained between the analytical and experimental results. These preliminary investigations confirmed the possibility of objective evaluating the influence of cleft palate, which results in the forced nasalization of originally non-nasal speech sounds, upon the formant

*) This work was performed within the framework of the nodal problem 10.4.3.

structure of the oral vowels. This influence proved to be measurable and amenable to analytical treatment. It establishes many possibilities for the clinical application of acoustical methods in phoniatric diagnostics for the objective evaluation of pathological and postoperative states in cleft palate and for the objective control and documentation of the rehabilitation process.

The results of preliminary experiments, although encouraging from the cognitive point of view, proved to have limited clinical usefulness because of the simplifications introduced. The aim of the present work is to extend the previous results and to increase their accuracy by elaborating a universal articulatory model of the speech organ, which is adapted to computer simulation. This aim may be achieved by:

– reproducing the anatomy of the vocal tract in the articulation of different oral vowels (firstly for the Polish basic vowels [i, y, e, a, o, u]) under steady state conditions with all the possible individual variants, determined by the individual features of the patient's voice in both normal physiological states and pathological conditions;

– taking into account the losses existing in the vocal tract and the radiation impedance of the mouth and/or nose orifice;

– introducing an additional parameter, which expresses in a quantitative manner the degree of nasalization, which depends on the extent of cleft palate or on its actual advance.

2. An acoustical model of the vocal tract in the articulation of voiced sounds with glottal excitation

A simplified acoustical model of the supraglottal part of the speech organ is shown in Fig. 1a. The vocal pharynx-mouth tract ($G+U$) may be approximated as a tube of irregular cross-sectional area given by the function $A(x)$, where x is the distance from the glottis. Under the normal anatomical conditions of an adult male the length $l_k = l_g + l_u$ of the pharynx-mouth tract, ending at the point $x = l_k$ by the mouth orifice, is of the order 16-19 cm, and its cross-sectional area ranges from some fraction of a square centimetre to about 15 cm². The pharynx tract is excited at the point $x = 0$ by the larynx generator with an internal impedance Z_g, which produces a volume velocity U_g in the glottal orifice. At the point $x \approx 8$ cm the nasal tract N is attached to the pharynx-mouth tract. The nasal tract has the form of an irregular tube of the length $l_n \approx 12.5$ cm and total volume equal to about 50 cm³. The movable soft palate (med.: *epiglottis*) acts as a valve which controls the degree of acoustic coupling between the nasal and mouth tracts, i.e. the amount of nasalization of originally non-nasal speech sounds.

The acoustic energy of the speech wave is emitted through the radiation impedance Z_{pu} of the mouth and/or Z_{pn} of the nose. The corresponding volume velocities are U_u and U_n. These volume velocities produce, at a distance r

on the axis of symmetry of the radiating system, an acoustic pressure $p(r)$ in the speech wave, which is a linear superposition of two waves radiated through the mouth and nasal orifices[1]). It may be expressed as

$$p(r) = \frac{j\omega\varrho}{4\pi r}(U_u + U_n)e^{-j2\pi r/\lambda} \tag{1}$$

(for a spherical wave, $2\pi r \leqslant \lambda$), as is explained in [12].

The acoustic model of the articulatory effectors shown in Fig. 1a may be presented in the form of the equivalent electrical circuit shown in Fig. 1b, in which the continuous cross sectional area function $A(x)$ of the pharynx-mouth tract is approximated by a definite number q of uniform elementary segments, e.g. T-type four-poles, each of which represents a segment of the tract of length $l = l_k/q$ in the form of a cylindrical tube of cross-section A_i $(i = 1, 2, \ldots, q)$. The accuracy of the approximation increases as the length l

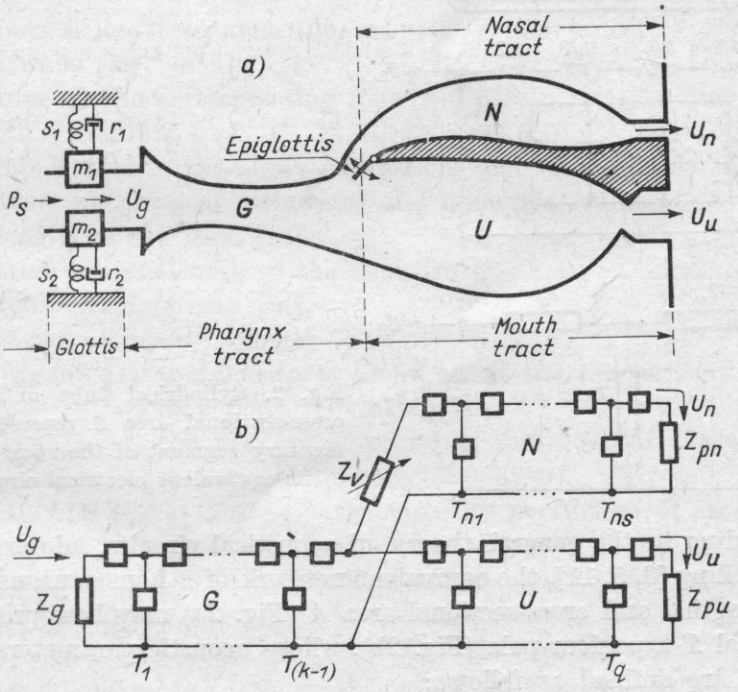

Fig. 1. Simplified acoustical model of the supraglottal part of the speech organ (a) and its electrical equivalent circuit (b)

P_s – subglottal pressure, (m_1, s_1, r_1) and (m_2, s_2, r_2) – vocal folds, U_g – volume velocity of the larynx source, Z_g – acoustic impedance of the glottis (larynx source), G – pharynx tract, U – mouth tract, N – nasal tract, U_u – volume velocity through the mouth orifice, U_n – volume velocity through the nostrils, Z_{pu} – radiation impedance of the mouth orifice, Z_{pn} – radiation impedance of the nose orifice, Z_v – acoustic impedance of the nose-to-mouth coupling

[1]) The component originating from the acoustic pressure wave, radiated by the external surface of the vibrating walls of the vocaltract, has been neglected.

of each cylindrical elementary segment decreases. The limiting condition is $l = \lambda_{min}/8$, where λ_{min} is the wave-length at the highest frequency in the considered range, i.e. $f_{max} \approx 5000$ Hz. It follows that the representation of the area function $A(x)$ of the pharynx-mouth tract in terms of $q \approx 20$ discrete elementary uniform segments of length $l = 1$ cm is quite satisfactory for modelling purposes, since the resulting «quantization noise» may be neglected. The electrical equivalent circuit in Fig. 1b constitutes the theoretical basis and starting point for the design and construction of many articulatory speech synthesizers, both static [2, 3, 17] and dynamic, i.e. controlled by electrical signals, initially analogue [6, 16] and subsequently digital [1].

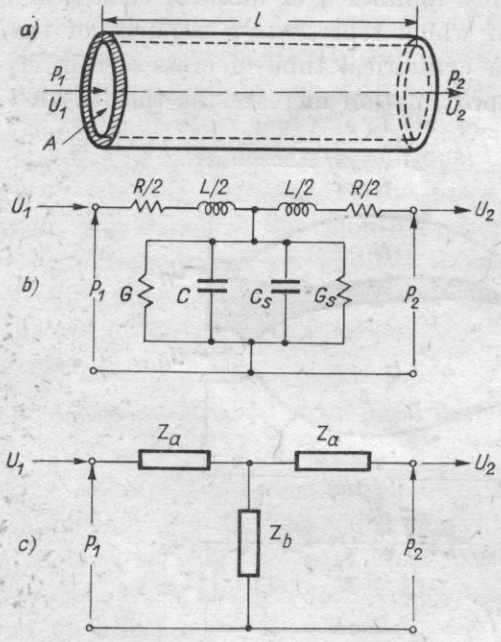

Fig. 2. Cylindrical tube of length l and cross-sectional area A representing an elementary segment of the vocal tract (a) and its equivalent electrical circuits (b, c)

According to the general theory of acoustical circuits, adopted for vocal tract modelling [3, 5, 11], the equivalent network of a homogeneous cylindrical tube of length l and cross-sectional area A (Fig. 2a) may be represented as a symmetrical T-type four-pole (Fig. 2b), whose acoustic parameters L, C, R, G, C_s and G_s are defined as follows:

$$L = L_j l = \varrho l/A$$

— the acoustic mass of the air in the tube [kg m^{-4}];

$$C = C_j l = Al/\varrho c^2$$

— the acoustic compliance of the air in the tube [kg^{-1} m^4 s^2];

$$R = R_j l = \frac{S}{A^2}\sqrt{\omega \varrho \mu/2}\, l$$

– the acoustic viscous loss resistance at the tube walls [kg m^{-4} s^{-1}];

$$G = G_j l = S \frac{\eta - 1}{\varrho c^2} \sqrt{\frac{\lambda \omega}{2 c_p \varrho}}\, l$$

– the acoustic loss conductance due to heat conduction at the tube walls [kg^{-1} m^4 s];

$$C_s = C_{sj} l = -\frac{m_s S}{r_s^2 + \omega^2 m_s^2}\, l$$

– the negative acoustic compliance equivalent to the reciprocal acoustic mass of the vibrating vocal tract walls [kg^{-1}m^4 s^2];

$$G_s = G_{sj} l = \frac{r_s S}{r_s^2 + \omega^2 m_s^2}\, l$$

– the reciprocal loss resistance (i.e. acoustic conductance) of the vibrating vocal tract walls [kg^{-1} m^4 s].

The value of the corresponding four-pole parameters per unit length are labelled with subscripts j (L_j, C_j, R_j, G_j, C_{sj} and G_{sj}).

The symbols used in the above expressions and subsequently in this paper have the following physical definitions and numerical values:

l – the length of the tube [m],

A – the cross-sectional area of the tube [m^2],

S – the tube circumference [m],

$\omega = 2\pi f$ – the angular frequency [s^{-1}],

$\varrho = 1.14$ [kg m^{-3}] – the density of moist air at the temperature of the human body ($t = 37°$C),

$c = 350$ [m s^{-1}] – the velocity of sound in moist air at the temperature of the human body ($t = 37°$C),

$\mu = 1.86 \cdot 10^{-5}$ [N m^{-2} s] – the viscosity coefficient of air ($t = 20$ °C, $P_0 = 0.76$ m Hg),

$\lambda = 55 \cdot 10^{-4}$ [cal m^{-1} s^{-1} °C^{-1}] – the thermal conductivity of air,

$c_p = 0.24 \cdot 10^3$ [calkg^{-1} °C^{-1}] – the specific heat of air at constant pressure $P_0 = 0.76$ m Hg,

$\eta = 1.4$ – the adiabatic constant of air,

$r_s \approx 16 \cdot 10^3$ [kg m^{-2} s^{-1}] – the loss resistance of the tissue of the vocal-tract walls per unit area[2]),

$m_s \approx 15$ [kg m^{-2}] – the mass of the tissue of the vocal-tract walls per unit area[2]).

---

[2]) The numerical values of r_s and m_s were determined by direct measurements of the surface impedance of the vocal-tract wall tissue [9], and then experimentally verified in model investigations [15].

The physical meaning of the parameters C_s and G_s calls for a brief explanation. The vocal-tract walls are not infinitely stiff, as was assumed in simplified considerations (see e.g. [12]), but have a finite acoustic impedance per unit area z_s. When exposed to local acoustic pressure in the vocal tract, they are forced to vibrate. The reactive component of the impedance z_s at frequencies above the resonance frequency of the walls, i.e. when $f > 100$ Hz, has an inertial character. The impedance z_s may thus be represented by a series connection of the loss resistance r_s and the mass m_s,

$$z_s = r_s + j\omega m_s, \tag{2}$$

where z_s, r_s and m_s are quantities per unit area. The impedance z_s, which corresponds to the energy dissipated by the wall vibrations, ought (on formal grounds) to be included in the transverse branch of the four-pole T in Fig. 2b. It is thus convenient to present z_s as the parallel connection of a real (resistive) and an imaginary (reactive) component in the form of an admittance per unit area $y_s = z_s^{-1}$:

$$y_s = \frac{r_s}{r_s^2 + \omega^2 m_s^2} - j\omega \frac{m_s}{r_s^2 + \omega^2 m_s^2}. \tag{3}$$

Multiplying y_s by the tube circumference S, we obtain the acoustic admittance per unit length of the walls,

$$Y_{sj} = \frac{r_s S}{r_s^2 + \omega^2 m_s^2} - j\omega \frac{m_s S}{r_s^2 + \omega^2 m_s^2} = G_{sj} + j\omega C_{sj}, \tag{4}$$

where G_{sj} is the acoustic conductance per unit length of the walls, and C_{sj} is the negative acoustic compliance, equivalent to the reciprocal mass, per unit length of the tube walls.

If the walls of the tube are infinitely hard, $m_s = \infty$ and $r_s = \infty$; thus $Y_{sj} = 0$ and the transverse admittance of the four-pole T in Fig. 2b is simply $Y = G + j\omega C$.

Using the general transmission line theory, we can transmute the equivalent circuit of a tube of length l, shown in Fig. 2b as a T-type four-pole, to the form presented in Fig. 2c. Expressing the acoustic impedance per unit length of the longitudinal branches of the four-pole as

$$Z_j = R_j + j\omega L_j = \frac{S}{A^2}\sqrt{\omega\varrho\mu/2} + j\omega\frac{\varrho}{A} \tag{5}$$

and the acoustic admittance per unit length of the transverse branch of the four-pole as

$$\begin{aligned} Y_j &= (G_j + G_{sj}) + j\omega(C_j + C_{sj}) \\ &= \left(S\frac{\eta - 1}{\varrho c^2}\sqrt{\frac{\lambda\omega}{2c_p\varrho}} + \frac{r_s S}{r_s^2 + \omega^2 m_s^2}\right) + j\omega\left(\frac{A}{\varrho c^2} - \frac{m_s S}{r_s^2 + \omega^2 m_s^2}\right), \end{aligned} \tag{6}$$

we obtain the following expressions for the characteristic impedance Z_0 and the propagation constant γ of the tube[3]):

$$Z_0 = \sqrt{Z_j/Y_j} = \sqrt{(R_j+j\omega L_j)/[(G_j+G_{sj})+j\omega(C_j+C_{sj})]}, \tag{7}$$

$$\gamma = \alpha+j\beta = \sqrt{Z_j Y_j} = \sqrt{(R_j+j\omega L_j)[(G_j+G_{sj})+j\omega(C_j+C_{sj})]}. \tag{8}$$

From expressions (7) and (8) for the impedance parameters z_a and z_b of the four-pole T in Fig. 1c we get

$$z_a = Z_0 \operatorname{tgh}(\gamma l/2) \approx Z_0 \gamma l/2 = Z_j \frac{l}{2} = (R_j+j\omega L_j)\frac{l}{2} = \frac{1}{2}(R+j\omega L) = \frac{1}{2}Z, \tag{9}$$

$$\begin{aligned} z_b &= \frac{Z_0}{\sinh(\gamma l)} \approx \frac{Z_0}{\gamma l} = \frac{1}{[(G_j+G_{sj})+j\omega(C_j+C_{sj})]l} \\ &= [(G+G_s)+j\omega(C+C_s)]^{-1} = Y^{-1}, \end{aligned} \tag{10}$$

where $Z = Z_j l$ and $Y = Y_j l$ (cf. (5) and (6)).

The approximations applied in expressions (9) and (10), which consist in replacing the hyperbolic functions $\operatorname{tgh} x$ and $\sinh x$ by the values of the argument x, that is by the first term of their series expansions, are valid for small values of the argument x. This condition is fulfilled in the case considered of a short section of tube, where $l \leqslant \frac{1}{8}\lambda_{\min}$.

3. The pharynx-mouth tract including the shunting effect of the nasal tract

According to the earlier assumptions, the pharynx-mouth tract can be approximated by a cascade connection (or chain) of q T-type four-poles (cf. Fig. 1b) each of which represents an elementary segment of the vocal tract in the form of a cylindrical tube of length $l = 1$ cm and cross-sectional area A_i $(i = 1, 2, \ldots, q)$. The quantities q and A_i are considered as variable parameters which describe the actual geometric configuration of the vocal tract depending on the personal physio-pathological features of the patient's speech organ and on the temporary articulation conditions expressed in terms of the vocal-tract cross-sectional area function $A(x)$.

Using the scarce data which may be found in the literature (see e.g. [3, 15, 17] and adapting it to the articulatory conditions of Polish vowels, we have shown in Table 1 the discrete values of the $A(x)$ area function of the basic vowels [i, y, e, a, o, u] for different values of x in $\Delta x = 1$ cm steps. The data should be treated as preliminary information which corresponds to the average normal physiological conditions of an adult male and which calls for individual verification. The eventual correction should be based on an experimental comparison of the pole frequencies of the vocal-tract transfer function $H(\omega)$

[3]) Depending on the relative values of the parameters $R_j, L_j, G_j, G_{sj}, C_j$ and C_{sj}, some of them may be neglected in numerical calculations. In the particular case of a lossless tube, the parameters R_j, G_j, G_{sj} and C_{sj} are equal to zero, whence

$$Z_0 = \sqrt{L_j/C_j} = \varrho c/A, \quad \gamma = j\omega\sqrt{L_j C_j} = j\omega/c = j\beta, \quad \alpha = 0.$$

Table 1. Discrete values of the cross-sectional area function $A(x)$ of the vocal tract in the articulation of Polish oral vowels [i, y, e, a, o, u] (approximate data).

Vowel		[i]	[y]	[e]	[a]	[o]	[u]
Distance from glottis x [cm]		Cross-sectional area of the vocal tract $A(x)$ [cm^2]					
Pharynx tract	1	3.0	3.0	2.3	2.0	2.5	2.6
	2	2.0	3.0	1.4	1.0	1.5	2.0
	3	10.0	12.0	6.5	2.3	6.0	8.0
	4	13.0	12.0	6.5	3.0	5.0	10.0
	5	13.0	11.0	8.0	4.3	3.2	5.3
	6	12.8	8.5	10.5	5.0	1.4	2.5
	7	12.8	7.5	10.5	1.7	0.7	3.5
Bifurcation	8	11.7	2.0	9.5	2.0	1.4	2.5
Mouth tract	9	7.3	1.0	7.3	3.0	2.3	2.5
	10	3.6	1.0	5.7	5.0	2.6	2.2
	11	2.2	2.0	4.5	6.5	4.0	2.2
	12	0.6	6.0	4.0	8.0	5.7	2.5
	13	0.6	8.0	3.3	9.0	8.0	4.0
	14	0.5	6.0	3.3	9.5	10.3	7.5
	15	0.6	3.0	3.3	9.0	10.3	13.0
	16	1.5	5.0	5.0	8.8	10.3	13.0
	17	4.0	8.0	7.0	8.0	8.0	4.0
	18	—	—	—	—	4.8	0.3
	19	—	—	—	—	—	0.7

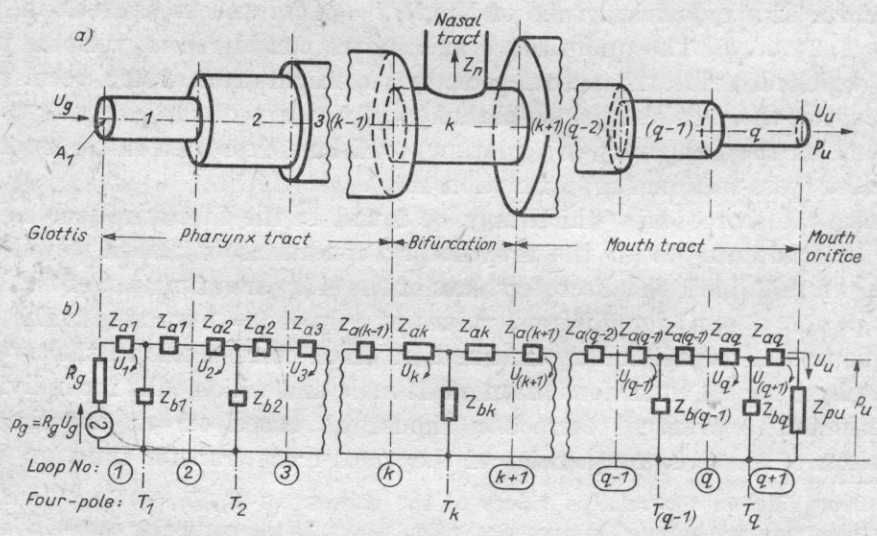

Fig. 3. Acoustical model of the vocal pharynx-mouth tract including the shunting effect of the nasal tract (a) and its equivalent electrical circuit (b)

with the formant frequencies of the different natural isolated vowels spoken by the subject under examination. The average values of the formant frequencies of Polish basic vowels measured by JASSEM may be found in [10].

An accustic model of the vocal pharynx-mouth tract (including the shunting effect of the nasal tract) and the corresponding electrical equivalent circuit are shown in Fig. 3. The k-th[4]) segment of the model and the corresponding equivalent T_k-section represent the bifurcation of the vocal tract (med.: *nasopharynx*), i.e. the space in which the input impedance Z_n of the nasal tract shunts the pharynx-mouth tract constituting the only path of energy flow in purely oral articulation. The section T_k differs with regard to its structure from other sections of the vocal tract and may be represented in the form shown in Fig. 4 which may be used to determine the longitudinal impendances z_{ak}

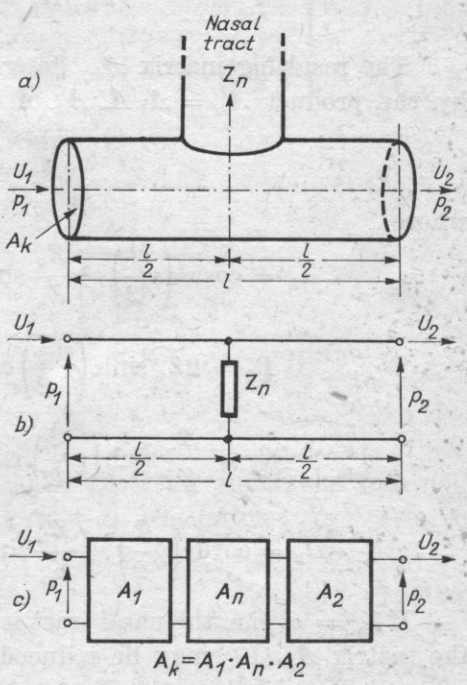

Fig. 4. Acoustical model of the vocal-tract's bifurcation (a) and its equivalent electrical circuits in the form of a transmission line shunted by the impedance Z_n (b) and in the form of a cascade connection of three four-poles described by their chain matrices $\boldsymbol{A}_1$, $\boldsymbol{A}_n$, $\boldsymbol{A}_2$ (c)

and the transverse impedance z_{bk} of the T_k four-pole, depending on the actual value of the shunting impedance Z_n. For this reason the k-th segment of length l and cross-sectional area A_k were replaced by two identical subsegments, each of length $l/2$, between which the shunting impedance Z_n, representing the input impedance of the nasal tract, is located (see Fig. 4b). In this way the equivalent electrical circuit of the bifurcation has been shown in the form of a cascade connection of three four-poles described by their chain matrices $\boldsymbol{A}$.

[4]) Provided that under normal physiological conditions the vocal tract bifurcates into the mouth and nasal tract at a distance $x \approx 8$ cm from the glottis, $k = 8$.

Two extreme four-poles represent the left and the right halves of the bifurcation segment, and their matrices $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$ are equal,

$$\boldsymbol{A}_1 = \boldsymbol{A}_2 = \begin{bmatrix} \cosh\left(\gamma \frac{l}{2}\right) & Z_0 \sinh\left(\gamma \frac{l}{2}\right) \\ \frac{\sinh\left(\gamma \frac{l}{2}\right)}{Z_0} & \cosh\left(\gamma \frac{l}{2}\right) \end{bmatrix}, \tag{11}$$

while the middle four-pole, which represents the shunting impedance Z_n, is described by a matrix $\boldsymbol{A}_n$ of the form

$$\boldsymbol{A}_n = \begin{bmatrix} 1 & 0 \\ 1/Z_n & 1 \end{bmatrix}. \tag{12}$$

The resulting matrix $\boldsymbol{A}_k$, describing the k-th segment, may be expressed by the product $\boldsymbol{A}_k = \boldsymbol{A}_1 \boldsymbol{A}_n \boldsymbol{A}_2$ of the three matrices as

$$\boldsymbol{A}_k = \begin{bmatrix} A_k & B_k \\ C_k & D_k \end{bmatrix}, \tag{13}$$

where

$$A_k = \cosh^2\left(\gamma \frac{l}{2}\right) + \frac{Z_0}{Z_n} \sinh\left(\gamma \frac{l}{2}\right) \cosh\left(\gamma \frac{l}{2}\right) + \sinh^2\left(\gamma \frac{l}{2}\right),$$

$$B_k = 2Z_0 \sinh\left(\gamma \frac{l}{2}\right) \cosh\left(\gamma \frac{l}{2}\right) + \frac{Z_0^2}{Z_n} \sinh^2\left(\gamma \frac{l}{2}\right),$$

$$C_k = \frac{2}{Z_0} \sinh\left(\gamma \frac{l}{2}\right) \cosh\left(\gamma \frac{l}{2}\right) + \frac{1}{Z_n} \cosh^2\left(\gamma \frac{l}{2}\right),$$

$$D_k = \cosh^2\left(\gamma \frac{l}{2}\right) + \frac{Z_0}{Z_n} \sinh\left(\gamma \frac{l}{2}\right) \cosh\left(\gamma \frac{l}{2}\right) + \sinh^2\left(\gamma \frac{l}{2}\right).$$

If $Z_n = \infty$ (i.e. the nasal tract is disconnected and there is no nasalization), the matrix $\boldsymbol{A}_k$ (13) may be reduced to the simple form

$$\boldsymbol{A}_k = \begin{bmatrix} \cosh(\gamma l) & Z_0 \sinh(\gamma l) \\ \frac{1}{Z_0} \sinh(\gamma l) & \cosh(\gamma l) \end{bmatrix}, \tag{14}$$

which describes the four-pole network of a homogeneous tube segment of length l. This may have been expected and confirms the theoretical assumptions and the correctness of mathematical manipulations.

Using the well-known formulae of matrix algebra applied in the theory of electrical four-poles (see e.g. [14]), we can determine the impedances z_{ak} of the longitudinal branches and the impedance z_{bk} of the transverse branch of the

four-pole T_k by formulae:

$$z_{ak} = \frac{A_k - 1}{C_k}, \tag{15a}$$

$$z_{bk} = \frac{1}{C_k}, \tag{15b}$$

where A_k and C_k are elements of the matrix $\boldsymbol{A}_k$ (formula (13)).

Using the previously justified approximations: $\sinh(\gamma l/2) \approx \gamma l/2$ and $\cosh(\gamma l/2) \approx 1$ which are valid provided $l \leqslant \lambda/8$, we can reduce the matrix $\boldsymbol{A}_k$ (formula (13)) to a simpler form,

$$\boldsymbol{A}_k = \begin{bmatrix} 1 + \left(\gamma \frac{l}{2}\right)^2 + \frac{Z_0}{Z_n}\left(\gamma \frac{l}{2}\right), & 2Z_0\left(\gamma \frac{l}{2}\right) + \frac{Z_0^2}{Z_n}\left(\gamma \frac{l}{2}\right)^2 \\ \frac{2}{Z_0}\left(\gamma \frac{l}{2}\right) + \frac{1}{Z_n}, & 1 + \left(\gamma \frac{l}{2}\right)^2 + \frac{Z_0}{Z_n}\left(\gamma \frac{l}{2}\right) \end{bmatrix}, \tag{16}$$

from which, according to formulae (15), we obtain

$$z_{ak} = Z_0\left(\gamma \frac{l}{2}\right) \frac{Z_0 + Z_n\left(\gamma \frac{l}{2}\right)}{Z_0 + 2Z_n\left(\gamma \frac{l}{2}\right)} \tag{17a}$$

and

$$z_{bk} = \frac{Z_0}{\gamma l} \frac{Z_n}{Z_n + Z_0/\gamma l}. \tag{17b}$$

As before, if $Z_n \infty$, the impedances z_{ak} and z_{bk} of the equivalent four-pole representing the bifurcation of the tract reduce to the forms

$$z_{ak} = Z_0 \frac{\gamma l}{2} \tag{18a}$$

and

$$z_{bk} = \frac{Z_0}{\gamma l} \tag{18b}$$

which express the longitudinal impedances z_a and the transverse impedance z_b of a symmetrical T-type four-pole replacing a homogeneous cylindrical tube of length l (cf. formulae (9) and (10)).

Without jeopardizing the accuracy of the numerical calculations, the k-th segment and the corresponding four-pole T_k, which represent the bifurcation of the vocal tract, may be considered as lossless circuits with $Z_0 = \varrho c/A_k$ and $\gamma = j\omega/c$. This implies further simplification of the structure of the four-pole

T_k, whose impedance elements z_{ak} (17a) and z_{bk} (17b) reduce to

$$z_{ak} = j\frac{\omega l\varrho}{2A_k}\frac{\varrho c/A_k + jZ_n\omega l/2c}{\varrho c/A_k + jZ_n\omega l/c} \tag{19a}$$

and

$$z_{bk} = -j\frac{\varrho c^2}{A_k\omega l}\frac{Z_n}{Z_n - j\varrho c^2/A_k\omega l}, \tag{19b}$$

where c is the velocity of sound in moist air at 37 °C (the temperature of the human body) and $\omega = 2\pi f$ is the angular frequency.

The larynx tone generator is shown in Fig. 3b as a voltage (pressure) source of electromotive force (acoustic pressure) $p_g = R_g U_g$ and internal resistance R_g. The source parameters U_g and R_g may be determined using the equivalent (a.c.) circuit of the larynx generator after FLANAGAN [4]. This takes into account only the a.c. components $U_g'(t)$, $A_g'(t)$ and $P_s'(t)$, of the flow (volume velocity) function

$$U_g(t) = U_{g0} + U_g'(t),$$

of the glottis area function

$$A_g(t) = A_{g0} + A_g'(t)$$

and of the subglottal pressure

$$P_s(t) = P_{s0} + P_s'(t)$$

for small signal amplitudes $\left(U_g'(t) \ll U_{g0},\ A_g'(t) \ll A_{g0},\ \text{and}\ P_s'(t) \ll P_{s0}\right)$. Under these conditions

$$U_g \approx \left(\frac{2P_{s0}}{\varrho}\right)^{1/2} A_g'(t) \tag{20}$$

and

$$R_g \approx (R_v + 2R_k)_{P_{s0}, A_{g0}}. \tag{21}$$

In equation (21)

$$R_v = \frac{12\mu d l^2}{A_{g0}^3}$$

is the classical viscous loss resistance in the glottal slit of depth d, length l and rest area A_{g0}, whereas

$$R_k = 0.875\frac{\varrho U_{g0}}{2A_{g0}^2} = 0.875\frac{(2\varrho P_{s0})^{1/2}}{2A_{g0}}$$

is the kinetic loss resistance due to the transformation of the pressure drop P_{s0} in the glottis into the kinetic energy of the air flow (cf. [13]).

Under the normal anatomical conditions of an adult male ($d \approx 3$ mm,

$l \approx 18$ mm, $A_{g0} \approx 5$ mm^2) and at average voice effort ($P_{s0} \approx 10$ cm H_2O) the source resistance R_g (21) is of the order $100 \cdot 10^5$ MKS acoustic ohms and is much greater than the vocal-tract input impedance Z_t [5]).

4. The transmission function of the pharynx-mouth tract

The electrical equivalent circuit of the vocal pharynx-mouth tract in Fig. 3b may be described by Kirchhoff's equations for $(q+1)$ independent loops of the circuit:

$$\text{Loop No. 1:} \quad Z_{11}U_1 + Z_{12}U_2 = R_g U_g, \tag{22a}$$

$$\text{Loop No. 2:} \quad Z_{21}U_1 + Z_{22}U_2 + Z_{23}U_3 = 0, \tag{22b}$$

$$\text{Loop No. 3:} \quad Z_{32}U_2 + Z_{33}U_3 + Z_{34}U_4 = 0, \tag{22c}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\text{Loop No. } i: \quad Z_{i(i-1)}U_{(i-1)} + Z_{ii}U_i + Z_{i(i+1)}U_{(i+1)} = 0, \tag{22d}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\text{Loop No. } (q+1): \quad Z_{(q+1)q}U_q + Z_{(q+1)(q+1)}U_{(q+1)} = 0, \tag{22e}$$

where

$$Z_{ii} = \underbrace{(z_{a(i-1)} + z_{b(i-1)}}_{z_{d(i-1)}} + \underbrace{(z_{ai} + z_{bi})}_{z_{di}} = z_{d(i-1)} + z_{di} \tag{23}$$

is the *self impedance* of the i-th loop, while

$$Z_{(i-1)i} = Z_{i(i-1)} = -z_{b(i-1)} \tag{24a}$$

and

$$Z_{i(i+1)} = Z_{(i+1)i} = -z_{bi} \tag{24b}$$

are the *mutual impedances* of the adjacent loops $(i-1)i$ (formula (24a)) or $i(i+1)$ (formula (24b)), respectively. The defining equations (23) and (24) are valid for loops numbered from 2 to q inclusively, i.e. for $2 \leqslant i \leqslant q$. On account of the assymetry of loops No 1 and $(q+1)$, the following notation was introduced for their self and mutual impedances:

$$Z_{11} = z_{a1} + z_{b1} + R_g = z_{d1} + R_g, \tag{25a}$$

$$Z_{12} = Z_{21} = -z_{b1} \tag{25b}$$

and

$$Z_{(q+1)(q+1)} = z_{aq} + z_{bq} + Z_{pu} = z_{dq} + z_{pu}, \tag{26a}$$

$$Z_{q(q+1)} = Z_{(q+1)q} = -z_{bq}, \tag{26b}$$

where Z_{pu} is the radiation impedance of the mouth orifice.

---

[5]) Except in the neighbourhood of the first ($F1$) and the second ($F2$) formant range of vowels (see e.g. [13]).

The loop equations (22) may be represented in the form of the matrix equation $\boldsymbol{Z}_{GU}\boldsymbol{U} = \boldsymbol{P}$, in which

$$\boldsymbol{Z}_{GU} = \left[\begin{array}{cccc:cc} Z_{11} & Z_{12} & 0 & 0 & 0 & 0 \\ Z_{21} & Z_{22} & Z_{23} & 0 & 0 & 0 \\ 0 & Z_{32} & Z_{33} & Z_{34} & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 & Z_{(q+1)q} & Z_{(q+1)(q+1)} \end{array}\right] \tag{27}$$

is the loop-impedance matrix of the $(q+1)$-th order. The elements lying on the main diagonal are the self impedances of the particular loops, and elements beyond the main diagonal are the mutual impedances of adjacent loops. The characteristic feature of this matrix is that in each i-th row $(1 \leqslant i \leqslant q+1)$, only the elements $Z_{i(i-1)}$, Z_{ii} and $Z_{i(i+1)}$ differ from zero.

From equations (22) the volume velocity U_i in an arbitrary i-th loop of the equivalent circuit in Fig. 3b may be directly determined and calculated according to the general formula

$$U_i = \frac{\Delta_{1i}}{\Delta} R_g U_g, \tag{28}$$

where $\Delta = \det \boldsymbol{Z}_{GU}$ is the characteristic determinant of the impedence matrix (27), and Δ_{1i} is the cofactor of the element Z_{1i}. The transmission function $H_u(\omega)$ of the pharynx-mouth tract (including nasalization due to the shunting effect of the nasal tract), expressed as the ratio of the volume velocity $U_{(q+1)}$ through the radiation impedance Z_{pu} of the mouth orifice to the volume velocity $U_g \approx U_1$ through the glottis, is thus equal to

$$H_u(\omega) = \frac{U_{(q+1)}}{U_1} = \frac{\Delta_{1(q+1)}}{\Delta_{11}}. \tag{29}$$

Similarly, the transmission function $H'_u(\omega)$, expressed as the ratio of the acoustic pressure $p_u = Z_{pu} U_{(q+1)}$ in the mouth orifice to the volume velocity $U_g \approx U_1$ through the glottis, is equal to

$$H'_u(\omega) = \frac{p_u}{U_g} = \frac{Z_{pu} U_{(q+1)}}{U_1} = \frac{\Delta_{i(q+1)}}{\Delta_{11}} Z_{pu}, \tag{30}$$

where Z_{pu} is the radiation impedance of the mouth orifice, simulated by a circular piston of area A_q vibrating in a infinitely large flat baffle. The impedance Z_{pu} is given by the approximate expression [5, 12]

$$Z_{pu} \approx \frac{\omega^2 \varrho}{2\pi c} + j\omega \frac{8\varrho}{3\pi\sqrt{\pi A_q}}, \tag{31}$$

where ϱ is the density of air, c — the velocity of sound, and $\omega = 2\pi f$ — the angular frequency.

The use of formulae (29) and (30) in numerical calculations and in computer simulation must be based on a knowledge of the numerical data for both the cross-sectional area function $A(x)$ of the vocal tract under definite articulation conditions (cf. Table 1) and the input impedance Z_n of the nasal tract shunting the vocal tract at the point of bifurcation.

5. The input impedance of the nasal tract

In view of the scarcity of sufficiently representative data for determining the anatomical structure of the nasal tract, e.g. from X-ray pictures of the human head (cf. [3], Fig. 2.4-2), its cross-sectional area function $A_n(x)$, as simulated in most analogue speech synthesizers, is usually extremely simplified (see [6, 7, 8, 17]). The simplification is reasonable since the shape of the area function $A_n(x)$ does not depend on the articulatory conditions and is constant except for differences between individuals.

Using data from the medical and phonetic literature, we approximated the nasal tract for the present work in the form of a cascade connection of five homogeneous cylindrical tubes of lengths l_i and cross-sectional areas A_{ni}, where $i = 1, 2, \ldots, 5$ (Fig. 5). In Table 2 the numerical values of the parameters

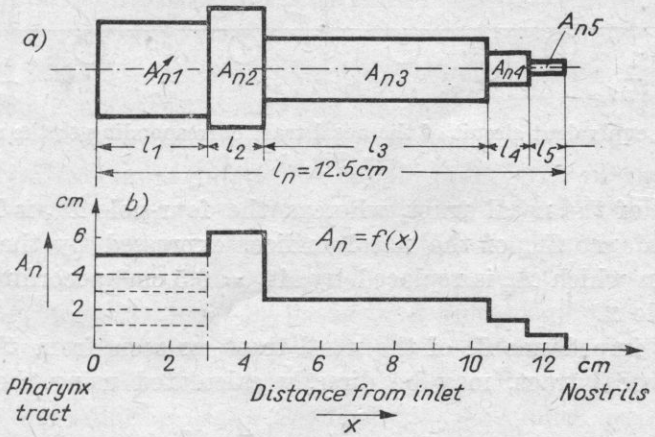

Fig. 5. Acoustical model of the nasal tract (a) and the area function $A_n(x)$ of its cross-section (b)

l_i and A_{ni} are presented and describe the geometric configuration of the model which corresponds to the average anatomical conditions of an adult male. The only variable parameter is the cross-sectional area A_{n1} of the first cylindrical segment, whose value determines the actual degree of nasalization of the vowel considered, which depends on the extent of cleft palate. In this way, the input impedance Z_n of the nasal channel, shunting the pharynx-mouth tract, is determined — ceteris paribus — by the value of A_{n1}.

Table 2. Numerical values of the parameters l_i, A_{ni} $(i = 1, 2, \ldots, 5)$ determining the geometrical configuration of the nasal tract in Fig. 5

Segment No	1	2	3	4	5
Length l_i [cm]	3.0	1.5	6.0	1.0	1.0
Cross-sectional area A_{ni} [cm^2]	variable	6.0	2.0	1.2	0.5

As for the pharynx-mouth tract considered above, the electrical equivalent circuit of the nasal tract, shown in Fig. 5, may be presented in the form of a cascade connection of five four-poles T_{ni} $(i = 1, 2, \ldots, 5)$ which simulate the successive segments of the acoustical model, as is shown in Fig. 6. The four-pole T_{ni} is supplied from a voltage source which represents the acoustic pressure

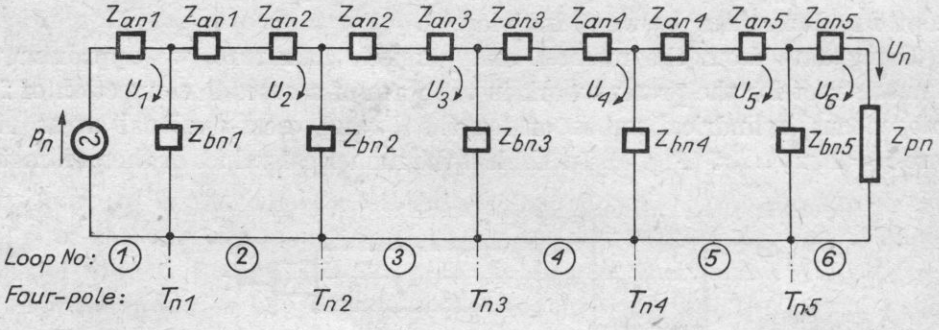

Fig. 6. Electrical equivalent circuit of the nasal tract corresponding to the acoustical model in Fig. 5

p_n at the inlet of the nasal tract, whereas the four-pole T_{n5} is loaded by the radiation impedance Z_{pn} of the nasal orifice, expressed by the approximate formula (31) in which A_q is replaced by $A_{n5} = 0.5$ cm^2 according to the data shown in Table 2.

The input impedance Z_n of the nasal tract, as seen from the bifurcation point of the vocal tract, may be directly calculated using the formula

$$Z_n = \frac{p_n}{U_1} = \frac{\Delta}{\Delta_{11}}, \tag{32}$$

in which Δ is the characteristic determinant, and Δ_{11} is the cofactor of the element Z_{11} of the loop-impedance matrix $\mathbf{Z}_N$,

$$\mathbf{Z}_N = \begin{bmatrix} Z_{11} & Z_{12} & 0 & 0 & 0 & 0 \\ Z_{21} & Z_{22} & Z_{23} & 0 & 0 & 0 \\ 0 & Z_{32} & Z_{33} & Z_{34} & 0 & 0 \\ 0 & 0 & Z_{43} & Z_{44} & Z_{45} & 0 \\ 0 & 0 & 0 & Z_{54} & Z_{55} & Z_{56} \\ 0 & 0 & 0 & 0 & Z_{65} & Z_{66} \end{bmatrix}, \tag{33}$$

which describes the model of the nasal tract in Fig. 6. The elements Z_{ii} ($i = 1, 2, \ldots, 6$), lying on the main diagonal of the matrix $\boldsymbol{Z}_N$, are, by analogy with matrix $\boldsymbol{Z}_{GU}$ (27), the self-impedances of the respective loops of the equivalent network in Fig. 6, whereas the elements Z_{ij} ($i, j = 1, 2, \ldots, 6$; $i \neq j$), beyond the main diagonal, are the mutual impedances of the adjacent loops, according to definitions (23) and (24). The physical meaning of the impedances z_{ani} of the longitudinal branches and the impedances z_{bni} of the transverse branches of the four-poles T_{ni} ($i = 1, 2, \ldots, 5$) is determined by the general formulae (9) and (10) derived earlier for the pharynx-mouth tract. The only difference is that in the case of the nasal tract the effect of energy dissipation due to vibration of the tract wall may be disregarded by putting $G_{sj} = C_{sj} = 0$ in equation (10).

The input impedance Z_n of the nasal tract, given by equation (32) for different values of the inlet area A_{ni} depending on the degree of nasalization, is a variable parameter occurring in equations (12), (13), (16), (17), (19) and subsequent ones. It is used for the computation of the transmission functions $H_u(\omega)$ (formula (29)) and $H'_u(\omega)$ (formula (30)) of the pharynx-mouth tract including nasalization.

6. The transmission function of the pharynx-nasal tract

As a result of the nasalization effect due to cleft palate, the sound wave is radiated through both the mouth and the nostrils. The contribution of the sound energy emitted through the nostrils to the total energy of the speech wave depends on the input impedance Z_n (formula (32)) of the nasal tract, and decreases to zero as $Z_n \to \infty$. In most pathological cases and in postoperative states this contribution is significant. In this connection, the application of the superposition theorem to acoustic pressure, $p(r) = p_u + p_n$, according to the general formula (1), must be based on a knowledge of the transmission function $H_n(\omega)$ of the pharynx-nasal tract ($G + N$), including the parallel impedance of the mouth-tract input impedance Z_u (see Fig. 1). The methodology and procedure of computation are analogous to those ones used in sections 3 and 4 for the pharynx-mouth tract.

The equivalent electrical circuit for the pharynx-nasal tract including the shunting effect of the mouth tract is shown in Fig. 7. The four-poles $T_1, T_2, \ldots, \ldots, T_{(k-1)}$, where $k = 8$[6]), represent, as in Fig. 3, the pharynx tract, whose cross-sectional area function $A(x)$ is determined — depending on the conditions of articulation — by the numerical values in the upper $(k - 1)$ rows of Table 1. The four-pole T'_k simulates the bifurcation of the vocal tract, i.e.

[6]) Provided that the bifurcation of the vocal tract is located at a distance $x = 8$ cm from the glottis.

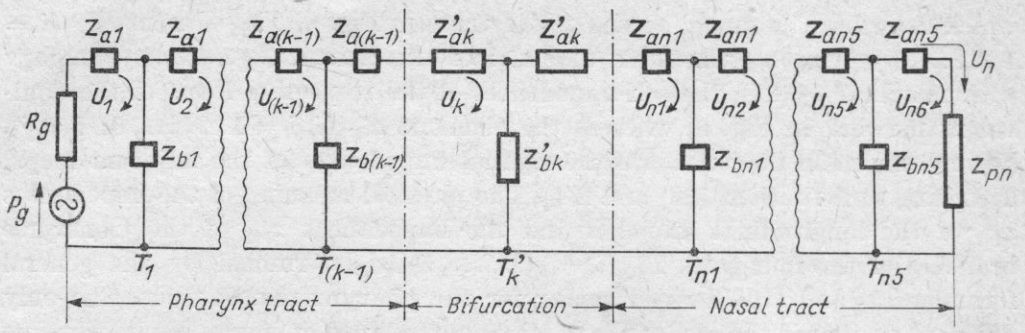

Fig. 7. Electrical equivalent circuit of the pharynx-nasal vocal tract including the shunting effect of the mouth tract

the place where the input impedance Z_u of the mouth tract shunts the pharynx-nasal tract. The configuration of the four-pole T'_k is identical with that of the four-pole T_k previously shown in Fig. 4, the only difference consisting in the shunting impedance whose role is now fulfilled by Z_u instead of Z_n. In full analogy to formulae (17) and (19), the longitudinal impedances z'_{ak} and the transverse impedance z'_{bk} of the four-pole T'_k may be expressed as

$$z'_{ak} = Z_0\left(\gamma\frac{l}{2}\right)\frac{Z_0+Z_u\left(\gamma\dfrac{l}{2}\right)}{Z_0+2Z_u\left(\gamma\dfrac{l}{2}\right)} \tag{34a}$$

and

$$z'_{bk} = \frac{Z_0}{\gamma l}\,\frac{Z_u}{Z_u+Z_0/\gamma l}. \tag{34b}$$

Disregarding the losses of the four-pole T'_k, they may be written as

$$z'_{ak} = j\,\frac{\varrho\omega l}{2A_k}\,\frac{\varrho c/A_k+jZ_u\omega l/2c}{\varrho c/A_k+jZ_u\omega l/c} \tag{35a}$$

and

$$z'_{bk} = -j\,\frac{\varrho c^2}{A_k\omega l}\,\frac{Z_u}{Z_u-j\varrho c^2/A_k\omega l}, \tag{35b}$$

where Z_0 and γ denote, as before, the characteristic impedance (7) and the equivalent propagation constant (8) of the four-pole. Z_u is the input impedance of the mouth tract, measured or calculated at the input terminals of the four-pole $T_{(k+1)}$, provided the four-pole T_q is loaded with the radiation impedance Z_{pu} (formula (31)) of the mouth orifice (cf. Fig. 3). As for the input impedance Z_n of the nasal-tract (32), the impedance Z_u may be computed from the formula

$$Z_u = \frac{p_u}{U_{(k+1)}} = \frac{\Delta_u}{\Delta_{(k+1)(k+1)}}, \tag{36}$$

where Δ_u is the characteristic determinant of the loop-impedance matrix $\mathbf{Z}_u$ of the mouth tract, which is presented in Fig. 3 as a cascade connection of the four-poles $T_{(k+1)}, T_{(k+2)}, \ldots, T_q$ under definite articulation conditions, and $\Delta_{(k+1)(k+1)}$ is the cofactor of the element $Z_{(k+1)(k+1)}$ of this matrix.

Finally, the transmission function $H_n(\omega)$ of the pharynx-nasal tract (including the shunting effect of the mouth tract), defined as the ratio of the volume velocity U_n through the radiation impedance Z_{pn} of the nostrils to the volume velocity U_g through the glottis, is given by the expression

$$H_n(\omega) = \frac{U_n}{U_g} = \frac{\Delta_{1n}}{\Delta_{11}}, \tag{37}$$

where Δ_{11} and Δ_{1n} are the cofactors of the elements Z_{11} and Z_{1n} of the loop-impedance matrix $\mathbf{Z}_{GN}$ of order $n = k+6$, which corresponds to the Kirchhoff equations for $k+6$ independent loops of the network, shown in Fig. 7, which represents the pharynx-nasal tract. Similarly, the transmission function $H'_n(\omega)$ of the tract, defined as the ratio of the acoustic pressure $p_n = Z_{pn} U_n$ in the nose outlet orifice to the volume velocity U_g through the glottis, is equal to

$$H'_n(\omega) = \frac{p_n}{U_g} = \frac{Z_{pn} U_n}{U_g} = \frac{\Delta_{1n}}{\Delta_{11}} Z_{pn}. \tag{38}$$

7. Conclusions

The resulting transmission function $H(\omega)$ of the complex, i.e. pharynx-mouth-nasal vocal tract, defined as the ratio of the volume velocity or acoustic pressure at the measuring point on the axis of symmetry of the system, which radiates the sound wave through the mouth and nose orifices, to the volume velocity of the larynx generator, is given, in accordance with the superposition theorem (1), by the expressions

$$|H(\omega)| = |H_u(\omega)| + |H_n(\omega)| \tag{39}$$

or

$$|H'(\omega)| = |H'_u(\omega)| + |H'_n(\omega)|, \tag{40}$$

where $H_u(\omega)$, $H'_u(\omega)$, and $H_n(\omega)$, $H'_n(\omega)$ describe radiation from the mouth (29), (30), and the nose (37), (38), respectively.

The pole-zero distribution in the frequency domain of the model transmission function $|H(\omega)|$ describes the formant-antiformant structure of the particular speech sound, e.g. an oral vowel, whose phonetic features are uniquely determined by the cross-sectional area function $A(x)$ of the vocal tract, and the degree of nasalization Z_n. In this way, using the model and its mathematical description, one can compute the spectral characteristic of an arbitrary speech sound under definite articulatory conditions (which depends on the

vocal-tract geometry), thus simulating the anatomy of the real (i.e. biological) system.

The simulative model of the vocal tract presented is both universal and versatile. It can simulate the articulatory conditions of arbitrary speech sounds with glottal excitation, taking into account any personal differences which result from either the individual anatomical features of the speech organ or the pathological anomalies, including the forced nasalization of oral vowels due to cleft palate.

The superiority of the simulative model, over analogue models, arises from its high accuracy in reproducing the real physio-pathological conditions of the biological system. This is a consequence of taking into account both the losses in the vocal tract and the radiation impedances of the mouth and nose orifices, which cannot be taken into consideration in analogue systems [12]. An additional advantage of the model is the variable parameter Z_n which corresponds to the nasal-tract input impedance, constituting a quantitative measure of the nasalization effect which depends on the extent of the cleft palate. It enables us to determine the quantitative relations between the spectral structure of the directly measured diagnostic signal (i.e. the acoustic pressure of the speech wave) and the physio-pathology of the vocal tract. These relations not only facilitate and make phoniatric diagnosis objective, but also are useful in the control and documentation of the rehabilitation process.

The clinical verification of the simulative model presented constitutes the subject of active research being conducted at the Department of Cybernetic Acoustics, Institue of Fundamental Technological Research, Polish Academy of Sciences, in close cooperation with the Otolaryngological Clinic of the Institute of Surgery of the Medical Academy in Warsaw. This will be preceded by the elaboration of the model in a form convenient to digital presentation and computer simulation, in order to obtain a rational compromise between the accuracy of reproducing the physio-pathological features of the biological system and the operative capabilities of moderately-sized computers currently available (e.g. the ODRA 1300).

References

[1] J. B. Dennis, *Computer control of an analog vocal tract*, Proc. Stockholm Speech Comm. Seminar, Royal Inst. Techn., Stockholm 1962.

[2] H. K. Dunn, *The calculation of the vowel resonances and an electrical vocal tract*, J. Acoust. Soc. Am., **22**, 740–753 (1950).

[3] G. Fant, *Acoustic theory of speech production*, s'-Gravenhage, Mouton and Co., 1960.

[4] J. L. Flanagan, *Some properties of the glottal sound source*, J. Speech and Hearing Res., **1**, 99–116 (1968).

[5] J. L. Flanagan, *Speech analysis, synthesis and perception*, Springer Verlag, 2nd Edition, Berlin-Heidelberg-New York 1972.

[6] M. H. L. HECKER, *Studies of nasal consonants with an articulatory speech synthesizer*, J. Acoust. Soc. Am., **34**, 179–188 (1962).

[7] A. S. HOUSE, *Analog studies of nasal consonants*, J. Speech and Hearing Disorders, **22**, 190–204 (1957).

[8] A. S. HOUSE, K. N. STEVENS, *Analog studies of the nasalization of vowels*, J. Speech and Hearing Disorders, **21**, 218–232 (1956).

[9] K. ISHIZAKA, J. C. FRENCH, J. L. FLANAGAN, *Direct determination of vocal-tract wall impedance*, J. Acoust. Soc. Am., **55**, 879 (A) (1974).

[10] W. JASSEM, *Fundamentals of acoustic phonetics*, Polish Scientific Publishers, PWN, Warsaw 1973 [in Polish].

[11] J. KACPROWSKI, *Theoretical bases of the synthesis of Polish vowels in resonance circuits*, Speech Analysis and Synthesis, Vol. 1, 219–287, Polish Scientific Publishers, PWN, Warsaw 1968.

[12] J. KACPROWSKI, W. MIKIEL, A. SZEWCZYK, *Acoustical modelling of cleft palate*, Archives of Acoustics, **1**, 2, 137–158 (1976).

[13] J. KACPROWSKI, *Physical models of the larynx source*, Archives of Acoustics, **2**, 1, 47–70 (1977).

[14] J. LAGASSE, *Étude des circuits électriques*, Editions Eyrolles, Paris 1962.

[15] M. MRAYATI, B. GUERIN, *Études des characteristiques acoustiques des voyelles françaises par simulation du conduit vocal avec pertes*, Revue d'Acoustique, **9**, 36, 18–32 (1976).

[16] G. ROSEN, *Dynamic analog speech synthesizer*, J. Acoust. Soc. Am., **30**, 201–209 (1958).

[17] K. N. STEVENS, S. KASOWSKI, G. FANT, *An electrical analog of the vocal tract*, J. Acoust. Soc. Am., **25**, 734–742 (1953).

Received on 14th May 1977

ARCHIVES OF ACOUSTICS
2, 4, 257 (1977)

VIBRATO AND VOWEL IDENTIFICATION

J. SUNDBERG

Royal Institute of Technology, S-100 44 Stockholm 70 Center for Speech Communication Research and Musical Acoustics

The influence of vibrato on vowel identification is studied in synthesized vowel sounds with fundamental frequencies between 300 and 1000 Hz. Phonetically trained subjects were asked to identify these stimuli presented with and without vibrato as any of 12 Swedish long vowels. Small and occasional effects are observed on the mean formant frequencies and the scatter of the responses.

However, the effects are greater and more frequent than could be expected to occur by chance. In the majority of the cases, where the vibrato affected the responses, the vowel identification became somewhat harder when the stimulus was presented with vibrato. It is assumed that the vibrato may be used to conceal some of the deviations from the vowel qualities in speech which are typically made in singing.

1. Introduction

Vibrato occurs in music performed on many types of instruments, such as string and wind instruments, and in singing. Physically, it corresponds to a quasi sinusoidal modulation of the fundamental frequency. Provided that the rate and extent of the vibrato is kept within certain limits, the vibrato is generally considered as useful means of musical expression in a variety of musical styles.

Acoustic characteristics typical of music sounds can be assumed to serve some purpose in musical communication. It is frequently assumed that the vibrato covers up small errors in the fundamental frequency. In a previous investigation no support was found for this hypothesis as far as the pitch of single tones – and thus not chords – is concerned (SUNDBERG 1972). Here, another frequently assumed explanation of the vibrato will be tested, namely that the vibrato facilitates vowel identification. This hypothesis seems plausible, because as the partials move in frequency (e.g. owing to a vibrato) their amplitudes scan small parts of the vocal-tract transfer function, and more

information concerning the formant frequencies is communicated. The effect would be particularly strong in the soprano range where the fundamental frequency may be even higher than 1000 Hz. In such cases the information on the formant frequencies would be hard to decode from the acoustic spectrum.

2. Experiment

A series of high-pitched sounds was synthesized with and without vibrato using a terminal analogue. Six formant-frequency combinations were used, each of which corresponds to a Swedish long vowel: [u, o, ɑ, e, i, y]. The vowel formant frequencies (see Table 1) were found in a soprano singing at a funda-

Table 1. Formant frequency values (in Hz) used for the synthesis of vowel stimuli and ascribed to response vowels

Vowel	Vowel stimuli				Response vowels		
	F_1	F_2	F_3	F_4	F_1	F_2	F_3
u	325	700	2700	4000	320	565	2775
o	430	650	2820	4000	415	725	2775
a	640	1090	2800	4000	700	1065	3000
ɑ					855	1200	2820
æ					815	2030	3000
ɛ					625	2305	3040
e	400	2000	2700	4000	375	2790	3360
i	270	1900	1950	4000	275	2675	3655
y	250	1950	2510	4000	275	2555	3210
ʉ					300	1920	2745
ø					390	2040	2725
œ					565	1290	2690

mental frequency of 262 Hz (SUNDBERG 1975). Contrary to the effect which had been observed in this soprano, each of these 6 combinations of formant frequencies was used «without changes» for 4 different pitches within the soprano range: 300, 450, 675, and 1000 Hz. In this way the material included sounds ranging from maximum ease to maximum difficulty as regards vowel identification. The vibrato consisted of a sinusoidal modulation of the fundamental frequency. The modulation rate was 6 undulations per second, and the deviation was ±50 cents. These values are found in normal singing, even though the deviation is normally sligthly narrower.

Four samples of each vowel stimulus were tape recorded at a constant overall amplitude. The time and the shape of the onsets and decays were all similar. The stimuli samples were arranged in random order avoiding more than one repetition of a given stimulus in sequence. The entire tape contained a total of 192 vowel samples: 6 formant-frequency combinations × 4 fundamental

frequencies $\times 4$ presentations $\times$ 2 cases (with and without vibrato). Each stimulus had a duration of 1 s and was followed by a silent interval of 4 s.

The stimuli were presented through head-phones (Sennheiser HD 414) at an overall *SPL* of 70 dB, approximately, to 10 phonetically trained subjects. Their task was to decide which one of 12 given Swedish long vowels [u, o, ɑ, a, æ, ɛ, e, i, y, ʉ, ø, œ] sounded most similar to the stimulus they heard and to write that vowel in phonetic symbols on a test chart. These vowels, suggested by the subjects as interpretations of the stimuli, will be referred to as «response vowels». The test, including 3 short breaks, lasted about 30 min.

3. Evaluation of responses

It seems reasonable to hypothesize that the process of identifying an isolated stimulus as a specific vowel involves three major steps. The first is a purely sensory one in which the acoustic stimulus is converted into some kind of «sensory information». The second step involves the extraction of an ensemble of «sensory characteristics» from this sensory information. In the third step a «comparison» is made between these sensory characteristics and corresponding characteristics of vowels stored in the subject's internal reference. The difficulty in identifying a sound as a specific vowel would depend on the degree of similarity between the ensemble of sensory characteristics and the characteristics in the internal reference.

The vibrato may affect the sensory information in various ways. Let us assume that it affects the extraction of sensory characteristics. If so, the responses for a given vibrato stimulus should differ from those obtained when the same stimulus was presented without vibrato. In the test responses the effect would then be that the mean formant frequencies of the response vowels for a given stimulus would be different in the cases with and without vibrato. In that case the identification difficulties may be affected in various ways, depending on whether the vibrato makes the sensory characteristics (1) more similar or (2) less similar to the subject's internal reference. The cases (1) and (2) would differently affect the scatter of the responses.

The above considerations demonstrate the need for some measures of the average and of the scatter of the response vowels. In order to obtain such measures a set of formant frequencies was ascribed to each of the 12 response vowels in accordance with data on female speakers (FANT 1973). Thereby the fourth formant, having negligeable influence on the perceived vowel identity, was disregarded. Those values are listed in Table 1. These formant frequency values were then expressed in Mels and for each stimulus the averages and the standard deviations of each of the three lowest formant frequencies of the response vowels obtained from that stimulus were computed. The resulting set of three mean formant frequencies was considered to represent the «average response vowel» of that stimulus.

The scatter between the responses obtained from a given stimulus was estimated in the following way. Recently, LINDBLOM (1975), following ideas of PLOMP (1970), established that the perceptual distance D between any two Swedish vowels can be approximated as

$$D = \sqrt{\Delta M_1^2 + \Delta M_2^2 + \Delta M_3^2},$$

where ΔM_n is the frequency difference in Mels in the nth formant between the two vowels. This equation was used to calculate for each stimulus the perceptual distance between the average response vowel and each of the individual responses to the same stimulus. Then, the average of these distances over subjects was determined for each stimulus. The resulting values were assumed to reflect the «scatter of the responses», i.e. the difficulty with which the stimulus was identified as a specific vowel.

4. Results

One of the subjects refused to give any interpretation of the stimuli, all of which happened to lack vibrato. The inter-subject differences regarding the internal reference appeared to be rather small.

The formant frequencies of the average responses are shown in Fig. 1, and the corresponding standard deviations are listed in Table 2. The formant frequencies, used in the synthesis of the stimuli, are also indicated in Fig. 1. Hence, the relationships between these formant frequencies and those of the average responses can be examined. The stimuli synthesized with [ɑ]-formants offer an illustrative example. The formant frequencies of the average response drop with rising fundamental frequency up to 675 Hz. The reason for this is probably that, in order to maintain the vowel quality when the fundamental is raised, it is necessary to raise the formant frequencies slightly. SLAWSON (1968) found an average increase of 10 % in the formant frequencies to be required per octave rise of the fundamental.

If the fundamental is raised while the formant frequencies are kept constant, as was done in our experiment, the vowel quality will shift towards a vowel with lower formant frequencies. In this case the subjects probably identified the two lowest formant frequencies correctly, and the formant frequencies ascribed to the response vowels contain an error which depends on the fundamental frequency. A correction would be needed, but it is not known exactly how to make this correction. For instance, SLAWSON's correction of 10 % increase in formant frequency per octave rise in the fundamental does not fully compensate the effect. It should also be mentioned that all stimuli with a fundamental frequency of 675 Hz gave almost invariously response vowels in which one of the (assumed) formants matched one of the stimulus partials in frequency almost perfectly.

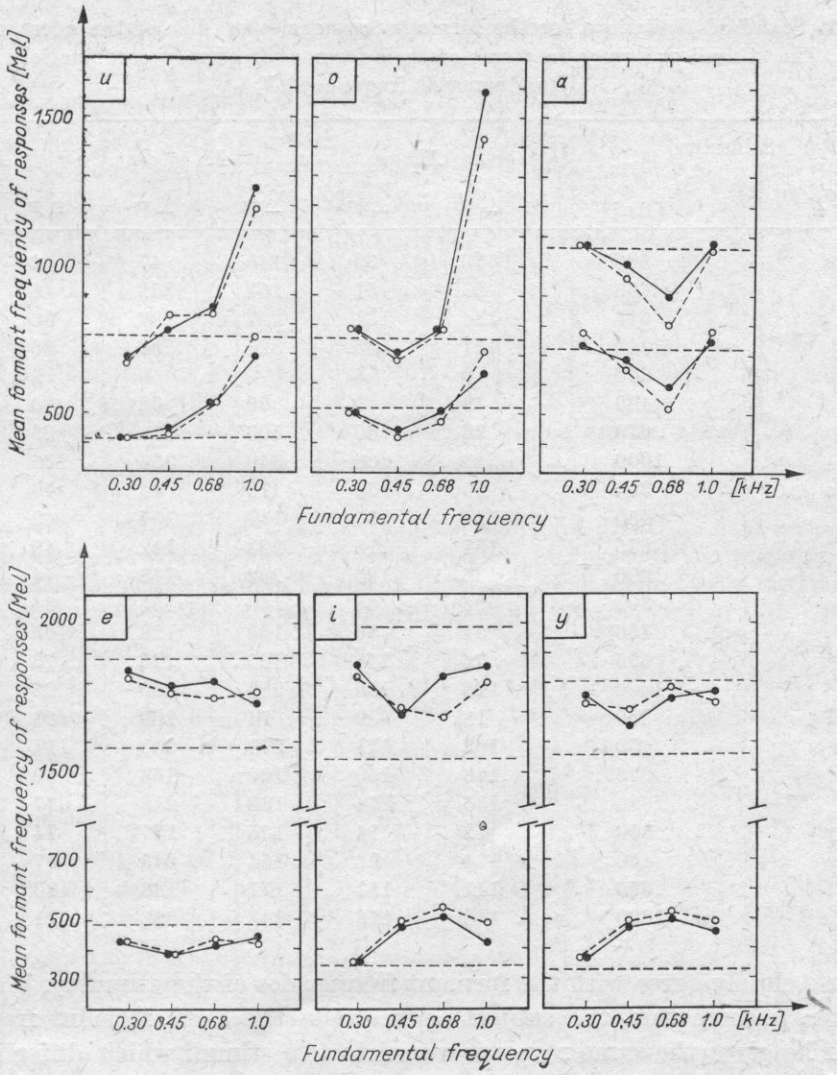

Fig. 1. Evaluation of the response vowels in terms of the means of their assumed two lowest formant frequencies expressed in Mels. Filled and open circles pertain to stimuli presented with and without vibrato, respectively. Horizontal dashed lines show the formant frequencies which were used in the synthesis of the stimuli

This may indicate that, in very high-pitched vowels, subjects tend to interpret the frequency of a spectrum partial as the frequency of a formant. This would lend support to the assumption that the formant frequencies which we ascribed to the response vowels may agree with the formant frequencies «perceived» by the subjects in very high-pitched vowels. Additional support for the same speculation may be found in the case of the [ɑ]-formant stimulus at 1000 Hz fundamental frequency. Here, the formant frequencies of the average

Table 2. Standard deviations for the formant frequencies in Mels of the average response vowels. The stimuli are given in terms of their formant-frequency combinations (F_N) and fundamental frequency (F_0)

Stimulus		M_1		M_2		M_3	
F_N	F_0 [Hz]	$-V$	$+V$	$-V$	$+V$	$-V$	$+V$
u	300	30	33	146	47	2	0
	450	59	57	251	351	6	42
	675	134	131	227	187	36	36
	1000	167	145	324	290	86	53
o	300	0	0	0	0	0	0
	450	49	43	69	62	0	0
	675	96	66	105	158	26	14
	1000	213	206	340	354	106	99
ɑ	300	60	55	73	44	46	33
	450	124	130	147	154	49	47
	675	151	118	203	142	40	31
	1000	135	125	202	148	44	35
e	300	65	65	150	155	84	94
	450	51	50	133	128	105	83
	675	96	141	120	226	76	101
	1000	125	139	296	236	95	102
i	300	18	9	74	109	81	110
	450	102	121	273	208	111	101
	675	146	150	120	163	80	84
	1000	135	174	132	216	77	106
y	300	12	14	125	137	74	81
	450	104	92	154	243	97	114
	675	141	153	177	136	110	90
	1000	146	176	243	208	111	91

response closely agree with the formant frequencies of the stimulus. We cannot however, assume that the subject «perceived» the same formant frequencies even if he gives the same response vowel to two stimuli which differ in pitch. Therefore, the formant frequencies of the average responses should not be compared between stimuli differing in pitch, and our material cannot elucidate the relationships between the stimulus spectra differing in pitch and the «perceived» formant frequencies.

A comparison of the mean formant frequencies of the response vowels obtained with and without vibrato would inform us on the effect that the vibrato may have on the extraction of sensory characteristics, as mentioned. Such comparisons can be made in Fig. 2. It shows the difference in the mean formant frequencies of the response vowels. Negative values indicate that the mean dropped when the stimulus had vibrato. Differences exceeding two times the standard error of difference were found in 13 cases (circled dots in Fig. 2). These cases pertain to 10 different pairs of stimuli. Thus, under certain

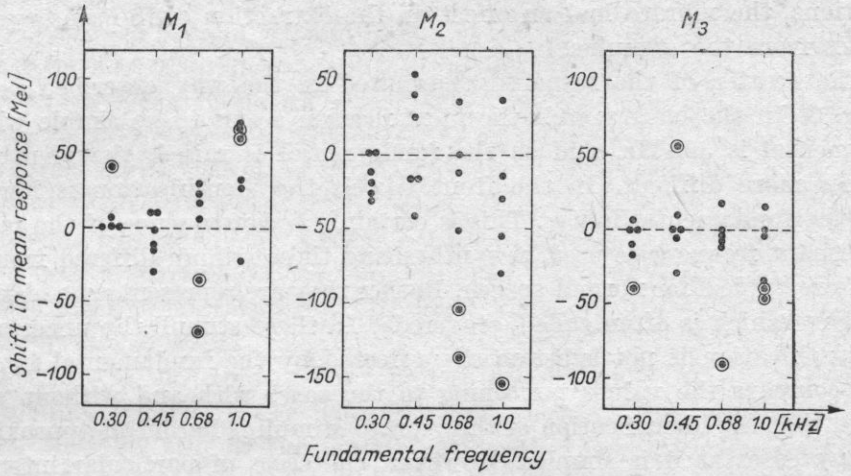

Fig. 2. Difference in the first (left), second (middle), and third (right) formant frequencies in Mels of the response vowels collected from a given stimulus when it was presented with and without vibrato. The circled dots represent differences which exceed twice the standard error of difference

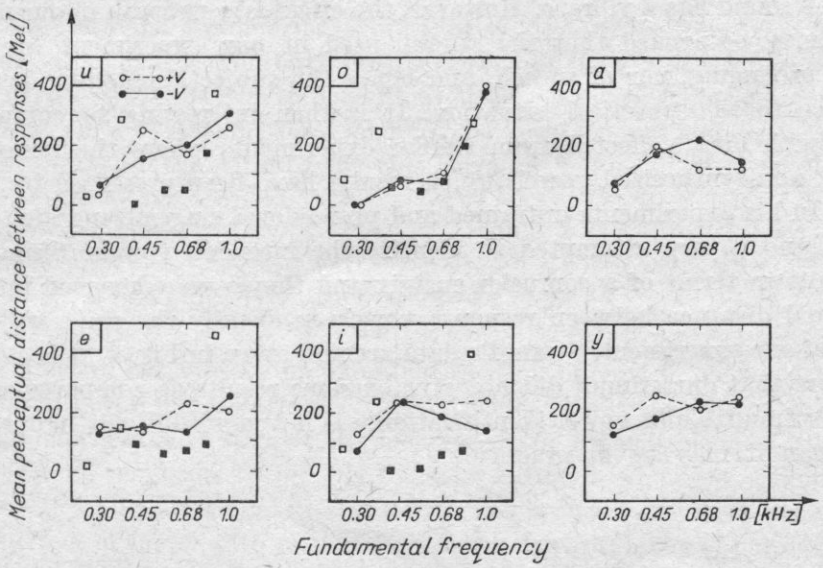

Fig. 3. Scatter of response vowels in terms of the average distance between the individual responses and the average response. Filled and open circles pertain to responses collected when the stimulus was presented with and without vibrato, respectively. The squares give values reported by STUMPF (1926). They were observed when subjects identified vowels sung by two untrained singers (open squares) and one professional soprano (filled squares)

conditions, the vibrato had an effect on the extraction of the sensory characteristics from the stimulus.

The scatter of the responses estimated in the way described is shown in Fig. 3. In the back vowels the identification seems to be simple when the fundamental is 300 Hz, and as the fundamental is raised, the identification becomes more difficult. In the front vowels the identification is hard even when the fundamental is low. This is certainly a consequence of the fact that the formant frequencies used in synthesizing these stimuli differed from those which are typical for normal speech. Rather they are representative of the type of singing which is often called «covered». In these stimuli the uncertainty of the identification is not substantially affected by the fundamental frequency. If we compare the values pertaining to the cases with and without vibrato, we find that the identification of the vibrato stimuli is harder in approximately as many cases as it is simpler. However, the cases of particular interest are those where the responses changed significantly in one or more formant frequencies when the vibrato was added to the stimulus. This happened in 10 cases, and in these cases the sensory characteristics can be assumed to have changed owing to the vibrato. Out of these 10 cases the identification became more uncertain in 7 cases. This seems to support the assumption that the vibrato may have a slight effect on the difficulty with which a sound is identified as a vowel, and the effect is that the identification is somewhat harder to perform when the sound has a vibrato. However, the effect is very weak in our material.

It may be argued that the stimuli used in our experiment are typical neither of singing, nor of speech, and hence the subjects' reactions have little relevance to the practical situation. It is then interesting to compare our results with data collected from similar experiments where the stimuli were natural, non-synthesized vowels. Such results have been presented by STUMPF (1926). In his experiments untrained and professional singers sang high-pitched vowels, and a jury attempted to identify the intended vowels. Stumpf gave his results in terms of a confusion matrix and they were converted into mean perceptual distance between response vowels in exactly the same way as the results of our experiment. Stumpf's data are also given in Fig. 3. It is clear from the figure that our stimuli did not give extreme results as compared with the natural stimuli. Thus, vowel identification was not more difficult in our experiment than it may be in practice.

5. Discussion

Our experiments have shown that the vibrato may have a slight effect on vowel identification. It occasionally changes the interpretation of the acoustic stimulus somewhat and it seems to obscure the vowel identity slightly. This result may seem rather unexpected. However, it is in agreement with previous findings of CARLSON et al. (1973). They studied the vowel identification at the

boundary between [e] and [i], and used a constant as well as a gliding fundamental frequency. The vowel identification seemed to be somewhat harder to perform when the fundamental was gliding. This appears to suggest that the difference limen for formant frequency is somewhat larger when the fundamental is not kept constant (cf. FLANAGAN 1955). The primary question seems to be to what extent we can derive information on the formant frequencies of a vowel sound from the phase relation between the frequency and the amplitude of the spectrum partials; if the amplitude of a partial rises with increasing pitch, this means that this partial is lower in frequency than the formant, and vice versa.

The effect of the vibrato was found to be very slight in our experiment. It is probably related to the extraction of sensory characteristics of the stimulus. In some stimuli the determination of what is characteristic may be very simple regardless of whether or not the stimulus has a vibrato. In other stimuli it may be hard to determine what is characteristic and in these cases the vibrato is more likely to play a role. Probably a singer would be intuitively aware of this possibility and make systematical use of it in singing. Then, the effect of the vibrato might be greater in practice than in our experiment.

According to our results, the vowel identification is somewhat harder to perform when the stimulus has a vibrato. This is to say that the vibrato tends to conceal the singer's formant frequencies. But why should a singer try to conceal the formant frequencies? Substantial deviations from the formant frequencies of normal speech have been observed at professional singers (SUNDBERG 1970; 1975). It may be important to obscure the consequences for vowel quality that such deviations may lead to, and professional singers may make systematical use of the vibrato for this purpose.

6. Conclusions

The effect of the vibrato on the identification of vowels seems to be very slight and to occur under certain conditions only. However, the vibrato may change the vowel quality slightly. In such cases, the identification of the sound as a specific vowel seems to be slightly more difficult when the stimulus has a vibrato. It is believed that a singer may in practice profit systematically from this effect of the vibrato so as to reduce the perceptability of her deviations from the formant frequencies of normal speech.

Acknowledgement. Stefan PAULI is acknowledged for his assistance in computer programming. The work was supported by The Bank of Sweden Tercentenary Foundation.

References

[1] R. Carlson, G. Fant, B. Granström, *Two-formant models, pitch and vowel perception*, paper presented at the Symposium on Auditory Analysis and Perception of Speech (Leningrad 1973), Academic Press, London 1975.

[2] G. Fant, *Speech sounds and features*, MIT-Press, Cambridge, Mass., 1973, p. 36 and p. 84.

[3] J. L. Flanagan, *A difference limen for vowel formant frequency*, J. Acoust. Soc. Am., **27**, 3, 613 (1955).

[4] B. Lindblom, *Experiments in sound structure*, paper presented at the 8th Int. Congr. of Phonetic Sciences, Leeds, Aug. (1975).

[5] R. Plomp, *Timbre as a multidimensional attribute of complex tones*, in *Frequency Analysis and Periodicity Detection in Hearing*, ed. by R. Plomp and G. F. Smoorenburg, A. W. Sijthoff, Leiden 1970, p. 397.

[6] A. W. Slawson, *Vowel quality and musical timbre as functions of spectrum envelope and fundamental frequency*, J. Acoust. Soc. Am., **43**, 1, 87 (1968).

[7] C. Stumpf, *Die Sprachlaute*, Springer, Berlin 1926, p. 77–85.

[8] J. Sundberg, *Formant structure and articulation in spoken and sung vowels*, Fol. Phon., **22**, 1, 28 (1970).

[9] J. Sundberg, *Pitch of synthetic sung vowels*, STL-QPSR, 1, 34 (1972).

[10] J. Sundberg, *Formant technique in a professional female singer*, Acustica **32**, 3, 90 (1975).

Received on 5th July 1976
Revised 12 May 1977

ARCHIVES OF ACOUSTICS
2, 4, 267 (1977)

THE EQUIVALENT AREA AND ACTIVE MASS OF MICROPHONE MEMBRANES

RUFIN MAKAREWICZ, JACEK KONIECZNY

Chair of Acoustics, A. Mickiewicz University (Poznań)

The paper presents a proposed method for determining the equivalent area and active mass of microphone membranes. It is based on an energetic definition of these parameters and assumes a flat piston. It requires a knowledge of the phase distribution and also of the displacement amplitude and velocity over the whole surface of the real membrane. The methodology of measuring these quantities is the subject of a separate publication. In these considerations the effect of the wavelength, the direction of incidence and the effect of the membrane shape on the active parameters are taken into account. Indirectly, the manner of mounting the membrane at its border has also been considered.

1. Introduction

The equivalent diagrams of electroacoustic transducers are generally constructed in the form of analogue systems with lumped parameters. For the analysis of microphones the active parameters of the membrane are determined by means of lumped parameters, including the equivalent area and the active mass.

In this paper analytical relations are derived which permit the determination of both parameters assuming that the shape of the membrane, its surface density, the phase, displacement amplitude and velocity distributions are known.

The theoretical arguments used to obtain the relations mentioned above were based on the equivalence of the work of the acoustic field and the kinetic energy of the real membrane, assuming a flat piston. This approach is not new but, to the authors' knowledge, the equivalent area and the active mass have not been consistently defined up to the present time, according to the assumed equivalence with the simultaneous assumption that the phase distribution, the displacement amplitudes and velocities are given *a priori*. (It is assumed very frequently for example that the equivalent area of a membrane is its cross-section in a plane normal to the axis of symmetry).

2. The equivalent area of a membrane

The *equivalent area* S_e of a membrane is the area of a vibrating flat piston with an amplitude of vibration (A_m) equal to the maximum amplitude of displacement of the membrane, normal to its surface. It is of such an area that the mechanical force exerted upon it by an evenly distributed acoustic pressure field does the same work as the forces actually acting upon the membrane.

This definition differs from the definition of the equivalent area formulated by ŻYSZKOWSKI [5] in that instead of the formulation «equal to the maximum displacement of the membrane normal to its surface ...» there is the expresion «... equal to the amplitude at the centre of the membrane». This difference can be explained by the fact that the latter considers a particular case of a vibrating membrane: that for which the amplitude decreases with distance from the axis of symmetry of the membrane. In the general case which will be considered in this paper, the maximum normal displacement amplitude need not occur at the centre of the membrane.

The proposed definition uses the normal component of the displacement amplitude because, when the viscosity of the air is neglected, the work to be performed by the tangential component of the displacement of the membrane is equal to zero.

It is assumed that the shape of the fixed membrane is described by the function

$$z = f(r, \varphi), \tag{1}$$

and the orientation of any element δS of the membrane is determined by the angle of inclination $\alpha(r, \varphi)$ of the tangent at the point (r, φ, z) (Fig. 1).

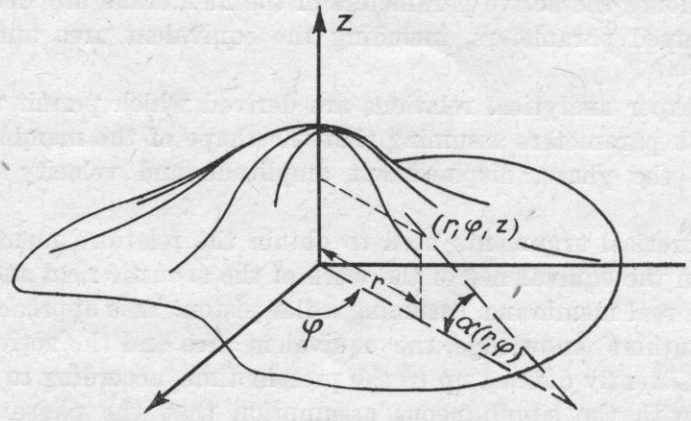

Fig. 1. The membrane whose shape at rest is described by the function $z = f(r, \varphi)$ (in cylindrical coordinates). $\alpha(r, \varphi)$ denotes the angle of inclination of the tangent to the membrane at the point (r, φ, z)

The border of the membrane will be assumed to be in the plane (r,φ) and is described by $r = R(\varphi)$ (Fig. 2). This delineates the boundary of the area defined by the function z (formula (1)).

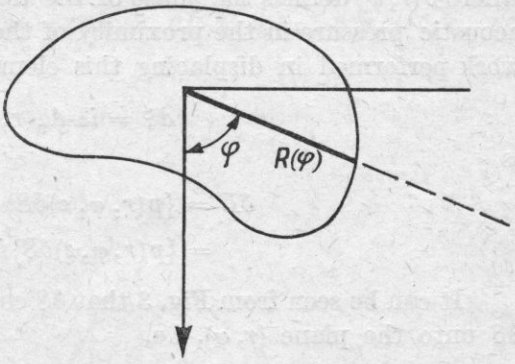

Fig. 2. The border of the membrane in the plane (r, φ) described by the function $r = R(\varphi)$

Let us calculate the work performed by the vibrating membrane. In the general case the displacement amplitude A of the element δS depends on its position: $A = A(r, \varphi)$. (The dependence on z can be neglected because of the dependence $z \leftrightarrow (r, \varphi)$ via formula (1).)

If we neglect the influence of the viscosity of the air, then the force acting on the element δS will only be related to the normal component of the deflection amplitude $A_n(r, \varphi)$ (Fig. 3).

When a wave of angular frequency ω falls onto the membrane, then the instantaneous value of the displacement of the element δS from its equilibrium

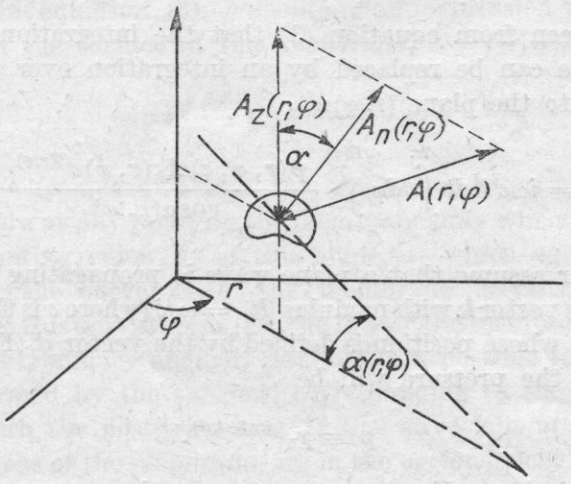

Fig. 3. Components of the displacement amplitude of an element of the membrane δS located at the point (r, φ, z), in the normal direction $A_n(r, \varphi)$ and in a direction $A_z(r, \varphi)$ parallel to the z-axis

position in a normal direction (Fig. 3) is a function of the form

$$\xi = A_n(r, \varphi) e^{i\{\omega t + \theta(r,\varphi)\}},$$

where $\theta(r, \varphi)$ defines the phase of the displacement of the element δS. If the acoustic pressure in the proximity of the element δS is denoted by p, then the work performed in displacing this element through a distance

$$d\xi = i\omega A_n(r, \varphi) e^{i\{\omega t + \theta(r,\varphi)\}}$$

is

$$\begin{aligned} \delta L &= \{p(r, \varphi, z)\delta S\} d\xi = \\ &= \{p(r, \varphi, z)\delta S\} i\omega A_n(r, \varphi) e^{i\{\omega t + \theta(r,\varphi)\}} \end{aligned} \tag{2}$$

It can be seen from Fig. 3 that $\delta S \cos \alpha(r, \varphi)$ is the projection of the surface δS onto the plane (r, φ), i.e.

$$\delta S \cos \alpha(r, \varphi) = r\, dr\, d\varphi,$$

whence

$$\delta S = \frac{r\, dr\, d\varphi}{\cos \alpha(r, \varphi)}. \tag{3}$$

The work dL done by forces acting on the whole membrane during a time dt is the sum of the components δL obtained for all the individual elements δS. Thus expression (2) must be integrated over the whole surface of the membrane S:

$$dL = i\omega e^{i\omega t} dt \iint_S p(r, \varphi, z) A_n(r, \varphi) e^{i\theta(r,\varphi)} \delta S.$$

It can be seen from equation (3) that the integration over the surface of the membrane can be replaced by an integration over the area which is its projection onto the plane (r, φ):

$$dL = i\omega e^{i\omega t} dt \int_0^{2\pi} d\varphi \int_0^{R(\varphi)} \frac{p(r, \varphi, z) A_n(r, \varphi) e^{i\theta(r,\varphi)}}{\cos \alpha(r, \varphi)} r\, dr. \tag{4}$$

If we further assume that a plane wave is propagating in a direction defined by the wave vector $\boldsymbol{k}$ with modulus $|\boldsymbol{k}| = \omega/c$ (where c is the sound velocity), then at the point whose position is defined by the vector $\boldsymbol{d}$ (Fig. 4) the instantaneous value of the pressure will be

$$p = p_0 e^{i\{\omega t + \boldsymbol{k}\boldsymbol{d}\}}. \tag{5}$$

In a cylindrical coordinate system (r, φ, z) the components of the vector $\boldsymbol{d}$ are the following:

$$d_x = r \cos \varphi, \quad d_y = r \sin \varphi, \quad d_z = z.$$

Using the notation of Fig. 4 the components of the vector $\boldsymbol{k}$ take the form

$$k_x = \frac{\omega}{c} \sin\delta \cos\gamma, \qquad k_y = \frac{\omega}{c} \sin\delta \sin\gamma, \qquad k_z = \frac{\omega}{c} \cos\delta.$$

Using the definition of the scalar product $\boldsymbol{ab} = a_x b_x + a_y b_y + a_z b_z$ we obtain

$$\boldsymbol{kd} = \frac{\omega}{c} (r \cos\varphi \sin\delta \cos\gamma + r \sin\varphi \sin\delta \sin\gamma + z \cos\delta)$$

$$= \frac{\omega}{c} \{r \sin\delta \cos(\varphi - \gamma) + z \cos\delta\}.$$

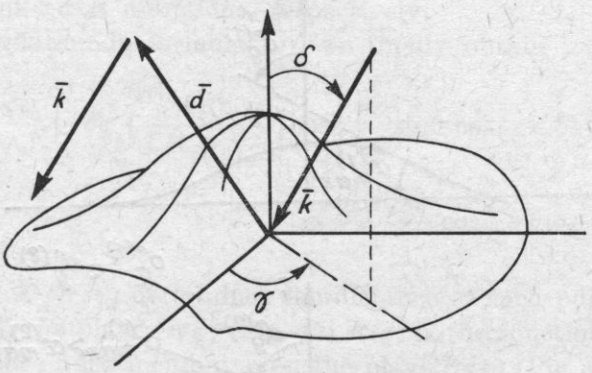

Fig. 4. The angles γ and δ describing the direction of propagation of the plane wave, i.e. the direction of the wave vector $\boldsymbol{k}$

Substituting in equation (5), we obtain an expression for the acoustic pressure acting on the surface of the membrane, $z = f(r, \varphi)$, as

$$p(r, \varphi) = p_0 e^{i\omega t} \exp i \left\{ \frac{\omega r}{c} \sin\delta \cos(\varphi - \gamma) + \frac{\omega f(r, \varphi)}{c} \cos\delta \right\}. \tag{6}$$

It can be seen from Fig. 4 that this expression gives the magnitude of the acoustic pressure at any point on the membrane only when the angle δ is not too large. The limiting value δ_g of this angle, at which equation (6) is still valid, depends on the curvature of the membrane described by the angle $\alpha(r, \varphi)$. The smaller this curvature is (i.e. the flatter the membrane), then greater is the value of the boundary angle δ_g (Fig. 5). It is possible to define δ_g as the angle which is formed by the tangent, corresponding to the maximum value of the angle α, with the positive z-axis. If the wave falls at an angle greater than δ_g, certain areas of the membrane are in the «geometrical shade». To define the value of the pressure there, it is necessary to use the complicated diffraction theory. In this paper we shall not consider such conditions, assuming that the angle of incidence is always smaller than the boundary angle δ_g.

Returning to equation (4), describing the work performed by the real membrane in a time dt, we may substitute the pressure $p(r, \varphi)$ given by equation (6) and obtain

$$dL = i\omega e^{2i\omega t} dt \int_0^{2\pi} d\varphi \int_0^{R(\varphi)} \frac{A_n(r, \varphi)}{\cos \alpha(r, \varphi)} \exp i \left\{ \omega r/c \sin\delta \cos(\varphi - \gamma) + \right.$$

$$\left. + \frac{\omega f(r, \varphi)}{c} \cos\delta + \theta(r, \varphi) \right\} r\, dr. \qquad (7)$$

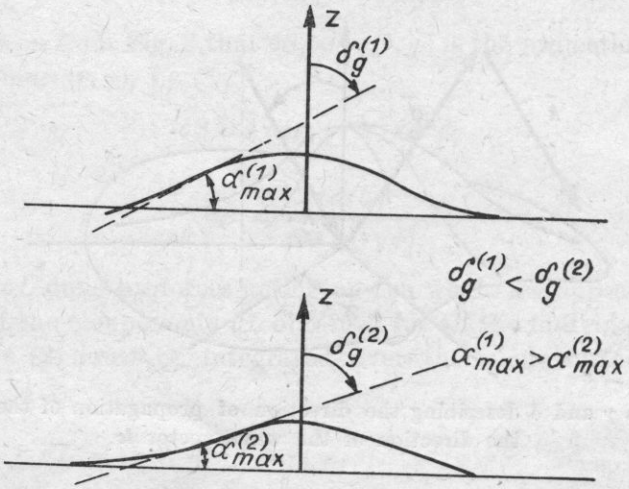

Fig. 5. The effect of the boundary angle δ_g on the maximum value of the angle of inclination of the tangent to the membrane α_{max}

In view of the definition accepted for the equivalent area, we calculate in turn the work performed by the moving flat piston. It will be assumed that this piston has an area S_c (this is the equivalent area we wish to evaluate) and vibrates with an amplitude equal to the maximum displacement amplitude A_m under the action of the incident plane wave of amplitude p_0.

The instantaneous value of the acoustic pressure on its surface is $p = p_0 e^{i\omega t}$, and the instantaneous value of the displacement is $\eta = A_m e^{i\omega t}$. The work done by the forces acting on the piston in time dt is

$$dL = \{p S_c\} d\eta = i\omega p_0 e^{2i\omega t} S_c A_m. \qquad (8)$$

According to the definition accepted for the equivalent area S_c, this work should be equal to the work performed by the forces acting on the membrane.

From (7) and (8) we obtain

$$S_c = \int_0^{2\pi} d\varphi \int_0^{R(\varphi)} \frac{A_n(r,\varphi)}{A_m \cos\alpha(r,\varphi)} \cos\left\{\frac{\omega r}{c}\sin\delta\cos(\varphi-\gamma)+\frac{\omega f(r,\varphi)}{c}\cos\delta+\right.$$

$$\left.+\theta(r,\varphi)\right\} r\,dr.$$

It can be seen from Fig. 3 that

$$A_n(r,\varphi) = A_z(r,\varphi)\cos\alpha(r,\varphi), \tag{10}$$

where $A_n(r,\varphi)$ and $A_z(r,\varphi)$ denote components of the displacement amplitude in the normal and z-axis directions, respectively.

Using this equation in formula (9), we finally obtain

$$S_c = \frac{1}{\cos\alpha(r_0,\varphi_0)} \int_0^{2\pi} d\varphi \int_0^{R(\varphi)} \frac{A_z(r,\varphi)}{A_z(r_0,\varphi_0)} \cos\left\{\frac{\omega r}{c}\sin\delta\cos(\varphi-\gamma)+\right.$$

$$\left.+\frac{\omega f(r,\varphi)}{c}\cos\delta+\theta(r,\varphi)\right\} r\,dr. \tag{11}$$

In this formula $f(r,\varphi)$ determines the distance of each point of the membrane at rest from the plane (r,φ) (Fig. 1); $R(\varphi)$ is the function describing the shape of the border of the membrane on the plane (r,φ) (Fig. 2). The angles δ and γ determine the direction of the incident wave (Fig. 4), and $A_z(r,\varphi)$ denotes the amplitude of deflection of the membrane in the direction of z-axis, measured in the neighbourhood of any point of the membrane (r,φ). $A_z(r_0,\varphi_0)$ and $\alpha(r_0,\varphi_0)$ denote, respectively, the deflection amplitude in the direction of z-axis and the angle of inclination of the element δS (Fig. 3) in the proximity of the point which is vibrating with maximum amplitude in the normal direction A_m. Thus, assuming that $A_z(r,\varphi)$ is measurable, in order to determine the coordinates of the point (r_0,φ_0) and, consequently, $A_z(r_0,\varphi_0)$ and $\alpha(r_0,\varphi_0)$, it is necessary to use equation (10), since only in this way we can find the maximum value of the normal component A_m. Equation (11) contains the quantity $\theta(r,\varphi)$ which denotes the phase of vibration of individual points of the membrane.

In order to define the equivalent area of the membrane, induced to vibrate by a plane wave of angular frequency ω, which falls onto the membrane from a direction determined by the angles (γ,δ) (Fig. 4), we must know the distribution of vibration amplitudes $A_z(r,\varphi)$ and phases $\theta(r,\varphi)$ at each point of the membrane, whose shape is described by the functions $f(r,\varphi)$ and $R(\varphi)$ (the angle $\alpha(r,\varphi)$ can be determined from $f(r,\varphi)$).

In the particular case of the membrane having axial symmetry or its being excited by a plane wave propagating along the axis of symmetry, knowledge of the phases is superfluous. It can be assumed that $\theta = 0$ for each point,

i.e. that the instant value of the displacement is

$$\xi = A(r)e^{i\omega t}.$$

If the axis of symmetry of the membrane coincides with z-axis (Fig. 6), then $\delta = 0$ and from (11) we obtain

$$S_c = \frac{2\pi}{\cos\alpha(r_0)} \int_0^R \frac{A_z(r)}{A_z(r_0)} \cos\left\{\frac{\omega f(r)}{c}\right\} r\,dr, \tag{12}$$

where R is the radius of the circle constituting the border of the membrane, $f(r)$ – the function describing its shape in the plane (r, z), and r_0 – the point at which the values of the normal component $A_n(r_0) = A_z(r_0)\cos\alpha(r_0)$ is the greatest ($\alpha(r_0)$ is the angle of inclination of the tangent at the point r_0).

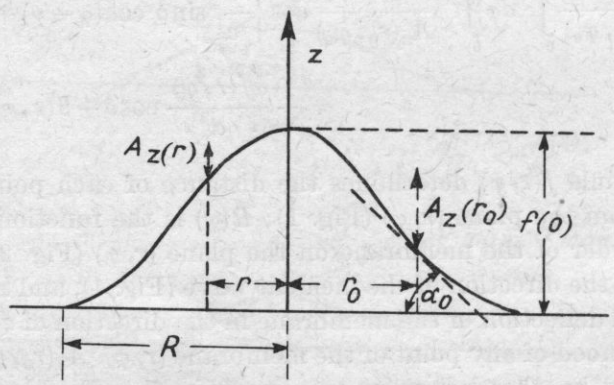

Fig. 6. The cross-section of an axially symmetrical membrane in the plane (r, z); $f(0)$ is the function describing the shape of the membrane for $r = 0$

If the wavelength λ is considerably greater than the «height» of the membrane $f(0)$ (Fig. 6), then

$$\frac{\omega}{c} f(r) = 2\pi \frac{f(r)}{\lambda} \ll 1.$$

In this case, after a series expansion of the function $\cos\{\omega f(r)/c\}$, it follows from (12) that

$$S_c \approx \frac{2\pi}{\cos\alpha(r_0)} \int_0^R \frac{A_z(r)}{A_z(r_0)} r\,dr - \frac{\pi}{\cos\alpha(r_0)} \int_0^R \left\{\frac{\omega f(r)}{c}\right\}^2 \frac{A_z(r)}{A_z(r_0)} r\,dr, \tag{13}$$

he second term being considerably smaller than the first.

If the maximum value of the amplitude occurs at the centre of the mem-

brane ($r_0 = 0$), then $\cos\alpha(r_0) = 1$ and we can make the approximation

$$S_c \approx 2\pi \int_0^R \frac{A_z(r)}{A_z(0)}\, r\, dr. \tag{14}$$

As was stated at the beginning, it is very often assummed that the equivalent area of the membrane S_c is equal to the area of its projection onto the plane (r, φ), i.e. $S_c = \pi R^2$ (Fig. 6). From the relations which have been derived from the definition of the equivalent area, on the assumption of the equivalence of the work performed by the real membrane and a flat piston, it can be seen that the equivalent surface cannot be equal to the projected one. For instance, in considering equation (13) we see that $S_c = \pi R^2$ only when the membrane is flat ($f(r) = 0$), and each point is vibrating with the same amplitude $A(r) =$ const. It is obvious that the last condition cannot be satisfied because of the necessity of mounting the membrane by its edges.

3. The active mass of the membrane

It is assumed that the active mass m_c of the membrane is that mass which, when vibrating with a velocity amplitude equal to the maximum velocity amplitude of the membrane v_m, has a kinetic energy equal to the total kinetic energy of the membrane.

Similarly as in the case of the equivalent area, this definition differs from the definition stated in the monograph of ŻYSZKOWSKI [4] in that it does not give the centre of the membrane the value v_m because in the general case the maximum velocity amplitude can be recorded at any point.

If the instantaneous value of the velocity of vibration of the element δS is

$$v = v(r, \varphi)\, e^{i\{\omega t + \Psi(r,\varphi)\}},$$

then the time average of the kinetic energy is equal to

$$\delta E_k = \frac{1}{4}\varrho(r, \varphi)\delta S v^2(r, \varphi),$$

where $\varrho(r, \varphi)$ is the surface density of the membrane.

Summing over the whole surface S, we obtain the total kinetic energy of the membrane:

$$E_k = \frac{1}{4}\iint_S \varrho(r, \varphi) v^2(r, \varphi)\delta S.$$

Using equation (3) we subsequently obtain

$$E_k = \frac{1}{4}\int_0^{2\pi} d\varphi \int_0^{R(\varphi)} \frac{\varrho(r, \varphi) v^2(r, \varphi)}{\cos\alpha(r, \varphi)}\, r\, dr. \tag{15}$$

Let us assume that the velocity amplitude $v(r_0, \varphi_0)$ at the point $(r_0, {}_0\varphi)$ is the maximum. According to the definition of the active mass m_c, for the mean kinetic energy we have

$$E_k = \frac{1}{4} m_c v^2(r_0, \varphi_0). \tag{16}$$

From equations (15) and (16) we get

$$m_c = \int_0^{2\pi} d\varphi \int_0^{R(\varphi)} \frac{\varrho(r, \varphi) v^2(r, \varphi)}{\cos \alpha(r, \varphi) v^2(r_0, \varphi_0)} r dr. \tag{17}$$

In this expression $\varrho(r, \varphi)$ denotes the surface density of the membrane which, in general, may change from point to point; $R(\varphi)$ is the function describing the shape of the border of the membrane (Fig. 2), while $v(r_0, \varphi_0)$ and $v(r, \varphi)$ denote the maximum velocity amplitude and the velocity amplitude of the membrane vibrations, measured in the proximity of any point (r, φ). The methods of measuring these quantities and the displacement amplitude will be presented in a subsequent paper.

For the particular case of axial symmetry and a homogeneous membrane (Fig. 6) we have from equation (17),

$$m_c = 2\pi\varrho \int_0^R \frac{v^2(r)}{\cos \alpha(r) v^2(r_0)} r dr. \tag{18}$$

If the angle of inclination of the tangent to the surface of the membrane $\alpha(r)$ is not too great (i.e. membrane is almost flat), then using the series expansion

$$\frac{1}{\cos \alpha(r)} = 1 + \frac{1}{2} \alpha^2(r) + \ldots$$

the right-hand side of equation (18) can be expanded as a series

$$m_c = 2\pi\varrho \int_0^R \frac{v^2(r)}{v^2(r_0)} r dr + \pi\varrho \int_0^R \alpha^2(r) \frac{v^2(r)}{v^2(r_0)} r dr + \ldots$$

It can be seen that in this case it is not necessary to know the geometry of the membrane (the function $f(r)$) in order to define the active mass m_c, since the succeeding terms of the series are very small in comparison with the first.

Similarly as in the case of the equivalent area, we can see from the foregoing equations that the active mass m_c can be equal to the mass of the membrane only when the whole surface of the membrane is vibrating with the same amplitude ($v(r, \varphi) = \text{const}$). Such a case is excluded because of the necessity of mounting the membrane by its border.

4. The equivalent area in a diffuse field

Formula (11), defining the equivalent area of a membrane in a plane wave field, is based on the assumption that the wave falls at an angle δ (Fig. 4) which is smaller than the boundary angle δ_g(Fig. 5). This formula is in agreement with the definition of the equivalent area S_g. In this section a formula describing the active surface of a membrane in a diffuse field will be derived. It differs somewhat from the definition of S_c accepted above as regards both its form and its generality (compared with the formula (11)), and is therefore derived separately.

From the assumed diffusivity of the field, i.e. the probability that a plane wave is incident on any element of the membrane δS in a direction determined by the angles (γ, δ) (Fig. 4), remaining at the same time within the solid angle

$$d\Omega = \sin\delta d\delta d\gamma,$$

is equal to

$$dp(\gamma, \delta) = \frac{d\Omega}{2\pi} = \frac{\sin\delta d\delta d\gamma}{2\pi}. \tag{19}$$

In this case 2π is the value of the solid angle which fills the half-space $z > 0$.

It can be assumed that the equivalent area of the membrane S_c within the field of the plane wave whose direction of propagation is determined by the angles (γ, δ) is a random variable of these two quantities. This means that the mean value S_c (the expected value) at any moment is equal to

$$\bar{S}_c = \iint_{2\pi} dp(\gamma, \delta) S_c(\gamma, \delta).$$

Substituting for S_c and dp from formulae (11) and (19), and subsequently integrating with respect to γ and δ so as to cover the whole of the solid angle 2π, we obtain

$$S_c = \frac{1}{\cos\alpha(r_0, \varphi_0)} \int_0^{2\pi} d\varphi \int_0^{R(\varphi)} \frac{A_z(r, \varphi)}{A_z(r_0, \varphi_0)} H(r, \varphi) r dr, \tag{20}$$

where

$$H(r, \varphi) = \int_0^{\pi/2} \sin\delta d\delta \int_0^{2\pi} \cos\left\{\frac{\omega r}{c} \sin\delta \cos(\varphi - \gamma) + \frac{\omega f(r, \varphi)}{c} \cos\delta + \theta(r, \varphi)\right\} d\gamma. \tag{21}$$

In assuming the limits of integration for δ(0, to $\pi/2$), and for γ(0 to 2π), we introduce some error since waves from all possible directions can reach only the elements δS located on the top of the membrane. For any other location there exist values of the angles γ and δ for which the corresponding

plane waves do not reach the element δS. For example, points in the «geometrical shade» (Fig. 5) have waves corresponding to values of the angle δ greater than δ_g.

In order to have a convenient form of the formula determining the mean value of the active surface S_c, we assume the integration limits of equation (21). It is, however, to note that this formula is valid only if the membrane is not too convex, i.e. if $\delta_g \to \pi/2$ (Fig. 5).

It we subsequently assume that the membrane has the axial symmetry and that its border is a circle of radius R lying in the plane (r, φ) (Fig. 6), we may expect that in a diffuse field the phase of vibration θ of all points of the membrane will be equal (assuming, as in section 2, that $\theta = 0$).

The integration relative to γ will be carried out using the identity $\cos\alpha = \frac{1}{2}(e^{i\alpha} + e^{-i\alpha})$:

$$\int_0^{2\pi} \exp\left\{\pm i\frac{\omega r}{c}\sin\delta\cos(\gamma-\varphi)\right\}d\gamma = 2\int_0^{\pi} \exp\left\{\pm i\frac{\omega r}{c}\cos\gamma\right\}d\gamma.$$

According to WHITTAKER [2]

$$\int_0^{\pi} e^{\pm iz\cos\gamma}d\gamma = \pi J_0(z),$$

where $J_0(z)$ is a zero-order Bessel function. Using these equalities we obtain

$$\int_0^{2\pi} \cos\left\{\frac{\omega r}{c}\sin\delta\cos(\varphi-\gamma)+\frac{\omega f(r)}{c}\cos\delta\right\} = 2\pi\cos\left\{\frac{\omega f(r)}{c}\cos\delta\right\}J_0\left\{\frac{\omega r}{c}\sin\delta\right\}.$$

Thus, the function H in (21) (with $\theta = 0$ and the axial symmetry assumed for the membrane) has the form

$$H(r) = 2\pi\int_0^{\pi/2} \sin\delta\cos\left\{\frac{\omega f(r)}{c}\cos\delta\right\}J_0\left(\frac{\omega r}{c}\sin\delta\right)d\delta.$$

Substituting $\cos\delta = x$, we obtain

$$H(r) = 2\pi\int_0^1 \cos\left\{\frac{\omega f(r)}{c}x\right\}J_0\left(\frac{\omega r}{c}\sqrt{1-x^2}\right)dx.$$

Using integral transformation tables [1] we finally obtain

$$H(r) = 2\pi\frac{\sin\left\{\frac{\omega}{c}\sqrt{r^2+f^2(r)}\right\}}{\frac{\omega}{c}\sqrt{r^2+f^2(r)}}.$$

Substituting this function into equation (20) we obtain an expression for the mean value of the equivalent area of an axially symmetrical membrane in a diffuse field:

$$\bar{S}_c = \frac{4\pi^2}{\cos\alpha(r_0)} \int_0^R \frac{A_z(r)}{A_z(r_0)} \frac{\sin\left\{\frac{\omega}{c}\sqrt{r^2+f^2(r)}\right\}}{\frac{\omega}{c}\sqrt{r^2+f^2(r)}}. \tag{22}$$

The notation in this formula is the same as in formula (12), which describes the equivalent area of an axially symmetrical membrane induced to vibrate by a plane wave.

As has already been stated, formula (22) is valid only if the membrane is not too convex, i.e. if $f(r) \to 0$ ($\delta_g \to \pi/2$). Thus in the conclusion we shall refer only to the formulae derived in sections 2 and 3.

5. Conclusions

The method proposed in this paper for the determination of active parameters, together with the procedure of measuring the vibrations of a membrane in a plane wave field, give a precise determination (according to the definition) of the equivalent area and the active mass of microphone membranes. Using the method of trial and error we can, for instance, construct membranes of various shapes. By measuring their active parameters we can also find membranes which possess such values of the parameters as are required by the designer. An asset of the method of describing the equivalent area (section 2) and the active mass (section 3), as proposed in this paper, is that it is based consistently on the equivalence of work and energy for the real membrane and the equivalent piston. It results from the relations obtained that the active parameters depend on the shape of the membrane and the quantities describing the acoustic field. Some of the previous formulae describing the equivalent area and the active mass do not consider the real shape of the membrane (for instance, by assuming that the equivalent area is equal to the projection of the surface). These formulae also do not contain quantilics describing the acoustic field (e.g. the wavelength and the angle of incidence), which are taken into consideration in the method described in this paper.

In many papers devoted to the problem of equivalent diagrams, attempts have been made to introduce into the definitions of the active parameters the velocity and displacement amplitudes which are the solutions of the correspondingly formulated boundary problems. This trend in the investigations constitutes the next step, after the method of trail and error (to which we have already referred), on the way to the optimal design of microphone membranes. The results of these investigations would be analytical relations between the active

parameters and the quantities describing the material properties of the membrane, the manner of mounting, etc. Due to the complexity of the problem, the determination of the distribution of the velocity and displacement amplitudes of a real microphone membrane, excited by an acoustic wave, is possible only with substantial simplification, thus considerably reducing the validity of the results. In this situation there is only one approach to finding the relations between the active parameters and the quantities describing the material properties of the membrane, the manner of mounting, etc.: by experimental investigations. These relations can be found with the aid of the method of trial and error provided that the algorithm defining the active parameters is based consistently on their energetic definition. This condition is satisfied by the method proposed in this paper.

The authors of this paper are very thankful to Dr. E. HOJAN, Mr. M. NIEWIAROWICZ (M. Sc. Eng.) and Mr. J. FLORKOWSKI (M. Sc.) for discussion and helpful remarks.

References

[1] H. BATEMAN, *Tables of integral transforms*, Vol. I., N. Y., Mc Graw-Hill, 1954.

[2] E. T. WHITTAKER, G. N. WATSON, *A course of modern analysis*, University Press, Cambridge 1927.

[3] S. ZIMMERMANN, M. BRUNEAU, A. M. BRUNEAU, *Sur les modèles des haut-parleurs à diaphragme conique*, Acustica **35**, 2, 127–131 (1976).

[4] Zb. ŻYSZKOWSKI, *Fundamentals of electro-acoustics* [in Polish], WNT (Scientific and Technical Publishers), Warszawa 1966, p. 183–188.

Received on 18th December, 1976. Final version Revised 20th April 1977.

ARCHIVES OF ACOUSTICS
2, 4, 281 (1977)

DETECTABILITY OF PULSE DISTORTION IN AN ACOUSTIC LOUDSPEAKER FIELD

EDWARD HOJAN

Department of Acoustics, A. Mickiewicz University (Poznań)

This paper presents results of the evaluation of the distortion of transient signals produced by a set of loudspeakers. They were obtained on the basis of certain objective signal parameters and subjective tests. The separate objective parameters provided complex information on the degree of signal distortion and were correlated with the psychoacoustic ability of the sense of hearing to detect distortions. These parameters were related to the rising transients with complex spectra and to the decay transients with simple spectra.

1. Introduction

It was shown [7] that the behaviour of differential thresholds in time depends very little on the duration of transient for $t \geqslant 1$ s, while for shorter durations the threshold values increase. This is true for both rising and decaying transients of tone pulses.

The differential thresholds are independent of changes both in the tone pulse intensity, over a range from 40 to 100 phons, and for component frequencies in the pulsed signals above 400 Hz. From the detailed analysis of the psychoacoustic evaluation of differential thresholds of pulse tones with transient duration smaller than 1 s, it can be concluded that the organ of hearing detects differences in the transient durations primarily on the criterion of the rate of rise or decay of the signal intensity. A secondary criterion is the detection of differences in the intensity of audible signals. Similar results were obtained [7, 11] for complex tone and noise pulses.

Another important factor affecting the psychoacoustic evaluation of sound pulses is the shape of their envelope. Investigations of sound pulses with rising and decaying transients of the same duration but with different envelopes [2, 3, 4], provided initial information on the ability of the organ of hearing to detect, changes in the signal envelope. It was shown [7] that the differentiation of sound pulses with rising transients longer than 10 ms is related to the detectability of differences in the rising transients. For tran-

sients shorter than 10 ms, the detectability is associated with differences in the signal timbre. Similar experiments were performed for sound pulses with a "distorted" non-exponential envelope for the decay transient The decay signal transient should be either extended or shortened (at a level 40 phons below the maximum) by 10 to 20% if these changes are to be detectable. There must be a time interval 70 to 300 ms [3, 4] between two successive maxima in the decaying transient envelope described by $\sin x/x$, in order to make them detectable. The amplitudes of these maxima are then 1/2 and 1/3 of their values in the steady state. The longer this time interval is, the higher the detecting ability becomes, with a peak detectability for a 250 ms time interval between the maximum values. In the author's opinion, this is related to the occurrence of a minimum in the auditory threshold of amplitude modulation (the modulation frequency being 2 Hz).

The results presented in this paper show that the transients play an important role in the psychoacoustic evaluation of sound pulses. They also permit the determination, in a general way, of the transient parameters (duration, envelope shape which are of vital importance in this evaluation.

2. The purpose of the paper

Within the framework of research on the dependence of the detectability of changes in sound pulse forms on their objective parameters, we have made an attempt to determine the parameters that would be capable of providing information on the degres of distortion of the entire signal or of selected parts. The primary selection of these parameters was aimed at obtaining detailed information on the role of the rising and decaying transients in the psychoacoustic evaluation of signals taken as a whole.

The problems, which relate the psychoacoustic evaluation of sound pulses with the objective parameters of the pulses, constitute a group of problems dealing in general with the subjective evaluation of complex acoustic fields in which the effects of diffraction, interference, and amplitude and phase nonuniformities produce major distortions in the acoustic signals.

We describe here the results of an attempt to determine the signal parameters by evaluating the acoustic field of a set of loudspeakers[1]) (2 broadband loudspeakers, 1 high-pitch loudspeaker enclosed in a housing 0.250 m³). We have objectively determined the degree of distortion of the sound pulses in both transient and steady states at particular points in the field and analysed the psychoacoustic detection of these distortions. From these results conclusions have been drawn on the interrelation between the psychoacoustic evaluation and the objective parameters of the signals under consideration.

---

[1]) A set of loudspeakers produced by firm Klein-Hummel, type Studio-Regielautsprecher 02.

3. Measurements

3.1. The measuring system

The schematic diagram of the measuring system employed for the assumed investigation program is shown in Fig. 1.

The set of loudspeakers was excited by various acoustic generators in accordance to a selected measuring method. The signal from the loudspeakers was received simultaneously by two microphones, M-2 being at a fixed point,

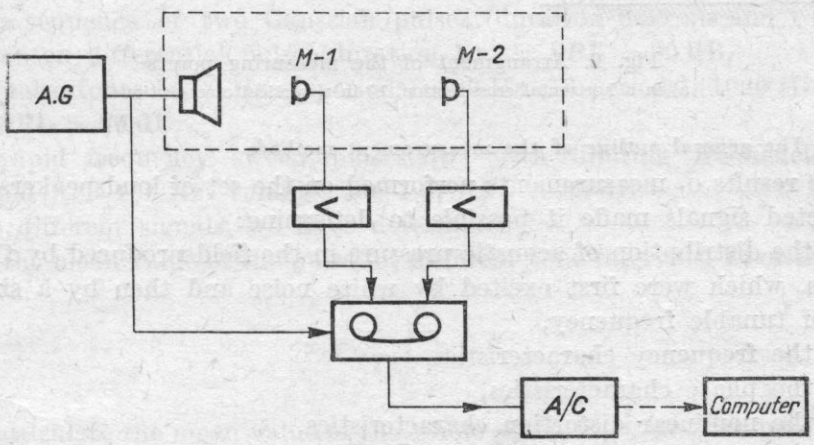

Fig. 1. Block diagram of the measuring apparatus
AG – acoustic generator, *M*-1, *M*-2 – measuring microphones, *AC* – analog-to-digital converter, *COMPUT* – IBM computer

while M-1 was moved from one place to another as the measurements were in progress. The signals received by the microphones were amplified and fed to the recording systems. In the case of pulse methods, these signals were first recorded on magnetic tape and transformed by means of an A/C converter. The time–response of the loudspeaker of a given pulse excitation was obtained on paper tape at the output of punch unit.

3.2. Measuring conditions

All tests were performed in an anechoic chamber. Fig. 2 shows the arrangement of the measuring points and the location of the different loudspeakers in the set. The measuring points were located along three axes. Two axes *a* and *c* were coincident with the axes of the loudspeakers, while the third *b* axis was symmetrical to the axes of the two loudpseakers. The signals which excited the loudspeakers were selected in accordance with the programme of investigation.

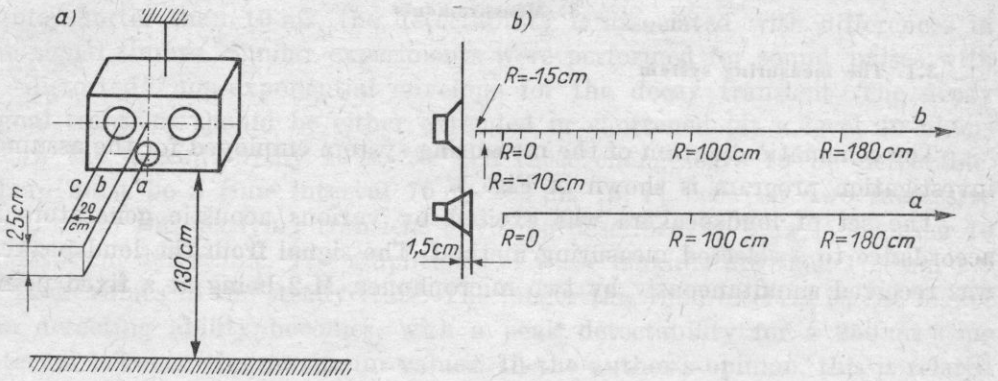

Fig. 2. Arrangment of the measuring points
a) in a horizontal cross-section, b) in a vertical cross-section

3.3. The general outline of the measurement methods

The results o measurements performed on the set of loudspeakers excited by selected signals made it possible to determine:

— the distribution of acoustic pressure in the field produced by the loudspeakers, which were first excited by white noise and then by a sine-wave signal of tunable frequency,

— the frequency characteristics,

— the phase characteristics,

— the nonlinear distortion characteristics,

— the directional characteristics.

Since measuring methods for the above quantities are well known, they will not be described in this paper.

The reaction of the loudspeakers excited by pulse signals were recorded on the basis of the following three signals:

signal 0 — obtained from a generator and recorded by a purely electrical method,

signal 1 — (acoustic) received by microphone M-1 located at different points of the acoustic field,

signal 2 — (acoustic) received by microphone M-2 located at a fixed measuring point of the acoustic field (a, $R = 180$ cm). The magnetic tape recordings of the reaction of the loudspeaker to a pulse excitation were used to prepare psychoacoustic tests. A series of measurements were made for a single location of microphones in order to eliminate possible inaccuracy related to the positions of the measuring points.

3.3.1. Methods employed for the unsteady state

From the theoretical point of view, the maximum information on a transmission system is obtained by employing a Dirac delta pulse signal [8]. However such a experiment is very difficult to perform if an acoustic pressure level

$SPL = 70$ to 90 dB is required at a distance 1-2 m from the loudspeaker plane. For this reason, experimenters select pulsed signals of a finite duration strictly related to the frequency under test [10]. The other parameters of the exciting pulse such as rise and decay times, and shape and amplitude of the envelope, are selected in accordance with the testing programme of the transmitting system [5]. The following pulses were used in these experiments:

– a pulse with a rectangular envelope, of durations 1 ms or 0.25 ms; providing acoustic pressure levels $SPL = 70$ and $SPL = 90$ dB at a distance of 1.80 m from the loudspeaker plane,

– pulses of Gaussian envelope, durations 1 ms (rise and decay times 0.3 ms), and 1.5 ms (rise and decay times 0.5 ms); $SPL = 80$ dB.,

– a sequence of two Gaussian pulses (duration 0.25 ms and 1 ms) and one Thomson differential pulse (duration 1 ms); $SPL = 80$ dB,

– pulse tones at 80 Hz (3 periods), 800 Hz (16 periods), 1020 Hz (16 periods); SPL = 80 dB,

– rapid frequency sweep pulses [6] with limiting frequencies $f_1 = 30$ Hz and $f_2 = 150$ Hz; tuning time 1 s; $SPL = 80$ dB.

For the different signals, we have determined:

– the mean value of the pulse for different time intervals, T, of the signal

$$\bar{x} = \frac{1}{T}\int_0^T x(t)\,dt.$$

To calculate the mean value of the whole signal $\bar{x}_{11}$[2]), we took into account all the data, while for the rising transient $\bar{x}_{01}$ only the data from 0 to N was used, the mean value being determined for a given kind of pulse recorded at a the measuring point a, $R = 180$; and being related to the maximum walue of the amplitude after the rising transient has heen completed. The mean value of the decay transient $\bar{x}_{13}$ was determined for the range from N to Z, where Z was measured at the point a, $R = 180$ and related to the background level. The mean values of the transients of the signals $\bar{x}_{22}$, $\bar{x}_{02}$, $\bar{x}_{23}$ recorded at the remaining measuring points were calculated for unchanged values of N and Z for a givan type of pulse. With this approach, it was possible to compare the corresponding mean values of signals recorded at different measuring points by determining the differences $\Delta\bar{x}$.

The standard deviation S from the mean value within the time intervals mentioned above is given by:

$$S = \sqrt{\frac{1}{T}\int_0^T [x(t)-\bar{x}]^2\,dt}.$$

[2]) Indices used for signal parameters are described in subsection 4.2.

The standard deviation S was assumed to be a measure of the variation of the process. The values of S are equal to zero only if all the numerical data describing the time interval of a pulse are identical. The higher the value of S (which is always positive), the greater is the dispersion of the numerical data describing a given signal. As for the mean values, the differences ΔS for the standard deviations were also determined.

The coefficient of skewness Q for time interval T is given by:

$$Q = \frac{1}{T} \frac{\int_0^T [x(t) - \bar{x}]^3 dt}{S^3}.$$

The coefficient of skewness of a distribution Q provides information on the shift of the distribution with respect to the zero-axis. If the greater part of the distribution lies on the positive side of the axis, then Q is positive, since the weighted sum of the third powers of large positive deviations is greater than the sum of the third powers of negative deviations. In this case, the skewness is positive. A negative skewness corresponds to such a distribution that the greater part of the process lies below the axis. For the symmetrical distribution, the sum of the positive third powers is equal to the sum the negative third powers and thus $Q = 0$. ΔQ was also determined for this parameter.

By direct comparison of signals 1 and 2, we have determined their correlation coefficient ϱ,

$$\varrho_{12} = \frac{\mathrm{cov}_{12}}{S_1 \cdot S_2},$$

where cov_{12} is a covariance

$$\mathrm{cov}_{12} = \frac{1}{n+1} \sum_{i=0}^{n} [x_i(t) - \bar{x}_1] \cdot [x_i(t) - \bar{x}_2]$$

and S_1, S_2 are the standard deviations.

The correlation coefficient is positive if the two characteristics either decrease or increase at the same time, and is negative if one characteristic increases while the other decreases. If the characteristics are independent of each other, the correlation coefficient is zero. This coefficient takes the value -1 or $+1$, if the two characteristics are interrelated in the form of a simple linear function. If a number od values of x_2 (t) correspond to a fixed characteristic $x_1(t)$, the correlation coefficient is either a fraction or zero. The correlation coefficient was determined here for the time intervals mentioned above of the signals to be compared.

3.3.2. The methods employed for psychoacoustic tests

As has already been mentioned in section 2, the psychoacoustic tests were aimed at determining the relation between the changes in particulas objective parameters characterizing the degree of distortion of the pulses recorded at different points of the acoustic field, and the ability of the human sense of hearing to detect these changes.

The psychoacoustic test involved successive pairs of pulses which were always repeated three times. The observer was asked to listen to a pulse pair repeated three times and to give an answer to the question: Are the signals of the pulse pairs identical? The answer could be either «yes» or «no». If the observer hesitated he should answer «no». Fig. 3 presents the time diagram

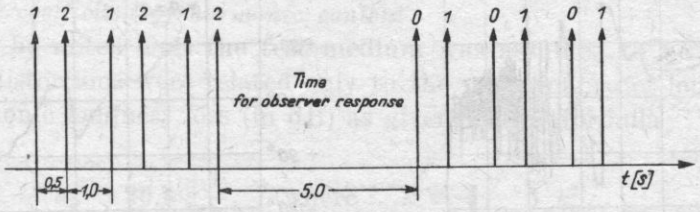

Fig. 3. Succession of acoustic signals in a psychoacoustic evaluation test with a 5 s break for the observes answer

0 – acoustic signal from generator, electrically recorded on magnetic tape, 1 – acoustic signal received by microphone M-1 positioned at successive points of the loudspeaker radiation field, 2 – acoustic signal received by microphone M-2 positioned at a fixed measuring point (a – axis, distance between microphone and loudspeaker R = 180 cm)

of the test. The succession of signals in a pair, selection of signals and signal reception points were determined at random. For the set of loudspeakers, the tape recordings were taken at the five points marked in Fig. 2 (c, $R = -15$; b, $R = -15$; a, $R = 0$; b, $R = 100$; c, $R = 180$) and this provides $2 \times 5 \times 10$ signal pairs. These signals were heard by three normally hearing observers over 10 successive days using dynamic earphones type HD 414 produced by the firm of Sennheiser.

4. The results of the investigations

The changes in the sound signal structure in the acoustic field produced by a set of loudspeakers were evaluated with loudspeakers excited by both continuous and pulsed signals.

4.1. The continuous wave excitation

4.1.1. Amplitude and phase characteristics

The amplitude and phase characteristics were measured at different points of the acoustic field produced by the set of loudspeakers (Fig. 4). It was found that these characteristics depend significantly on the position of the measuring point.

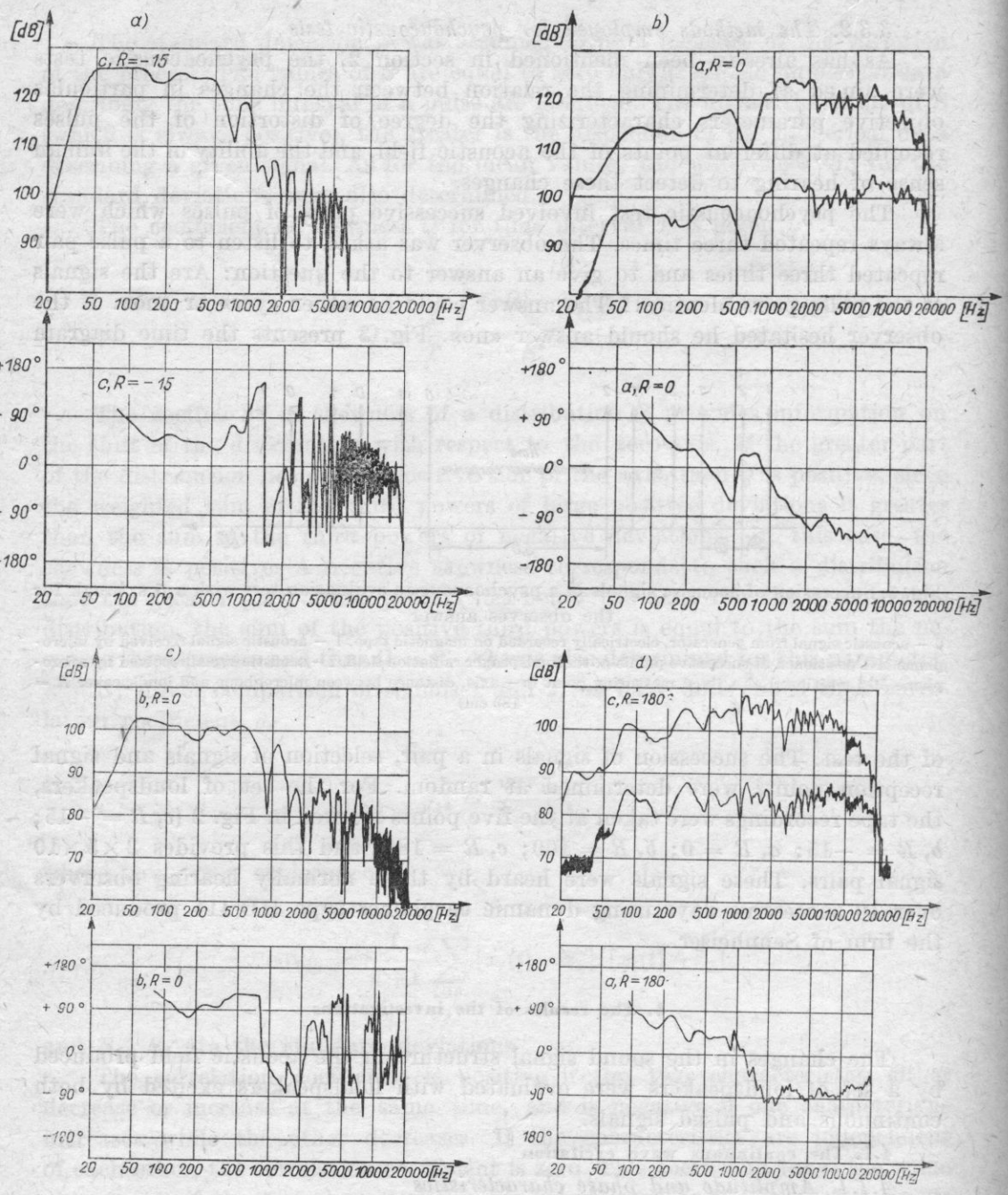

Fig. 4. Amplitude and phase characteristics of the tested set of loudspeakers. Measuring points were located as follows

a) c-axis, distance $R = -15$ cm; b) a-axis, $R = 0$ cm; c) b-axis, $R = 0$ cm; d) c-axis, $R = 180$ cm
The diagrams 4b and 4d represent amplitude characteristics for two intensities corresponding to acoustic pressure levels of $SPL = 70$ dB (the lower curve) and $SPL = 90$ dB (the upper curve) at the point a, $R = 180$; the remaining curves (Figs 4a, 4c) correspond to the case $SPL = 90$ dB

Differences occur not only between the near and far fields but also between axes at a fixed distance from the measuring points and the front plane of the set of loudspeakers.

The selected amplitude and phase characteristics shown in Fig. 4 indicate that there are major differences in both the band width and nonuniformity of the transfer function. It is known that the nonuniformity of the transfer function results in amplitude distortions. Its effect on frequencies is obvious. Nonuniformity also causes major distortions in the wave envelope (distortion of the pulse shape [9]). It follows from analysis of the phase characteristics that they are nonlinear. Thus different frequency components are subject to different attenuation and this results in additional distortion of the signal

4.1.2. The coefficient of harmonic content

It should be noted that the test medium was assumed to be linear and all nonlinear distortions were related only to the performance of loudspeakers.

The harmonic loudness loss (in dB) as given by the formula

$$a_k = 20 \log \frac{1}{k}, \quad \text{where} \quad k = \sqrt{\sum_{n=2}^{N} k_n^2}, \tag{1}$$

was measured at points located on different axes of the set of loudspeakers and at different distance from its front plane. In (1), k is the coefficient of harmonic content (on a linear scale) and n is the degree of distortion. From the data obtained we have

$$a_k = \begin{cases} z(L) & \text{with} \quad f = \text{const}, \\ g(f) & \text{with} \quad L = \text{const}, \end{cases} \tag{2}$$

where f is the signal frequency, and L is the acoustic presure level for a measuring point located on the axis of the set loudspeakers (a, $R = 180$ cm).

We also have the relations of the type

$$a_k = \begin{cases} S(R), & f = \text{const}, \quad L = \text{const}, \end{cases} \tag{3a}$$

$$\begin{cases} j(r), & f = \text{const}, \quad L = \text{const}, \end{cases} \tag{3b}$$

where R is the distance between the measuring point and the front plane of the set of loudspeakers, and r is the distance between the measuring point and the axis of the set of loudspeakers.

It is known in general how the harmonic loudness loss depends on the acoustic pressure level and signal frequency. The dependence of the quantity a_k, and the parameters related to the position of the measuring point within the radiation field of the set of loudspeakers, will be discussed.

Specimen results of the measurement of a_k as a function of R (Table 1) indicate that the relation described by (3a) is appropriate, especially at higher frequencies.

Table 1. Figures of nonlinear distortion of the 3rd order coefficient as a function of the frequency of the loudspeaker excitation signal and the position of the measuring points

SPL [dB]	*f* [Hz]	Axis	*R* [cm]	a_{k3} [dB]
90	80	*C*	180	48
			10	45
			−10	46
			−15	45
	1*k*	*C*	180	50
			10	43
			−10	48
			−15	47
	5*k*	*C*	180	55
			10	50
			−10	> 70
			−15	> 70

To interprete the processes described by expression (3), it is possible to utilize the factors affecting the forms of amplitude and phase characteristics.

The recorded differences in the amplitude and phase characteristics are reflected in the changes of the coefficient of harmonic content in view of the related signal distortions. In the author's opinion, numerous problems arise in the determination of the relationship between the psychoacoustical detectability of changes in the form of the acoustic processes and the objective parameters related to the type of signals in a piece of music or speech. These problems involve testing material, its description by means of objective methods and the ability of the observers to memorize the compared test samples. For this reason, pulse methods were used for the tests described, which are regarded as a first step to the whole problem under consideration. The set of loudspeakers was excited by pulses with different envelope forms, different times of rise and decay, and different spectra. It was intended to use the signals most frequently encountered in practice, and to perform a detailed analysis of the selected pulse parameters by objective methods.

4.2. Pulse signal excitation

In subsection 3.3.1 the detailed information is provided on the pulse signals used in the tests. The diagrams given in Fig. 5 illustrate the responses of the set of loudspeakers for various pulse excitations recorded at particular points in the acoustic field. These responses were processed in the A/C converter and in the computer. All the calculations described in subsection 3.3.1 were carried out each of the responses.

Specimen computer results for the sound pulses which were then used for the psychoacoustic tests are given in Table 2. The three groups of columns

Table 2. Values of the calculated parameters and the psychoacoustic evaluation of pulse signals from the set of loudspeakers

Kind of pulses	Axis	R	Data of total signal				Data of rising signal transient				Data of decaying signal transient				Psychoacoustic evaluation Answer «yes» in %
			$\bar{x}_{11}/\bar{x}_{22}$	S_{11}/S_{22}	Q_{11}/Q_{22}	ϱ	$\bar{x}_{01}/\bar{x}_{02}$	S_{01}/S_{02}	Q_{01}/Q_{02}	ϱ_{01}	$\bar{x}_{13}/\bar{x}_{23}$	S_{13}/S_{23}	Q_{13}/Q_{23}	ϱ_{02}	
Sequence of pulses	*a*	180	510	108	−0.88	—	518.0	194	−0.71	—	485	54	−2.5	—	—
	b	−15	511	100	−1.1	0.85	530	185	−94	0.87	481	26	−0.08	0.62	60
	b	−15	507	160	−0.55	0.44	596	215	0.17	0.54	490	93	0.33	0.08	20
	d	0	510	122	−0.20	0.78	540	221	−0.48	0.81	470	37	0.82	0.54	20
	b	100	509	97.6	−0.67	0.91	526	181	−0.64	0.94	488	35	−0.37	0.68	100
	c	180	512	102	−0.21	0.92	533	187	−0.43	0.94	490	46	−0.82	0.74	90
Gauss shape	*a*	180	509	66.8	−1.35	—	628	103.4	0.41	—	492	105	−0.97	—	—
	c	−15	510	97.6	−0.44	0.64	571	69.2	0.97	0.92	510	158.9	−0.30	0.67	90
	b	−15	509	152.2	0.09	0.43	532	18.4	0.72	0.94	547	201.5	−0.01	0.62	0
	a	0	511	85.4	−1.07	0.87	643	90.8	−0.01	0.63	506	132.7	−0.81	0.86	30
	b	100	510	68.8	−1.57	0.93	653	110.4	0.17	0.95	501	105.7	−1.29	0.93	90
	c	180	509	62.8	−1.33	0.95	631	97.5	0.26	0.95	500	96.6	−1.03	0.95	90
Tone pulse (80 Hz)	*a*	180	523	118.6	0.37	—	734	112.7	−0.47	—	501	152	0.25	—	—
	c	−15	524	81.9	2.18	0.83	730	109	−0.49	0.98	514	86.2	1.31	0.81	90
	b	−15	526	120.6	0.24	0.62	716	102.8	−0.46	0.98	521	158.5	−0.03	0.52	60
	a	0	522	89.8	0.90	0.76	712	100.4	−0.49	0.98	513	109.0	0.11	0.70	40
	b	100	523	106.9	1.20	0.88	746	119.1	−0.49	0.98	511	128.4	0.71	0.88	80
	c	180	522	106.4	0.53	0.95	719	102.2	−0.49	0.98	506	131.6	0.46	0.97	100
Tone pulse (800 Hz)	*a*	180	513	21.2	4.75	—	646	105.8	−0.58	—	518	45	0.36	—	—
	c	−15	514	23.8	4.90	0.58	571	36.6	−0.82	0.85	536	60.3	1.31	0.58	—
	b	−15	513	58.2	1.24	0.48	617	79.5	−0.63	0.85	518	141.2	0.86	0.52	90
	a	0	511	32.2	2.76	0.72	705	146.5	−0.47	0.85	516	77.0	−0.12	0.61	0
	b	100	514	23.5	5.73	0.52	648	132.1	−0.34	0.83	523	51.6	1.77	0.34	90
	c	180	514	25.8	−2.63	0.45	644	111.9	−0.45	0.85	525	46.0	1.17	0.31	90

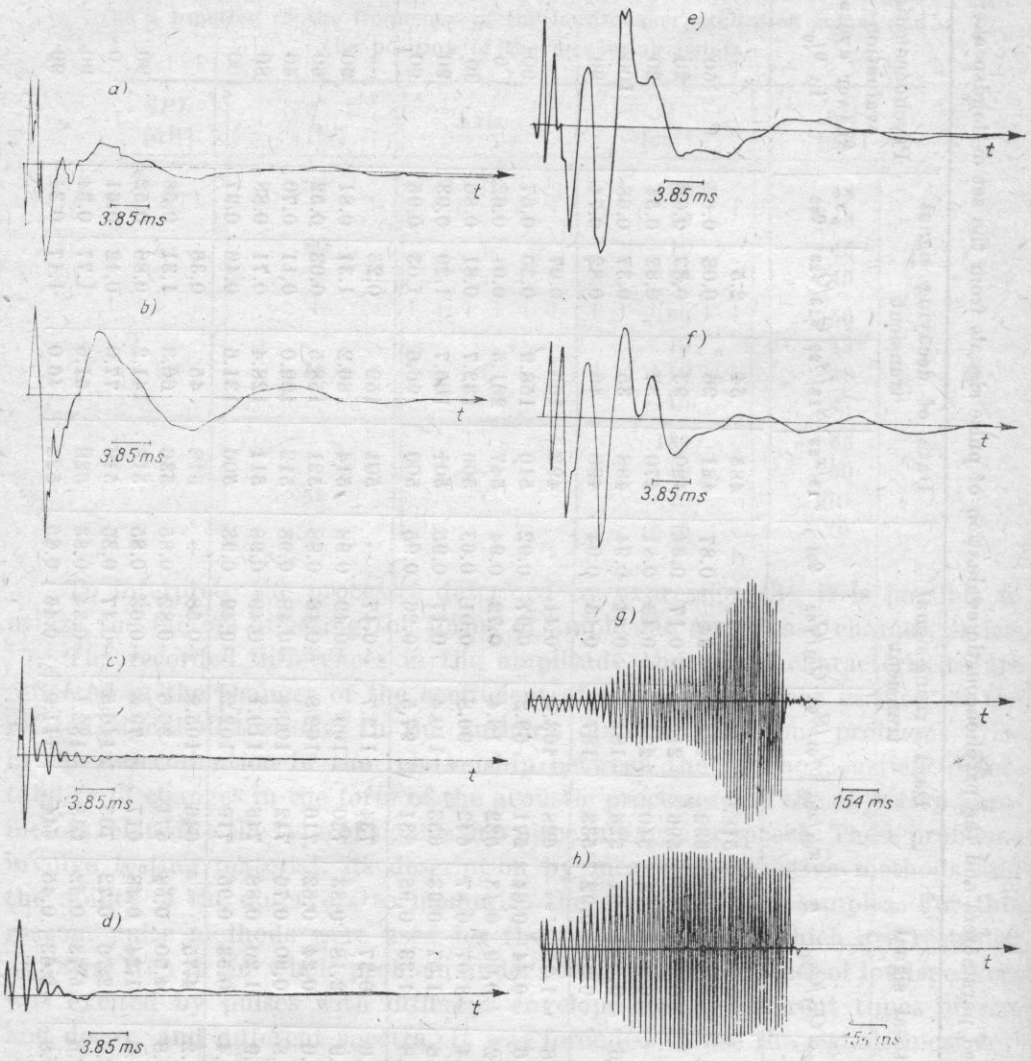

Fig. 5. Responses of the set of loudspeakers to various types of pulse excitation; these responses were recorded at selected points in the acoustic field

Exciting pulses: a) rectangular pulse, duration 1 ms, measuring point a, $R = 0$; b) rectangular pulse, duration 1 ms, measuring point b, $R = 10$; c) tone pulse of frequency 900 Hz, measuring point a, $R = 0$; d) tone pulse of frequency 800 Hz, measuring point c, $R = 100$; e) sequence of pulses, measuring point a, $R = 0$, f) sequence of pulses, measuring point a, $R = 100$; g) rapid frequency sweep pulses with limiting frequencies $f_1 = 30$ Hz and $f_2 = 150$ Hz, measuring point a, $R = 0$; h) rapid frequency sweep pulses with limiting frequencies $f_1 = 30$ Hz and $f_2 = 150$ Hz; measuring point a, $R = 180$

include the results calculated for the whole signal, and its rising and decaying transients recorded at the measuring point a, $R = 180$, and at corresponding points in the acoustic field of the set of loudspeakers. The parameters $\bar{x}$, S, Q characterize each signal independently, while the parameters $\Delta\bar{x}$, ΔS, ΔQ and ϱ describe the differences between the signals recorded at different points

of the acoustic field and their relation to the reference signal recorded at the measuring point a, $R = 180$. The indexes of the parameters are defined as:

11 – the whole signal recorded at the measuring point (a, $R = 180$),

22 – the whole signal recorded at particular measuring points

01 – the rising transient of the signal recorded at the measuring point (a, $R = 180$),

02 – the rising transient of the signal recorded at particular measuring points,

13 – the decaying transient of the signal recorded at the measuring point (a, $R = 180$),

23 – the decaying transient of the signal recorded at particular measuring points.

The numerical values of the selected objective parameters of the pulse signals recorded at various points in the acoustic field of the set of loudspeakers have been compared with the recorded distortions of these signals. It can be concluded:

1. The mean values $\bar{x}_{11}$ and $\bar{x}_{22}$ determined for the total duration of the time varying signals recorded at different measuring points are very close to each other; they provide no information on the signal distortion.

2. The differences in the mean values $\Delta\bar{x}_{02} = \bar{x}_{02} - \bar{x}_{01}$ related to the rising transients and $\Delta\bar{x}_{13} = \bar{x}_{23} - \bar{x}_{13}$ related to the decaying transients, of the reference and other signals provide initial information on the distortion of the rising and decaying transients.

3. The numerical values of the differences ΔS and ΔQ for succesive rising and decaying transients of the different signals provide information on the degree of distortion, the sign of the difference ΔQ determining the direction of deviation of the distortion.

4. The correlation coefficients ϱ, ϱ_{01}, ϱ_{02} related respectively to the entire time intervals of the different signals, and their rising and decaying transients, contribute general information on the degree of distortion only for signals with complex spectra (e.g. a sequence of pulses, Gaussian pulses).

4.3. The psychoacoustic evaluation of pulse distortion

The selected sound pulses were used for psychoacoustic evaluation on the basis of the test described in subsection 3.3.2. It was intended to investigate the question of the extent to which the objective parameters are related to the ability of the organ of hearing to detect distortion of the pulses. Analyzing the detailed measured results, part of which are given in Table 2, we have assumed the number of «yes» answers $z \geqslant 70\,\%$ of all answers, to be the criterion of significance. In this way, if $z \geqslant 70\,\%$ the observer considers the pulses contained in a pair to be identical and he detects no distortion in either of the two compared signals. Using this criterion, it is possible, for instance, to state that the signals produced by a set of loudspeakers excited by a sequence

of pulses and recorded at various points in the accoustic field are not identical (one has been distorted) for both $z = 60\,\%$ and $z = 20\,\%$ (see Table 2).

It follows from the table that the selected objective parameters x, S, Q, ϱ assume such values that for the compared «sequence of pulses», psychoacoustic evaluation can be based on the following statements:

1. The value of the correlation coefficient ϱ connects the entire durations of the compared signals and for $\varrho \geqslant 0.90$ the «yes» answers were estimated at $z \geqslant 90\,\%$ (i.e., if $\varrho \geqslant 90$, then the compared signals are considered to be identical).

2. The correlation coefficients ϱ_{01}, ϱ_{02} relate the corresponding rising and decaying transients $z \geqslant 70\,\%$ was determined for $\varrho_{01} \geqslant 0.94$ and $\varrho_{02} \geqslant 0.68$ (if $\varrho_{01} \geqslant 0.94$ and $\varrho_{02} \geqslant 0.68$ the compared signals were considered to be identical).

3. The psychoacoustic evaluation used the differences of the mean values $\Delta\bar{x}_{02} = \bar{x}_{02} - \bar{x}_{01}$, $\Delta\bar{x}_{13} = \bar{x}_{23} - \bar{x}_{13}$ and the coefficients of skewness $\Delta Q_{02} = Q_{02} - Q_{01}$, $\Delta Q_{13} = Q_{23} - Q_{13}$. The «yes» answers $z \geqslant 70\,\%$ were recorded for $\Delta\bar{x}_{02} \leqslant 15$ and $\Delta Q_{02} \leqslant 0.28$, and $z \geqslant 70\,\%$ for $\Delta\bar{x}_{13} \leqslant 5$ and $\Delta Q_{13} \leqslant 1.7$.

It should be remarked that the numerical values are not unique. Considering the parameters as functions of the position of the measuring points in the accoustic field of the set of loudspeakers, it can be found that they depend on the distance from the set of loudspeakers and from the measuring axis (e.g. for c, $R = -15$ wo have $\varrho = 0.85$ and $z = 60\,\%$, and for b, $R = -15$, $\varrho = 0.44$, $z = 20\,\%$). It follows from the psychoacoustic evaluation of signals recorded at points close to, and far away from the plane of the set of loudspeakers, that the signals have little similarity ($z \leqslant 60$ for $R \leqslant 0$ and $z \geqslant 90\,\%$ for $R \geqslant 100$).

From a similar analysis performed for Gaussian pulses (duration 1.5 ms, rise and decay times 0.5 ms) it can be concluded that:

1. For «yes» answers $z \geqslant 70\,\%$, the experiments give $\varrho \geqslant 90$; this relation is not unique since it was recorded that for $z = 90\,\%$, the other parameters were $\varrho = 0.64$, $\varrho_{01} = 0{,}92$ nad $\varrho_{02} = 0.67$. These responses were then evaluated psychoacouslicly mainly on the basis of the physical similarity of their rising transients (very high ϱ_{01}). This thesis is confirmed for example by high values $\varrho = 0.82$ for $z = 30\,\%$ and with $\varrho_{01} = 0.63$ and $\varrho_{02} = 0.86$. Thus, in spite of high values of ϱ and ϱ_{02} a smaller ϱ_{01}, the psychoacoustic evaluation in «no» (the compared signals are not identical), i.e. it is stated that one pulse is distorted.

2. The dispersion of the values $\Delta\bar{x}_{02}$ is not expressed in the values of ϱ_{01}, but changes in $\Delta\bar{x}_{13}$ affect ϱ_{02}.

3. Changes in ϱ_{01} and ϱ_{02} are related to changes in ΔS_{02} and ΔS_{13}.

From the specimen measured results obtained for tone pulses with frequencies of 80 Hz (3 periods) and 800 Hz (16 periods), it is possible to draw the following conclusions:

1. For higher frequency signals $f = 800$ Hz, the coefficients ϱ are small and their values are not uniquely related to the results of the psychoacoustic evaluation; such relations were found for low frequency signals ($f = 80$ Hz).

2. The value of ϱ_{01} is approximately constant for all the tested signals ($\varrho_{01} \geqslant 0.80$) and is not related to the psychoacoustic evaluation of the signals. This value is independent of changes in the parameters $\Delta\bar{x}_{02}$, ΔS_{02}, ΔQ_{02}.

3. For lower frequencies ($f = 80$ Hz), the psychoacoustic evaluation is strictly related to ϱ_{02} (the number of «yes» answers $z \geqslant 70\,\%$ is determined for $\varrho_{02} \geqslant 0.80$). The relation $\varrho = \text{const.}$ is obvious the pulses are sections of sine-wave tones; only major distortions of these signals changed the values of ϱ (e.g. for $f = 80$ Hz) and this was reflected in the psychoacoustic evaluation of the signals.

4. The differences in the number of «yes» answers are related to changes of $\Delta\bar{x}_{02}$ for $f = 800$ Hz.

5. Conclusions

From the results described in this paper it is possible to draw the following final conclusions:

1. The correlation coefficients ϱ are most strongly related to the psychoacoustic evaluation of the different signals.

2. It was found that the observes are most able to detect differences in the sound pulses with a wide frequency spectrum, for correlation coefficients $\varrho \leqslant 0.90$.

3. The high values of the correlation coefficient ϱ for sound pulses with a wide frequency spectrum cause high values of the coefficients ϱ_{01} which are related to the rising transients of these signals.

4. For tone pulses with a spectrum, in which one frequency is clearly determined, the correlation coefficients ϱ provide no information of the signal. However, signals of this type are distorted by a set of loudspeakers (especially, in the low frequency range) and, consequently, the correlation coefficient of the signal is $\varrho \neq \text{const}$. In the psychoacoustic evaluation, the distortion of a signal is detected for a correlation coefficient $\varrho \leqslant 0.80$.

5. The high value of the correlation coefficient ϱ for a tone pulse which has been distorted for example in transmission causes high values of the correlation coefficient ϱ_{02} which is related to the decaying transients of the signal.

Based on conclusions 1 to 5, it is possible to formulate two more, general, conclusions:

6. In the psychoacoustic evaluation of sound pulses with complex spectra, the rising transients play the most important role, while in the case of tone pulses with simple spectra (limited, moreover, to the low frequency range), the decisive role is played by the decaying transients.

7. It is true that the objective parameters and the results of psychoacoustic evaluation depend on the positions of the measuring points in the acoustic field as may be expected. However, this dependence is not a simple function of the distance between the measuring points and the plane of the set of loudspeakers. (This problem will be discussed by the author in another paper «Analysis and the principles of perception of signals with a varying structure in a complex acoustic field» which is in preparation.)

Acknowledgement. I am very grateful to Prof. Dr. H. RYFFERT for the discussion on the present paper.

I am also indebted to prof. dr E. ZWICKER and to the Humboldt Foundation for the assistance in conducting the research in the Institute of Electroacoustics in Münich (München).

References

[1] S. BEKESY, *Über die Hörsamkeit der Ein- und Ausschwingvorgänge mit Berücksichtigung der Raumakustik*, An. Physik **16**, 5, 844 (1933).

[2] E. HOJAN, *Subjective evaluation of transients with different envelopes as a function of their duration* [in Polish], XI Seminar of Acoustic, Poznań 1964

[3] E. HOJAN, A. ROZSYPAL, *Subjective evaluation of the hearing of transients with different envelopes as a function of the frequency*, Archiwum Akustyki, **2**, 3, 267 (1967) [in Polish].

[4] E. HOJAN, *Investigation of the hearing perception of transients as a function of their envelope and duration* [in Polish], Dissertation, Poznań 1969.

[5] E. HOJAN, J. FLORKOWSKI, M. NIEWIAROWICZ, U. KOKOWSKA, *Utilization of pulse methods for testing loudspeakers* [in Polish], Rozprawy Elektrotechniczne, **23**, 1, 169 (1977).

[6] E. HOJAN, *Rapid frequency sweep method in testing electro-acoustic transducers* [in Polish], to appear.

[7] R. E. KIRK, *Difference limen for tone diminutiobn*, JASA **30**, 10, 915 (1958).

[8] LATHI, *Theory of telecommunication signals and systems*, PWN, Warszawa 1970, [in Polish].

[9] K. KÜPFMÜLLER, *Die Systemtheorie der elektrischen Nachrichtenübertragung*, S. Hirzel Verlag, Stuttgart 1974.

[10] H. RYFFERT, *Analysis of nonstationary vibrations* [in Polish], III Seminar of Acoustic, Olsztyn 1959.

[11] W. TÜRK, *Über die physiologisch-akustischen Kennzeiten von Ausgleichsvorgängen*, Akustische Zeitschrift **5**, 2, 129 (1940).

Received on 29th December 1977

ARCHIVES OF ACOUSTICS
2, 4, 297 (1977)

THERMAL EFFECTS IN SOFT TISSUES DEVELOPED UNDER THE ACTION OF ULTRASONIC FIELDS OF LONG DURATION *)

LESZEK FILIPCZYŃSKI

The Department of Ultrasounds of the Institute of Fundamental Technological Research (Warszawa)

The thermal effect arising from the sonification of soft tissues by an ultrasonic beam, cylindrical in shape, with the assumption of an even distribution of heat sources in the beam, has been considered.

The analysis is based on the solution of the thermal conductivity equation, using the Laplace transformation as in the author's paper [1]. The formulae obtained permit determination of the rise in and distribution of temperature inside and outside the ultrasonic beam for sonification times longer than 20 s.

The formulae have been applied to estimate the temperature changes encountered in ultrasonic continuous wave Doppler methods used in medical diagnosis. For example, with a cylindrical ultrasonic beam of radius 2.2 mm, frequency 5 MHz, mean spatial intensity of 0.1 W/cm^2 and sonification time of 100 s, the estimated value of the temperature increase at the centre of the beam was 1.8 °C.

The values obtained are overestimated since they do not consider the transfer of heat by the circulating blood or the thermal conductivity along the ultrasonic beam, which is particularly evident for higher frequencies.

1. Introduction

In the previous work [1] the author considered thermal effects in soft tissues due to the action of focused ultrasonic fields of short duration (i.e. from microseconds to seconds). The microsecond duration ultrasonic fields occur in impulse ultrasonography used for the visualization of internal organs of the human body, e.g. in gynaecology and obstetrics.

However, in ultrasonic Doppler methods of blood flow measurements, considerably longer sonification is used, sometimes lasting up to several minutes.

---

*) This paper is written within the framework of the problem MR I-24.

In order to estimate the thermal effects which may occur as a result of such a long sonification, it was decided to use the results obtained in [1], extending them to the considerably longer times of sonification.

2. The temperature at the focus of the ultrasonic field

It will be assumed that the focus (of the beam) has the shape of an infinite cylinder of radius R (cm). It is further assumed that heat sources of strength $\dot{Q}_v$[cal s^{-1} cm^{-3}][1]) are located within this volume. According to [1], the Laplace transformation for the increment T_i of temperature increase within the focus can be written in the form

$$\bar{T}_i = \frac{\dot{Q}_v}{s^2 \varrho c_w}\left[1+\frac{\pi}{2}\sqrt{\frac{s}{a}}RH_1^{(1)}\left(i\sqrt{\frac{s}{a}}R\right)J_0\left(i\sqrt{\frac{s}{a}}r\right)\right], \tag{1}$$

where s is the complex variable, ϱ – the density of medium [g cm^{-3}], c_w – the specific heat of the medium [cal g^{-1} °C^{-1}], $a = \lambda/\varrho c_w$, λ is the coefficient of thermal conductivity [cal cm^{-1} °C^{-1}], $i = \sqrt{-1}$, $H_n^{(1)}$ is a Hankel function of the first order equal to $H_n^{(1)} = J_n + i\,N_n$, J_n – the Bessel function, and N_n – the Neumann function. The subscripts at H, J and N denote the order of functions.

The inverse transform of expression (1) will be determined by a series expansion of the Bessel and Hankel functions. From the similarity theorem [5] between the transformed function f and its transform $\bar{f}$,

$$L[f(at)] = \frac{1}{a}\bar{f}\left(\frac{s}{a}\right), \tag{2}$$

it can be seen that large arguments of the transformed function f correspond to small arguments $i\sqrt{s/a}x$ $(x = R, r)$ of the transform $\bar{f}$. This is of interest to us due to the longer time of sonification and we shall thus use the series expansions [3, 5] (valid for all α)

$$J_0(i\alpha) = 1+\frac{(\alpha/2)^2}{1!\,1!}+\frac{(\alpha/2)^4}{2!\,2!}+\frac{(\alpha/2)^6}{3!\,3!}+\ldots, \tag{3}$$

$$J_1(i\alpha) = i\left[\frac{\alpha/2}{0!\,1!}+\frac{(\alpha/2)^3}{1!\,2!}+\frac{(\alpha/2)^5}{2!\,3!}+\ldots\right], \tag{4}$$

$$N_0(i\alpha) = \frac{2}{\pi}\left[\left(\ln\frac{\gamma i\alpha}{2}\right)J_0(i\alpha)-\frac{(\alpha/2)^2}{1!\,1!}-\right.$$
$$\left.-\left(1+\frac{1}{2}\right)\frac{(\alpha/2)^4}{2!\,2!}-\left(1+\frac{1}{2}+\frac{1}{3}\right)\frac{(\alpha/2)^6}{3!\,3!}-\ldots\right], \tag{5}$$

[1]) Since this article is a continuation of the author's paper [1] which used the CGS-system of units, the same system will also be used here.

$$N_1(ia) = \frac{2}{\pi}\left[\left(\ln\frac{\gamma ia}{2}\right)J_1(ia) - \frac{1}{ia} - i\left(1-\frac{1}{2}\right)\frac{a/2}{0!\,1!} - \right.$$

$$\left. - i\left(1+\frac{1}{2}-\frac{1}{4}\right)\frac{(a/2)^3}{1!\,2!} - i\left(1+\frac{1}{2}+\frac{1}{3}-\frac{1}{6}\right)\frac{(a/2)^5}{2!\,3!} - \ldots\right], \quad (6)$$

where $\gamma = 1.781072\ldots$ is a constant.

Substituting expansions (3)-(6) into (1), using the formula

$$\ln\frac{\gamma}{2} i\sqrt{\frac{s}{a}}\, x = \frac{\pi}{2} i + \frac{1}{2}\ln s\,\frac{\gamma^2 x^2}{4a} \quad (7)$$

and neglecting small order quantities, we finally obtain

$$\bar{T}_i = \frac{\dot{Q}_v}{\varrho c_w}\left[\frac{1}{4as}(R^2-r^2) + \frac{1}{64\,a^2}(5\,R^4+4\,R^2r^2-r^4) - \right.$$

$$\left. - \frac{R^2}{4as}\ln s\,\frac{\gamma^2R^2}{4a} - \frac{1}{32a^2}(R^4+2R^2r^2)\ln s\,\frac{\gamma^2R^2}{4a}\right]. \quad (8)$$

Using the relations [5]

$$L^{-1}\left[\frac{1}{s}\ln ks\right] = -\ln\frac{\gamma t}{k}, \quad (9)$$

$$L^{-1}[\ln ks] = -\frac{1}{t}, \quad (10)$$

$$L^{-1}\left[\frac{1}{s}\right] = 1, \quad (11)$$

$$L^{-1}[s^n] = 0, \qquad n \geqslant 0, \quad (12)$$

we calculate the inverse transform of expression (8) and obtain the final temperature within the focal volume ($r \leqslant R$) as

$$T_i = \frac{\dot{Q}_v}{\varrho c_w}\left[\frac{1}{4a}(R^2-r^2) + \frac{R^2}{4a}\ln\frac{4a}{\gamma R^2}\,t + \frac{R^4+2R^2r^2}{32a^2}\,\frac{1}{t}\right]. \quad (13)$$

From equation (13) we can determine the maximum value T_M of the temperature, which occurs at the centre of the focus ($r = 0$) to be

$$T_M = \frac{\dot{Q}_v}{\varrho c_w}\left[\frac{R^2}{4a}\left(1+\ln\frac{4a}{\gamma R^2}\,t\right) + \frac{R^4}{32a^2}\,\frac{1}{t}\right]. \quad (14)$$

At the boundary of the focus ($r = R$) we have

$$T_i = \frac{\dot{Q}_v}{\varrho c_w}\left[\frac{R^2}{4a}\ln\frac{4a}{\gamma R^2}\,t + \frac{3R^4}{32a^2}\,\frac{1}{t}\right]. \quad (15)$$

Formulae (13)-(15) fail for $a \to 0$ and for $t \to 0$ due to the higher order terms neglected in expansions (3)-(6) of the Bessel and Hankel functions.

3. The temperature distribution outside the focal region

The temperature outside the focal region will be determined on the basis of formula (3) of paper [1]. Thus we have.

$$\bar{T}_o = \frac{\dot{Q}_v}{s^2 \varrho c_w} \frac{\pi}{2} \sqrt{\frac{s}{a}} R J_1\left(i\sqrt{\frac{s}{a}} R\right) H_0^{(1)}\left(i\sqrt{\frac{s}{a}} r\right). \tag{16}$$

The inverse transform of this expression will be determined as before using expansions (3)-(6). Substituting these into expression (16) and neglecting the higher order terms, we obtain the transform of the temperature outside the focal region as

$$\bar{T}_o = \frac{\dot{Q}_v}{\varrho c_w}\left[\frac{R^2 r^2}{8a^2} - \frac{R^2}{4sa} \ln s\frac{\gamma^2 r^2}{4a} - \frac{R^2(R^2+2r^2)}{32a^2} \ln s\frac{\gamma^2 r^2}{4a}\right]. \tag{17}$$

Using relations (9)-(12) we can calculate the inverse transform in the form

$$T_o = \frac{\dot{Q}_v}{\varrho c_w}\left[\frac{R^2}{4a} \ln\frac{4a}{\gamma r^2} t + \frac{R^2(R^2+2r^2)}{32a^2}\frac{1}{t}\right]. \tag{18}$$

For $a \to 0$, $t \to 0$ and $r \to \infty$ formula (18) fails for the same reason as before.

It can easily be seen that on the boundary of the focal region ($r = R$) we obtain from formula (18), as expected, the expression already given by formula (15).

The temperature distributions, calculated from formulae (13) and (18) inside and outside the focal region, are shown in Fig. 1. Fig. 2 shows the temperature at the centre of the focus ($r = 0$) as a function of the time of sonification, calculated on the basis of formula (14) for times $t > 20$ s and $t < 10$ s. The latter were taken from [1]. These latter distributions were calculated for the conditions encountered in pulse-echo ultrasonic obstetric and gynaecological diagnostic investigations [2]. The focus of the ultrasonic beam was in this case approximated by a cylinder of diameter of 0.125 cm and length 6.2 cm.

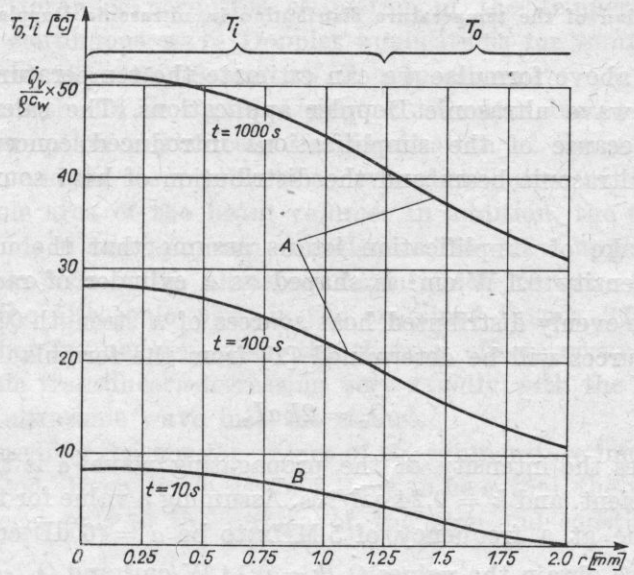

Fig. 1. The distribution of the temperature inside ($r < 1.25$ mm) and outside ($r > 1.25$ mm) the focal region at different times of sonification t as a function of r, calculated from formulae (13) and (18) (curve A), and taken from [1] (curve B)

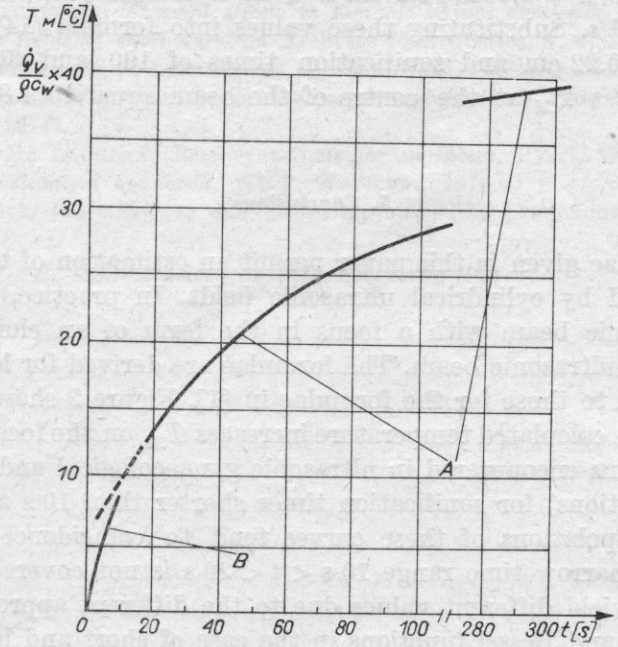

Fig. 2. The temperature at the centre of the focus of the ultrasonic beam, for sonification times $t \geqslant 20$ s, calculated on the basis of formula (14) (curve A) and taken from [1] for sonification times $t \leqslant 10$ s (curve B)

4. The estimation of the temperature distribution in ultrasonic Doppler applifications

From the above formulae we can estimate the temperatures encountered in continuous wave ultrasonic Doppler applications. The calculation will be an estimate because of the simplifications introduced concerning the geometry of the ultrasonic beam and the distribution of heat sources within the beam.

For the sake of simplification let us assume that the ultrasonic beam of uniform intensity 0.1 W/cm^2 is shaped as a cylinder of radius 2.2 mm in which there are evenly distributed heat sources of a strength $\dot{Q}_v$. The strength of the heat sources will be determined [1] from the formula

$$\dot{Q}_v = 2k\alpha I, \tag{19}$$

where I denotes the intensity of the propagating wave, α is the pressure absorption coefficient, and $k = 0.24$ cal/Ws. Assuming a value for the coefficient α of muscle tissue at a frequency of 5 MHz to be $\alpha = (6 \text{ dB/cm}) \times (8.67 \text{ dB})^{-1} = 0.69 \text{ cm}^{-1}$, we obtain the values $\dot{Q}_v/k = 0.14$ W/cm^3 and $\dot{Q}_v/\varrho c_w = 0.033$ °C/s.

In view of the fact that water constitutes 75 % of the content of the cells of soft tissues, we assume that the thermal conductivity is the same as that for water. Thus $\lambda = 0.00038$ cal/cm·s °C at a temperature of 30°C [4] and $a = 0.00038$ cm^2/s. Substituting these values into formula (14), with a beam radius of $R = 0.22$ cm and sonification times of 100 and 300 s, we obtain temperature rises T_M at the centre of the beam equal to 1.8 °C and 2.8 °C, respectively.

5. Conclusions

The formulae given in this paper permit an estimation of the temperature increases caused by cylindrical ultrasonic fields. In practice, this may be a focused ultrasonic beam with a focus in the form of an elongated cylinder or a cylindrical ultrasonic beam. The formulae are derived for long sonification times compared to those for the formulae in [1]. Figure 2 shows a comparison of these with the calculated temperature increases T_M on the focal axis (obtained for the conditions encountered in ultrasonic gynaecological and obstetric diagnostic investigations) for sonification times shorter than 10 s and longer than 20 s. The extrapolations of these curves tend to coincidence. However, the comparatively narrow time range $10 \text{ s} < t < 20 \text{ s}$ is not covered by the calculations, which yield different values due to the different approximations used for the Hankel and Bessel functions in the case of short and long sonification times.

The character of the temperature distribution inside and outside the focal region (Fig. 1) for long times of sonification (curves A) corresponds to the character of the curves calculated for short times of sonification (curves B).

The above formulae permitted estimation of the temperature rises T_M in the case of continuous wave Doppler applications for sonification times t equal to 100 and 300 s as 1.8 and 2.8 °C, respectively.

These values should be regarded only as estimates due to the number of simplifying assumption introduced: the idealization of the geometry of the ultrasonic beam and the assumption of an even distribution of the heat sources within the whole area of the beam volume. In addition, the transfer of heat by circulating blood has not been considered. The same applies to the flow of heat along the axis of the ultrasonic beam which increases with increasing absorption of the ultrasonic wave in the examined tissues. This is especially evident at higher frequencies, when the thermal effects occur at the surface of the ultrasonic transducer, decreasing very rapidly with the depth of penetration of the ultrasonic wave into the tissues.

For the foregoing reasons the values of the temperature increases obtained are overestimated. Even if we assume them to be actual, the values of a few degrees may occur only at the surface of the skin and constitute no danger to the patient.

References

[1] L. Filipczyński, *Thermal effects in the soft tissues developed under the action of focused ultrasonic fields of short duration* Archives of Acoustics, **1**, 4, 309–321 (1976).

[2] L. Filipczyński, G. Łypacewicz, J. Sałkowski, *Intensity determination of focused ultrasonic beams by means of electrodynamic and capacitance methods*, Proc. Vibration Problems, **15**, 4, 297–305 (1974).

[3] N. W. Mc Lachlan, *Bessel functions for engineers*, PWN, Warszawa 1964.

[4] *Physico-chemical handbook*, WNT, Warszawa 1974.

[5] H. Tautz, *Wärmeleitung und Temperaturausgleich*, Akademieverlag, Berlin 1971.

Received on 25th February, 1977

ARCHIVES OF ACOUSTICS
2, 4, 305 (1977)

ULTRASONIC PULSE DOPPLER METHOD IN BLOOD FLOW MEASUREMENT

ANDRZEJ NOWICKI

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

The phenomenon of the scattering of ultrasonic waves by the morphotic elements of blood is discussed. An analysis is presented of a correlation system for receiving Doppler signals, with consideration being given to the problem of the simultaneous measurement of the position and velocity of the blood corpuscles. The principles of operation of the device UDIMP, developed by IPPT PAN, is described and its basic technical parameters are given.

1. Introduction

Among the present diagnostic methods for the evaluation of the blood flow velocity in blood vessels, the ultrasonic Doppler method of blood flow [4] appears to be particularly interesting in promising further rapid development.

The Doppler phenomenon consists in the fact that a signal, reflected (or scattered) from a moving object, changes its frequency f_0 relative to the frequency of the transmitted signal f_n by a frequency f_d (termed subsequently the «Doppler frequency») which is proportional to the linear velocity of the object, according ot the expression

$$f_d = \frac{f_n v}{c}(\cos\theta_N + \cos\theta_0) \quad \text{for } \frac{v}{c} \ll 1, \tag{1}$$

where v denotes the velocity of the moving object (a blood corpuscle), c is the velocity of ultrasound in blood, θ_n – the angle between the direction of the blood velocity and the normal to the transmitting transducer, and θ_0 – the angle between the direction of the blood velocity and the normal to the receiving transducer.

In normal physiological conditions, the distribution of blood velocity across the vessel is a complicated function of the vessel radius, the flow pulsation and of the blood viscosity [6]. Thus each particle in the flowing blood produces a signal with a different frequency, so that the component signal,

received by the receiving transducer, will contain a spectrum of Doppler frequencies over a range from $f_d = 0$ — for the particles at the walls of the vessel — to $f_d = f_{d\max}$ for the particles moving in the middle of the vessel.

The determination of the mean Doppler frequency $f_{d\text{śr}}$ and, subsequently, the determination of the mean blood flow velocity, corresponding to the above frequency throughout the whole cross-section of the blood-vessel, is troublesome and sometimes quite imposible, especially for turbulent flow. Additional difficulties appear when we measure the velocity through the skin, beacuse the angles θ_n and θ_0 (Fig. 1) are unknown.

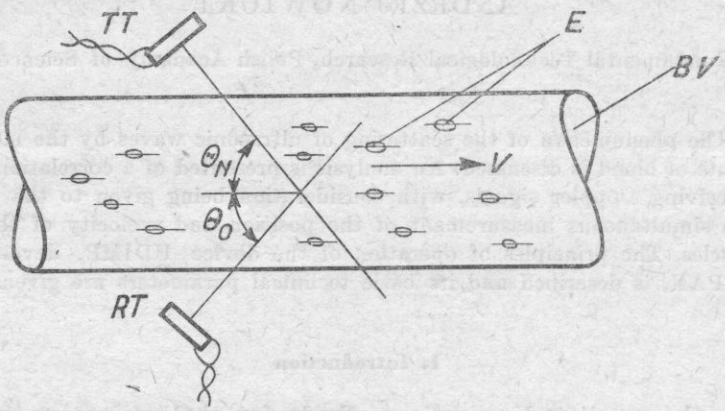

Fig. 1. The principle of operation of the ultrasonic blood flowmeter
TT – transmitting transducer, *RT* – receiving transducer, *E* – morphotic elements of blood, *WV* – blood vessel

For clinic diagnosis, the mean volume blood flow velocity Q, described by the relation

$$Q = v_{\text{śr}} A, \tag{2}$$

is more important than the mean instantaneous blood flow velocity, $v_{\text{śr}}$ being the instantaneous mean blood flow velocity, A — the cross-sectional area of the vessel, and Q — the instantaneous volume blood flow velocity.

The continuous wave Doppler method does not permit measurement of the vessel diameter and thus cannot be used for determination of the volume velocity of the blood flow.

The transition from qualitative to quantitative investigations has become possible due to a method which combines an ultrasonic Doppler technique with a pulse-echo technique. It permits, for example, experimental determination of the profiles of the blood flow velocity in different phases of the working cycle of the heart, the mean velocity, and also the volume blood flow velocity [11, 12] (Fig. 2).

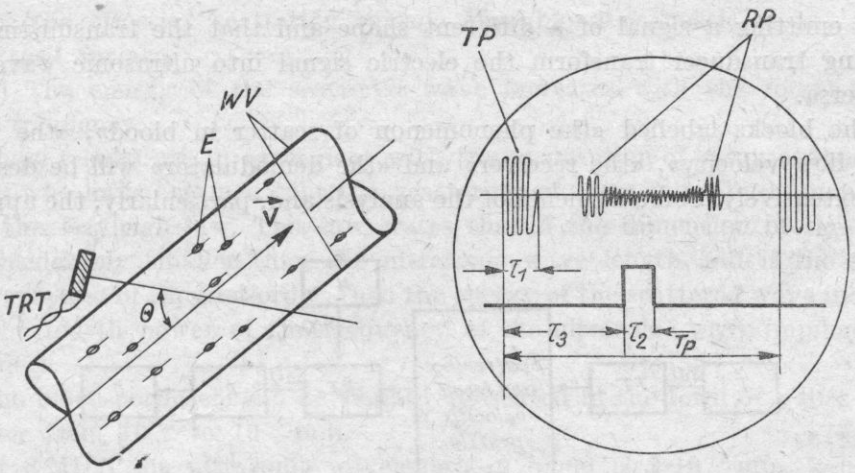

Fig. 2. The principle of operation of the pulsed ultrasonic Doppler flowmeter for measuring the blood flow velocity profile

TRT – transmitting and receiving transducer, *WV* – the walls of the vessel, *E* – morphotic elements of blood, *TP* – transmitted pulse, *RP* – pulses reflected from the walls of the vessel

The piezoelectric transducer emits, in the direction of the blood vessel, high-frequency pulses E of duration τ_1, frequency f_n and repetition frequency rate $F_p (F_p = 1/T_p)$. The ultrasonic wave is partially reflected and scattered by both the walls of the vessel and the blood corpuscles. As a result, during the time T_p between two subsequent transmitted pulses echoes return to the transducers corresponding to reflection from the walls of the blood-vessel and to the scattering from the blood corpuscles.

Knowing the angle θ between the direction of the incident ultrasonic wave and the blood-vessel, and also the propagation velocity of the ultrasound in tissue and blood, we can determine the transverse dimension of the vessel [2, 9]. Simultaneously, due to the use in the receiver of an electronic gate of adjustable duration τ_2 and time delay τ_3 relative to the transmitting pulse, it is possible to select the echoes that correspond to the scattering from the moving blood corpuscles at a selected depth.

2. The general model of the system for measuring blood flow velocity

Fig. 3 shows a general model of the system for measuring blood flow, consideration having been given to the macroscopic physical properties of blood and the nature of its flow in the circulating system.

The blocks labelled «the transmitter», «the transmitting transducer» and «the receiving transducer» do not require detailed discussion. Consideration will only be given in this paper to the problems related with the emission of a sequence of coherent pulses of high frequency. Nevertheless, it can be assumed that, without particular technical problems, it is possible to develop a trans-

mitter emitting a signal of a different shape and that the transmitting and receiving transducer transform the electric signal into ultrasonic waves and vice versa.

The blocks labelled «the phenomenon of scatter in blood», «the profile of the flow velocity», «the receiver» and «the demodulator» will be described more extensively. Some elements of the analysis and, particularly, the approach

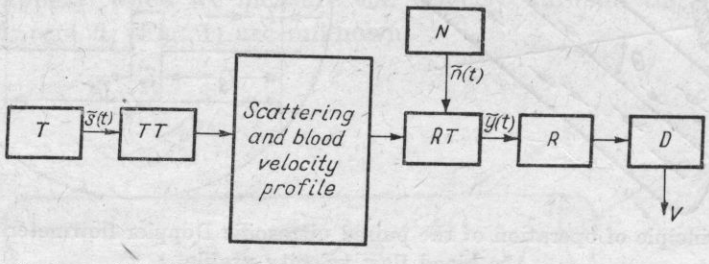

Fig. 3. A model of the system for measuring the blood flow velocity
T – transmitter, TT – transmitting transducer, N – noise, RT – receiving transducer, R – receiver, D – demodulator, v – the signal proportional to the flow rate

to the subject were taken from the theory of the radar techniques [16]. However, the differences encountered in blood flow measurement, compared with measurements using radar technique for determining the velocity of flying objects, make it impossible to consider ultrasonic Doppler flowmeters as analogous to Doppler radar.

From the physical point of view blood is a suspension of cells in a liquid plasma. These cells are red blood corpuscles (erythrocytes), white corpuscles (leucocytes), and blood lamellae (thrombocytes). Their mean number per cubic milimetre of blood is [1] about $5 \cdot 10^6$, 10^4, $3 \cdot 10^5$, respectively.

The experimental investigations carrried out by REID [13] concerning the scattering uf ultrasonic wave by human blood over a frequency range from 4 MHz to 16 MHz gave the following results:

(a) The main source of scattering is the erythrocytes. At a frequency of 5 MHz the energy of the ultrasonic wave scattered by the thrombocytes was about 1000 times less than the energy scattered by the erythrocytes. It is practically undetectable for normal physiological structures.

(b) The quantity (acoustic pressure) of the scattered wave is proportional to the density (hematocryt) over the range from 7 % to 40 %, i.e. in the physiological range of blood density.

The hematocryt is defined by the expression:

$$\text{Hematocryt} = \frac{\text{volume of erythrocytes} \times 100\,\%}{\text{total volume of erythrocytes and plasma}}.$$

(c) The scattering is isotropic.

(d) The effective scattering surface of a blood corpuscle is 10^{-4} of its geometrical surface.

(e) The energy of the scattered wave increases with the fourth power of the frequency.

These results are in agreement with the assumption of many authors (e.g. [3, 17]) who have assumed that the scattering of ultrasound by human blood obeys the Rayleigh law. This law states that if the dimension of the object are considerably smaller than the ultrasonic wave length and if the scatter is isotropic and of the first order, then the energy of the scattered wave increases with the fourth power of the frequency of the ultrasonic wave impinging on the object.

The blood corpuscle can be visually presented in the form of a disc of the diameter from 10^{-3} to 10^{-2} mm.

At 8 MHz, the ultrasonic wavelength in blood is $2 \cdot 10^{-1}$ mm, hence the ratio of the ultrasonic wavelength to the geometrical size of the blood corpuscle varies from about 20 to 200 and this permits, according to the Raleigh theory, its consideration as a point scatterer.

It was then assumed that the scattering is a random process. This means that at any moment the number of particles that scatter the ultrasonic wave will be subject to a particular function of probability density.

The assumptions of the first order and random scattering of the blood particles permit the supposition that the scattering process is subject to the Poisson distribution. With such an assumption, the probability that N of the blood corpuscles is contained in the volume V, defined by the ultrasonic field, is given by the formula

$$P(N/V) = P(N \text{ blood corpuscles in a volume } V) = \frac{(\varrho V)^N e^{-\varrho v}}{N!}, \tag{3}$$

where ϱ denotes the mean number of blood corpuscles per cubic millimetre.

The expected value of the number of blood corpuscles in the volume V is equal to the first moment of the distribution $P(N/V)$, and amounts to ϱV.

The description of the first orderly mathematical model of blood flow was presented by WOMERSLEY [6] who introduced expressions describing the instantaneous velocity $v_{\text{śr}}$ and the volume velocity Q for a pressure gradient varying sinusoidally in time along the axis of an elastic tube (assuming that only radial motion of the tube walls occurs). WOMERSLEY also proved that the flow velocities, described in this manner, are identical with the corresponding flow velocities in a straight tube with rigid walls,

$$v(y) = \frac{M}{j\omega\varrho_0}\left[1 - \frac{J_0(\alpha j^{3/2} y)}{J_0(\alpha j^{3/2})}\right] e^{j\omega t}, \tag{4}$$

$$v_{śr} = \frac{M}{j\omega\varrho_0}\left[1 - \frac{2J_1(\alpha j^{3/2})}{j^{3/2}\alpha J_0(\alpha j^{3/2})}\right]e^{j\omega t}, \tag{5}$$

$$Q = V_{śr}A, \tag{6}$$

where $j = \sqrt{-1}$, $y = r/R$ (R is the radius of the tube, r – the radial coordinate in the cylindrical coordinate system (r, θ, z) with its origin at the centre of the tube), $\alpha = R\sqrt{\omega/\nu}$ (ν is the kinematic viscosity, ω – the pulsation of the pressure wave, ϱ_0 – the specific mass of liquid, J_0 and J_1 are the Bessel functions of the zero and first order, M is the pressure gradient, and A – the tube cross-sectional area.

Expressions (4), (5) and (6) were determined on the basis of the following assumptions:

1. we have an incompressible Newtonian liquid;
2. the system is axially symmetric;
3. the flow is laminar;
4. the thickness of the tube walls is thin relative to the radius of the tube, i.e. $h/R \leqslant 0{,}1$;
5. the transverse and longitudinal deformations of the tube wall during flow are negligible;
6. the tube wall is isotropic and homogeneous;
7. the tube is infinitely long and of uniform shape (the transverse components of the velocity are thus only related to the transverse motion of the walls);
8. the velocity of the pressure wave c_0 is considerably higher than the flow velocity v, $v/c_0 \leqslant 1$.

Not all of these assumptions are satisfactorily fulfilled by the blood system. The critical Reynold's number ($\approx$2300) can be considerably exceeded during systole particularly in the aorta, thus preventing laminar flow [6]. Newertheless, in the lower vessels (carotid, femoral artery, descending aorta etc.) we may generally assume laminar flows [1, 6]. The condition $h/R \leqslant 0.1$ is satisfied only for large vessels h/R varies from 0.1 to 0.15).

Arteries are not infinitely long but, for the large vessels, the segments between successive bifurcations can be regarded as cylindrical tubes, identical throughout their length. The condition $v/c_0 \leqslant 1$ is generally met. The mean velocity c_0 of the pressure wave in the blood system lies between 8 and 12 m/s and this, for the range of instantaneous velocities v of blood flow varying from about 1 cm/s to about 200 cm/s (the latter in the ascending aorta), gives a value of the ratio v/c_0 between 0.001 and 0.2.

A more general characteristic of blood flow was proposed by BRODY [3] who introduced the scattering function $p(v, r)$ describing the probability density of the appearance of a scattering blood corpuscle at a position r moving at a speed v (at the assumption that the position and velocity of the blood corpuscle are random variables).

The scattering function can be represented in the form

$$p(v, r) = p(v/r)\, p(r), \tag{7}$$

where $p(v/r)$ denotes the conditional probability density function for a blood corpuscle scattering at a speed v, provided that this blood corpuscle appears in the position r, $p(r)$ is the probability density function for the blood corpuscle in position r.

The function $p(r)$ describes the geometrical distribution of particles in the blood vessel. For a homogeneous medium and a cylindrical vessel we have

$$p(r) = p(r, \theta, z). \tag{8}$$

The condition

$$\int_{-\infty}^{+\infty} p(r)\, dr = 1$$

implies that for a homogeneous distribution of blood corpuscles we have

$$p(r) = \varrho \frac{r}{N}, \tag{9}$$

where ϱ denotes the density of blood corpuscles, and N is the total number of blood corpuscles in the region in which the probability density function $p(r)$ is defined.

Equation (9) states that the distribution of particles across the vessel is linear. Nevertheless, it should be pointed out here that, as LIH [5] has already shown, the blood corpuscles tend to migrate towards the middle of the vessel.

When the motion of blood corpuscles is caused only by the convection of plasma, the first factor in expression (7) can be written in the form

$$p(v/r) = \delta[v - v(r)], \tag{10}$$

where δ is the Dirac distribution. This distribution is thus referred to as the velocity distribution across the vessel, i.e. the velocity profile $v(r)$.

Diffusive flow, occurring in the vessel, determines the displacement of the blood corpuscles due to random collisions (Brownian motion) and can cause some fluctuation of the value of $v(r)$, but it does not contribute quantitatively to the blood flow velocity.

In conclusion, it should be stressed that a knowledge of $p(v, r)$ permits satisfactory determination of all flow parameters (velocity profile, flow rate etc.).

The function $p(v, r)$ can obviously be represented in the form of a function in which the independent variables are the quantities measured by a pulsed ultrasonic Doppler flowmeter. Thus, when the analogue of v is the Doppler frequency f_d and the position v corresponds to t_0, the transmitting transducer — blood corpuscle — receiving transducer delay.

Assuming new variables, the scattering function can be written in the form $p(f_d, t_0)$.

3. The correlative receiver

It will be assumed that a signal $\tilde{s}(t) = s(t)e^{j\omega_0 t}$ of the duration τ_1 and the repetition time T_p is emitted in the direction of a fixed object. After a time t_{01} the reflected (or scattered) wave returns from the fixed object and has the form $\tilde{y}(t) = \tilde{ks}(t-t_{01})+\tilde{n}(t)$, where $\tilde{n}(t)$ is the noise with a random Gaussian distribution of the amplitude in the same frequency band as that of the transmitting signal and noise power density in watts per hertz equal to b.

The determination of t_{01} is reduced to the problem of finding the time t_0 for which the expression

$$C(t_0) = \frac{k}{b}\int_{\tau_1} \tilde{s}(t-t_0)\tilde{s}(t-t_{01})\,dt + \frac{k}{b}\int_{\tau_1} \tilde{s}(t-t_0)n(t)\,dt, \qquad C(t_0) = C_u(t_0)+C_b(t_0) \tag{11}$$

attains its maximum.

The term $C_u(t_0)$ of expression (11) is (with a coefficient $k\tau_1/b$) the auto-correlation function of the envelope of the transmitted signal $\tilde{s}(t)$ and attains its minimum for $t_0 = t_{01}$. This confirms *a priori* obvious fact that in the absence of noise the most probable position of the measured object is its real position.

The shape of the auto-correlation function $C_u(t_0)$ (i.e. its distribution about the time t_{01}) depends only on the square of the modulus of the Fourier transform, $|F(f)|^2$, of the signal $\tilde{s}(t)$, since the Fourier transform of the function $C_u(t_0)$ is equal to $|F(f)|^2$.

The maximum of the function $C_u(t_0)$ for $t_0 = t_{01}$ is equal to

$$[C_u(t_{01})]_{\max} = \frac{k}{b}\int_{\tau_1} = [\tilde{s}(t_0-t_{01})]^2 dt, \tag{12}$$

where $[\tilde{s}(t-t_{01})]^2$ is equal to the power of the received signal.

The integral of the power in time τ_1 is equal to the energy E of the received signal, thus

$$C_u(t_0) = \frac{E}{b} = R, \tag{13}$$

where R determines the signal/noise energy ratio.

The term $C_b(t_0)$ in expression (11) shows the effect of noise on the accuracy of the measurement of the position of the object reflecting (scattering) the transmitted signal $\tilde{s}(t)$.

After suitable transformations we can determine the mean standard deviation $\sqrt{\overline{\Delta t^2}}$ of the measurement of the time t_{01} in the function R and also characteristic parameters of the transmitted signal (16),

$$\sqrt{\overline{\Delta t^2}} = \frac{1}{2\pi B\sqrt{R}}, \tag{14}$$

where

$$B^2 = \frac{\int_{\Delta f} (f-f_n) F^2 (f-f_n) df}{\int_{\Delta f} F^2 (f-f_n) df}. \tag{15}$$

If the signal $\tilde{s}(t)$ has the form of a sequence of high frequency pulses, B is approximately equal to

$$B = 0.43/\tau_1. \tag{16}$$

Remembering that

$$t = \frac{2d}{c}, \tag{17}$$

where d denotes the distance of the transducer from the reflecting object, and c is the propagation velocity of the ultrasonic wave in the medium, we can determine from expressions (14) and (16) the mean standard deviation $\sqrt{\overline{\Delta d^2}}$ of the measurement of the distance d:

$$\sqrt{\overline{\Delta d^2}} = 0.18 \frac{c\tau_1}{\sqrt{R}}. \tag{18}$$

Thus, for a given signal to noise ratio, the precision with which we can measure the position of a reflecting object is inversely proportional to the duration of the transmitted pulse.

If the scattering or the reflection of the ultrasonic wave occurs from the moving object, then we are concerned with the Doppler-Fizeau effect which is described in the introduction to this paper.

In the case of the emission of a continuous signal, the measurement of the frequency f_d and also of the velocity v can be performed with great precision, depending on the type of frequencymeter used. However, in this case it is impossible to determine the position of the scattering object.

For pulsed emission where, as a result of the scattering by the moving object, the whole spectrum of the transmitted signal is subjected to a Doppler shift, the precision of the measurement of f_d, with the simultaneous measurement of t_0, is subject to a limitation defined generally by the ambiguity function $\chi(\theta, f_d)$. It should be stressed that the maximum Doppler frequency $f_{d\max}$, measured explicitly, must satisfy the condition

$$f_{d\max} \leqslant \left| \frac{F_p}{2} \right|, \tag{19}$$

where $F_p = 1/T_p$ denotes the repetition function of high frequency pulses.

This relation is consistent with the Shannon law stating that in sampling systems the maximum measured frequency must be at last two times lower than the sampling frequency.

According to WOODWARD [18], in pulsed emission the mean standard deviation of the measured frequency f_d is expressed by the formula

$$\sqrt{\overline{\Delta f_d^2}} = \frac{3}{\pi \tau_1 R}, \tag{20}$$

and this corresponds to a mean standard deviation, for the measurement of the velocity v, of

$$\sqrt{\overline{\Delta v^2}} = \frac{c\sqrt{3}}{2\pi f_n \tau_1 R}. \tag{21}$$

For a coherent system, i.e. a system in which the phase of the subsequent pulses radiated into the medium is constant during the measurement, formulae (20) and (21) are modified to

$$\sqrt{\overline{\Delta f_d^2}} = \frac{\sqrt{3}}{\pi \tau_1 \sqrt{nR}} \tag{22}$$

and

$$\sqrt{\overline{\Delta v^2}} = \frac{c\sqrt{3}}{2\pi f_n \tau_1 \sqrt{nR}}, \tag{23}$$

where n denotes the number of pulses contained in the time during which the velocity of the object is quasi-constant.

Thus we see that the precision, with which the velocity of the object can be measured, is directly proportional to the duration of the transmitted pulse, the inverse being the case with the precision of measurement of the position. The ambiquity function is introduced into the theory of the radar and sonar systems as a measure of the quality of the signal repetition system. The ambiguity function takes the form

$$\chi(\theta, f_d) = \int_{-\infty}^{+\infty} \tilde{s}(t)\tilde{s}^*(t-\theta)e^{2\pi j f_d t}\,dt, \tag{24}$$

where $\theta = t_0 - t_{01}$ is the two-dimensional correlation function for the delay time t_0 and the Doppler frequency of the signal scattered by a particle which is at a distance d_1 (t_{01} is the delay time for the ultrasonic transducer — particle — ultrasonic transducer path) from the receiver and moving with a velocity v that corresponds to a change of the transmitter frequency f_n by a Doppler frequency f_d, delayed in time by t_0.

The simplified diagram of operation of the correlative receiver is shown in Fig. 4. The process is described by

$$\tilde{u}(t) = \tilde{y}(t) * \tilde{s}(t) = \int_{\tau_2} \tilde{y}(t)\tilde{s}^*(t-t_0)\,dt, \tag{25}$$

where $\tilde{y}(t) = \tilde{s}(t-t_{01})$, when noise and interference are neglected.

In the case where the signal is scattered by only one particle, we can replace $\tilde{u}(t)$ by $\tilde{u}_1(t)$.

Substituting

$$\tilde{s}(t-t_0) = s(t-t_0)e^{j\omega_n(t-t_0)} \tag{26}$$

and

$$\tilde{y}(t) = s(t-t_0)e^{j\omega_1(t-t_{01})}, \tag{27}$$

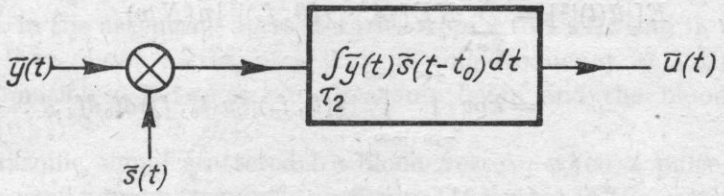

Fig. 4. Simplified diagram of the correlative receiver

and also introducing new variables $u = t-t_{01}$ and $\theta_1 = t_0 - t_{01}$, we obtain

$$\tilde{u}_1(t) = e^{j\omega_n\theta_1} \int_{\tau_2} s(u)s^*(u-\theta)e^{2\pi j f_{d1}u}du, \tag{28}$$

where

$$2\pi f_{d1} = \omega_1 - \omega_n.$$

The envelope of the signal output from the correlator takes the form $|\tilde{u}_1(t)|$, where

$$|\tilde{u}_1(t)|^2 = \chi(\theta_1, f_{d1})\chi^*(\theta_1, f_{d1}). \tag{29}$$

It can be seen that the signal output from the correlator can be determined from the ambiguity function.

In order to generalize the problem to N particles scattering the transmitted signal $\tilde{s}(t)$, we must introduce additional assumptions:

1. A signal scattered by any particle takes the form

$$y_i(t) = \tilde{A}_i s(t-t_{0i})e^{j2\pi f_i(t-t_{0i})}, \tag{30}$$

where $\tilde{A}_i$ is a complex scattering factor.

2. The scattering is isotropic and of the first order (i.e. independence of events), thus for N particles

$$\tilde{u}^N(t) = \sum_{i=1}^{N} \tilde{u}_i(t). \tag{31}$$

3. The scattering is a random process for which the mean value (expected value) of the envelope of the signal output is $E[u^N(t)]$.

When we assume that the number of scattering particles is a random variable, the square of the envelope of the global signal $|\tilde{u}(t)|^2$ takes the form

$$|\tilde{u}(t)|^2 = \sum_{N=1}^{\infty} |\tilde{u}^N(t)|^2 p(N/v), \tag{32}$$

while its expected value is

$$\begin{aligned} E[|\tilde{u}(t)|^2] &= \sum_{N=1}^{\infty} NE[|\tilde{A}^2 \cdot |\chi(\theta, f_d)|^2] p(N/v) \\ &= k\varrho v \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} |\chi(\theta, f_d)|^2 p(t_0, f_d)\, dt_0\, df_d, \end{aligned} \tag{33}$$

where

$$k = E[|\tilde{A}|^2], \qquad \varrho v = \sum_{N=1}^{\infty} NP(N/v),$$

$p(t_0, f_d)$ denotes the probability density function of the appearance of a particle at a time t_0 which is scattering with the frequency f_d.

Expression (33) represents the relation between the signal output from the correlator, the ambiguity function and the scattering function.

The most suitable for the purposes of profilometry (the determination of the velocity v as a function of the radius r of the blood vessel) would be an ambiguity function with the Dirac distribution, thus having the form $|\chi(t_0, f_d)|^2 = \delta(t_0, f_d)$, since expression (33) would then become

$$E[|\tilde{u}(t)|^2] = k\varrho v p(t_0, f_d). \tag{34}$$

This case, however, is not physically feasible.

Several systems have an ambiguity function with a distribution to ensure the required precision of both t_0 and f_d in one measurement, e.g.: noise emission with a Gaussian distribution, pseudo-noise with a band limited Gaussian distribution («coloured» noise), and the coherent emission of high-frequency pulses.

4. The demodulation and measurement of the Doppler frequency

It has been shown in section 2 that a satisfactory condition for the determination of the instantaneous velocity flow profile, and thus also of the instantaneous mean flow velocity and the delivery, there is the determination of the function $p(f_d, t_0)$. The measurement of the delay time t_0 has been explained in section 3.

The problem of the measurement of the Doppler frequency f_d requires a separate discussion.

It can be generally assumed that the blood volume, as seen by the receiving transducer in a time τ_2 equal to the duration of the analyzing gate, has a finite

value and that the erythrocytes, contained in this blood volume, are moving at different speeds. Consequently, the signal scattered by the erythrocytes consists of components with different frequencies. It is thus necessary to replace the discrete Doppler frequency f_d by a frequency spectrum Δf_d.

When using a continuous wave Doppler ultrasonic flowmeter, the problem of the measurement of the mean Doppler frequency is very complicated. This follows from the fact that the spectrum of the signal scattered by the blood-corpuscles flowing across the whole cross-section of the vessel is very wide. For example, in the ascending aorta it varies from 0 to 7 kHz and in the carotid artery from 0 to about 4 kHz, at a transmitting frequency of 8 MHz and an angle of inclination θ between the ultrasonic beam and the blood vessel of about 60° [10].

The ultrasonic signal scattered by blood vessels, when a pulse technique is used, has a very narrow frequency spectrum [10], since in the section selected for measurement (which is determined by the duration of the transmitted pulse and the analyzing gate) the distribution of the blood flow velocity is small.

The received signal is a random one with a Gaussian distribution whose amplitude is stationary (for continuous flow) or quasi-stationary (for real blood flow).

Thus, according to the definition of RICE [14] of the mean frequency of a signal with a given finite, fixed, continuous frequency spectrum, the mean frequency of the signal scattered by the flowing blood is given by

$$f_{\text{śr}} = \frac{\int_{-\infty}^{+\infty} fP(f)\,df}{\int_{-\infty}^{+\infty} P(f)\,df}, \tag{35}$$

where f denotes the frequency of the signal scattered, while $P(f)$ is the signal spectral density.

The mean Doppler frequency is given by the expression

$$\Delta f_{d\text{śr}} = f_{\text{śr}} - f_n. \tag{36}$$

Substituting $f_n + \Delta f$ for f in (35), and then substituting expression obtained in (36), we obtain

$$\Delta f_{d\text{śr}} = \frac{\int_{-\infty}^{+\infty} \Delta f P(f_n + \Delta f)\,d\Delta f}{\int_{-\infty}^{+\infty} P(f_n + \Delta f)\,d\Delta f}. \tag{37}$$

On the basis of the results obtained be REEVROS [15] an original measuring system (Fig. 5) was developed which directly measures the mean Doppler frequency given by expression (37) [8].

The system measuring the mean Doppler frequency has not so far found widespread use due to its complex structure.

Electronic frequency meters generally make use of the techniques for counting the number of crossings of the measured signal through the zero level. This is the so-called zero-crossing technique.

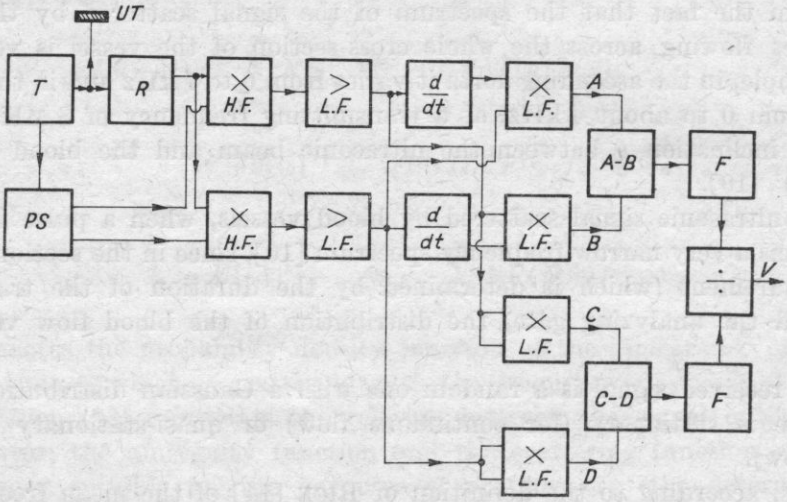

Fig. 5. Block diagram of the system for measuring the mean Doppler frequency
T – transmitter, UT – ultrasonic transducer, R – receiver, PS – phase shifter, F – filter, v – the signal proportional to the flow rate

As a result of mixing the total transmitted signal with the received signal which has been scattered by the flowing blood corpuscles, information is lost on the sign of the Doppler frequency f_d, and thus on the direction of the blood flow relative to the direction of the incident ultrasonic wave.

In order to obtain this information, McLEOD [7] and PERONNEAU [10] independently proposed a zero-crossing system with phase-sensitive detection.

The measurement of the frequency by means of the zero-crossing technique is, however, burdened with an error. RICE [14] has shown that the expected value $E[N]$ of crossings through the zero level of a signal with a random Gaussian distribution of amplitude is given by the formula

$$E[N] = 2\left[\frac{\int_{-\infty}^{+\infty} f^2 P(f)\,df}{\int_{-\infty}^{+\infty} P(f)\,df}\right]^{1/2}, \tag{38}$$

and the measured Doppler frequency is determined by the relation

$$f_d = \tfrac{1}{2}E[N]. \tag{39}$$

The response of the zero-crossing system to an input signal with a finite spectrum is thus proportional to the second moment of the spectrum of the measured signal.

Additional errors in the measurement of the Doppler frequency by the zero-crossing technique are accounted for by interference, by the noise of the input stage of the high frequency amplifier, the level of the release of the Schmitt flip-flop multivibrators, and also by the ratio of the ultrasonic bandwidth to the cross-section of the blood vessel.

5. The block diagram and description of the operation of the pulsed ultrasonic Doppler flowmeter

The results of the theoretical analysis permitted the construction and development of a pulsed ultrasonic Doppler flowmeter, model type UDIMP-1 whose block diagram is shown in Fig. 6.

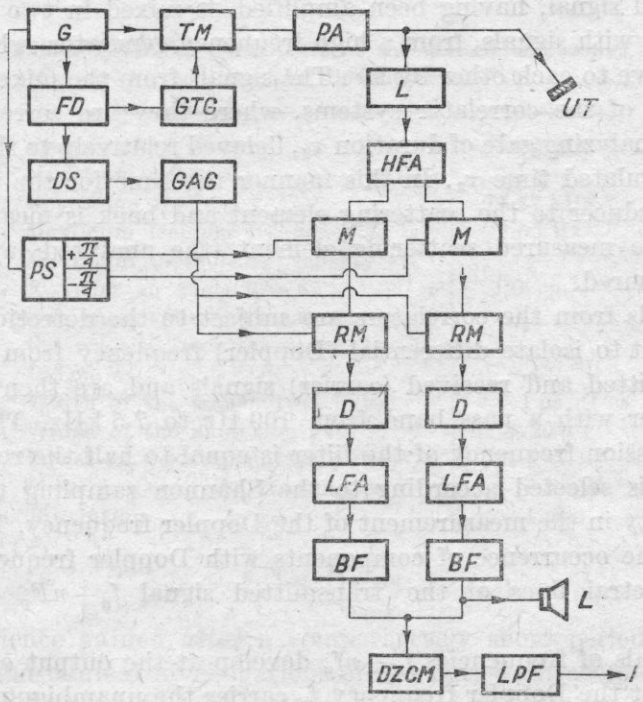

Fig. 6. Block diagram of the pulsed ultrasonic flowmeter type UDIMP-1

G – generator, *FD* – frequency divider, *GTG* – the generator of the transmitter gate, *TM* – the modulator of the transmitter, *PA* – power amplifier, *L* – limiter, *HFA* – high frequency voltage amplifier, *M* – mixer, *PS* – phase shifter, *DS* – delay system, *GAG* – the generator of the analyzing gate, *RM* – the modulator of the receiver, *D* – detector, *LFA* – low frequency amplifier, *BF* – band-pass filter, *DZCH* – directional zero-crossing meter, *LPF* – low-pass filter, *R* – recorder, *L* – loudspeaker, *UT* – ultrasonic transducer

A signal from the high frequency generator $f_N (f_N = 7{,}8$ MHz) is supplied to the frequency divider system (division by 2^9) to produce a signal in the form of a sequence of pulses with a repetition frequency $F_p (F_p \cong 15.2$ kHz).

The high frequency signal from the generator is modulated in the modulator by a signal from the transmitter gate generator. The high frequency pulses f_n, with a repetition frequency of F_p obtained in this manner, having been amplified in a wideband aperiodical power amplifier, are supplied to the ultrasonic transducer where they are transformed into ultrasonic signals and emitted in the direction of the blood vessel. These signals, having been reflected from the walls of the blood vessel and scattered by the blood corpuscles, return to the ultrasonic transducer where they are transformed into electric signals.

When the ultrasonic wave is scattered by the flowing blood, the frequency of the scattered signal changes proportionally to the velocity of the moving blood corpuscles. This signal is supplied via a limiter to a high frequency amplifier. The limiter used in the input of the high frequency amplifier prevents from the saturation of the receiver during the transmitting pulse (the so called dead zone).

The useful signal, having been amplified, is mixed in two identical summation mixers with signals, from a high frequency generator, which are shifted in phase relative to each other by 90°. The signals from the mixers are supplied to two tracks of the correlator systems, where they are correlated with the pulses of the analyzing gate of duration τ_2, delayed relatively to the transmitting pulse by a regulated time τ_3. In this manner the time for the pulse to travel from the transducer to the scattering element and back is measured, and the position of the measured scattering element (the analyzed layer of flowing blood) is measured.

The signals from the correlators are subject to the detection in a sample and hold circuit to isolate differential (Doppler) frequency from the frequency of the transmitted and received (carrier) signals and are then supplied to a band-pass filter with a pass band from 100 Hz to 7.5 kHz. The upper limit of the transmission frequency of the filter is equal to half the repetition frequency F_p and is selected according to the Shannon sampling theory, and to avoid ambiguity in the measurement of the Doppler frequency. This ambiguity results from the occurrence of components with Doppler frequencies near the successive spectral lines of the transmitted signal $f_n + nF_p$ (where $n = 0, \pm 1, \pm 2, \ldots$).

Thus signals of frequencies $f_d \pm nf_p$ develop at the output of the detector, one of which at the Doppler frequency f_d carries the unambiguous information on the flow velocity.

The attenuation of the filter above 7.5 kHz is 80 dB/octave.

The attenuation below the lower limiting transmission frequency of the filter is not critical and is selected experimentally to optimize isolation from the industrial interference of the mains and the Doppler frequency which

results from the reflection of the ultrasonic wave from the pulsating walls of the blood vessel. This limits the lowest measurable flow velocity to about 1 cm/s.

The signals from the band-pass filters are supplied to a phase-sensitive zero-crosing system. In this system the frequencies of the Doppler signal are transformed into a voltage whose amplitude is proportional to the measured Doppler frequency, with a positive or negative sign depending on the direction of the blood flow in the vessel. This signal, after filtration in the band 0-25 Hz, is supplied to a $y-t$ recorder that records the constant component.

The oscilloscope coupled to the device permits simultaneous observation of the echoes of the signal reflected from the walls of the blood vessel and of the position of the pulse of the analyzing gate.

Provision has also been made in the device to permit audiomonitoring of the Doppler frequency signal and thus of the blood flow.

The main technical parameters of the described apparatus are given in the Table.

Table. Technical parameters of the pulsed ultrasonic Doppler flowmeter

Frequency of receiver	7.8 MHz
Maximum range d_{max}	5.1 cm
Repetition frequency F_p	15.23 kHz
Maximum Doppler frequency f_d	7.6 kHz
Maximum measurable velocity V_{max}(for an angle $\theta = 67°$)	185 cm/s
Non-linearity in the measurement of the velocity over the range 2 cm/s-185 cm/s	~1%
Width of the transmitted pulse	1 μs
Width of the analyzing gate	1-20 μ
Power in the transmitted pulse	1-7 W
Mean power of the transmitter	15-100 mW
Resolution	~1 mm

6. Conclusions

The experience gained after a comparatively short period (2 1/2 years) of laboratory and clinical investigations of the pulsed ultrasonic Doppler flowmeter type UDIMP-1 has fully confirmed its practicability in the diagnosis of vessel disease.

Among the particular advantages of the device let us note its ability to make transcutaneous measurements of the blood flow velocity and of the diameter of the examined vessel, and thus of the volume blood flow in the blood vessel.

The investigations have confirmed the suitability of the device as a means of investigating the volume flow capacity of the vessels and vessel protheses, and also in the recognition and location of artery and vein fistulae.

It should be stressed that the technique presented here permits information to be obtained on the flow in a selected vessel situated deep in the patient's body when in the way of the ultrasonic beam another vessel occurs which, if a continuous wave Doppler method was used, would produce a masking effect, thus preventing from the measurement of velocity in the chosen deeper vessel.

References

[1] D. H. BERGEL, *Cardiovascular fluid dynamics*, vol. 1., Academic Press, London 1972.

[2] K. BORODZIŃSKI, K. FILIPCZYŃSKI, A. NOWICKI, T. POWAŁOWSKI, *Quantitative transcutaneous measurements of blood flow in carotid artery by means of pulse and continuous wave Doppler methods*, Pergamon Press, Ultrasound in Medicine and Biology, **2**, 189–193 (1976).

[3] W. R. BRODY, I. D. MEINDL, *Theoretical analysis of the C.W. Doppler ultrasonic flowmeter*, IEEE Trans. on Biomedical Engng. **21**, 3, 183–192 (1974).

[4] D. L. FRANKLIN, W. A. SCHLEGAL, R. F. RUSHMER, *Blood flow measured by Doppler frequency shift of backscattered ultrasound*, Science, 134, 564–565 (1961).

[5] M. M. LIH, *Axial migration of flow of suspended particles: interaction between particles*, Proc. 8th ICMBE, 117–7, July 20–25, Chicago, Illinois 1969.

[6] D. S. MdDONALD, *Blood flow in arteries*, Publ. Edward Arnold Ltd. London 1974.

[7] F. Jr. MCLEOD, *A directional Doppler Flowmeter*, Digest 7th Int. Conf. Med. and Biol. Engng, Stockholm 1967, p. 271.

[8] A. NOWICKI, P. A. PERONNEAU, Y. P. BOURNAT, A. BARBET, *The evaluation of a new method of frequency detection for Doppler blood flow velocity techniques* (in Polish.) Proc. 20th Open Seminar on Acoustics, September, Poznań – Mierzyn 1973.

[9] A. NOWICKI, K. BORODZIŃSKI, *Quantitative steady flow measurements with pulsed ultrasonic Doppler velocity meter*, Proc. Sciantific. Conf. of the Faculty of Electr. Engng. Tech. Univ. Brno, B-57, 193–202, June 3–5, Brno 1975.

[10] P. A. PERONNEAU, F. LEGER, M. XHAARD, M. PELLET, SCHWARTZ P. V., *Velocimetre sanguine a effect Doppler a emission ultrasonore pulse*, L'onde Electrique, **50**, 5, 369–389 (1970).

[11] P. A. PERONEAU, M. XHAARD, A. NOWICKI, M. PELLET, Ph. DELOUCHE, J. HINGLAIS, *Pulsed Doppler ultrasonic flowmeter and flow patterns analysis*, in: *Blood Flow Measurements*, Colin Roberts Ed., Publ. Sector Publishing Ltd., London 1972, p. 24–28.

[12] P. A. PERONNEAU, M. XHARD, A. NOWICKI, M. PELLET, Ph. DELOUCHE, J. HINGLAIS. *The application of a pulsed ultrasonic Doppler flowmeter in the analysis of the velocity profile of a fluid in tubes with rigid walls* [in Polish], Archiwum Akustyki, **8**, 2, 311–320 (1973).

[13] J. M. REID, R. A. SIEGELMANN, M. G. NASSER, D. W. BAKER, *The scattering of ultrasound by human blood*, Proc,. 8th ICMBE, July 20–25, Chicago-Illinois 1969.

[14] S. O. RICE, *Mathematical analysis of random noise*, Bell System Tech. J., **23**. 282–332 (1944).

[15] J. M. J. G. ROEVROS, *Analogue processing of C. W. Doppler flowmeter signals*

to determine average frequency shift momentaneously without the use of a wave analyser, in: *Cardiovascular applications of ultrasound*, Ed. R. S. Reneman, North Holland Publ. Comp,. Amsterdam 1974, p. 43–54.

[16] M. I. SKOLNIK, *Introduction to radar system*, Publ. McGraw Hill Comp. Inc., New York 1962.

[17] J. P. WOODCOCK, *The transcutaneous ultrasonic flow velocity meter in the study of arterial blood velocity*, in: *Ultrasonics in biology and medicine*, Ed L. Filipczyński, PWN, Warszawa–Poznań 1972, 251–260.

[18] P. M. WOODWARD, *Introduction to the theory of information and its application to radar systems* [in Polish], PWN, Warszawa 1959.

Received on 24th February 1977

ARCHIVES OF ACOUSTICS
2, 4, 325 (1977)

THE DEPENDENCE ON PRESSURE OF THE VELOCITY OF ULTRASONIC WAVES IN LIQUIDS AND THE AVAILABLE VOLUME OF MOLECULES *)

EUGENIUSZ SOCZKIEWICZ

Institute of Physics, Silesian Technical University (Gliwice)

Using the Kittel formula for the pressure coefficient of the velocity of propagation of ultrasonic waves in liquids and the Kuczera formulae for the temperature coefficients of the velocity of ultrasound, a relation has been derived between the «available volume» of the liquid molecules and the repulsion exponent in the Lennard-Jones potential of the molecularinteraction. The values of the available volume» for a number of liquids have been determined. It has also been shown that the degree of space filling increases in a homologous series of saturated hydrocarbons with the increasing number of the homologue.

1. Introduction

An important quantity describing some acoustical properties of liquids, for instance in addition to the velocity of propagation, the attenuation of ultrasonic waves, and their dependence on temperature and pressure, is the so-called free available molar volume Va of liquid molecules [3]. According to SCHAAFFS [8], p. 242, it can be defined as a difference between the molar volume V of the liquid and the specific volume B of the molecules per mole of the liquid:

$$V_a = V - B. \tag{1}$$

This volume should not be identified with the so-called «free volume» encountered in model theories of liquids, defined as the region in which the centre of mass of a molecule can move.

KITTEL [3] taking advantage of the Tonks equation of state

$$pV(1-\theta^{1/3}) = RT, \tag{2}$$

*) The paper is written in the framework of the problem MR I-24.

where p denotes the pressure, R is the gas constant, T — the absolute temperature, and $\theta = B/V$, has derived a relation between the velocity of propagation of ultrasonic waves in a liquid w and the «available volume»,

$$w = \frac{V}{V_a}\left(\frac{3\varkappa_c}{\varkappa_g}\right)^{1/2} w_g \tag{3}$$

where $\varkappa_c$ denotes the ratio of the specific heats of a liquid at constant pressure and constant volume, $\varkappa_g$ is the ratio of the specific heats at constant pressure and constant volume of the given substance in the gaseous state, while w_g is the velocity of ultrasound in the gas,

$$w_g = \left(\frac{\varkappa_g RT}{M}\right)^{1/2}, \tag{4}$$

where M denotes the molar mass.

2. The dependence of the velocity of propagation of ultrasound in a liquid on the pressure

For the pressure coefficient of the velocity of propagation of the ultrasonic waves in a liquid, Kittel derived from (3) the formula

$$\frac{1}{w}\left(\frac{\partial w}{\partial p}\right)_T = \left(\frac{V}{V_a} - 1\right)\beta_T, \tag{5}$$

where β_T denotes the isothermal coefficient of compressibility of the liquid, and the remaining symbols are as previously.

The pressure coefficient of the velocity of ultrasound can be related to the temperature coefficients, at constant pressure and at constant volume, of the propagation velocity of ultrasonic waves in a liquid [9].

Writing $(\partial w/\partial T)_p$ in the form of a Jacobian $\partial(w, p)/\partial(T, p)$ (cf. [5], p.69) and using the transformation

$$\frac{\partial(w, p)}{\partial(T, p)} = \frac{\partial(w, p)}{\partial(T, V)}\,\frac{\partial(T, V)}{\partial(T, p)}, \tag{6}$$

we obtain

$$\left(\frac{\partial w}{\partial T}\right)_p = \left(\frac{\partial w}{\partial T}\right)_V - \left(\frac{\partial w}{\partial p}\right)_T \left(\frac{\partial p}{\partial T}\right)_V. \tag{7}$$

Since

$$\left(\frac{\partial p}{\partial T}\right)_V = \frac{\partial(p, V)}{\partial(p, T)}\,\frac{\partial(p, T)}{\partial(T, V)} = \frac{\alpha}{\beta_T}, \tag{8}$$

where α denotes the cubic expansion coefficient of the liquid, from (7) and (8) we obtain

$$\frac{1}{w}\left(\frac{\partial w}{\partial p}\right)_T = \left[\frac{1}{w}\left(\frac{\partial w}{\partial T}\right)_V - \frac{1}{w}\left(\frac{\partial w}{\partial T}\right)_p\right]\frac{\beta_T}{\alpha}. \tag{9}$$

3. The dependence between the available volume and the intermolecular potential

We assume that intermolecular action in the liquid is described by the generalized Lennard-Jones potential

$$\Phi(r) = -\frac{A}{r^6} + \frac{D}{r^n}, \tag{10}$$

where r denotes the intermolecular distance, n — the individual exponent for each liquid describing the magnitude of the repulsion forces of the molecules, and A and D are constants.

Using the KUCZERA [4] relations for the temperature coefficients of the velocity of ultrasound:

$$\frac{1}{w}\left(\frac{\partial w}{\partial T}\right)_p = -\frac{n}{6}\alpha, \tag{11}$$

and

$$\frac{1}{w}\left(\frac{\partial w}{\partial T}\right)_V = \frac{7}{6}\alpha, \tag{12}$$

we obtain from formula (9) the following expression for the pressure coefficient of the velocity of ultrasonic waves:

$$\frac{1}{w}\left(\frac{\partial w}{\partial p}\right)_T = \frac{n+7}{6}\beta_T. \tag{13}$$

Comparing formulae (5) and (13) we see that there is a relation between the available volume of the liquid molecules and the exponent n of the repulsion in the intermolecular potential:

$$V_a = \frac{6}{n+13}V. \tag{14}$$

The values of the available molar volume V_a, evaluated from the above formula, are given in the Table, the data on the velocity of ultrasonic waves and the temperature coefficients of the velocity being taken from [1, 6], and the values of the cubic expansion coefficients α from [7]. The values of pressure coefficients of the velocity of propagation of ultrasonic waves $w^{-1}(\partial w/\partial p)_T$, evaluated from the Kittel formula (5) by substituting in it the ratios V_a/V determined in this paper, are in good agreement with the experimental values of these coefficients.

Table. Values of the available volume of the liquid molecules obtained from formula (14) for $T = 293$ K

Substance	$V \times 10^2[\mathrm{m}^3]$	$a \times 10^5$ $[\mathrm{K}^{-1}]$	$w[\mathrm{ms}^{-1}]$	$-\left(\frac{\partial w}{\partial T}\right)_p$	n	V_a/V	$V_a \times 10^6$ $[\mathrm{m}^3]$
n-pentane	115.202	161	1008	4.2	15.5	0.210	24.19
n-hexane	130.770	135	1116	4.4	17.5	0.197	25.76
n-heptane	146.520	124	1154	4.0	16.8	0.201	29.45
n-octane	162.720	116	1192	3.9	16.9	0.200	32.54
n-nonane	173.440	107	1234	3.8	17.3	0.198	34.34
n-decane	194.670	101	1255	3.7	17.5	0.197	38.43
n-dodecane	226.760	96	1300	3.7	17.8	0.195	44.22
n-tetradecane	260.200	89	1331	3.6	18.2	0.192	49.96
n-hexadecane	291.930	80	1358	3.5	19.3	0.186	54.30
benzene	88.968	123	1324	4.8	17.5	0.197	17.53
bromoform	87.463	91	931	2.2	15.3	0.212	18.54
chloroform	80.281	127	1001	3.4	16.1	0.206	16.54
cyclohexane	108.161	119	1277	4.6	18.2	0.193	20.87
carbon tetrachloride	96.567	123	938	3.0	16.1	0.206	19.89

4. Conclusions

The values of the admitted volume of the liquid molecules are determined by the values of the repulsion exponent in the intermolecular potential. The Table shows that in the case of a homologous series of saturated hadrocarbons the ratio V_a/V decreases with increasing length of the chain that forms a molecule of a given chemical compound.

This indicates that the degree of space filling increases with increasing number of the homologue and confirms the conclusions of SCHAAFFS [8], p. 254, and BONDI [2] on space filling in a homologous series of saturated hydrocarbons, derived by other means. The author of this paper has shown [10], p. 104, that the ratio of the free volume to the molar volume also decreases in the homologous series of saturated hydrocarbons with increasing number of the homologue.

References

[1] A. BERGMAN, *Ultrasound and its application in science and technology* in [Russain], Izdatelstwo Inostrannoj Literatury, Moscow 1967, p. 238.

[2] A. BONDI, *Free volumes and free rotation in simple liquids and liquid saturated hydrocarbons*, J. Phys. Chem., **58**, 929 (1954).

[3] Ch. KITTEL, *Ultrasonic propagation in liquids. II. Theoretical study of the free volume model of the liquid state*, J. Chem. Phys., **14**, 614 (1946).

[4] F. KUCZERA, *Thermic properties of the sound velocity in liquids of constant density* [in Russian], Prim. ultraak. k issled. vescestva, **13**, 277 (1969).

[5] L. LANDAU, E. LIFSZYC, *Statistical physics* [in Polish], PWN, Warszawa 1959.

[6] LANDOLT-BÖRNSTEIN, *Zahlenwerte und Funktionen aus Naturwissenschaften und Technik*, Gruppe II, Band 5, Springer Verlag, Berlin 1967.

[7] LANDOLT-BÖRNSTEIN, *Zahlenwerte und Funktionen aus Physik, Chemie, Astronomie, Geophysik und Technik*, Band 2, 1 Teil, Springer Verlag, Berlin 1971.

[8] W. SCHAAFFS, *Molekularakustik*, Springer Verlag, Berlin 1963, s. 242.

[9] W. SCHAAFFS, ibid., s. 254.

[10] E. SOCZKIEWICZ, *The generalized Lennard-Jones potential and the acoustic properties of a liquid* [in Polish], Thesis, IPPT PAN, Warszawa 1973.

[11] E. SOCZKIEWICZ, *A new acoustical method of the determination of the free volume of a liquid* [in Polish], Proceedings of the 20th Open Seminar on Acoustic, Poznań-Mierzyn 1973, part II.

Received on 1st June 1977

C H R O N I C L E

Book Review

Musical Acoustics, Part I, Violin Family Components
Musical Acoustics, Part II, Violin Family Functions

Edited in the series: Benchmark Papers in Acoustics, Carleen M. Hutchins, editor. Publisher: Dowden, Hutchinson and Ross; Halsted Press (Wiley) USA, Part I, 1975, p. 476, Part II, 1976, p. 379

These two volumes on musical acoustics are devoted to the acoustics of stringed instruments. The editor of both volumes and the author of a part of one book is Carleen M. Hutchins, Permanent Secretary of the Catgut Acoustical Society.

This efficient and active Society links both scientists and artists interested in research on musical instruments.

The volumes under review were conceived as a selection of the most important contributions dealing with the acoustics of the violin and other related instruments.

Part I, *Violin Family Components*, consists of 29 articles arranged in six subgroups.

1. *General Papers on Violin Acoustics* by HUTCHINS 1962, SAUNDERS 1937, MEINEL 1957, ITOKAWA and KUMAGAI 1952, SCHELLENG 1963, CREMER 1971.

2. *The Bowed String* by RAMAN 1916, 1918, 1921, BLADIER 1931, 1964, SCHELLENG 1973, CREMER 1972, 1973, FIRTH and BUCHANAN 1973.

3. *The Bridge* by MINNAERT and VLAM 1937, BLADIER 1960, REINICKE 1973.

4. *The Soundpost* by SAVART 1840, SCHELLENG 1971.

5. *Wood for the Violin Family* by KRÜGER and ROHLOFF 1938, ROHLOFF 1940, BARDUCCI and PASQUALINI 1958, FUKEDA 1950, 1951, BELDIE 1968, GHELMEZIU and BELDIE 1972, FRYXELL 1965.

6. *Varnish* by SCHELLENG 1968.

Part II, *Violin Family Functions*, consists of 33 articles grouped into an introduction and three subparts.

Introduction by HUTCHINS 1973.

1. *Body Vibrations* by SAVART 1840, BACKHAUS 1930, 1931, MEINEL 1937, EGGERS 1959, BELDIE 1969, AGREN and STETSON 1972, REINICKE and CREMER 1970, HUTCHINS, STETSON and TAYLOR 1971, JANSSON, MOLIN and SUNDIN 1970, JANSSON 1973.

2. *Radiation* by SCHELLENG 1968, MEINEL 1937, SAUNDERS 1946, 1953, HUTCHINS, HOPPING and SAUNDERS 1960, LOTTERMOSER and MEYER 1957, 1962, LOTTERMOSER 1958, 1968, MEYER 1972, LEIPP and MOLES 1959, KOHASI and TOKITA 1954, FLETCHER, BLACKHAM and GEERSTEN 1965, FLETCHER and SANDERS 1967, JANKOWSKI 1966, MATHEWS and KOHUT 1973.

3. *Musical Focus* by HUTCHINS 1967, HUTCHINS and SCHELLENG 1967, SAUNDERS 1940, ROHLOFF 1964, BOOMSLITER and CREEL 1972.

The selection presented is only a small part of a vast literature devoted to the accoustics of stringed instruments. The selection is relevant, presents large range of topics and permits acquaintance with the works which, while often referred to, are practically unavailable to interested readers, e.g. the remarkable works of Savart from the first half of the last century and some fundamental articles by BACKHAUS and SAUNDERS. Each article or group of articles is provided with editorial comments. Concise information about authors as well bibliographical references of their works dealing with related subjects are given. Also the references to selected literature by other authors are cited.

The original language of 42 of the articles (out of 62) is English, some are translations into English (including Japanese works). The remainder are written in German (15), French (4) and Italian (1). The se papers have summaries or shortened versions in English.

Among the various famous works published in these volumes, the most striking ones are the pioneer works of F. SAVART which present valuable methods and observations, the fundamental work of SAUNDERS, and a remarkable article on wood properties by I. BARDUCCI and G. PASQUALINI.

It should be also pointed out that the interest in acoustics of stringed instruments has much increased in the 1970's. New outstanding papers included are e.g. by J. C. SCHELLENG, L. CREMER, E. V. JANSSON, C. M. HUTCHINS, and others.

This book, being an excellent and carefully compiled reader, is on the one hand evidence for the growing interest in the field and, on the other hand, it may be considered as an enlightening guide for those who intend to direct their research to the field of the acoustics of musical instruments.

Andrzej Rakowski (Warszawa)