ARCHIVES OF ACOUSTICS
6, 4, 329-338 (1981)

A RECURSIVE METHOD FOR THE DETERMINATION OF THE OUTPUT POWER SPECTRUM OF SOME NONLINEAR SYSTEMS

ANDRZEJ GABOR, JAN ZARZYCKI

Institute of Telecommunication and Acoustics, Technical University of Wrocław
(50-317 Wrocław, ul. B. Prusa 53/55)

This paper presents a new method for the determination of the output power spectral density of different types of cascading a linear system with memory and a nonlinear system without memory, with a stationary Gaussian input.

This method permits uniform determination of the output spectrum for each type of the systems mentioned above. The recursive form of the resulting expressions also permits fast calculation of the output spectral density for higher orders of nonlinearity.

1. Introduction

This paper presents a new method for the determination of the output power spectral density for different types of cascading a linear system with memory and a nonlinear no-memory system with a stationary Gaussian input. Different methods for calculating the power spectral density in the cases mentioned above have already been discussed [2, 6, 7]. They are relatively complex, however, and for each type of the cascade system a different approach has been used, becoming essentially more complicated with the increasing nonlinearity of the system.

The method presented here is based on the known [1] method for the derivation of the output power spectral density of a general nonlinear system with memory with a stationary Gaussian input. The method proposed gives simple and, equally importantly, uniform determination of the power spectral density.

2. The relation between the output power spectral density and the multidimensional transfer functions of cascade-systems

The starting point is a general nonlinear system with memory [1, 3, 4] described by the following set of multidimensional transfer functions

$$\{H_n(\boldsymbol{f}^n)\}, \quad n = 1, 2, \ldots, \tag{1}$$

where $\boldsymbol{f}^n = \{f_1, \ldots, f_n\}$.

The functions $H_n(f_1, \ldots, f_n)$ are symmetrical with respect to their arguments. They are generalizations of the transfer function $H_1(f)$ of a linear-memory system to include the case of a nonlinear-memory system [1, 3-5]. This generalization results directly from the Volterra series approach to the problem of a nonlinear-memory system representation.

When the input $x(t)$ is a stationary ergodic zero-mean Gaussian with the two-sided power spectral density $W_x(f)$, the two-sided power spectral density $W(f)$ of the nonlinear-memory system output $y(t)$ is given by the following expression [1]

$$W_y(f) = |b_0|^2 \delta(f) + \sum_{n=1}^{\infty} \frac{1}{n!} \int_{-\infty}^{+\infty} d\boldsymbol{f}^n \, |b_n(\boldsymbol{f}^n)|^2 \delta(f - \boldsymbol{f}^n) \prod_{i=1}^{n} W_x(f_i), \tag{2}$$

where

$$\int_{-\infty}^{+\infty} d\boldsymbol{f}^n = \int_{-\infty}^{+\infty} df_1 \ldots \int_{-\infty}^{+\infty} df_n, \tag{3}$$

$$\delta(f - \boldsymbol{f}^n) = \delta(f - f_1 - \ldots - f_n), \tag{4}$$

$$b_0 = \sum_{m=1}^{\infty} \frac{1}{m! 2^m} \int_{-\infty}^{+\infty} d\boldsymbol{f}^m H_{2m}(\boldsymbol{f}^m, -\boldsymbol{f}^m) \prod_{i=1}^{m} W_x(f_i); \tag{5}$$

$$b_n(\boldsymbol{f}^n) = H_n(\boldsymbol{f}^n) + \sum_{m=1}^{\infty} \frac{1}{m! 2^m} \int_{-\infty}^{+\infty} d\boldsymbol{k}^m H_{2m+n}(\boldsymbol{f}^n, \boldsymbol{k}_m - \boldsymbol{k}^m) \prod_{i=1}^{m} W_x(k_i),$$
$$n = 1, 2, \ldots, \tag{6}$$

$$H_{2m}(\boldsymbol{f}^m, -\boldsymbol{f}^m) = H_{2m}(f_1, -f_1, \ldots, f_m, -f_m), \tag{7a}$$

$$H_{2m+n}(\boldsymbol{f}^n, \boldsymbol{k}^m, -\boldsymbol{k}^m) = H_{2m+n}(f_1, \ldots, f_n, k_1, -k_1, \ldots, k_m, -k_m). \tag{7b}$$

Now, the forms and properties of multidimensional transfer functions will be considered for different types of cascading the linear-memory system *LI* and the nonlinear no-memory system *NB*. The formulae for the two-sided output power spectral density will then follow for each type of the cascade-system. The cascade-systems considered here consist of the linear-memory systems *LI* and *LI'* with the transfer functions $H(f)$ and $H'(f)$, respectively,

and of the nonlinear no-memory system *NB* whose nonlinearity is given by the power series

$$y(t) = \sum_{n=1}^{\infty} a_n x^n(t), \tag{8}$$

with $x(t)$ and $y(t)$ being the system input and output, respectively, and a_n denote the coefficients of the series.

A. *The system LI-NB-LI*

The multidimensional transfer functions (1) of the *LI-NB-LI* cascade-system (Fig. 1) are given by [4, 5, 8]

$$H_n(\boldsymbol{f}^n) = a_n H'\Big(\sum_{i=1}^{n} f_i\Big) \prod_{j=1}^{n} H(f_j). \tag{9}$$

It can be easily shown that the multidimensional transfer functions (9) have the following properties

$$H_{2m}(\boldsymbol{f}^m, -\boldsymbol{f}^m) = a_{2m} H'(0) \prod_{i=1}^{m} H(f_i) H(-f_i), \quad m = 1, 2, \ldots, \tag{10}$$

$$H_{2m+n}(\boldsymbol{f}^n, \boldsymbol{k}^m, -\boldsymbol{k}^m) = a_{2m+n} H'\Big(\sum_{i=1}^{n} f_i\Big) \prod_{j=1}^{n} H(f_j) \prod_{l=1}^{m} H(k_l) H(-k_l), \quad m = 1, 2, \ldots \tag{11}$$

Using (9)-(11) in (2), (5), (6), the two-sided output power spectral density of the *LI-NB-LI* cascade-system can be expressed as (see the Appendix)

$$W_y(f) = c_0 |H'(0)|^2 \delta f + |H'(f)|^2 \sum_{n=1}^{\infty} \frac{c_n}{n!} \int_{-\infty}^{+\infty} d\boldsymbol{f}^n \delta(f - \boldsymbol{f}^n) \prod_{i=1}^{n} W_x(f_i) |H(f_i)|^2, \tag{12}$$

where

$$c_n = \left| a_n + \sum_{m=1}^{\infty} \frac{1}{m! 2^m} a_{2m+n} I^m \right|^2, \quad a_0 = 0, \; n = 0, 1, \ldots, \tag{13}$$

$$I = \int_{-\infty}^{+\infty} W_x(f) |H(f)|^2 df.$$

B. *The system LI-NB*

The two-sided output power spectral density for the cascade-system *LI-NB* follows from (12), if $H'(f) = 1$ for all values of the argument f, and is given by

$$W_y(f) = c_0 \delta(f) + \sum_{n=1}^{\infty} \frac{c_n}{n!} \int_{-\infty}^{+\infty} d\boldsymbol{f}^n \delta(f - \boldsymbol{f}^n) \prod_{i=1}^{n} W_x(f_i) |H(f_i)|^2. \tag{14}$$

The coefficients c_n in formula (14) are defined by (13).

C. *The system NB-LI*

Assumption that $H(f) = 1$ for all values of the argument f in expressions (12) and (13) yields the two-sided output power spectral density for the cascade-system *NB-LI*

$$W_y(f) = d_0|H'(0)|^2\delta(f) + |H'(f)|^2 \sum_{n=1}^{\infty} \frac{d_n}{n!} \int_{-\infty}^{+\infty} df^n \delta(f-f^n) \prod_{i=1}^{n} W_x(f_i), \tag{15}$$

where

$$d_n = \left| a_n + \sum_{m=1}^{\infty} \frac{1}{m!2^m} a_{2m+n} J^m \right|^2, \quad J = \int_{-\infty}^{+\infty} W_x(f)df, \quad n = 0, 1, 2, \ldots, a_0 = 0. \tag{16}$$

D. *The system NB*

The two-sided output power spectral density for the system *NB* can be obtained from (12), assuming that $H(f) = H'(f) = 1$ for all f, and is given by

$$W_y(f) = d_0\delta(f) + \sum_{n=1}^{\infty} \frac{dn}{n!} \int_{-\infty}^{+\infty} df^n \prod_{i=1}^{n} W_x(f_i)\delta(f-f^n), \tag{17}$$

where the coefficients d_n are expressed by (16).

E. *The system LI-LI*

If $a_1 = 1$ and $a_n = 0$, $n = 2, 3, \ldots$, in (13), then

$$W_y(f) = |H(f)|^2|H'(f)|^2 W_x(f). \tag{18}$$

F. *The system LI*

The well-known formula

$$W_y(f) = |H(f)|^2 W_x(f) \tag{19}$$

results from (12) if $a_1 = 1$, $a_n = 0$, $n = 2, 3, \ldots$ and $H'(f) = 1$ for all f.

3. A recursive method for the calculation of power spectral density

The calculation of multiple integrals is necessary when formulae (12)-(17) of the previous section are used for the calculation of the power spectral density. This is rather troublesome, especially when computer-calculations are used. To avoid this inconvenience, a recursive method for the calculation of the power spectral density mentioned will be presented in this section.

The following observation is the basis for the recursive method. For $n = 2, 3, \ldots$ let

$$B_n(f) = \int_{-\infty}^{+\infty} df^n \delta(f-f^n) \prod_{i=1}^{n} A(f_i),$$

then

$$B_n(f) = \underbrace{A(f)* \dots *A(f)}_{n} = B_{n-1}(f)*B_1(f), \tag{20}$$

where

$$B_1(f) = A(f) = |H(f)|^2 W_x(f)$$

and $*$ denotes convolution.

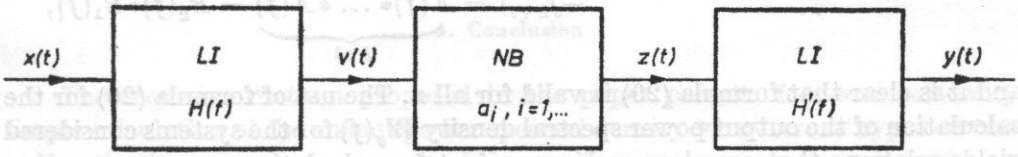


Fig. 1. The *LI-NB-LI* cascade-system

Property (20) can be proven by induction. For $n = 2$

$$B_2(f) = \int_{-\infty}^{+\infty} df^2\, \delta(f - f^2) \prod_{i=1}^{2} A(f_i).$$

Since

$$\delta(f - f^n) = \delta(f - f_1)* \dots *\delta(f - f_n), \quad n = 2, 3, \dots, \tag{21}$$

then

$$\delta(f - f^2) = \int_{-\infty}^{+\infty} \delta(l - f_1)\, \delta(f - f_2 - l)\, dl.$$

Hence

$$B_2(f) = \int_{-\infty}^{+\infty} dl \int_{-\infty}^{+\infty} A(f_1)\, \delta(l - f_1)\, df_1 \int_{-\infty}^{+\infty} A(f_2)\, \delta(f - l - f_2)\, df_2 =$$

$$= A(f)*A(f) = B_1(f)*B_1(f).$$

Now, suppose that for $n = k$

$$B_k = \int_{-\infty}^{+\infty} df^k\, \delta(f - f^k) \prod_{i=1}^{k} A(f_i) = \underbrace{A(f)* \dots *A(f)}_{k} = B_{k-1}(f)*B_1(f).$$

Then

$$B_{k-1}(f) = B_k(f)*B_1(f).$$

Indeed, since

$$B_{k+1}(f) = \int_{-\infty}^{+\infty} df^{k+1}\, \delta(f - f^{k+1}) \prod_{i=1}^{k+1} A(f_i)$$

and

$$\delta(f-f^{k+1}) = \int_{-\infty}^{+\infty} \delta(l-f^k)\,\delta(f-l-f_{k+1})\,dl,$$

then

$$B_{k+1}(f) = \int_{-\infty}^{+\infty} dl \int_{-\infty}^{+\infty} df^k\,\delta(l-f^k) \prod_{i=1}^{k} A(f_i) \int_{-\infty}^{+\infty} df_{k+1} A(f_{k+1})\,\delta(f-l-$$

$$-f_{k+1}) = \underbrace{A(f)*\ldots*A(f)}_{k+1} = B_k(f)*B_1(f),$$

and it is clear that formula (20) is valid for all n. The use of formula (20) for the calculation of the output power spectral density $W_y(f)$ for the systems considered yields relations that are clear and convenient for calculations.

A. *The system LI-NB-LI*

Application of formula (20) in (12) yields

$$W_y(f) = c_0|H'(0)|^2\,\delta(f) + |H'(f)|^2\left[c_1 B_1(f) + \sum_{n=2}^{\infty} \frac{c_n}{n!} B_{n-1}(f)*B_1(f)\right]. \tag{22}$$

The coefficients c_n are given by (13).

B. *The system LI-NB*

Application of formula (20) in (14) gives

$$W_y(f) = c_0\,\delta(f) + c_1 B_1(f) + \sum_{n=2}^{\infty} \frac{c_n}{n!} B_{n-1}(f)*B_1(f), \tag{23}$$

with the coefficients c_n expressed by (13).

C. *The system NB-LI*

It follows from equation (15) that

$$W_y(f) = d_0|H'(0)|^2\,\delta(f) + |H'(f)|^2\left[d_1 B_1(f) + \sum_{n=2}^{\infty} \frac{d_n}{n!} B_{n-1}(f)*B_1(f)\right], \tag{24}$$

where the coefficients d_n are given by (16).

D. *The system NB*

The resulting form of equation (17) is

$$W_y(f) = d_0\,\delta(f) + d_1 B_1(f) + \sum_{n=2}^{\infty} \frac{d_n}{n!} B_{n-1}(f)*B_1(f) \tag{25}$$

and the coefficients d_n are expressed by (16).

E. *The system LI-LI*
Equation (18) becomes

$$W_y(f) = |H'(f)|^2 B_1(f). \tag{26}$$

F. *The system LI*
Equation (19) yields

$$W_y(f) = B_1(f). \tag{27}$$

4. Conclusion

The recursive method proposed in the paper permits fast and relatively simple calculation of the output power spectral density for different types of cascading the liner-memory systems and the nonlinear no-memory system (with any nonlinearity order) with Gaussian inputs. The considerations in section 2 permit uniform calculation of the output power spectral density for any cascade-system. The recursive form of the final formulae, as given in section 3, permits the avoidance of troublesome complexity of formulae that arises with increasing the nonlinearity order. The calculation of multiple integrals may also be avoided.

The cascade-systems mentioned are commonly used models of nonlinear devices. These models can be implemented, for example, for the analysis of a modulator or a detector cascaded with a linear filter – the well known subsystems of radio-devices.

References

[1] E. Bedrosian, S. O. Rice, *The output properties of Volterra systems (nonlinear systems with memory) driven by harmonic and gaussian inputs*, Proc. IEEE, **59**, 12, 1688-1707 (1971).

[2] G. Bonnet, *Transformations des signaux aléatoires á travers les systèmes non linéaires, sans mémoire*, Ann. Telecom., **19**, 9-10, 203-220 (1964).

[3] A. Bagor, J. Zarzycki, *The measurements of electroacoustic system nonlinear distortion-the reasons of the present failures*, Proc. of the Int. Symp. on Measurements in Telecommunications, URSI – CNET, Lannion 1977, 621-626.

[4] A. Gabor, J. Zarzycki, *Electroacoustic systems nonlinearity and memory analysis using Volterra series*, Prace XXII Seminarium z Akustyki, Wrocław 1975, 54-58.

[5] A. Gabor, J. Zarzycki, *A new method of the electroacoustic system nonlinear distortion evaluation*, Institute of Telecommunication and Acoustics, Technical University of Wrocław, 128/K-075/77.

[6] B. R. Levin, *Tieorieticzeskije osnowy statisticzeskoj radiotechniki*, Sow. Radio, Moscow 1975.

[7] V. I. Tichonow, *Statisticzeskaja radiotehnika*, Sow. Radio, Moscow 1966.

[8] J. Zarzycki, A. Gabor, *Random signal distortion analysis using multidimensional transfer functions*, Proc. of the 9-th ICA, Madrid 1977, R15.

Received on November 2, 1979; revised version on March 11, 1981.

Appendix

1. *Derivation of formula (9)*

The relation between the input $x(t)$ and output $y(t)$ of a general nonlinear system with memory can be expressed as follows [1]

$$y(t) = \sum_{n=1}^{\infty} y_n(t), \tag{A1.a}$$

$$y_n(t) = \int_{-\infty}^{+\infty} d\mathbf{s}^n h_n(\mathbf{s}^n) \prod_{i=1}^{n} x(t-s_i), \tag{A1.b}$$

where $h_n(\mathbf{s}^n)$ is an n-dimensional inverse Fourier transform of the function $H_n(\mathbf{f}^n)$, i.e. the n-dimensional impulse response of the system.

The corresponding relation can be written for the *LI-NB-LI* cascade-system (Fig. 1). Aplication of the configuration of the system (Fig. 1) and consideration that $h_1(\cdot)$ and $h_1'(\cdot)$ are the inverse Fourier transforms of $H(f)$ and $H'(f)$, respectively, yields

$$y(t) = \sum_{n=1}^{\infty} y_n(t), \tag{A2.a}$$

$$y_n(t) = \int_{-\infty}^{+\infty} d\mathbf{s}^n \int_{-\infty}^{+\infty} dr\, a_n h_1'(r) \prod_{i=1}^{n} h(u_i - r) \prod_{i=1}^{n} x(t-u_i), \tag{A2.b}$$

where $u_i = r + s_i$.

Comparison of formulae (A1) and (A2) gives the expression describing the n-dimensional impulse response of the *LI-NB-LI* cascade-system:

$$h_n(\mathbf{u}^n) = a_n \int_{-\infty}^{+\infty} h_1'(r) \prod_{i=1}^{n} h_1(u_1 - r)\, dr. \tag{A3}$$

The n-dimensional Fourier transform of formula (A3) yields equation (9):

$$H_n(\mathbf{f}^n) = a_n H_n'\left(\sum_{i=1}^{n} f_i\right) \prod_{i=1}^{n} H_1(f_i).$$

Formulae (10) and (11) follow directly from (9), after substitution of the relevant arguments and indices.

2. *Derivation of formula (12)*

Application of property (10) in (5) yields

$$b_0 = H'(0) \sum_{m=1}^{\infty} \frac{a_{2m}}{m!\,2^m} I^m. \tag{A4}$$

Combination of formulae (9) and (11) with (6) gives

$$b_n(\boldsymbol{f}^n) = a_n H'\Big(\sum_{j=1}^{n} f_j\Big)\prod_{i=1}^{n} H(f_i) + \\ + \sum_{i=1}^{\infty}\frac{1}{m!\,2^m}\int_{-\infty}^{+\infty} d\boldsymbol{k}^m \prod_{i=1}^{m} W_x(k_i)\prod_{j=1}^{n} H(f_j)H'\Big(\sum_{l=1}^{n} f_l\Big)\, a_{2m+n} \times$$

$$\times \prod_{p=1}^{m} H(k_p)H(-k_p) = \prod_{i=1}^{n} H(f_i)H'\Big(\sum_{j=1}^{n} f_j\Big)\Big[a_n + \sum_{m=1}^{\infty}\frac{1}{m!\,2^m}\, a_{2m+n} I^m\Big]. \qquad \text{(A5)}$$

Application of expressions (A4) and (A5) in formula (2) and the use of the relation $f = \sum_{i=1}^{n} f_i$ leads to

$$W_y(f) = |H'(0)|^2\Big|\sum_{m=1}^{\infty}\frac{1}{m!\,2^m}\, a_{2m} I^m\Big|^2 \delta(f) + \\ + \sum_{n=1}^{\infty}\frac{1}{n!}\int_{-\infty}^{+\infty} d\boldsymbol{f}^n \prod_{i=1}^{n} W_x(f_i)\Big|\prod_{j=1}^{n} H(f_j)H'\Big(\sum_{l=1}^{n} f_l\Big)\Big[a_n + \sum_{m=1}^{\infty}\frac{1}{m!\,2^m}\, a_{2m+n} I^m\Big]\Big|^2 \times$$

$$\times \delta(f - \boldsymbol{f}^n) = |H'(o)|^2\Big|\sum_{m=1}^{\infty}\frac{1}{m!\,2^m}\, a_{2m} I^m\Big|^2 \delta(f) + \\ + \Big|a_1 + \sum_{m=1}^{\infty}\frac{1}{m!\,2^m}\, a_{2m+1} I^m\Big|^2 W_x(f)\,|H(f)|^2\,|H'(f)|^2 + \\ + \sum_{n=2}^{\infty}\frac{1}{n!}\Big|a_n + \sum_{m=1}^{\infty}\frac{1}{m!\,2^m}\, a_{2m+n} I^m\Big|^2 \int_{-\infty}^{+\infty} d\boldsymbol{f}^{n-1} \prod_{i=1}^{n-1} W_x(f_i)\,|H(f_i)|^2 \times \\ \times W_x\Big(f - \sum_{j=1}^{n-1} f_j\Big)\Big|H\Big(f - \sum_{j=1}^{n-1} f_j\Big)\Big|^2 |H'(f)|^2.$$

The last expression can be written in the desired form of (12).

ARCHIVES OF ACOUSTICS
6, 4, 339-358 (1981)

SOUND POWER AND RADIATION EFFICIENCY OF A CIRCULAR PLATE

STEFAN CZARNECKI

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

ZBIGNIEW ENGEL, RYSZARD PANUSZKA

Institute of Mechanics and Vibroacoustics, Academy of Mining and Metallurgy
(30-059 Kraków, ul. Mickiewicza 30)

This paper considers the problem of estimating the sound radiation by vibrating surfaces in the case of a circular plate clamped on the circumference. The aim of this paper is to verify the values of the equivalent surfaces of the plate, which were obtained theoretically, with the experimental values obtained in the free field and reverberant field conditions. The results obtained show that there is a good agreement of results for $k_{mn}a < 5$. For $k_{mn}a > 5$ an increase in the sound radiation of the plate can be observed, which results both from a complex character of the vibration of the plate and from the effect of the acoustic field on the vibration conditions of the plate.

1. Introduction

There are many methods of decreasing the noise and vibration level in the working environment. The most efficient method of noise control is decreasing the emission of sources. Decreasing or limiting the emission of sources is related to the development of effective methods of location of noise sources in machinery or devices. In the Institute of Mechanics and Vibroacoustics, Academy of Mining and Metallurgy, and in the Institute of Fundamental Technological Research Polish Academy of Sciences, of the research has long been done on the identification of the sources of vibroacoustic energy. This research is concerned with developing methods of identification of sound and vibration sources

in machinery, in order to estimate the sound power radiated by the particular parts. This estimation permits the dominating sources and their frequency response to be determined. The early investigations were concerned with the use of a correlation method in the near field conditions to estimate the sound power of surface sources [5, 7].

The practical investigations were related to the estimation of the sound power of complex mechanical systems in the chamber airless shotblasting machines.

Another problem related to the methods of investigation of the emission of sources and to the assessment of the character of the emission is the investigation of relations between the sound power radiated and the structure-borne vibration.

From the practical point of view this problem is particularly important in the case of plates used as enclosures of machinery, which are often themselves the sources of noise. In the real conditions, the plate elements have a complex boundary condition as well as a complex shape with holes, ribbing or other irregularities which make their mathematical description impossible. However, in all the cases the resonance vibrations of these plates have a dominant significance. The classical theory of plate vibration permits the determination of the resonance vibration frequency for different boundary conditions. It is, however, much more difficult to estimate the sound power emitted or the radiation efficiency of the vibrating plates. These quantities are usually determined experimentally.

There are a number of methods of measurement of the sound power of plates, of which the most significance has been recently gained by the two-microphone method based on the determination of the imaginary part of the mutual spectral power density of the sound pressures in the near field conditions. This method is very precise but requires, however, complex and expensive instrumentation.

The need, therefore, arises for a method which can preliminarily give an approximate estimation of the sound power of a plate by means of simple measurements. This estimation is possible if a piston model of a plate is assumed, for which the values of the equivalent surface of the plate were previously determined. The values of the amplitude of the vibration velocity should, however, be determined from measurement at a chosen point of the plate.

Thus on the basis of simple measurements of the plate vibration, it could be possible to estimate an approximate characteristic of its sound power as a function of vibration frequency. The method of equivalent surface determination was given by MORSE in [11] and concerned the first vibration mode of a circular membrane.

The aim of this paper is to verify through measurements the relation mentioned above and to extend it to include the higher vibration modes. Since calculations of the sound power radiated by plates of complex shape and different

clamping are rather difficult, this paper is restricted to the analysis of a circular plate clamped on the circumference. Using the classical fundamentals of plate vibration theory, calculations can be made in this case, and verified experimentally later on. This approach can contribute to deeper knowledge of plate radiation over a wide frequency range, which can provide the basis for expanding the investigations to include the more complex systems.

2. Theoretical approach

It is known that flexural vibrations of plates are characterized by resonance frequencies, the so-called vibration modes, which result from the solution of the differential equation

$$\nabla^4 \xi + \frac{12\varrho(1-\nu^2)}{Eh^2}\frac{\partial^2 \xi}{\partial t^2} = 0. \tag{1}$$

where ξ is the displacement of the vibrations of the plate, ϱ is the density of the plate material, ν is the Poisson ratio, E is the modulus of the longitudinal elasticity, and h is the plate thickness. For a circular plate excited uniformly over the whole surface clamped on the circumference, i.e. satisfying the boundary conditions $\xi = 0$ and $\partial\xi/\partial r = 0$ for $r = a$ (where r is the current radius of the plate and a is the value of the radius on the circumference of the plate), the characteristic equation can be given as

$$I_m(\gamma_a)\left(\frac{d}{dr}\right)J_m(\gamma_r) - J_m(\gamma_a)\frac{d}{dr}I_m(\gamma_r) = 0. \tag{2}$$

Solution of this equation gives a series of the values γ_{mn}, which define the eigenfrequencies f_{mn}

$$f_{mn} = \frac{\pi h}{4a^2}\sqrt{\frac{E}{3\varrho(1-\nu^2)}}\,(\beta_{mn})^2, \qquad \beta_{mn} = \gamma_{mn}\frac{a}{\pi}, \tag{3}$$

where the indices m and n denote the order of the axially nonsymmetrical and axially symmetrical modes, respectively.

Since the sound power of vibrating plates is determined mainly by the resonance vibrations, the further considerations will concern the sound power measurements for the particular vibration modes dependent on the product $k_{mn}a$, where $k_{mn} = 2\pi f_{mn}/c$ is the wave number, while c is the wave propagation velocity in the air.

One of the methods of calculating the sound power of circular plates for the particular vibration modes is the method based on the measurement of the vibration velocity v on the surface of the plate.

Considering that for higher vibration modes a complex vibration distribution occurs on the surface of the plate, it is necessary to determine the mean value of the velocity

$$\overline{v}^2 = \frac{1}{S_r} \iint_{S_r} v^2(r, \varphi)\, dr\, d\varphi, \tag{4}$$

where S_r is the surface of the plate.

The sound power radiated by the plate is thus

$$W_{\overline{v}} = \overline{v}^2 R_p S_r, \tag{5}$$

where R_p is the radiation resistance of the plate, which for a circular plate in an infinite baffle is

$$R_p = \varrho_0 c \left[1 - \frac{J_1(2k_{mn}a)}{k_{mn}a}\right], \tag{6}$$

where $J_1(k_{mn}a)$ is a Bessel function of the first kind of the first order, and ϱ_0 is the air density.

The sound power of a circular plate can also be obtained on the basis of the determination of its vibration velocity at a given point and of the equivalent surface S_{eq} so that the sound power thus determined is equal to the power expressed by formula (5).

For axially symmetrical vibrations, the best point for which the vibration velocity of the plate can be determined is its centre. Designating as v_0 the *r.m.s.* vibration velocity of the plate centre, the following expression can be written for the sound power radiated by the plate

$$W_v = v_0^2 R_p S_{eq}. \tag{7}$$

The ratio of the equivalent surface of the plate, S_{eq} and its real surface S_r can be called the coefficient of the vibration distribution of the plate

$$\varkappa = \frac{S_{eq}}{S_r}. \tag{8}$$

Hence, considering realtions (5) and (6), the expression for the equivalent surface of the plate takes the form

$$S_{eq} = \frac{W_v}{v_0^2 \left[1 - \dfrac{J_1(2k_{mn}a)}{k_{mn}a}\right]}. \tag{9}$$

The equivalent surface can also be estimated theoretically from the relation

$$S_{eq} = 2\pi \int_0^a R(r) r\, dr, \tag{10}$$

where the function $R(r)$ can be defined from the solution of equation (1).

In the case of the flexural vibrations of a thin circular plate clamped on the circumference of the radius $r = a$, the function $R(r)$ takes the form

$$R(r) = CJ_m(\gamma r) + DI_m(\varrho r), \tag{11}$$

where C and D are constant, and the quantity γ is defined by the solutions of equation (2).

Insertion of equation (11) into relation (10) gives after integration the following expression for the equivalent surface of the plate

$$S_{eq} = \frac{2\pi a}{\sqrt{k_{mn}}cD}\left[J_1(B\sqrt{k_{mn}}ca) - \frac{J(B\sqrt{k_{mn}}ca)}{I_0(B\sqrt{k_{mn}}ca)} I_1(B\sqrt{k_{mn}}ca)\right], \tag{12}$$

where

$$B = \sqrt[4]{\frac{12\varrho}{Eh^2}(1-\nu^2)}. \tag{13}$$

The equivalent surface can also be determined experimentally by the measurements of the vibration velocity at the centre of the plate (for the axially symmetrical modes) and by the determination of the sound power by one of the known methods.

3. Experimental set-up and procedure

An experimental set-up was constructed for the investigation of the radiation of acoustic energy of thin circular plates (Fig. 1). It included two steel rings which permitted the plate tested to be clamped. The lower ring *1* was clamped to the base *2*, while the upper ring *3* was axially pressed down by a system of six pneumatic servos *4*. The vibration of the plate *5* was forced by the piston excited, in turn, by the head (BK 4813) of the exciter *6* (BK 4801). The plate vibration was forced by a vibrating air layer between the plate and the circular piston. As a result of this, according to the theoretical assumptions, the whole plate was excited to uniform vibration.

In order to assure the conditions of a plate vibrating in an infinite baffle, the system included an acoustic baffle and the encasing of the lower part containing the exciter.

The plate tested was clamped in annular jaws which pressed it down uniformly over the circumference. The mean value of the unitary pressure on the circumference of the plate was $1570\ \mathrm{Nm}^{-1}$. The value of the mean unitary pressure on the circumference of the plate could be adjusted by changing the value of the air pressure in the adjusting chamber which fed the system of 6 servos.

The exciting system used (Fig. 2) permitted the flexural vibrations of the plate to be forced. The exciting source was an electrodynamic vibration exciter fed with a sinusoidal signal from a generator over the frequency range 20-1200 Hz.

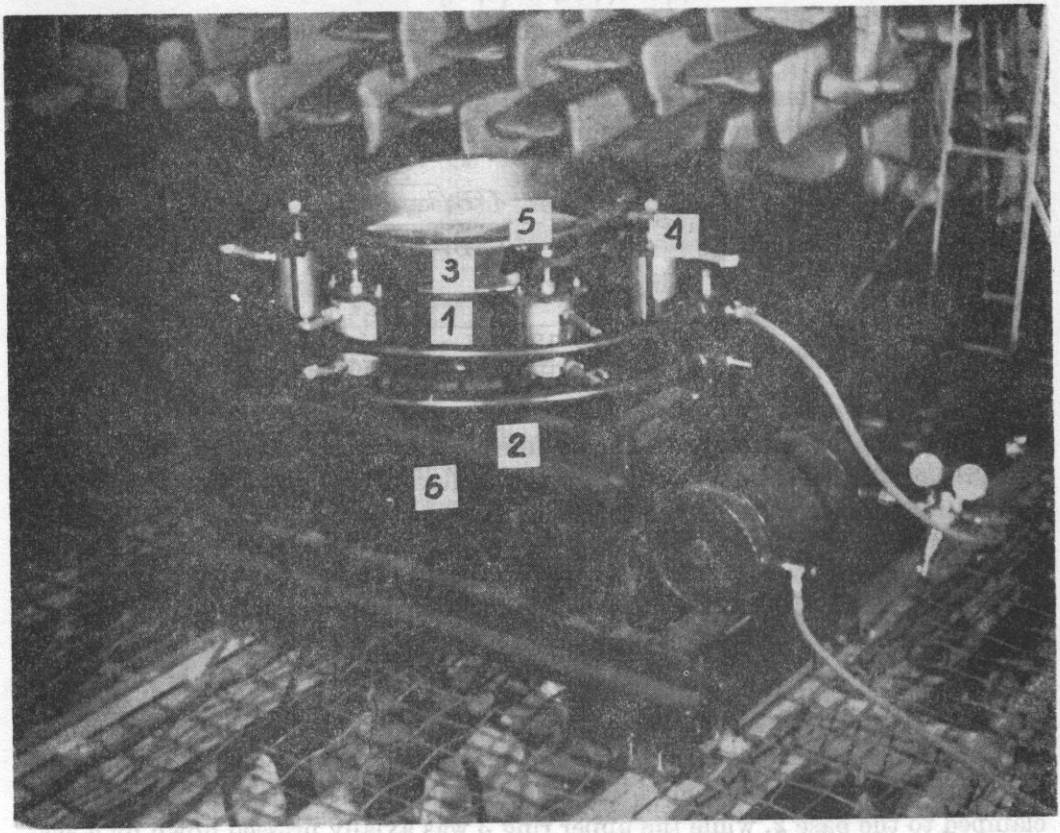


Fig. 1. The system for investigating vibroacoustical plates

1 — the lower clamping ring, *2* — the base, *3* — the upper clamping ring, *4* — a system of clamping servos, *5* — the tested plate, *6* — the head of a vibration exciter

The experimental determination of the equivalent surface of the plate required the assessment of the sound power radiated by the plate and the amplitude of the vibration velocity in the anti-node of the plate vibrations. Since the basic aim of these investigations was the essessment of radiation at the eigenfrequencies of the plate, it was necessary to determine preliminarily these frequencies by defining the vibration modes. The modes were visualized by the method of Chladni figures, consisting in covering the vibrating plate surface with a fine material layer [8]. As a result of the plate vibrations, this material gathered in the vibration nodes. The radiation of acoustic energy by the plate was investigated in the free field conditions (in an anechoic chamber) and in the reverberant field conditions (in a reverberation chamber).

The procedure consisted of three main stages. The first stage involved preliminary investigation aimed at the determination of the frequencies of the vibration nodes and their comparison with the calculated values. These investigations

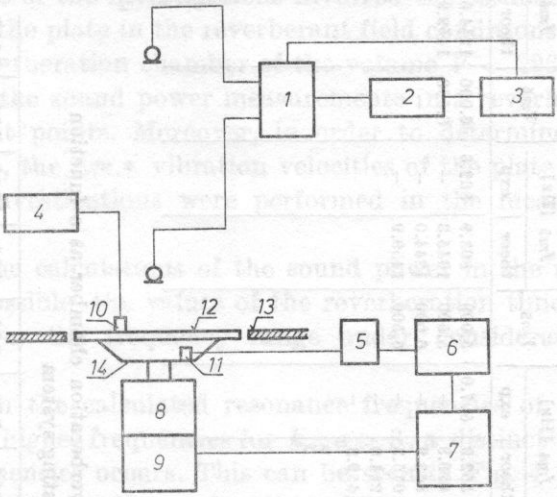


Fig. 2. The block diagram of the system for measuring and exciting vibration of the plate, which was used in the investigations in the free and reverberant field conditions

1 – a *BK* 4408 two-channel microphone selector, *2* – a *BK* 2607 measuring amplifier, *3* – a *BK* 2305 recorder, *4* – a *BK* 2607 measuring amplifier, *5* – a *BK* 2650 precision preamplifier, *6* – a *BK* 1026 generator, *7* – a *BK* 2707 power amplifier, *8* – the head of a *BK* 4813 exciter, *9* – the encasing of a *BK* 4801 exciter, *10* – an *HM* 0002 contactless electromagnetic vibration velocity sensor, *11* – a *BK* 8305 accelerometer, *12* – the plate tested, *13* – the baffle, *14* – the exciting piston

were performed in an anechoic chamber. In order to assure regular vibration distributions on the plate, plates were chosen from the point of view of good isotropic properties. As a result of this selection, a steel isotropic plate of the thickness $b = 0.9$ mm and the diameter $2a = 497$ mm was used in the further investigations. Table 1 shows the values of the eigenfrequencies of the plate, both obtained experimentally and from theoretical calculations according to relation (3). The following numerical values were taken for the calculations: $E = 2.06 \cdot 10^{11}\ \mathrm{N \cdot m^{-2}}$, $\nu = 0.27$, $\varrho = 7.86 \cdot 10^{3}\ \mathrm{kg \cdot m^{-3}}$. The values of the serfies γ_{mn} were taken after MORSE [11].

The second stage of the investigations involved the estimation of the sound power radiated by the plate at the resonance frequencies. These investigations were performed in the free field conditions, where the plate was excited to vibrate at the frequencies of axially symmetrical and axially nonsymmetrical resonances. The examples of Chladni figures for chosen eigenmodes are shown in Fig. 3.

The sound power was determined on the basis of the directional characteristic by measuring the sound pressure level as a function of angle for a constant distance from the plate centre. This distance was 2 m. In order to determine experimentally the equivalent surface of the plate for the axially symmetrical

Table 1. The calculated and experimentally determined values of the eigenfrequencies of the plate

m	γ_{m1}	f_{m1} [Hz]		γ_{m2}	f_{m2} [Hz]		γ_{m3}	f_{m3} [Hz]		γ_{m4}	f_{m4} [Hz]		γ_{m5}	f_{m5} [Hz]		γ_{m6}	f_{m6} [Hz]	
		theor	exp		theor	exp		theor	exp		theor	exp		theor	exp		theor	exp
0	1.015	35.6	28.1	2.007	139.2	123.2	3.000	311.0	286.3	4.000	552.9	507.0	5.000	863.9	862.6	6.000	1244.0	–
1	1.468	74.5	–	2.483	213.0	–	3.490	420.9	409.0	4.500	699.8	–	5.500	1045.3	–	6.500	1459.9	–
2	1.879	122.0	–	2.992	309.3	–	4.000	552.9	520.7	5.000	863.9	–	6.000	1244.0	–	–		
3	2.290	181.2	155.5	3.501	423.5	–	4.500	699.8	–	5.500	1045.3	–	6.500	1459.9	–			
4	2.700	251.9	–	4.010	555.7	–	5.000	863.9	–	6.000	1244.0	–						
5	3.110	334.2	295.0	4.519	705.7	–	5.500	1045.3	–	6.500	1459.9	–						
6	3.523	428.9	–	5.028	873.6	–	6.000	1244.0	–									
7	3.934	534.8	–	5.537	1059.4	–	6.500	1459.9	–									
8	4.345	652.4	–	6.046	1263.1	–												
9	4.756	781.6	–	6.5	1459.9	–												
10	5.167	922.6	–															
11	5.578	1075.2	–															
12	6.0	1244.0	–															
13	6.5	1459.9	–															

Table 2. The values of the reverberation time of the reverberation chamber as a function of frequency after mounting the testing system

f [Hz]	100	200	250	300	350	400	450	500	600	700	800	900	1000	1100	1200
T [s]	6.6	7.9	7.0	7.3	6.7	6.1	6.2	6.8	6.0	5.6	5.1	5.2	5.0	4.6	4.3

modes, the *r.m.s.* vibration velocity of the plate was also measured at its centre. The directional characteristics were measured using a BK 3922 turn table and registered with a BK 2305 level recorder.

The third stage of the investigations involved the estimation of the sound power radiated by the plate in the reverberant field conditions. This estimation took place in a reverberation chamber of the volume $V = 196$ m^3. The estimation was based on the sound power measurements in a reverberation chamber at six measurement points. Moreover, in order to determine the equivalent surface of the plate, the *r.m.s.* vibration velocities of the plate centre were also measured. These investigations were performed in the measurement system shown in Fig. 1.

In order for the calculations of the sound power in the reverberant field conditions to be possible, the values of the reverberation time of the chamber were measured over the frequency range under consideration. These are given in Table 2.

It follows from the calculated resonance frequencies of the plate shown in Table 1, that for higher frequencies for $k_{mn}a > 8$, a distinct concentration of the resonance frequencies occurs. This can be seen in Fig. 4 which shows the number of vibration modes N as a function of frequency, corresponding to the frequency range 20 Hz. The vibration distribution is much more complex then, as a result of which, the vibration velocity measurements at the plate centre may be insufficient. The measurements were expanded, therefore, to include the vibration velocities at several points of the plate in order to provide the mear value.

After the preliminary investigations, the measurements were decided to be taken at 5 points, one of which was at the centre of the plate, while the other four were on the circumference at the distance $r = 80$ mm from the centre of the plate.

From these measurements the mean square value of the vibration velocity was calculated from the relation

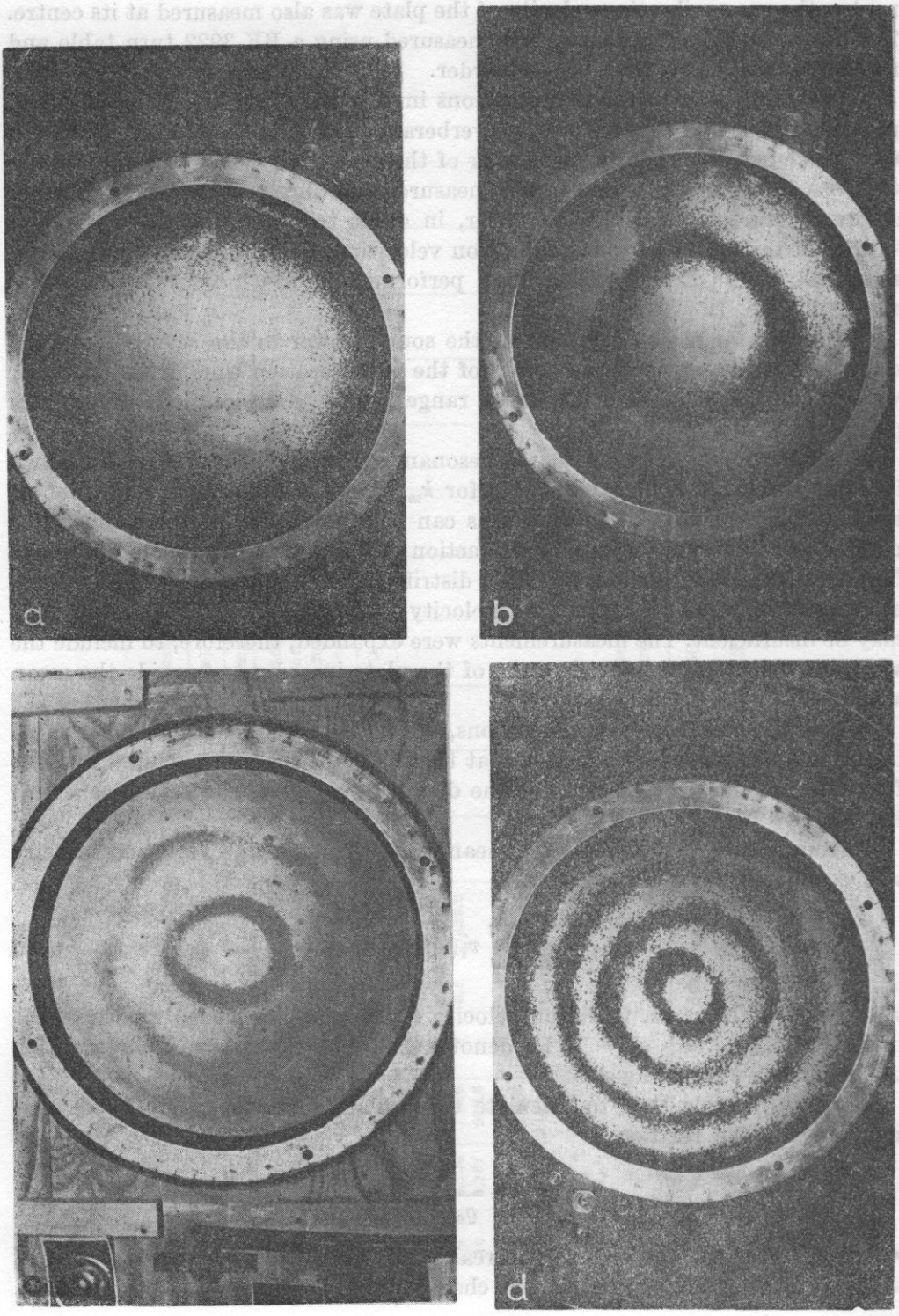
$$\bar{v}^2_{(\Delta f)} = \frac{1}{n}\sum_{i=1}^{n} v_i^2(\Delta f), \qquad n = 5, \tag{14}$$

where $v_{i(\Delta f)}$ is the *r.m.s.* vibration velocity of the plate at the ith measurement point ($i = 1,2, \ldots, 5$), $\Delta f = 10$ Hz denotes the frequency bandwidth over which the amplitude was averaged.

It was possible thus to calculate the radiation efficiency [1, 2], defined according to the relation

$$\sigma_{(\Delta f)} = \frac{W_{T(\Delta f)}}{\varrho_0 c \bar{v}^2_{(\Delta f)} S_r}, \tag{15}$$

where S_r is the surface area of the real plate, and W_T is the sound power of the plate measured in a reverberation chamber.


a
b
c
d

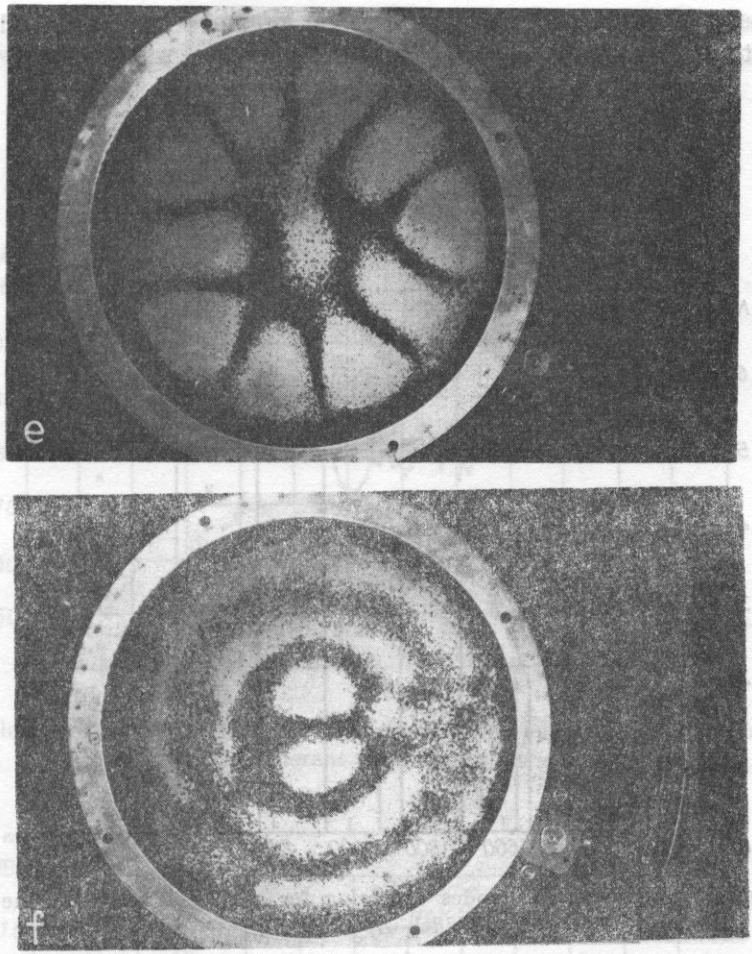


Fig. 3. Examples of the Chladni figures obtained for several vibration modes of a circular plate clamped on the circumference (with diameter $2a = 497$ m, thickness $b = 0.9$ mm)

a) $f_{01} = 28.1$ Hz; b) $f_{02} = 123.0$ Hz; c) $f_{03} = 286.3$ Hz; d) $f_{04} = 507.9$ Hz; e) $f_{11} = 295.0$ Hz, f) $f_{13} = 499.0$ Hz

4. Estimation of the radiated sound power and the equivalent surface of the plate

a. The measured and calculated results obtained from the investigations performed in the free field conditions

On the basis of the directional characteristics determined, the squared values of the sound pressure were calculated at a given frequency. From the measurement of the distribution of the sound pressures measured on a hemisphere at the distance $r_m = 2$ m from the centre of the plate, 38 values of the sound pressures p_i were determined for averaging the pressure at each of the frequencies

under investigation. Hence, the mean sound pressure levels were determined according to the relation

$$L_{\bar{p}} = 10 \log \bar{p}^2 / p_0^2,$$

where

$$\bar{p}^2 = \frac{1}{38} \sum_{i=1}^{38} p_i^2, \quad p_0^2 = 4 \cdot 10^{-10} \frac{\mathrm{N}^2}{\mathrm{m}^4}. \tag{16}$$

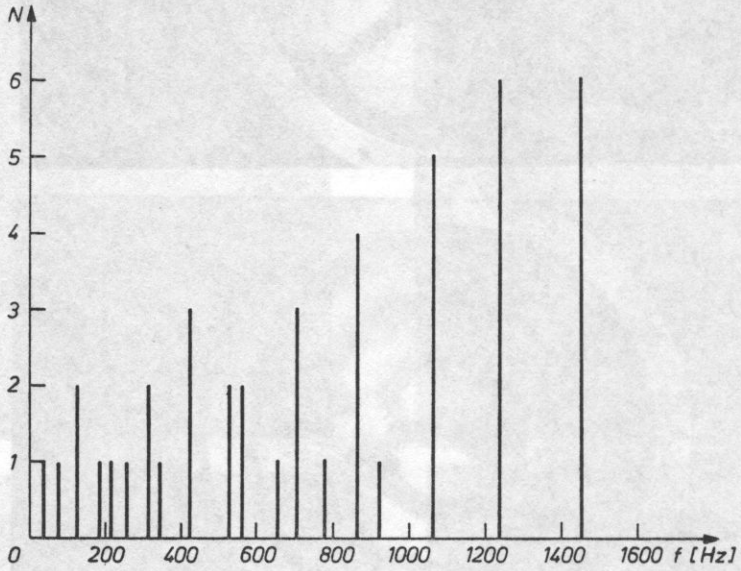


Fig. 4. The number of vibration modes occurring for a circular plate in the frequency bandwidth $\Delta f = 20$ Hz

The sound power radiated was calculated from the relation

$$W_{\bar{p}} = \frac{\bar{p}^2}{\varrho_0 c} S_p, \tag{17}$$

where $\varrho_0 c = 415\ \mathrm{kg \cdot m^{-2} \cdot s^{-1}}$ is the characteristic impedance of the air, and $S_p = 2\pi r_m^2 = 25.1\ \mathrm{m^2}$ is the area of the hemisphere for the measurements of sound pressure. Hence, the sound power level was expressed in the form

$$L_W = 10 \log \frac{W_{\bar{p}}}{W_0} \quad [\mathrm{dB}], \tag{18}$$

where $W_0 = 10^{-12}$ W.

The values of the average sound pressure levels and of the sound power are shown in Figs. 5 and 6.

On the basis of the measured *r.m.s.* vibration velocities at the centre of the plate, the values of their levels were calculated, according to the relation

$$L_{v0} = 10 \log \frac{v_0^2}{v_{\text{ref}}^2}, \tag{19}$$

where $v_{\text{ref}} = 5 \cdot 10^{-8}\ \text{m} \cdot \text{s}^{-1}$. These values are given in Fig. 7.

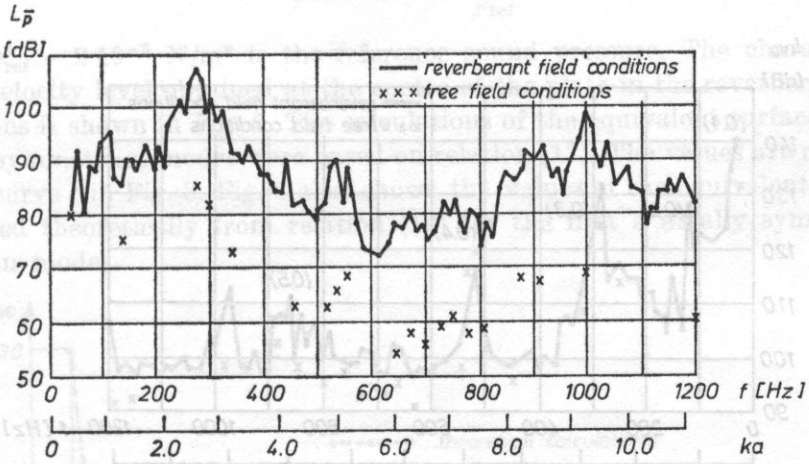


Fig. 5. The levels of the averaged sound pressure radiated by the plate in the free field and reverberant field conditions

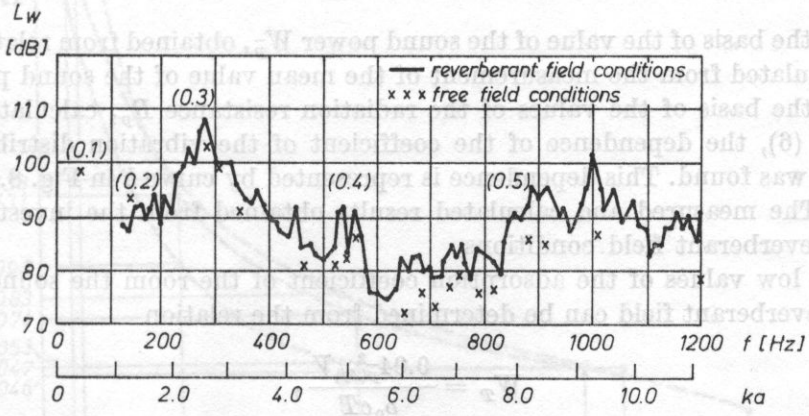


Fig. 6. The levels of the measured sound power radiated by the plate in the free field and reverberant field conditions

Relation (7) in the following form was used for the calculation of the equivalent surfaces

$$S_{\text{eq}} = \frac{W_{\bar{p}}}{R_p v_0^2}. \tag{20}$$

After division of the value obtained for the plate surface S_{eq} by its real surface S_r, the coefficient of the vibration distribution, defined by formula (8), was calculated,

$$\varkappa = \frac{W_{\bar{p}}}{R_p S_r v_0^2}, \tag{21}$$

where $S_r = 0.25$ m².

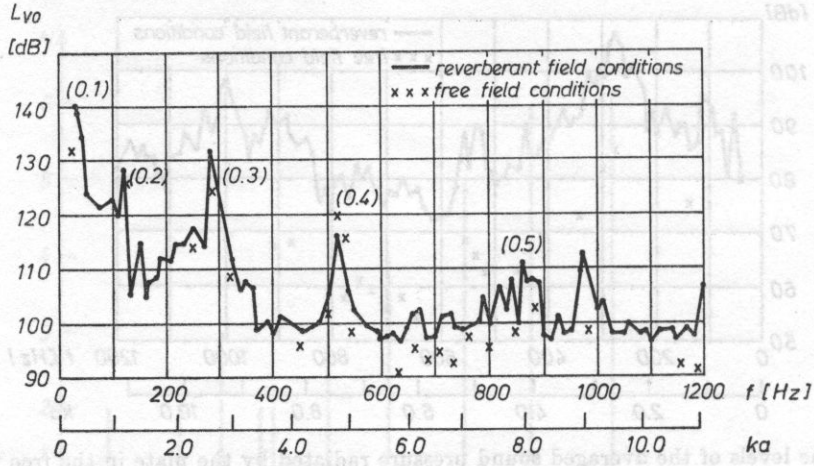


Fig. 7. The levels of the vibration velocity at the centre of the plate tested, measured in the free field and reverberant field conditions

On the basis of the value of the sound power $W_{\bar{p}}$, obtained from relation (17) and calculated from the measurement of the mean value of the sound pressure, and on the basis of the values of the radiation resistance R_p, calculated from formula (6), the dependence of the coefficient of the vibration distribution $\varkappa$ on $k_{mn}a$ was found. This dependence is represented by curve *b* in Fig. 8.

b. The measured and calculated results obtained from the investigations in the reverberant field conditions

For low values of the adsorption coefficient of the room the sound power in the reverberant field can be determined from the relation

$$W_T = \frac{0.04\overline{p}^2_{(\Delta f)} V}{\varrho_0 c T}, \tag{22}$$

where T is the reverberation time of the room, V is its volume, and $\overline{p}^2_{(\Delta f)}$ is the mean square value of the sound pressure, determined by averaging the results obtained at 6 positions of the microphone at the frequency band $\Delta f = 10$ Hz:

$$\overline{p}^2_{(\Delta f)} = \frac{1}{n} \sum_{i=1}^{n} p^2_{i(\Delta f)}, \quad n = 6. \tag{23}$$

Then, the sound pressure level in the reverberant field can be calculated from the relation

$$L_{WT(\Delta f)} = L_{\overline{p}(\Delta f)} - 10\log T + 10\log V \quad 14 \quad [\mathrm{dB}], \tag{24}$$

where

$$L_{\overline{p}(\Delta f)} = 10\log \frac{\overline{p}^2_{(\Delta f)}}{p^2_{\mathrm{ref}}}, \tag{25}$$

where $p^2_{\mathrm{ref}} = 2 \cdot 10^{-5}$ N/m² is the reference sound pressure. The characteristic of the velocity level obtained at the centre of the plate in the reverberant field conditions is shown in Fig. 7. The calculations of the equivalent surface for the axially symmetrical modes were based on relation (17). The values are represented by curve *c* in Fig. 8. Fig. 8 also shows the values of the equivalent surface, calculated theoretically from relation (12) for the first 5 axially symmetrical vibration modes.

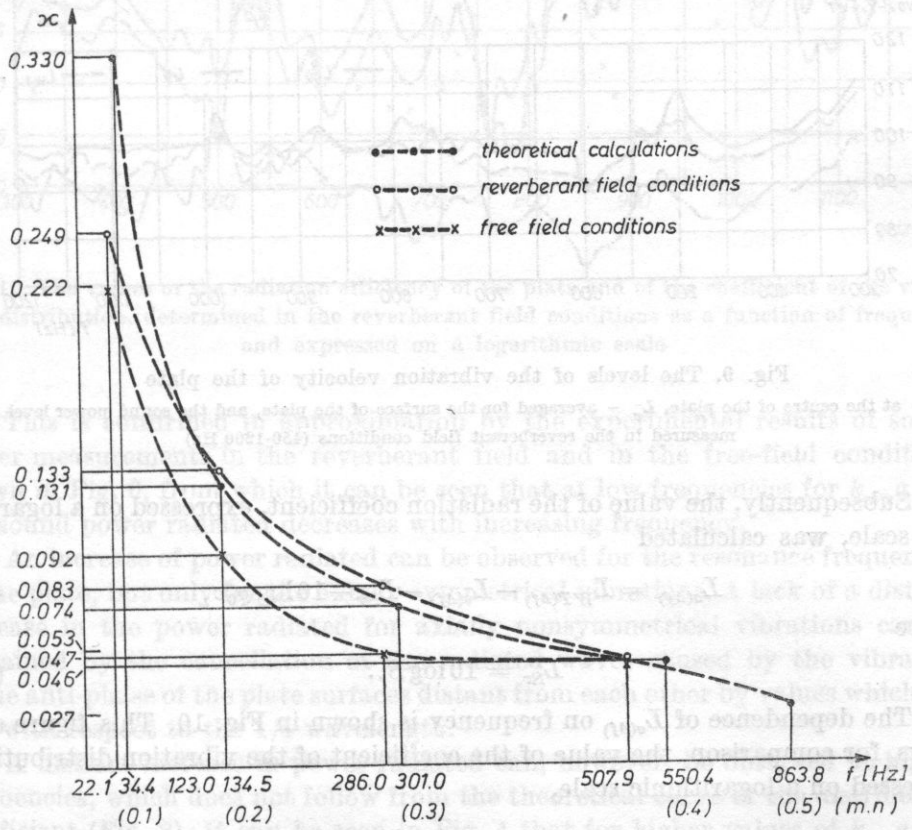


Fig. 8. The variation of the coefficient of the vibration distribution $\varkappa$ of a circular plate clamped on the circumference, as a function of the axially symmetrical vibration modes
a) calculated from formulae (8) and (13), b) measured in the reverberant field conditions, c) measured in the free field conditions

5. Measured and calculated values of the radiation efficiency of the plate

In order to verify the usefulness of the equivalent surface for the higher vibration modes, for which the vibration distribution of the plate is more complex, the radiation efficiency, defined by formula (15), was measured. The mean value of the vibration velocity was calculated from formula (14).

Subsequently, the levels of the mean velocity

$$L_{\bar{v}(\Delta f)} = 10 \log \frac{\bar{v}^2_{(\Delta f)}}{v_{\mathrm{ref}}}, \tag{26}$$

were measured, whose dependence on frequency is shown in Fig. 9. This figure also shows the vibration velocity level at the centre of the plate $L_{v0(\Delta f)}$ and the sound power level $L_{WT(\Delta f)}$ determined from measurements in the reverberant field, according to formula (24).

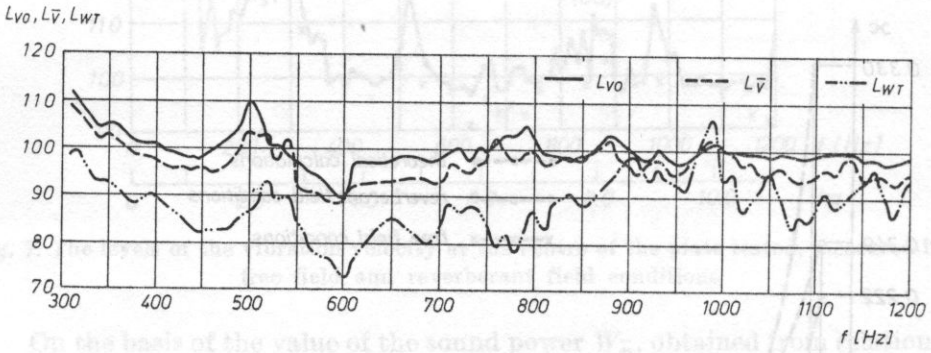


Fig. 9. The levels of the vibration velocity of the plate
L_{v0} – at the centre of the plate, $L_{\bar{v}}$ – averaged for the surface of the plate, and the sound power level L_{WT} measured in the reverberant field conditions (450-1200 Hz)

Subsequently, the value of the radiation coefficient, expressed on a logarithmic scale, was calculated

$$L_{\sigma(\Delta f)} = L_{WT(\Delta f)} - L_{\bar{v}(\Delta f)} - L_{Sr} - 10 \log \varrho_0 c, \tag{27}$$

where

$$L_{Sr} = 10 \log S_r. \tag{28}$$

The dependence of $L_{\sigma(\Delta f)}$ on frequency is shown in Fig. 10. This figure also shows, for comparison, the value of the coefficient of the vibration distribution, expressed on a logarithmic scale,

$$L_{\varkappa} = 10 \log \frac{S_{\mathrm{eq}}}{S_r}, \tag{29}$$

where S_{eq} is the equivalent surface of the plate, determined from relation (20).

6. Discussion of the measured and calculated results

It follows from theoretical analysis that there are two opposite tendencies affecting the plate radiation. On the one hand, the coefficient of the vibration distribution, $\varkappa$, decreases with increasing frequency, which results from Fig. 8; on the other, the radiation resistance increases as the frequency increases, and for $k_{mn}a > 2$ it is equal to $\varrho_0 c$.

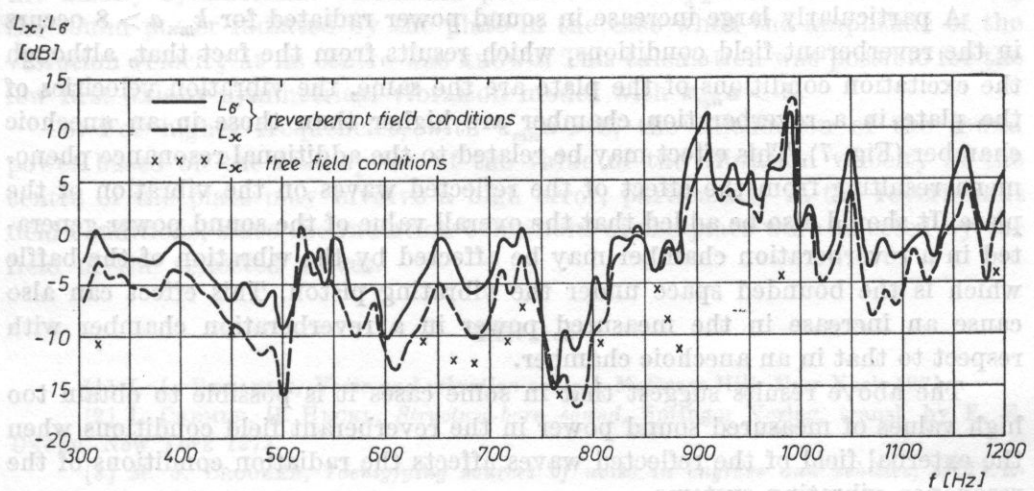


Fig. 10. The values of the radiation efficiency of the plate and of the coefficient of the vibration distribution, determined in the reverberant field conditions as a function of frequency and expressed on a logarithmic scale

This is confirmed in approximation by the experimental results of sound power measurements in the reverberant field and in the free-field conditions shown in Fig. 6, from which it can be seen that at low frequencies for $k_{mn}a < 5$ the sound power radiated decreases with increasing frequency.

An increase of power radiated can be observed for the resonance frequencies of the plate, but only for the axially symmetrical vibrations. A lack of a distinct increase in the power radiated for axially nonsymmetrical vibrations can be explained by the cancellation of the radiated waves caused by the vibration in the anti-phase of the plate surfaces distant from each other by values which are low with respect to the 1/4 wavelength.

A distinct increase in power radiated can, however, be observed at higher frequencies, which does not follow from the theoretical curve of the distribution coefficient (Fig. 8). It can be seen in Fig. 4 that for higher values of $k_{mn}a > 8$ the resonance frequencies cumulate over narrow frequency ranges. The plate vibrates in this case in several modes, which increases considerably its equivalent surface, since the distances between the vibration anti-nodes of the particular

parts of the plate are comparable to 1/4 wavelength, which decreases the cancellation effect mentioned above.

This is confirmed by the results shown in Fig. 10, which give large values of the coefficient $\varkappa$, expressed on a logarithmic scale for some frequencies over the range 900-1200 Hz. Similar results are obtained for the radiation coefficient $\varkappa$, expressed on a logarithmic scale, which is obtained from the measurements of the mean vibration velocity of the plate by averaging the results from several points.

A particularly large increase in sound power radiated for $k_{mn}a > 8$ occurs in the reverberant field conditions, which results from the fact that, although the excitation conditions of the plate are the same, the vibration velocities of the plate in a reverberation chamber are higher than those in an anechoic chamber (Fig. 7). This effect may be related to the additional resonance phenomena resulting from the effect of the reflected waves on the vibration of the plate. It should also be added that the overall value of the sound power generated in a reverberation chamber may be affected by the vibration of the baffle which is the bounded space under the vibrating piston. This effect can also cause an increase in the measured power in a reverberation chamber with respect to that in an anechoic chamber.

The above results suggest that in some cases it is possible to obtain too high values of measured sound power in the reverberant field conditions when the external field of the reflected waves affects the radiation conditions of the resonance vibrating systems.

These investigations show that it is possible to measure approximately the sound power of the plate by means of measuring the vibration velocity at its centre, using the theoretically calculated or experimentally found equivalent surface, S_{eq} — only, however, over the low frequency range for $k_{mn}a < 5$. Over the higher frequency range the values of S_{eq} can increase considerably, which requires the measurements of the vibration velocity at several points and the application of their mean value.

7. Conclusions

1. The highest values of the sound power level radiated by the plate occurred for the axially symmetrical resonance vibrations. In the case of the axially nonsymmetrical resonance vibrations, and also the frequency range between the resonances, the values of the sound power level radiated were lower by about 10 dB.

2. It was found from experimental investigations that over the higher frequency range ($k_{mn}a > 8$) a concentration of the resonance frequencies of the plate occurred, causing an increase in the radiation efficiency of the plate at some frequencies above 900 Hz.

3. The results of experimental investigations of the sound power in the free field and reverberant field conditions showed a good agreement over the low

frequency range. Over the higher frequency range the values of the sound power measured in the reverberant field conditions were higher than those measured in the free field conditions.

This might have resulted from the effect of the reflected waves which caused an increase in the vibration velocity of the plate at some frequencies, and also from the radiation of the lower part of the plate and the baffle.

4. The values of the equivalent surface of a circular plate determined for the axially symmetrical vibrations permitted an approximate estimation of the sound power radiated by the plate in the case when the amplitude of the vibration velocity at its centre was known. This calculation was possible for the few first axially symmetrical vibration modes with $k_{mn}a < 5$.

5. For higher frequencies, with $k_{mn}a > 8$, the calculation of the sound power based on the assumption of the value of the vibration velocity at the centre of the plate may involve a high error, particularly in the reverberant field conditions, since the radiation conditions of the plate were affected by the field of the reflected waves.

Rerences

[1] L. L. BERANEK, *Noise and vibration control*, McGraw Hill, New York 1971.

[2] L. CREMER, H. HECKL, *Structure-born sound*, Springer Verlag, transl. by E. E. UNGAR, New York 1973.

[3] M. J. CROCKER, *Identyfying sources of noise in engines and vehicles*, INTER-NOISE 79, vol. **I**, 347-356, 1979.

[4] S. CZARNECKI, Z. ENGEL, R. PANUSZKA, *Estimation of equivalent surface area for determination of the acoustic power of a circular plate*, INTER-NOISE 79, vol. **I**, 67-70, 1979.

[5] S. CZARNECKI, Z. ENGEL, R. PANUSZKA, *Correlation method of measurements of sound power in the near field conditions*, Archives of Acoustics, **1**, 2, 201-213 (1976).

[6] Ch. E. EBBING, T. H. HODGSON, *Diagnostic tests for locating noise sources*, Noise Control Eng., July-August 1974, 30-46.

[7] Z. ENGEL, R. PANUSZKA, *The method of the evaluation of vibroacoustical emission of complex mechanical systems on example of the rotor-plate system of the chamber airless shot-blasting machine* (in Polish), Archiwum Akustyki, **10**, 4, 345-356 (1975).

[8] Z. ENGEL, R. PANUSZKA, *Localization of the sources of machinery noise and vibration* (in Polish), Ergonomia, **2**, 1, 63-68 (1979).

[9] T. H. HODGSON, *Use of the correlation between surface velocity and near field pressure for measuring the noise of a large machine in situ*, 88-th Meeting of the Ac. Soc. of America, St. Louis 74.

[10] N. KAEMMER, H. J. CROCKER, *Sound power determination from surface intensity measurements on a vibrating cylinder*, NOISE-CON 79. West Lafayette, Indiana, 153-160.

[11] P. M. MORSE, K. U. INGARD, *Theoretical acoustics*, McGraw Hill, New York 1968.

[12] E. SKUDRZYK, *Simple and complex vibratory systems*, The Pensylvania State University, 1968.

[13] M. WHITE, *Noise radiation and machine design: using double skin integral panels*, INTER-NOISE 78, San Francisco 1978, 405-410.

[14] S. ŻYSZKOWSKI, *Fundamentals of electroacoustics* (in Polish), PWT, Warsaw 1966.

Received on May 5, 1981.

ARCHIVES OF ACOUSTICS
6, 4, 359-370(1981)

THE EFFECT OF CHOSEN PARAMETERS OF A TELEPHONE CHANNEL ON VOICE IDENTIFICATION

CZESŁAW BASZTURA, WOJCIECH MAJEWSKI

Institute of Telecommunication and Acoustics, Wrocław Technical University
(50-317 Wrocław, ul. B. Prusa 53/55)

The method and results of investigations aimed at the evaluation of the effect of chosen parameters of a telephone channel on the masking of the individual voice features are presented. In the experiments on voice samples of 10 speakers (men) a speech signal was represented by the amplitude correlation matrix and the distribution of time intervals between the zero-crossings of the speech signal. The effect of the frequency band of a telephone channel and of distortions determined by different signal to noise ratios on the probability of correct voice identification was investigated. The results obtained show the possibility of voice identification under the conditions of telephone transmission provided some definite values of the parameters of a telephone link are maintained.

1. Introduction

The continuous development of new, specialized computer generations has brought a situation where the peripheral equipment has bottle-necked the application of computer systems of increasingly greater calculation capacity. One of possible solutions to overcome this limitation is equipping computer systems with acoustic terminals. In this connection, in many countries wide research has for a number of years been done on the development of an acoustic output from the computer, i.e. on the solution of the problem of speech synthesis. Efforts to develop an acoustic computer input are equally intensive, which in the field of the transmission of linguistic information requires the solution of the problem of automatic speech recognition, and in the range of the transmission of individual information, that of automatic speaker recognition. Both aspects of the automatic computer input are essential in practice, and they sometimes occur jointly. An example of this may be the case when the access to some information stored in the computer memory is reserved for authorized persons

only, i.e. is available only when the identity of a person applying for this information has been checked on the basis of analysis of his voice sample.

The cheapest and simplest way of achieving a multi-access man-computer communication based on speech can be provided by the existing telephone network. This requires, on the one hand, a number of investigations on the effect of distortions and interferences caused by a telephone channel on the possibility of both speech and speaker recognition, and, on the other hand, needs the development of reliable telephone links of good parameters of speech signal transmission.

Little attention has been paid in the literature on speaker identification to the masking of the individual voice features as a result of a broadly understood effect of a telecommunication channel. Automatic speaker identification, which concerns the procedure of assigning an unknown utterance to a speaker from a given set of speakers, is the most difficult case of the general problem of automatic speaker recognition. It is now in the stage of laboratory investigations and it will be long before it can be used in practice under the conditions of telecommunication. ATAL's publication [1], which reviews the state of art in the investigations of automatic speaker identification, gives no results of research done under the conditions which occur in practice in a telephone conversation. Some information on the effect of a telecommunication channel on the results of identification when simulated under laboratory conditions can be found, however, in other communications in this field [4, 11]. It should be added here that simulation of the effect of a telecommunication channel under laboratory conditions differs from the conditions which occur in practical implementation of a system of automatic speaker identification in that in the first case the learning and the recognized sequences are registered under the same conditions, while in the second case the two sequence are as a rule registered under different conditions, thus additionally deteriorating the identification results.

Automatic speaker verification, which consists in deciding whether the unknown utterance belongs to a particular speaker or not, i.e. a procedure of making a binary decision, is a simpler case of the general problem of automatic speaker recognition, which is closer to practical implementation than speaker identification. Therefore, the ROSENBERG review [9] gives a larger number of works accounting for the effect of the conditions of speech signal transmission from speaker to processor on the verification results. It is of interest to note another ROSENBERG's paper [10] which describes the possibility of implementing a system of automatic speaker recognition with an error less than 5 procent for speech signal transmission via a telephone link.

On the basis of the facts given above and appreciating the necessity of investigating voice recognition on the phonetic and linguistic material of the Polish language, and, additionally, bearing in mind the possibility of using the procedures of automatic voice recognition not only in the generally accessible

telephone information systems but also in crime detection, the present authors set themselves the task of investigating the effect of chosen parameters of a telephone link on the masking of the individual voice features. The object of particular interest was the effect of the frequency bandwidth of a telephone channel on the possibility of voice identification and the effect of typical distortions caused by a telephone channel, represented by different levels of the signal to noise ratio.

2. Identification system

2.1. *Introduction.*

Automatic speaker recognition can be implemented in a system of coupled analogue and digital units, called a recognition system. The starting point for system selection is above all the establishing of the aim of recognition. The aim of the present investigation was to examine the effect of chosen parameters of a telephone channel on the possibility of speaker identification. A simple identification system, with a teacher, and the statistical criteria of decision-making [12], was selected for the implementation of this aim.

2.2. *Block diagram of the identification system.*

A simplified block diagram of the identification systems used is shown in Fig. 1. This identification system can be divided into the three basic units:

(a) the signal source unit;

(b) the unit for measurement, i.e. extraction, of parameter sets,

(c) the classification unit.

The signal source is provided by recorded utterances of M speakers saying the same test (the key phrase).

The measurement unit is an analogue-to-digital system for extraction from a speech signal of the parameters significant from the point of view of the individual features.

The classification unit consists of the information systems: the predetermined information and the "teacher", and the classifier proper including the subunit for making patterns and the identification algorithm with criteria and similarity measures. All the three units are interdependent on each other, i.e. the signal source influences the measurement unit, and the selection of the parameter set affects, in turn, both the way of making patterns and the kind of probability measure and the decision-making criterion of the identification algorithm.

Following the advantages of the model, a short-term recognition model [3] was used in this system.

In order to improve the reliability of the experimental results, two kinds of parameter sets were used. The first was the distribution of the time intervals between the zero-crossings of a speech signal [2, 3], the second was the amplitude

correlation matrix (ACM) [7], based on the short-term spectral analysis. Application of different parameter sets implies another way of pattern making and other similarity measures. Therefore, despite the retaining of one identification algorithm NM (nearest mean), the procedures differ in details and, accordingly, require separate discussion.

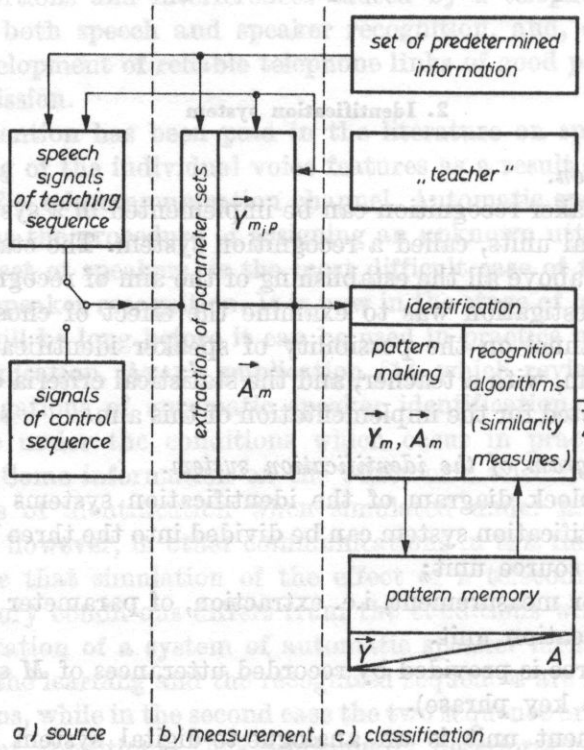


Fig. 1. A block diagram of the identification system

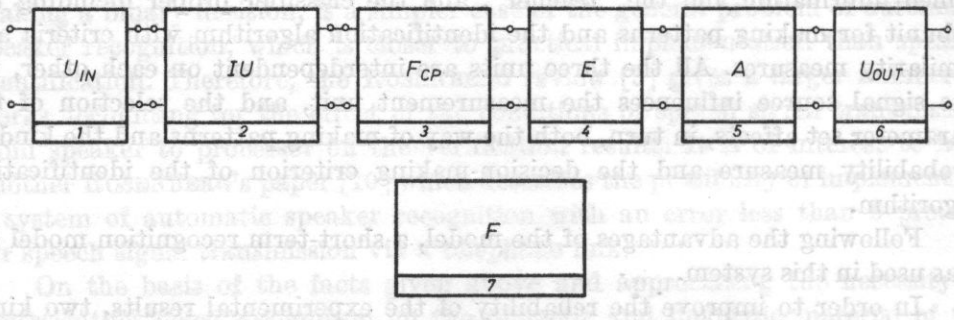


Fig. 2. A block diagram of the telephone channel model

1, 6 — input and output systems, *2* — interference unit, *3* — central-pass filter, *4* — frequency response equalizer, *5* — attenuator, *7* — feeder

2.3. *The distribution of the time intervals between the zero-crossings.*

Analysis of the previous investigations [2, 3] shows a fairly large effectiveness of representing the individual voice features by the parameter set with $\boldsymbol{Y}_{m,p}$, being the distribution of the time intervals between the zero-crossings of a speech signal

$$\boldsymbol{Y}_{m,p} = \{Y_{m,p,1}\, Y_{m,p,2} \ldots Y_{m,p,k} \ldots Y_{m,p,K}\}, \tag{1}$$

where $m = 1, 2, \ldots, M$ (M – the number of speakers), $p = 1, 2, \ldots, P_m$ (P_m – the number of utterance repetitions in the learning or the control sequence for the mth speaker), $k = 1, 2, \ldots, K$ (K – the number of time channels).

The stage following the parameter set making is the making of patterns of voice images. The basis for pattern making in this system is the learning sequence

$$\{Cu\} = (Y_{1,1}, \lambda_1), \ldots, (Y_{1,P_1}, \lambda_1), \ldots, (Y_{m,p_m}, \lambda_m), \ldots, (Y_{M,P_M}, \lambda_M), \tag{2}$$

where λ_m is the speaker class membership predetermined by the "teacher". In the heuristic NM algorithm the voice class patterns are the mean vectors from the repetitions $\boldsymbol{Y}_m$, and the Mahalanobis squared distance [3] was used as similarity functions.

2.4. *Amplitude correlation matrix*

The amplitude correlation matrix was used for speaker recognition in the long-term recognition model, by LI and HUGHES [7], for example. The essence of voice image description in this method consists in using the speaker-dependent correlations beetween the amplitudes for the particular frequency bands. Let a_t represent the vector (of the size K), being a set of the amplitudes for the individual $K \times T$ frequencies (Fig. 3).

Calculating the correlation matrix of amplitudes in the relation

$$A_t(i, j) = \left[(a_t(i) - s(i))\,(a_t(j) - a(j))\right] / \sigma_i \sigma_j, \tag{3}$$

where

$$a(l) = \frac{1}{T} \sum_{t=1}^{T} a_t(l), \qquad \sigma_l = \left[\frac{1}{T} \sum_{t=1}^{T} (a_t(l) - a(l))^2\right]^{1/2},$$

$$i, j, l = 1, 2, 3, \ldots, K, \qquad t = 1, 2, 3, \ldots, T,$$

it can be regarded as a representation of the utterance for a stationary speech signal in the long-term model [2, 4, 7], or for a definite utterance in the short-term model [3].

When using the identification algorithm NM analogously to the distributions of the time intervals, the voice image patterns of the learning sequence are based on the following relation

$$A^m(i, j) = \frac{1}{P_m} \sum_{p=1}^{P_m} A^{m,p}(i, j), \tag{4}$$

where

$$A^{m,p}(i,j) = \frac{1}{T}\sum_{t=1}^{T} A_t^{m,p}(i,j). \tag{5}$$

$A^{m,p}(i,j)$ is the correlation matrix of amplitudes of the size $K \times K$ representing the pth repetition of the mth speaker. Since the matrix $A(i,j)$ is a real, symmetrical matrix, it is sufficient for every second element of the matrix to be stored in the pattern memory, which permits a considerable saving of memory and calculation time.

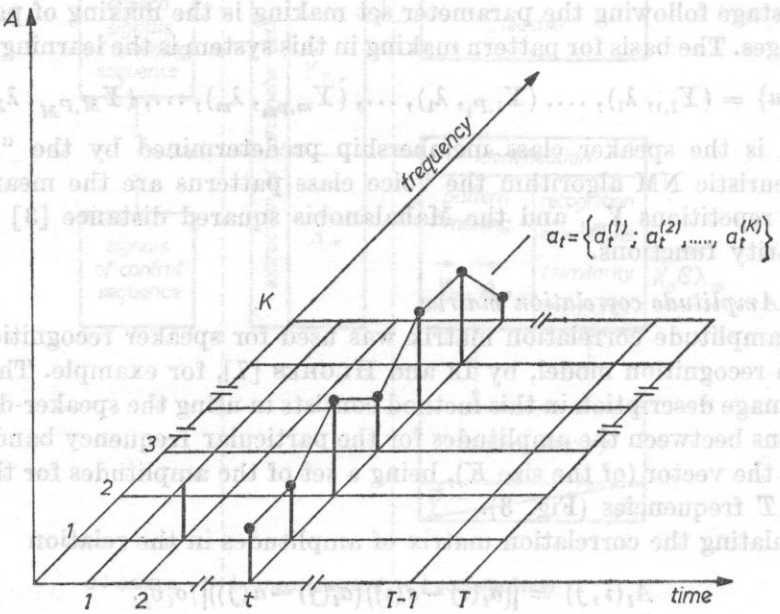


Fig. 3. Discrete representation of the short-term spectrum for the t-sample

The definition of the common decision-making rule of the NM algorithm was given in refs. [2, 3]. In view of the specificity of the parameter set $A(i,j)$, after the authors of ref. [7], two similarity functions were used experimentally. The first, defined as

$$d_{n,m}^{(1)} = \Big[\sum_{j=1}^{K}\sum_{i=j}^{K} B^n(i,j) - A^m(i,j)|^{1/2}\Big] \Big/ \Big[\sum_{j=1}^{K}\sum_{i=j}^{K} (B^n(i,j))^2\Big]^{1/2}, \tag{6}$$

is a normalized Euclidean distance between the two matrices, where $B^n(i,j)$ is, analogously to $A^m(i,j)$, the image of the nth recognized utterance. The second similarity function, defined as the average absolute difference between the matrices $A(i,j)$ and $B(i,j)$, can be calculated from the relation

$$d_{n,m}^{(2)} = \frac{2}{K(K+1)} \sum_{j=1}^{K}\sum_{i=j}^{K} |B^n(i,j) - A^m(i,j)|. \tag{7}$$

3. Identification experiment

3.1. *Introduction*

The considerations in sections 1 and 2 were the basis for the identification experiment aimed at answering the question whether and to what degree the basic parameters of a telephone channel, for which the frequency band and the signal to noise ratio were taken, affect the masking of the individual voice features. This requires sound material recorded under suitable technical conditions. In view of the desired measurement stability, particularly essential with the few identification statistics (the number of speakers, the lengths of the learning and the control sequences), the present investigations did not include measurements under the real conditions of telephone transmission, and these conditions were simulated using a model of telephone channel developed and built in the Institute of Telecommunication and Acoustics, Wrocław Technical University.

3.2. *Model of telephone channel*

The model of telephone channel (Fig. 2) is an analogue system implementing the predetermined physical parameters and the basic characteristics of a typical population of telephone channels. The variability range (the possibility of adjustment) covers the real range of variations in the parameters of the frequency response and of the distortions of typical telephone lines in natural telephony [5, 6]. The channel model consists of input and output systems and the five units of:

(a) additive distortions,
(b) the frequency response equalizer,
(c) the central-pass band filter,
(d) the attenuator,
(e) feeders.

The nominal signal level is 0 dB. The maximum input and output voltage is 3V. The input and output resistance is 600 Ω.

3.3. *Selection of the key phrase*

The present investigation used the text "jutro będzie ładny dzień" as the key phrase. This selection was justified by the easy pronunciation and the frequent use of the words in the phrase and the relatively good approximation of the mean statistics of the Polish language in terms of the frequency spectrum and the occurrence frequency of the phonemes [8].

3.4. *Population of speakers*

In view of the rather time-consuming calculations the population of speakers was limited to 10 persons-men from 20 to 35 years old. The key phrase was recorded in two sessions (I and II) at an interval of one month. The length of the learning sequence was 30, while that of the control sequence was 20, which required three repetitions of the key phrase by each of the speakers in session I

and two repetitions in session II. Irrespective of the kind of experiment the learning sequence was taken from session I, and the control sequence from session II.

3.5. *Preparation of the sound material*

The speakers' utterances were recorded in a listening studio of the attenuation of external distortions of — 30 dB. The speech signal was recorded on Superton C-60 Low-Noise cassettes using a M 601 SD Unitra cassette tape recorder manufactured by ZRK.

In order to introduce distortions and interferences the sound material recorded was fed to the channel model and when it passed through it was recorded again on the same recording equipment.

3.6. *Programme of the experiments*

The programme assumed for the experiments in speaker identification is shown in Fig. 4. Experiment 0 made for a speech signal of the form obtained directly from preliminary recordings, aimed at achieving some reference measure.

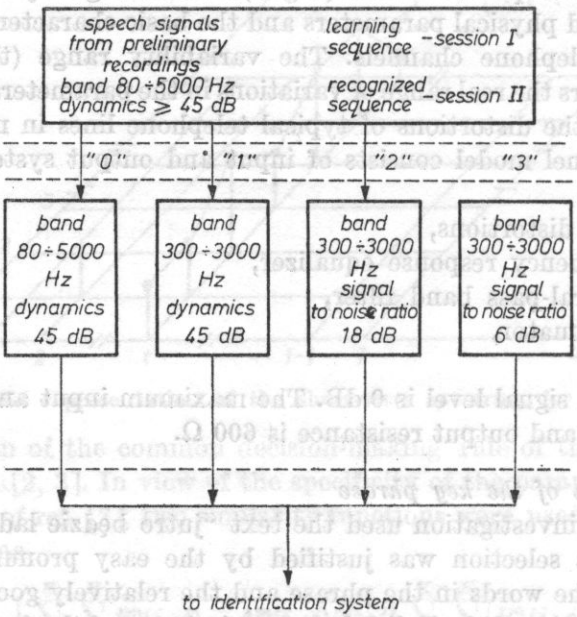


Fig. 4. Programme of the experiments

Experiments 1-3 were made for the parameters of the model representing typical parameters of telephone links and giving essential differences between the mean speech intelligibility measured subjectively (using 10 listeners and 400 *PB* nonsense words corresponding to each setting of the channel model).

The mean intelligibility was 96.3 per cent for experiment 0, while for experiments 1-3 it was 84.5, 73.0 and 66.1 per cent, respectively.

The principle and technical data of the programme implementing the extraction of the parameters Y_m, i.e. the distribution of the time intervals between the zero-crossings, were presented in greater detail in refs. [2, 3]. On the basis of these papers, 16 time channels $K = 16$ of the nominal boundary values shown in Table 1 were taken. In the identification experiments made according to the algorithm given in refs. [2, 3] 16-dimensional vectors of the parameters Y_m were thus used.

Table 1. The boundaries of time channels

k	t_{k-1} [ms]	t_k [ms]	k	t_{k-1} [ms]	t_k [ms]
1	2	3	1	2	3
1	0.15	0.189	9	0.969	1.216
2	0.189	0.238	10	1.216	1.535
3	0.238	0.301	11	1.535	1.937
4	0.301	0.380	12	1.937	2.445
5	0.380	0.478	13	2.445	3.085
6	0.478	0.605	14	3.085	3.893
7	0.605	0.764	15	3.893	4.913
8	0.764	0.969	16	4.913	6.200

The programme of the extraction of the correlation matrices of the amplitudes $A(i, j)$ consisted of two parts. The first stage was a FFT (fast Fourier transform) calculating the amplitude spectrum with a Hamming window of 6.4 m width and the gap between the windows $\Delta t = 20$ ms, which permitted spectral lines at an interval of 156.26 Hz. In view of the differences in duration of the registered utterances of the elements of the key phrase, which did not exceed 10-15 per cent of the mean duration, no time normalization was made and the number T of windows taken, i.e.

$$T = \frac{T_{\text{av}}}{\Delta t}, \tag{8}$$

where $T_{\text{av}} = 0.710$ s is the mean duration for the investigated set of $M \times P$ utterances. $K = 20$ first lines of the spectrum were used to make the pattern and recognized images.

3.7. *Experimental results*

The results obtained from the identification experiments for the two representations of voices Y and A are shown in Table 2.

Table 2. The identification results (percentage of correct decisions) for the parameter sets Y and A

Learning sequence -experiment	Control sequence -experiment	Reprezentation Y [%]	Representation A [%]
	0	90	80
	1	60	55
0	2	65	50
	3	35	25
1	1	85	75
2	2	85	80
3	3	40	40

NB. When the correlation matrix of amplitudes was used for voice description the same results of correct identification were obtained for both similarity functions $d^{(1)}_{n,m}$ and $d^{(2)}_{n,m}$.

4. Conclusion

The experiments permit the following observations and conclusions to be given:

1. Both parameter sets show an almost similarly effective representation of the individual voice features and an almost similar liability to the masking of these features by distortions and interferences caused by the telephone channel, except that slightly better identification results were obtained for the voice description using the distribution of the zero-crossings.

2. The experiments confirmed the thesis, proposed in chapter 1, of the essential effect of the agreement in the conditions of transmission between the signals forming the pattern sequences and those making up the recognized sequences, on the probability of correct voice identification. With this agreement, the identification results were distinctly better.

3. For a recognized sequence composed of a signal contained in the telephone band 300-3000 Hz a change in the value of the signal to noise ratio from 45 dB to 18 dB only slightly affected the identification results. When the pattern sequence was also composed of the signal defined above, the identification results were comparable to the results in the reference system, i.e. for the band 80-5000 Hz and the dynamics 45 dB. It follows therefore that the telephone band and the signal to noise ratio of the order of 18 dB or more do not consitute any essential obstacle to the achievement of safisfactory recognition results provided the agreement in the transmission conditions between the pattern and recognized signals is maintained.

4. Comparison of the identification results shown in Table 2 with the results of speech intelligibility measurements given in section 3.6 shows that distortions nad interferences caused by the telephone channel have a different effect on the masking of the individual voice features, compared to that on the intelligibility of speech. A rapid worsening of the identification results, i.e. weakening of the ability of the telephone channel to transmit individual information, occurred only for the channel parameters defined by experiment 3, while a weakening of the ability of the channel to transmit linguistic information in terms of speech intelligibility occurred in a smooth manner from the condition defined by experiment "0" to those defined by experiment "3". Accordingly, it can be expected that attempts at evaluation of the ability of the telecommunication channel to transmit individual information, based on the quality of the transmission of linguistic information and conversely, will not be succesful.

The future investigations of voice identification under the conditions of telephone transmission should concentrate mainly on the selection of such parameters of a speech signal, which being good carriers of the individual voice features, would not be liable to distortions and interferences caused by the telephone channel, or on the development of such methods of speech signal analysis as permit the compensation for the effect of unknown transmission conditions. Another important task is the definition of the numerical relations between the kind and value of distortions and interferences caused by the telephone channel and the probability of voice identification in the systems assumed.

The solution of these problems and others, difficult and time-consuming though it is, will permit a practical application of systems of automatic speaker recognition under the conditions of telephone communication. The authors hope that the present paper is a contribution to the implementation of this prospect.

References

[1] B. S. Atal, *Automatic recognition of speakers from their voices*, Proc. IEEE, **64**, 4, 460-475 (1976).

[2] C. Basztura, W. Majewski, *The application of long-term analysis of the zero-crossing of a speech signal in automatic speaker identification*, Archives of Acoustics, **3**, 1, 3-15 (1978).

[3] C. Basztura, J. Jurkiewicz, *The zero-crossing analysis of a speech signal in the short-term method of automatic speaker identification*, Archives of Acoustics, **3**, 3, 185-195 (1978).

[4] H. Hollien, W. Majewski, *Speaker identification by long-term spectra under normal and distorted speech conditions*, JASA, **62**, 4, 975-980 (1977).

[5] *Green book of* CCITT, vol. III-1, *Line transmission*, WKiŁ, Warsaw 1976.

[6] *Green book of* CCITT, vol. V, *Quality of telephone transmission, local networks and telephones*, WKiŁ, Warsaw 1976.

[7] K. P. LI, G. W. HUGHES, *Talker differences as they appear in correlation matrices of continuous speech spectra*, JASA, **55**, 4, 833-837 (1974).

[8] W. MAJEWSKI, *Phonetic test for subjective measurements of reference equivalent*, Przegląd Telekomunikacyjny, 8, 237-240 (1979).

[9] A. E. ROSENBERG, *Automatic speaker verification: a review*, Proc. IEEE, **64**, 4, 475-487 (1976).

[10] A. E. ROSENBERG, *Evaluation of an automatic speaker verification system over telephone lines*, Bell System Technical Journal, **55**, 6, 723-744 (1976).

[11] M. R. SAMBUR, *Speaker recognition using orthogonal linear prediction*, IEEE Trans. on Acoustics, Speech and Signal Proc. ASSP – **24**, 4, 283-289 (1976).

[12] V. A. SKRIPKIN, A. L. GORELIK, *Metody rozpoznavanja*, Moscow, Vizhshaya Shkola, 1977.

Received on September 26, 1980; revised version on April 7, 1981.

ARCHIVES OF ACOUSTICS
6, 4, 371-384(1981)

THE PERCEPTIBILITY OF MISTUNED MELODIC INTERVALS BY SCHOOL CHILDREN

H. HARAJDA, J. FYK

Pedagogical University (65-063 Zielona Góra, Plac Słowiański 5)

The aim of this study was to investigate the perceptibility of detuned (mistuned) melodic intervals of prime (1), minor second ($2m$), major second ($2w$) and minor third by school children, depending on the sound timbre and the age of persons examined. The investigation included isolated intervals and intervals in a melodic context. The results obtained for age groups from different centres were compared. It was found that a student of average musical talent tends to perceive detuned melodic intervals below a quarter tone. The perceptibility of these detunings is much higher in intervals composed of musical sounds ("coloured" sounds) occurring in a musical context than in isolated intervals composed of sinusoidal tones. The age of a student or the centre from which he comes were found not to have an effect on the perceptibility of the detunings.

1. Introduction

The perception of melodic intervals by groups of persons of high musical training were investigated by several researchers, WARD [14], DROBNER [3], TARNOCZY and SZENDE [13] and RAKOWSKI [11]. The investigations of this type performed on musically neglected children were described by LEWANDOWSKA [8]. There is a lack, however, of data obtained by acoustical investigation methods and related to the perceiving abilities of average gifted children. One of the basic conditions of sound pitch perception is the perceptibility of detuning (mistuning). A laboratory experiment was therefore conducted, which aimed at investigation of the abilities in this respect of children undergoing a general, nonprofessional musical training in school. The aim of the investigation was to determine the perceptibility of detunings of chosen melodic intervals below a semitone.

At the same time an attempt was made to investigate whether this perceptibility is related to the sound timbre and age of a child.

For this purpose, special tests composed of the following complex sounds were designed and made: piano sounds (called the pianoforte sounds, designated as *Pfp*), flute sounds (*Fl*) and violin sounds (*Vln*), called below as "coloured tests", and tests of sinusoidal sounds, defined in short as "sinusoidal tests". The coloured tests were used in three age groups including children from forms I and II (designated as I/II), III and IV (III/IV), and V and VI (V/VI). The pianoforte tests were used in all groups while the other were used only in groups III/IV.

Music material was taken from the repertory of songs in the school programme for the particular forms, i.e. from the following collections: POWROŹNIAK [19], LIPSKA and PRZYCHODZIŃSKA [17, 18], PRZYCHODZIŃSKA [20], STANKIEWICZ [21, 22], KACZURBINA [6]. The coloured tests were used to investigate the perceptibility of detunings of the intervals in a melodic context, while the sinusoidal tests were used in the case of detunings of isolated intervals.

The psychological position of a sound in a melody is defined by its tonal context, its absolute position on the pitch scale and the course of the melody [9]. Investigations in this field showed that a listener is hardly aware of intervals in a melody since his attention centres on the melody as a whole instead [1], and the melodic line functions independently of the order of interval sizes [2].

Moreover, even the same intervals occurring in different musical contexts are not equal in the psychological sense. The effect of the musical context on the perception of intervals was stressed by SAKHALTUEVA [12], CHMIELEWSKA [1], LEIPP [7], HARAJDA and FYK [4]. Therefore, the melodic phrases making up the music material of the tests take into account the distinct functional character of intervals under investigation and also their position in a melody.

The following intervals were selected:

— for the group of children from forms I/II; F^1-F^1 (the functional degrees I→I), F^1 sharp-G^1 (VII→VIII), D^1-C^1 (II→I), C^2-A^1 (V→III);

— for the group of children from forms III/IV, G^1-G^1 (I→I), F^1 sharp-G^1 (VII→VIII), A^1-G^1 (II→I), G^1-E^1 (V→III);

— for the group of children from forms V/VI; F^1-F^1 (I→I), D^1-E^1 flat (VII→VIII), A^1-G^1 (II→I), E^2-A^1 (V→III).

The intervals 1, *2m*, *3w*, *3m* were selected, because these intervals most frequently occur in school musical texts.

The domination of short intervals over long ones in melody was already stressed by ORTMANN [9], on the basis of analyzing 160 songs. This is also confirmed by the interval analysis of songs in music school-books. As an example, there were given results of the interval analysis performed on 45 songs contained in the currently used music school-book for form V [6].

The present experiment was carried out from 1976 to 1979 under laboratory conditions and was preceded by an experiment under natural conditions, in school. The children participating in both experiments represented the "average student group". The first experiment was to determine the abilities of perceiving

Table 1. The frequency of the occurrence of the melodic intervals in school songs in form V (45 melodies 2932 intervals)

Interval	Number	[%]	Interval	Number	[%]
1	579	19.8	*4zw*	9	0.3
2m	485	16.5	*5zmn*	19	0.7
2w	905	30.9	5	79	2.7
3zmn	1	0.0	*6m*	13	0.4
3m	402	13.7	*6w*	25	0.9
3w	218	7.4	*7m*	14	0.5
4zmn	1	0.0	*7w*	0	0.0
4	172	5.9	8	10	0.3

melodic intervals over an octave, within the solfeggio error, i.e. exceeding a semitone. It was carried out in groups of 8 to 15 persons in 24 centres, including 728 children. The centres included big cities (designated as *M*), small towns *m* and rural centres *W*. On the basis of carefully made subjective examinations by teachers and specially selected examiners, three children, the most representative ones for the "average student" level, were selected for the laboratory experiment from each group.

2. Composition of tests

a. Coloured tests. The composition of all the types of the tests is the same. They are composed of a set of tests presented in four series which correspond to the following intervals:

series I — prime (1), the repetition of the tonic,
series II — minor second (2 *m*) in direction VII→VIII degree,
series III — major second (2*w*) in direction II→I degree,
series IV — minor third (3*m*) in direction V→III degree.

Each of the series uses different musical-melodic material and is preceded by a longer and shorter piece of a given metody (Fig. 1). A series consists of

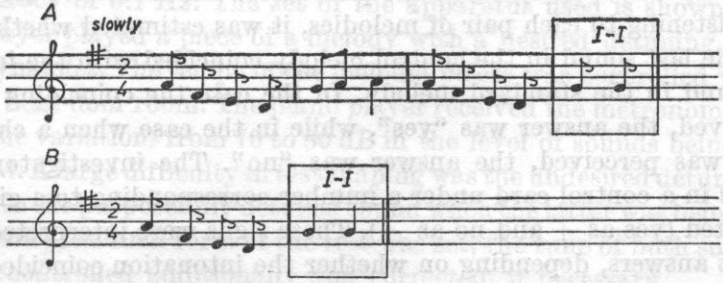


Fig. 1. The music material of series I of the test for forms III/IV; A — the piece of melody preceding the test, B — the test melody

12 pairs of a two-bar melody, i.e. of 10 pairs of melody pieces, which differ in the magnitude of the detuning of the last sound, and of 2 pairs in which the last sound is not detuned. In each pair the first melody is standard and not detuned while the second — the control melody — contains (in 40 cases) or does not contain (8 cases) a detuning of the last sound of a melody. Thus, each coloured test contains 48 tasks given in a quasi-stochastic order. The order of tasks in the pianoforte, flute and violin tests is given in Table 2. The detuning range

Table 2. The set of test tasks in pianoforte, flute and violin tests for the group from forms III/IV

Test task no	Detuning (ct)					
	pianoforte test		flute test		violin test	
1	+	100-75	+	100-75	+	100-75
2	−	50-25	−	50-25	−	50-25
3		0		0		0
4	+	50-25	+	50-25	+	50-25
5	−	15-10	−	15-10	−	25-15
6	+	25-15	+	25-15	+	25-15
7	−	100-75	−	100-75	−	100-75
8	+	25-10	+	15-10	+	15-10
9	−	75-50	−	75-50	−	75-50
10		0		0		0
11	+	75-50	+	75-50	+	75-50
12	−	25-15	−	25-15	−	15-10

+ and − denote the upward and downward mistunings, respectively

of the second sound of an interval varied in the following intervals; 100-75, 75-50, 50-25, 25-15 and 15-10 ct upwards and downwards. The tolerance range resulted from the restricted real possibilities of tuning in the particular sounds of instruments used.

The pause between each pair of tests was 5s. This pause was destined for the answer of the person tested. The pause within each pair of tasks was 1.5s (mostly hardly above 1.5s), while the pause between successive series was 10s.

After listening to each pair of melodies, it was estimated whether the intonation of the last sound in the control melody coincided with the intonation of the last sound in the standard melody. In the case the coinciding intonations were perceived, the answer was "yes", while in the case when a change in the intonation was perceived, the answer was "no". The investigator wrote the assessments in a control card under a number corresponding to a given melody pair estimated (yes as + and no as −). These signs were interpreted as correct or incorrect answers, depending on whether the intonation coincided in a given pair or the detuning took place. Each song on which the tests were drawn had been sung in class before.

b. Sinusoidal tests. A sinusoidal test had the same composition as the coloured tests, consisting of four series corresponding to the same intervals: series I — 1, II — $2m$, III — $2w$ and IV — $3m$. Each series contained 12 pairs of isolated intervals given in a quasi-stochastic order, where in 10 pairs the first interval was a standard, while the second was a mistuned one, and in 2 pairs both intervals were not detuned. The particular intervals were elaborated with keeping the same pitch register and the motion direction as in the coloured tests. The detunings of the second tone were 75, 50, 15 and 10 ct upwards and downwards. The duration of the interval tones was 1.00 s, the pause between the intervals in a pair was 1.67 s, while the response time was 1.77 s. The rise and decay times of the signals were 0.05 s.

3. Recording procedure

a. Sinusoidal tests. The tests of sinusoidal tones were programmed using a three — channel analog modulator [5]. The standard and control (variable) intervals were supplied from a sinusoidal tone generator at a grequency stability of the order of 10^{-4} per day. The signal frequency was measured with an electronic counter with an accuracy of 10^{-2} Hz. The durations of the intervals and of the pauses between the intervals were controlled by a unit of logical modules with an accuracy of 1 per cent. The sound material was recorded on a magnetic tape with a Nagra IV-*SJ* tape recorder at a tape shift rate of 0.19 m/s.

b. Coloured tests. In producing pianoforte tests, sounds were detuned by normal tuning of an instrument. A piano-tuner tuned or detuned given sounds to a desired pitch by ear. At the same time, the correctness of detuning was controlled by observing the Lissajous figures on an oscilloscope display. Control measurements showed that after some time the tuner remained in most cases within the predetermined detuning limits.

Electronic control consisted in precise comparison of the frequency of the fundamental tone of a given piano sound with a signal of the same frequency, with an accuracy of 0.1 Hz. The set of the apparatus used is shown in Fig. 2. The piano player played a piece of a melody with a desired detuning, at a fixed tempo and dynamics. The tempo of the melodic pieces was controlled with a metronome in a next-door room. The piano player received the metronome signal by earphones. The variations from 70 to 80 dB in the level of sounds being recorded were tolerated. A large difficulty in test shaping was the undesired detuning of sounds adjacent to the purposefully detuned sound when the latter was being detuned. Therefore, before the final form of the test was set, the tune of both sounds of an interval was controlled additionally and corrected, if necessary.

The procedure of recording tests composed of flute and violin sounds was similar to that of recording the pianoforte tests. The difference consisted in

changing the order of recording and control. The musician formed the test material for recordings a number of times, detuning a given sound on the basis of subjective auditory assessment. And only then was the real magnitude of the detuning controlled by objective measurements, and the chosen variant of the

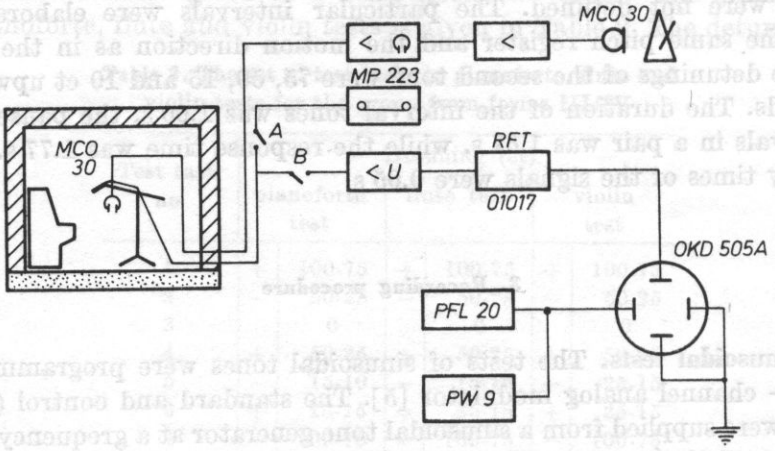


Fig. 2. A block diagram of the apparatus used for the recording and control of the music material for the coloured tests

whole melodic piece, which was assigned to a relevant group of detunings, to be edited further, was recorded. When all the detunings were ready, the whole series was edited. In this way, although the whole recording procedure was very long (particularly, in the case of the flute tests), the editing of a test within one melodic piece was avoided.

c. Control of tests. The experimental material of the violin and flute tests underwent a spot detailed check in terms of the frequency and the amplitude ($F0$ and $A0$) of sounds making up the intervals investigated. This check was based on the intonographic record made with an analogue-digital system [15]. The measurement set consisted of a MERA 30 mini computer (a digital convertor with memory), a DT-105 paper tape punch, and a CT-1001 punched tape reader.

4. Persons examined

On the basis of the results of the previously mentioned experiment performed under natural conditions, three children were chosen from each form in each centre. They represented the "average student" group in terms of musical training, with particular emphasis on the perception of melodic intervals. All

the children had normal physiological hearing, with a hearing loss tolerance of up to 30 dB.

There were the following age groups: the first consisted of 36 children from forms I and II (27 girls, 9 boys), including 3 children from the centres *M* (1+2), 20 children from the centres *m* (17+3), and 12 children from the centres *W* (9+3). The second group had 43 children from forms III and IV (28 girls, 15 boys), including 13 from the centres *M* (10+3), 18 from the centres *m* (7+11) and 12 from the centres *W* (11+1). The third group consisted of 63 children from forms V and VI (56 girls, 7 boys), including 22 from the centres *M* (21+1), 18 from the centres *m* (16+2) and 23 from the centres *W* (19+4).

The children selected for participation in the acoustic tests had earlier taken part in music testing in class. They were familiar with the laboratory conditions in the course of checking the audibility threshold curves. In order to prevent surprises in terms of organization, the children were instructed preliminarily about the investigation procedure. All the selected children volunteered for the tests, and were under the care of their music teacher. Four children had to be dropped in the course of the investigations, since they gave totally accidental answers, which was found in the preliminary control of answer repeatability. The total lack of repeatability may have been caused both by the bad perceptibility or the lack of attention.

5. Listenings

In order to avoid additional distortions and disturbances while reproducing tape recordings, caused by the amplification system, each time before the measurements followed, the nonlinear distortions of the power amplifier and the efficiency of the "contour" filter and the anti-noise filter were controlled. The same power level at the output of the amplification system was controlled before each listening.

The listenings took place in a GIG AU-1 audiometric booth, with SN-60 earphones, binaurally. The otter apparatus was installed in the room where the booth was. The connection diagram was the same as in paper [4].

The intensity level of the sound received corresponded approximately to the level on which the sound reached the microphone in the recordings. The student gave answers yes or no through a microphone in the booth, which permitted contact with the person recording the answers. A five-minute intermission was taken after two series. Each student was tested at least three times; with usually one session for each student per day; in the rare cases when the student lived far from Zielona Góra, two sessions were conducted. The duration of listening to each series was about 240 s. Each successive series was preceded with a special announcement.

6. Analysis of the results

The results were tabulated for each group, for the particular intervals individually. As an example, the results obtained for the children group of forms V and VI are shown in Tables 3-5. 75 per cent of the correct answers was taken as the criterion of interval detuning perceptibility. On the basis of the tables, histograms of the correct answers were elaborated in relation to the detuning degree of an interval (Figs. 3-5). The black column represents the percentage of the correct answers for a nondetuned interval.

a. Children group from forms I/II. Pianoforte tests. 36 children were tested. The maximum number of answers for each detuning in the intervals 1, $2m$, $2w$, $3m$ was 128 for each melody containing a detuned interval and 256 for a melody with a nondetuned interval.

The distinctly best perceptibility was found for the upward detuned prime interval (88 per cent of correct answers for 25-15 ct upwards), and minor third downwards (above 71 per cent of correct answers for 25-15 downwards). The perceptibility of detunings in the minor second and major second intervals was similar, including for the major second the detunings of 50-25 ct upwards and for the minor second and major second only 75-50 ct downwards*.

b. Children group from forms III/IV. Tests of all kinds. 43 children were tested with pianoforte tests. The maximum number of answers for each detuning in all the intervals was 151 for each melody containing a detuned interval and 302 for each melody containing a nondetuned interval.

The best perceptibility was found for the detuned minor third and prime, where, as in group I/II, the downward detuned minor third was perceived better than the upward one, while the detuned prime was recognized better for the upward case than the downward one. Even the least downward detuned minor third was perceived, i.e. 15-10 ct (81 per cent of correct answers), while the least perceptible prime detuning was 25-15 ct upwards (84 per cent of correct answers). The perceptibility of the detuned minor second and major second was similar and, as in group I/II, the detunings in these intervals were perceived up to the range 75-50 ct. Smaller detunings were not perceived.

33 children were tested with flute tests (27 girls, 6 boys). The maximum number of answers for each melody containing a detuned interval was 121, while for that with a nondetuned interval it was 242.

The best perceptibility of detunings was observed for the prime and major second intervals, where results obtained for the prime were similar to those in the pianoforte test. Both for the major second and the prime the upward detunings were perceived better than the downward ones. The least perceptible

* The test pair containing the mistuned interval $2m$ in the range 50-25 ct was damaged in the investigations.

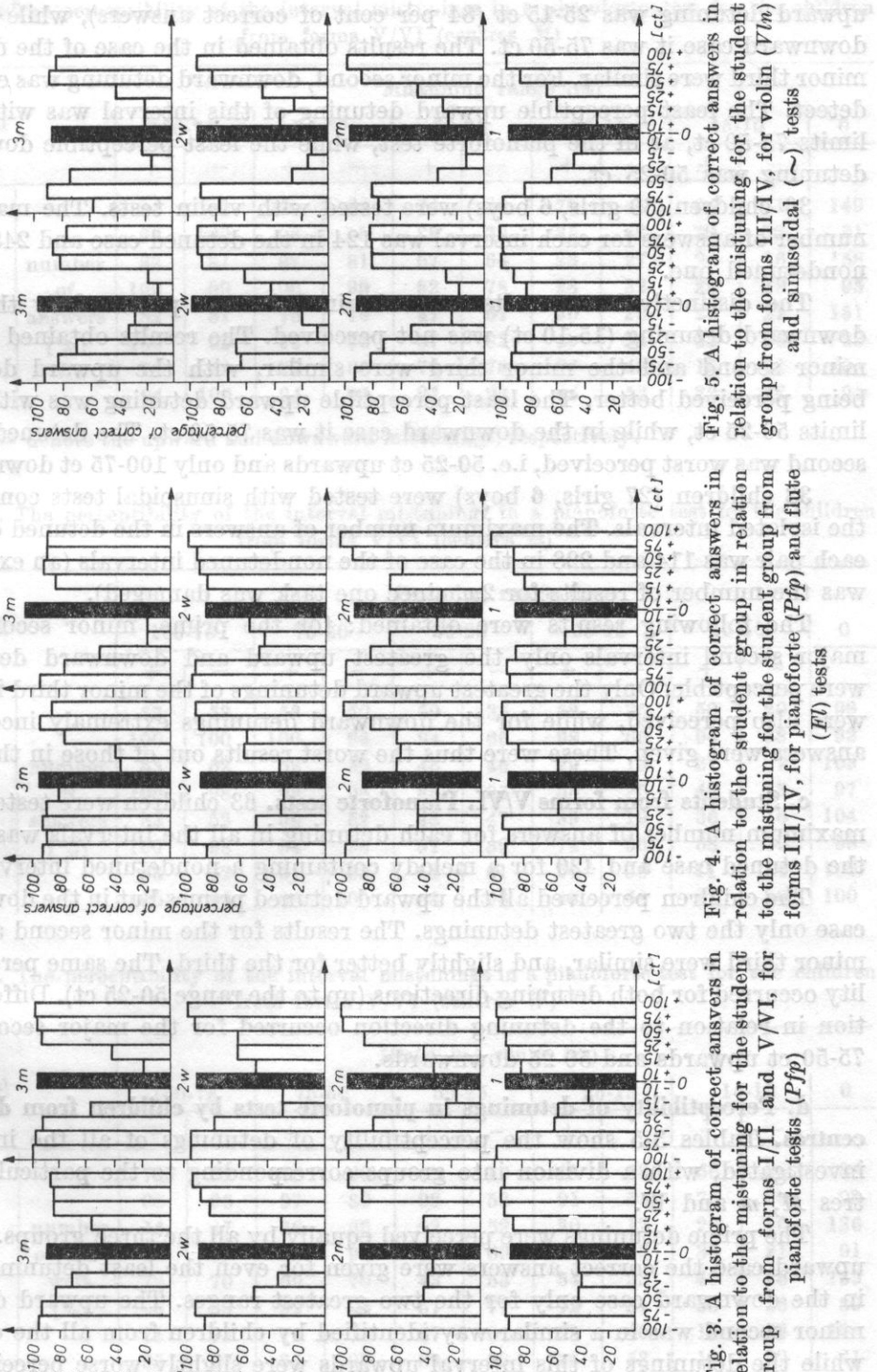


Fig. 3. A histogram of correct answers in relation to the mistuning for the student group from forms I/II and V/VI, for pianoforte tests (*Pfp*)

Fig. 4. A histogram of correct answers in relation to the student group in relation to the mistuning for the student group from forms III/IV, for pianoforte (*Pfp*) and flute (*Fl*) tests

Fig. 5. A histogram of correct answers in relation to the mistuning for the student group from forms III/IV, for violin (*VIn*) and sinusoidal (~) tests

upward detuning was 25-15 ct (84 per cent of correct answers), while for the downward case it was 75-50 ct. The results obtained in the case of the detuned minor third were similar. For the minor second, downward detuning was easier to detect. The least perceptible upward detuning of this interval was within the limits 75-50 ct, as in the pianoforte test, while the least perceptible downward detuning was 50-25 ct.

36 children (30 girls, 6 boys) were tested with violin tests. The maximum number of answers for each interval was 124 in the detuned case and 248 in the nondetuned one.

The distinctly best results were obtained for the prime. Only the least downward detuning (15-10 ct) was not perceived. The results obtained for the minor second and the minor third were similar, with the upward detuning being perceived better. The least perceptible upward detuning was within the limits 50-25 ct, while in the downward case it was 75-50 ct. The detuned major second was worst perceived, i.e. 50-25 ct upwards and only 100-75 ct downwards.

33 children (27 girls, 6 boys) were tested with sinusoidal tests concerning the isolated intervals. The maximum number of answers in the detuned case for each pair was 114 and 228 in the case of the nondetuned intervals (an exception was the number of results for $2w$, since one task was damaged).

The following results were obtained: for the prime, minor second and major second intervals only the greatest upward and downward detunings were perceptible. Only the greatest upward detunings of the minor third interval were also perceived, while for the downward detunings extremaly incoherent answers were given. These were thus the worst results out of those in the tests.

c. Students from forms V/VI. Pianoforte tests. 63 children were tested. The maximum number of answers for each detuning in all the intervals was 210 in the detuned case and 420 for a melody containing a nondetuned interval.

The children perceived all the upward detuned primes but in the downward case only the two greatest detunings. The results for the minor second and the minor third were similar, and slightly better for the third. The same perceptibility occurred for both detuning directions (up to the range 50-25 ct). Differentiation in relation to the detuning direction occurred for the major second, i.e. 75-50 ct upwards and 50-25 downwards.

d. Perceptibility of detunings in pianoforte tests by children from different centres. Tables 3-5 show the perceptibility of detunings of all the intervals investigated, with a division into groups corresponding to the particular centres M, m and W.

The prime detunings were perceived equally by all the three groups. In the upward case the correct answers were given for even the least detunings, and in the downward case only for the two greatest ranges. The upward detuned minor second was in a similar way identified by children from all the centres, while the detunings of this interval upwards were slightly worse perceived by

Table 3. The perceptibility of the interval mistunings in a pianoforte test for the children from forms V/VI (centres *M*)

Interval		Mistuning range (ct) 100-75 +	100-75 −	75-50 +	75-50 −	50-25 +	50-25 −	25-15 +	25-15 −	15-10 +	15-10 −	0
1		81	82	79	68	72	48	70	46	65	31	149
		99	100	96	83	88	58	85	56	79	38	91
2*m*	number	82	81	82	81	67	64	23	27	24	16	156
	of	100	99	100	99	82	78	28	33	29	19	95
2*w*	answers	82	81	76	79	47	67	40	20	29	21	151
	[%]	100	99	93	96	57	82	49	24	35	26	92
3*m*		77	82	77	80	76	76	37	43	18	21	155
		94	100	94	98	93	93	45	52	22	26	94

+ and − denote the upward and downward mistunings, respectively.

Table 4. The perceptibility of the interval mistunings in a pianoforte test for the children from forms V/VI (centres *m*)

Interval		Mistuning range (ct) 100-75 +	100-75 −	75-50 +	75-50 −	50-25 +	50-25 −	25-15 +	25-15 −	15-10 +	15-10 −	0
1		53	53	53	50	50	35	52	35	50	20	99
		100	100	100	94	94	66	98	66	94	38	93
2*m*	number	53	52	52	52	45	45	29	13	21	7	103
	of	100	98	98	98	85	85	55	25	40	13	97
2*w*	answers	53	52	50	52	43	47	39	19	36	16	104
	[%]	100	98	94	98	81	89	74	36	68	30	98
3*m*		53	52	51	53	52	51	24	36	11	14	106
		100	98	96	100	98	96	45	68	21	26	100

Table 5. The perceptibility of the interval mistunings in a pianoforte test for the children from forms V/VI (centres *W*)

Interval		Mistuning range (ct) 100-75 +	100-75 −	75-50 +	75-50 −	50-25 +	50-25 −	25-15 +	25-15 −	15-10 +	15-10 −	0
1		74	72	73	60	69	40	68	38	56	21	138
		98	96	97	80	92	53	91	51	75	28	92
2*m*	number	74	67	56	69	62	52	30	17	26	16	136
	of ans-	99	89	75	92	83	69	40	23	35	21	91
2*w*	wers	72	70	69	70	43	53	54	23	40	25	133
	[%]	96	93	92	93	57	71	72	31	53	33	89
3*m*		71	73	71	73	67	67	27	40	12	28	136
		95	97	95	97	89	89	36	53	16	37	91

children from the centres *W* (by one detuning group). In the case of the major second and minor third intervals the detunings were perceived best by the children from the W centres. Thus, no generally distinct difference was found between the results of children from the groups *M* and *W*; and the group *m* had the best results.

This, however, cannot be taken as a rule, since in the group from forms I/II the children from the centres *W* perceived the detunings better than the children from the group *m*.

No general relation was thus found to exist between the character of the centre and the perceiving abilities in terms of detunings, although it can be expected that in larger centres with the usually more intense musical life a greater musical training of children occurs, which is significant in the perception of detunings.

7. Conclusions

The following general conclusions can be drawn from the results of the investigations performed on all the student groups in all the tests.

1. The upward detuned interval 1 is best perceptible up to the range 25-15 ct, where in two cases: in a violin test for forms III/IV and in a pianoforte test for forms V/VI, even the least detunings were perceived. The best perceptibility of the detunings of the interval 1 was found for violin tests.

2. In all the tests and age groups the detuning of the interval 3 *m* was perceptible up to the range 50-25 ct, i.e. a detuning less than a quarter tone. An exception was made by the flute and violin tests for forms III/IV with the downward detunings (50 per cent and 69 per cent of correct answers, respectively).

3. The upward detunings of the interval 2*w* were worst perceptible.

4. No relation was found to exist between the ability of perceiving the intervals 1,2*m*, 2*w* and 3*m* in an "average student" and his age or the character of the centre from which he came.

5. A student of average musical ability can perceive the detunings in the melodic intervals below a quarter tone. This fact should be taken into account while programming musical training in school.

6. None of the three sound timbres (pianoforte, violin, flute) dominates unambiguously over other two in terms of the effect on the perceptibility of the detunings of the melodic intervals 1,2*m*, 2*w* and 3*m*.

7. The detuned intervals composed of sinusoidal tones are less perceptible than those composed of coloured sounds of such instruments as the piano, violin, or flute. It seems, therefore, that the sound timbre plays an essential role in perceiving the detuned melodic intervals. What may also be significant is the context in which the interval occurs (the sinusoidal intervals were presented

in isolation, while the "coloured" intervals were given in the context of a melodic phrase).

8. In the group of the intervals investigated the best perceptibility was found for the upward detuning of 1, and the worst for the downward detuning of $2w$.

9. A tendency for the zone of the interval 1 to expand was found, as a result of the tolerance of the upward detuning of this interval.

10. $2w$ shows a tendency to expand its magnitude, as a result of the tolerance of the downward detunings, which agrees with the detuning direction and the motion direction of the interval.

11. In moving from degree VII to degree VIII $2m$ shows a tendency to decrease its magnitude, as a result of the higher tolerance of the downward detuning.

12. $3m$ shows no tendency to either decrease or increase its magnitude.

Acknowledgements. The recordings of sinusoidal tests were performed on apparatus of Acoustics Department of Chopin Academy of Music, Warsaw. The measurements were taken on apparatus of Institute of Applied Linguistics, Warsaw University, in cooperation with Laboratory of Speech Acoustics, Institute of Fundamental Technological Research.

References

[1] E. CHMIELEWSKA, *Perception of isolated intervals and intervals in melodic structures* (in Polish), Mat. Pomocnicze dla Nauczycieli Szkół i Ognisk Artystycznych, COPSA, 99, 43-78 (1966).

[2] W. J. DOWLING, *Scale and contour: two components of a theory of memory for melodies*, Psychol. Review, **85**, 4, 341-354 (1978).

[3] M. DROBNER, *New results of the investigation of the zonality of the interval hearing* (in Polish), Zesz. Naukowe PWSM, Cracow, **5** (1967).

[4] H. HARAJDA, J. FYK, *Zonality of melodic intervals in utterances of average gifted child* (in Polish), Proc. XXV OSA 1978, 131-134.

[5] A. JAROSZEWSKI, A. RAKOWSKI, *Analogue modulator for psychoacoustical pulse measurements*, Archives Acoust. **1**, 1, 73-79 (1976).

[6] M. KACZURBINA, L. MIKLASZEWSKI, *Musical training, V form school-book* (in Polish), PZWS, Warsaw 1972.

[7] E. LEIPP, *Acoustique et musique*, Paris 1971.

[8] K. LEWANDOWSKA, *Development of musical abilities in school age children* (in Polish), WSiP, Warsaw 1978.

[9] O. ORTMANN, *On the melodic relativity of tones*, Psychological Monograph., **35**, 1, 1-47 (1926).

[10] A. RAKOWSKI, *Categorial perception of sound pitch in music* (in Polish), Chopin Academy of Music, Warsaw 1978.

[11] A. RAKOWSKI, K. MŁODZIEJEWSKA, W. MARUCHA, *Intoning isolated musical intervals* (in Polish), Zesz. Nauk. PWSM, Warsaw, **IV**, 155-166 (1979).

[12] O. Sakhaltueva, *O niekotoryh zachonomiernostiach intonirowanija w swiazi z formoj, dinamikoj i wadom*, Trudy Kaf. Teorii Muzyki, Moskowskoj Konserwatorii, Muzgiz, Moscow 1960.

[13] T. Tarnoczy, O. Szende, *Experiments on the musical interval sense*, Proc. 7th ICA Congress, Budapest, 3, 4, 7, 695-698, 1960.

[14 W. D. Ward, *Subjective musical pitch*, JASA, **26**, 3, 369-380 (1954).

[15] P. Żarnecki, *An analogue-digital system for multi-channel signal processing used for the measurements of time and frequency characteristics of acoustic systems* (in Polish), Proc. XXV OSA 1978, Poznań, pp. 379-382.

[16] E. Lipska, M. Przychodzińska, *Methodology of musical education in forms I-II of primary school* (in Polish), PZWS, Warsaw 1973, p. 164.

[17] E. Lipska, M. Przychodzińska, *Musical education, from VI* (in Polish), PZWS, Warsaw 1973, pp. 36, 69.

[18] E. Lipska, M. Przychodzińska, *Notes, my notes* (in Polish), WSiP, Warsaw 1973, pp. 8-9, 16-18.

[19] J. Powroźniak, *The children went into the garden* (in Polish), PWM, Cracow 1964, p. 18.

[20] M. Przychodzińska, *Methodology of musical education in forms I-IV* (in Polish), PZWS, Warsaw 1973, p. 35.

[21] Z. Stankiewicz, *A song in a nursery school* (in Polish), PZWS, Warsaw 1973, p. 63.

[22] *We learn the music, form III* (in Polish), PZWS, Warsaw 1970, p. 114.

Received on December 19, 1979; revised version on May 19, 1981.

ARCHIVES OF ACOUSTICS
6, 4, 385-400 (1981)

MAGNETOMECHANICAL COUPLING IN TRANSDUCERS

ZBIGNIEW KACZKOWSKI

Institute of Physics, Polish Academy of Sciences
(02-668 Warsaw, Al. Lotników 32/46)

The basic definitions of the coefficient of magnetomechanical coupling, k, based on generalized definitions of magnetic, mechanical (elastic) and piezomagnetic energies, and on the magnetic quantities in the case of transmitting transducers and the mechanical quantities in the case of receiving transducers, i.e. for the mechanomagnetic coefficient, are presented. The relations between the coupling coefficient and the coefficients occurring in each system of piezomagnetic equations and for different systems of equations and systems of units, and also for induction and magnetization, are derived.

1. Initial remarks

The properties of piezomagnetic (magnetostrictive) materials are characterized by piezomagnetic coefficients occurring in piezomagnetic equations [1-12, 14, 15] and by other basic quantities and physical parameters, such as e.g. the induction (magnetization M, I, J) of saturation B_s, the magnetostriction of saturation λ_s, the magnetocrystalline anisotropy constants $K_1, K_2, K_3, \ldots$, the density ϱ, the resistivity ϱ_{el}, the sound velocities and their temperature dependencies, the Curie temperature T_C or the Néel temperature T_N, the mechanical strength etc. [1-5, 8, 10-12, 14-15].

Of piezomagnetic coefficients occurring in piezomagnetic equations, such mechanical coefficients can be distinguished as the coefficients c or the moduli (s, E, G) of elasticity defined for the constant field intensity H or the constant magnetization (M, I, J), or the induction B, e.g. s_H, E_H, E_J, E_B; the magnetic coefficients (of permeability μ or of susceptibility $\varkappa$, or their reciprocals) defined for the constant mechanical stress T, σ, τ (of a free sample), or for the constant strains S, ε, λ, e.g. $\mu_T, \mu_S, \mu_\lambda, \varkappa_\sigma, \varkappa_\varepsilon$; and the magnetomechanical coefficients such as the piezomagnetic sensitivity d or λ, the constant (coefficient)

of magnetostriction h, λ or Γ, the coefficient of field stresses e and the coefficient of inductive strains g, e.g. [1-12, 14, 15]. Although each of these coefficients describes some property of piezomagnetic material as a function of magnetization, amplitude, or temperature, but none of them is itself a sufficient parameter for the usefulness of the piezomagnetic material to be satisfactorily determined. It is only the set of piezomagnetic coefficients occurring in at least one of the systems of piezomagnetic equations, i.e. at least 3 coefficients, e.g. E_H, μ_T and d, that gives almost enough information. None of them, however, can be a measure of the effectiveness of energy conversion.

This function is fulfilled by the coefficient of magnetomechanical coupling, k, which in addition to the mechanical and electrical quality factor, the electroacoustical efficiency, the vibration amplitude etc., permits the comparison of the properties of piezotronic materials and piezomagnetic transducers [1, 10-12, 14]. Its analogue in piezoelectric materials is the coefficient of electromechanical coupling [1, 10-12, 14].

Part of the energy supplied to the transmitting transducer is converted to the energy of elastic oscillation and radiated into the medium loading the transducer. The opposite process occurs in the receiving transducer. The mechanical energy E_e (the acoustic energy E_a), supplied to the receiver, is partly converted to the magnetic energy E_m (the electrical energy E_{el}). The measure of this transformation is the coupling coefficient: that of magnetomechanical coupling in the case of piezomagnetic materials and transducers, and that of electromechanical coupling in the case of piezoelectrical materials and transducers, e.g. [1, 14]. The coefficient of optomechanical coupling can be defined similarly in the case of piezooptic materials and transducers etc.

2. Definition of coupling

2.1. *Coupled electrical circuits*

Coupled electrical circuits are the systems permitting the transmission of the *ac* power from one circuit to another.

The quantity characterizing quantitatively the degree of coupling is the coupling coefficient which is in direct proportion to power transmitted, and in inverse proportion to the geometrical mean of the power of the circuits,

$$k = \frac{W_p}{\sqrt{W_1 W_2}} = \frac{\hat{Z}_{sp}}{\sqrt{\hat{Z}_1 \hat{Z}_2}}, \tag{1}$$

where W_1 and W_2 are the energies of the primary and the secondary circuits, respectively; W_p is the transmitted energy, $\hat{Z}_{sp}$, $\hat{Z}_1$ and $\hat{Z}_2$ are the characteristic impedances of the coupling element and the primary and the secondary circuits, respectively. As in the theory of electrical circuits, the definition of coupling is derived in the case of piezotronic materials and components.

2.2. *Generalized definitions of energies related to the coupling*

The part of energy which is transformed to another energy in magneto-electro- or opto-mechanical transformations, can be called the piezo-magnetic, piezo-electric, or piezo-optic energy, W_p.

The proper energies (energy densities), which participate in conversion in the case of piezomagnetic transducers, i.e. the magnetic energy W_m, the mechanical (elastic) energy W_e and the piezomagnetic energy W_p, can be defined in the following way:

$$W_m = \frac{1}{2}\sum_{i=1}^{i=3}\sum_{j=1}^{j=3} B_i H_j = \frac{1}{2}\sum_{i=1}^{i=3}\sum_{j=1}^{j=3} \mu_{Tij} H_i H_j, \tag{2}$$

$$W_e = \frac{1}{2}\sum_{p=1}^{p=6}\sum_{q=1}^{q=6} S_p T_q = \frac{1}{2}\sum_{p=1}^{p=6}\sum_{q=1}^{q=6} s_{Hpq} T_p T_q = \frac{1}{2}\sum_{p=1}^{p=6}\sum_{p=1}^{q=6} \frac{1}{c_{Hpq}} T_p T_q, \tag{3}$$

$$W_p = \frac{1}{2}\sum_{i=1}^{i=3}\sum_{q=1}^{q=6} d_{iq} H_i T_q, \tag{4}$$

where B_i, H_i and H_j are respectively 3 components of the axial vectors of the induction B and the field intensity H; S_p, T_p and T_q are respectively 6 components of the strain tensors S_{ij} and the mechanical stresses T_{ij}; μ_{Tij} are 9 components of the tensor of magnetic permeability at constant stresses (of a free sample); s_{Hpq} are 36 components of the tensor of the elastic coefficient at constant magnetic field, i.e. the proper state corresponding in the extreme to an open electrical circuit, which is in practice a system of constant current efficiency, i.e. very high impedance; and d_{iq} are 18 components of the pseudotensor of piezomagnetic sensitivity. The differential signs were neglected in all these quantities, but in should be borne in mind that these relations are valid over the range of small reversible lossless linear variations of the quantities discussed above.

2.3. *Components of magnetic, mechanical and piezomagnetic vectors and tensors*

The components of the vectors and tensors occurring in equations (2)-(4) in the most general expansion before simplification will have the following form,

$$B = B_i \quad (i = 1, 2, 3 \text{ or } x, y, z), \tag{5}$$

$$H = H_j \quad (j = 1, 2, 3 \text{ or } x, y, z), \tag{6}$$

$$S = S_{ij} \quad (i, j = 1, 2, 3 \text{ or } x, y, z), \tag{7}$$

$$T = T_{ij} \quad (i, j = 1, 2, 3 \text{ or } x, y, z), \tag{8}$$

$$\mu_T = [\mu_{Tij}] = \begin{bmatrix} \mu_{T11}, \mu_{T12}, \mu_{T13} \\ \mu_{T21}, \mu_{T22}, \mu_{T23} \\ \mu_{T31}, \mu_{T32}, \mu_{T33} \end{bmatrix} = \begin{bmatrix} \mu_0(1+\varkappa_{T11}), \mu_0\varkappa_{T12}, \mu_0\varkappa_{T13} \\ \mu_0\varkappa_{T21}, \mu_0(1+\varkappa_{T22}), \mu_9\varkappa_{T23} \\ \mu_0\varkappa_{T31}, \mu_0\varkappa_{T32}, \mu_0(1+\varkappa_{T33}) \end{bmatrix}, \tag{9}$$

$$d = d_{ijk} \quad (i,j,k = 1,2,3), \tag{10}$$

$$c = c_{Hijkl} \quad (i,j,k,l = 1,2,3), \tag{11}$$

$$s = s_{Hijkl} \quad (i,j,k,l = 1,2,3). \tag{12}$$

2.4. *Magnetic energy*

The energy stored in free material in the course of magnetization can be defined on the basis of work put in its magnetization. The increases in energy of a free unit volume (in the case of isothermal and reversible transformation) can be defined by the following relations:

$$dW_m = HdB_T = \mu_0 Hd(H+M_T) = \mu_0(1+\varkappa_T)HdH = \mu_T HdH \tag{13}$$

or

$$dW_m = H_i dB_{Ti} = \mu_{Tij} H_i dH_j, \tag{14}$$

which after expansion gives

$$dW_m = \mu_{T11}H_1dH_1 + \mu_{T12}H_1dH_2 + \mu_{T12}H_1dH_3 + \mu_{T21}H_2dH_1 + \\ + \mu_{T22}H_2dH_2 + \mu_{T23}H_2dH_3 + \mu_{T31}H_3dH_1 + \mu_{T32}H_3dH_2 + \mu_{T33}H_3dH_3. \tag{15}$$

Symmetricity of the tensor μ_{ij} can be proved in the following way. The differentials of the increases in energy with respect to the components of the fields H_1 and H_2 have the following form

$$\frac{\partial W_m}{\partial H_1} = \mu_{T11}H_1 + \mu_{T21}H_2 + \mu_{T31}H_3, \tag{16}$$

$$\frac{\partial W_m}{\partial H_2} = \mu_{T12}H_1 + \mu_{T22}H_2 + \mu_{T32}H_3. \tag{17}$$

Differentiation of equation (16) with respect to H_2 and that of (17) with respect to H_1 give

$$\frac{\partial}{\partial H_2}\left(\frac{\partial W_m}{\partial H_1}\right) = \mu_{T21} = \frac{\partial^2 W_m}{\partial H_1 \partial H_2}, \tag{18}$$

$$\frac{\partial}{\partial H_1}\left(\frac{\partial W_m}{\partial H_2}\right) = \mu_{T12} = \frac{\partial^2 W_m}{\partial H_1 \partial H_2}, \tag{19}$$

and thus

$$\mu_{T12} = \mu_{T21}, \tag{20}$$

and it can be shown similarly that

$$\mu_{T13} = \mu_{T31}, \tag{21}$$

$$\mu_{T23} = \mu_{T32}, \tag{22}$$

i.e. the tensor μ_{Tij} is symmetrical.

After integration of equation (15) and consideration of relations (20)-(22), the proper magnetization energy is

$$W_m = \frac{1}{2}\mu_{T11}H_1^2 + \mu_{T12}H_1H_2 + \mu_{T13}H_1H_3 + \frac{1}{2}\mu_{T22}H_2^2 + \mu_{T23}H_2H_3 + \frac{1}{2}\mu_{T33}H_3^2. \tag{23}$$

Using the Einstein summation notation saying that when a letter index occurs twice in the same term, summation should be done with respect to it, relations (2) and (23) can be written in the following way

$$W_m = \frac{1}{2}\mu_{Tij}H_iH_j \quad (i,j = 1,2,3). \tag{24}$$

2.5. *Elastic strain energy*

The change in the energy of elastic strain of a material unit volume, W_e, after consideration of the Hooke law, is defined by the following relations

$$dW_e = T_{ij}dS_{Hij} = s_{Hijkl}T_{ij}dT_{kl} \quad (i,j,k,l = 1,2,3). \tag{25}$$

In equation (25), 81 components occur in the most general form. In specific cases some simplifications can be introduced and in the extreme case of strains, e.g. longitudinal linear strains, 1 elastic coefficient, known as the Young's modulus Y_H or E_H, occurs. Similarly to the magnetic energy case for the symmetry condition, the number of different elastic coefficients may be reduced from 81 up to the maximum 36 [13, 16]. The components of the stress tensor, T_{ij}, define the forces acting on the elementary surfaces of an elementary solid cube. The first index: 1, 2, 3 or x, y, z denote the direction normal to the surface considered, e.g. x normal to the yz surface, while the second index denotes the direction of the stress component. With absent bulk forces and bulk torsional moments, and with uniform stresses, for the state of static equilibrium the tensor is symmetrical and has 6 components, i.e.

$$T_{ij} = T_{ji} \quad (i,j = 1,2,3). \tag{26}$$

The tensor of the strains S_{ij} is similarly symmetrical, i.e.

$$S_{ij} = S_{ji} \quad (i,j = 1,2,3). \tag{27}$$

The symmetry of elastic coefficients follows from the symmetry of the strain and stress tensors

$$c_{Hijkl} = c_{Hjikl} = c_{Hijlk} \quad (i,j,k,l = 1,2,3). \tag{28}$$

For an elastic medium, the strain energy is a function of state, independent of intermediate states, and then

$$c_{Hijkl} = c_{Hklij}. \tag{29}$$

The Hooke's law in tensor notation has the following forms

$$T_{ij} = c_{Hijkl} S_{kl} \quad (i,j,k,l = 1,2,3). \tag{30}$$

$$S_{ij} = s_{Hijkl} T_{kl} \quad (i,j,k,l = 1,2,3). \tag{31}$$

The flexibility or the coefficient (constant) of elasticity, c_{Hijkl}, is the inverse elastic modulus s_{Hijkl} and

$$s_{Hijkl} = s_{Hjikl} = s_{Hijlk} = s_{Hklij}, (i,j,k,l = 1,2,3). \tag{32}$$

The symmetry of the coefficients with respect to the first or second pair of indices permits a simpler matrix form to be used. Two-digit tensor notation can be replaced with single-digit matrix form, e.g. [13, 16], according to the following principle:

$$11 \to 1,\ 22 \to 2,\ 33 \to 3,\ 23 \to 4,\ 13 \to 5,\ 32 \to 4,\ 31 \to 5,\ 12 \to 6,\ 21 \to 6,$$

i.e. for example,

$$[T_{p.}] = [T_1, T_2, T_3, T_4, T_5, T_6], \quad [S_q] = [S_1, S_2, S_3, S_4, S_5, S_6],$$

and

$$s'_{Hpq} = s_{Hijkl} \begin{cases} p = i, & \text{when } i = j,\ p = m+3, \quad \text{when } i \neq j \neq m \\ & \qquad (i,j,k,l = 1,2,3), \\ q = k, & \text{when } k = l,\ q = m+3, \quad \text{when } k \neq l \neq m \\ & \qquad (p,q = 1,2,3,\ldots,6), \end{cases} \tag{33}$$

and in view of the symmetry of strains it can be assumed that

$$2s'_{Hij} = s_{Hij}, \quad \text{when } i \neq j = 1,2,3, \tag{34}$$

and the dash can be neglected in s_{Hii}, i.e.

$$s'_{Hij} = s_{Hij}, \quad \text{when } i = j = 1,2,3. \tag{35}$$

After the above simplifications in notation, integration, reduction and ordering, equation (25) for the proper energy of elastic strains becomes

$$\begin{aligned} W_e = {} & \frac{1}{2} s_{H11} T_1^2 + s_{H12} T_1 T_2 + s_{H13} T_1 T_3 + s_{H14} T_1 T_4 + s_{H15} T_1 T_5 + \\ & + s_{H16} T_1 T_6 + \frac{1}{2} s_{H22} T_2^2 + s_{H23} T_2 T_3 + s_{H24} T_2 T_4 + s_{H25} T_2 T_5 + s_{H26} T_2 T_6 + \\ & + \frac{1}{2} s_{H33} T_3^2 + s_{H34} T_3 T_4 + s_{H35} T_3 T_4 + s_{H36} T_3 T_6 + \frac{1}{2} s_{H44} T_4^2 + s_{H45} T_4 T_5 + \\ & + s_{H46} T_4 T_6 + \frac{1}{2} s_{H55} T_5^2 + s_{H56} T_5 T_6 + \frac{1}{2} s_{H66} T_6^2, \end{aligned} \tag{36}$$

which in shorter notation gives the form

$$W_e = \frac{1}{2} s_{Hpq} T_p T_q \quad (p, q = 1, 2, 3, 4, 5, 6), \tag{37}$$

i.e. formula (3) in the Einstein notation.

2.7. *Piezomagnetic energy*

The piezomagnetic energy of a transducer is the part of magnetic or mechanical energy which is transformed into other energy. The proper piezomagnetic energy is characterized, for example, as the product of additional strain caused by the intensity of the magnetic field H_k

$$S_i = d_{ijk} H_k \quad (i, j, k = 1, 2, 3) \tag{38}$$

and the mechanical stresses T_{ij}

$$W_p = \frac{1}{2} d_{ijk} T_{ij} H_k \tag{39}$$

or the additional induction B_i (magnetization M_i) caused by the external stresses T_{ij}

$$B_i = d_{ijk} T_{jk} \quad (i, j, k = 1, 2, 3), \tag{40}$$

and the field intensity H_i, i.e. again

$$W_p = \frac{1}{2} d_{ijk} H_i T_{jk}. \tag{41}$$

In the case of the symmetrical tensor T_{jk} (with the absence of the internal moments of a force), the piezomagnetic sensitivity, which functions here as the proportionality coefficient, is symmetrical, i.e.

$$d_{ijk} = d_{ikj} \tag{42}$$

and the number of its independent components reduces from 27 (see formula (10)) to 18.

After consideration of (42), formulae (39) or (41) in the expanded form are the following

$$\begin{aligned} W_p = {} & \frac{1}{2} d_{111} H_1 T_{11} + d_{112} H_1 T_{12} + d_{113} H_1 T_{13} + \frac{1}{2} d_{122} H_1 T_{22} + d_{123} H_1 T_{23} + \\ & + \frac{1}{2} d_{133} H_1 T_{33} + \frac{1}{2} d_{211} H_2 T_{11} + d_{212} H_2 T_{12} + d_{213} H_2 T_{13} + \frac{1}{2} d_{222} H_2 T_{22} + \\ & + d_{223} H_2 T_{23} + \frac{1}{2} d_{233} H_2 T_{33} + \frac{1}{2} d_{311} H_3 T_{11} + d_{312} H_3 T_{12} + d_{313} H_3 T_{13} + \\ & + \frac{1}{2} d_{322} H_3 T_{22} + d_{323} H_3 T_{23} + \frac{1}{2} d_{333} H_3 T_{33}. \end{aligned} \tag{43}$$

Substitution of the single-digit tensor notation for the two-digit form leads to 6 components of the stress tensor T_q and 18 components of the tensor of piezomagnetic sensitivity, d_{iq},

$$W_p = \frac{1}{2} d_{iq} H_i T_q, \tag{44}$$

i.e. formula (4) in the Einstein notation.

2.8. *Definition of the coefficient of magnetomechanical coupling*

The magnitude of magnetomechanical coupling depends on the products of the density ratios of the energy transformed to the supplied. As in the theory of circuits, the coefficient of magnetomechanical coupling, k, is defined as the ratio of the piezomagnetic (coupling) energy to the geometrical mean of magnetic and mechanical energy,

$$k = \frac{W_p}{\sqrt{W_m W_e}}, \tag{45}$$

which after insertion of relations (2)-(4), or simpler relations (24), (37) and (44), and squaring, gives

$$k^2 = \frac{(d_{iq} H_i T_q)^2}{\mu_{Tij} H_i H_j s_{Hpq} T_p T_q}. \tag{46}$$

For simple longitudinal or radial vibration, most components vanish and only one remains, e.g. k_{33}. Then indices and summations can be neglected, and

$$k^2 = \frac{d^2 H^2 T^2}{\mu_T H^2 s_H T^2} = \frac{d^2}{\mu_T s_H} = \frac{d^2 E_H}{\mu_T}. \tag{47}$$

3. Coefficients of magnetomechanical and mechanomagnetic coupling

3.1. *Coefficients of magnetomechanical coupling*

The coupling coefficient, defined in the previous section, is universal, since it applies both to magnetomechanical transducers, or magnetoacoustic transducers transforming magnetic energy into mechanical or acoustic energy, and to mechanomagnetic (acoustomagnetic) transducers under the assumption of idealized states, i.e. when no losses occur, for example, in magnetic, magnetomechanical, or mechanical energy, and of reversible behaviour. Under this assumption the direction of change is not specified and the definition is thus universal.

The coefficient of magnetomechanical coupling may be compared with the coefficient of electromechanical coupling which occurs in piezoelectric or electrostrictive transducers. The fact, however, that it is magnetic energy stored in

the transducer core that is converted, as opposed to electric energy in the other case, requires these two quantities to be differentiated.

The squared coefficient of magnetomechanical coupling in the case of transmitting transducers defines which part (W_p) of the magnetic energy W_m is converted into the mechanical (acoustic) energy W_e. The transformed piezomagnetic energy W_p is the difference in magnetic energy of a transducer between the two mechanical states: at a free state ($T = \text{const}$) and a clamped state ($S = \text{const}$), i.e.

$$W_p = W_T - W_S, \tag{48}$$

$$W_T = \frac{1}{2} HB_T = \frac{1}{2} \mu_T H^2, \tag{49}$$

$$W_S = \frac{1}{2} HB_S = \frac{1}{2} \mu_S H^2, \tag{50}$$

$$k^2 = \frac{W_p}{W_T} = \frac{W_T - W_S}{W_T} = \frac{\mu_T - \mu_S}{\mu_T} = 1 - \frac{\mu_S}{\mu_T}, \tag{51}$$

where H and B are the amplitudes of magnetic field and induction, and μ_T and μ_S are the permeabilities of a free and clamped, respectively.

3.2. *Coefficient of mechanomagnetic coupling*

The opposite case, with respect to the one discussed above, occurs for receiving transducers. In piezomagnetic receivers part of mechanical energy is converted to magnetic energy. The squared coefficient of mechanomagnetic coupling is defined in the following way

$$k^2 = \frac{W_p}{W_H} = \frac{W_H - W_B}{W_H}, \tag{52}$$

where W_H and W_B are the mechanical energies at the constant field intensity H (unloaded, open system) and the constant induction B (closed system), i.e.

$$W_H = \frac{1}{2} S_H T = \frac{1}{2} \frac{T^2}{c_H} = \frac{1}{2} \frac{T^2}{E_H} = \frac{1}{2} s_H T^2, \tag{53}$$

$$W_B = \frac{1}{2} S_B T = \frac{1}{2} \frac{T^2}{c_B} = \frac{1}{2} \frac{T^2}{E_H} = \frac{1}{2} s_B T^2, \tag{54}$$

$$k^2 = \frac{s_H - s_B}{s_H} = \frac{\dfrac{1}{c_H} - \dfrac{1}{c_B}}{1/c_B} = \frac{\dfrac{1}{E_H} - \dfrac{1}{E_B}}{1/E_H} = 1 - \frac{E_H}{E_B}. \tag{55}$$

4. Relations between the coefficient of magnetomechanical coupling k and the other piezomagnetic coefficients

Relation (46) between the coefficient of magnetomechanical coupling, k, and the piezomagnetic sensitivity d can be derived from the relevant system of piezomagnetic equations I B [5-10, 13]

$$\delta S = \frac{\delta T}{E_H} + d\delta H, \tag{56}$$

$$\delta B = \mu_T \delta H + d\delta T. \tag{57}$$

For a rigid sample (S = const) $\delta S = 0$, and insertion of δT from (56) into (57) gives

$$\delta B_S = \mu_T \delta H - d^2 E_H \delta H. \tag{58}$$

After differentiation with respect to H and division by μ_T (see formula (51))

$$k^2 = \frac{\mu_T - \mu_S}{\mu_T} = \frac{d^2 E_H}{\mu_T}. \tag{59}$$

For a system of equations II B and S = const [7], respectively,

$$\delta S = \frac{\delta T}{E_B} + g\delta B = 0, \tag{60}$$

$$\delta H_S = -g\delta T + \frac{\delta B}{\mu_T} = g^2 E_B \delta B + \frac{\delta B}{\mu_T}. \tag{61}$$

Differentiation with respect to B and transformations lead to a relation connecting the coefficient of coupling, k, with the coefficient of induction strains g:

$$k^2 = \frac{g^2 E_B \mu_T}{1 + g^2 E_B \mu_T}. \tag{62}$$

For the system of piezomagnetic equations III B and a free sample (T = const, $\delta T = 0$) the following relation occurs between the coefficient k and the magnetostrictive constant h:

$$\delta T = E_B \delta S - h\delta B = 0, \tag{63}$$

$$\delta H_T = \frac{\delta B}{\mu_S} - h\delta S = \left(\frac{1}{\mu_S} - \frac{h^2}{E_B}\right)\delta B, \tag{64}$$

$$k^2 = \frac{\mu_T - \mu_S}{\mu_T} = \frac{\mu_S h^2}{E_B}. \tag{65}$$

For the last (IV B) system of piezomagnetic equations [7] in the case of a free sample (T = const) the coefficients k and e are linked by the following relation

$$\delta T = E_H \delta S - e\delta H = 0, \tag{66}$$

$$\delta B_T = e\delta S + \mu_S \delta H = \left(\frac{e^2}{E_H} + \mu_S\right)\delta H, \tag{67}$$

$$k^2 = \frac{\mu_T - \mu_S}{\mu_T} = \frac{e^2}{E_H \mu_T} = \frac{e^2}{e^2 + E_H \mu_S}. \tag{68}$$

Formulae (59), (62), (65) and (68) relate the coefficient of magnetomechanical coupling k with the complete set of 3 coefficients occurring in each of 4 systems of piezomagnetic equations in the SI system. The coefficient k can be defined by 2 (see formulae (51) and (55)) or 3 other piezomagnetic coefficients occurring in different systems of piezomagnetic equations, for example,

$$k^2 = 1 - \frac{\mu_S}{\mu_T} = 1 - \frac{E_H}{E_B} = eg = \frac{dh}{1 + dh}, \tag{69}$$

$$k^2 = \mu_S g h = dg E_H = \mu_S g^2 E_B = \mu_T g^2 E_H, \tag{70}$$

$$k^2 = \frac{eh}{E_B} = \frac{ed}{\mu_T} = \frac{e^2}{\mu_T E_H} = \frac{e^2}{\mu_S E_B}. \tag{71}$$

5. Relations between the coefficients of magnetomechanical coupling defined from equations of type B and M

In piezomagnetic equations in which the variable magnetic quantities are the field intensity H and the magnetization M (or the magnetic polarisation I or J), i.e. in the M type equations, e.g. [1, 2, 6, 7, 10, 12, 15, 16], the coupling coefficient is defined as

$$k^2(M) = 1 - \frac{\varkappa_S}{\varkappa_T} = 1 - \frac{E_H}{E_M}. \tag{72}$$

Division of equation (51) by the first part of equation (72) gives the following ratio

$$\frac{k^2(B)}{k^2(M)} = \frac{(\mu_T - \mu_S)\varkappa_T}{\mu_T(\varkappa_T - \varkappa_S)} = \frac{(\varkappa_T - \varkappa_S)\varkappa_T}{(\varkappa_T + \mu_0)(\varkappa_T - \varkappa_S)} = \frac{\varkappa_T}{\varkappa_T + \mu_0}$$

$$= \frac{\varkappa_T}{\mu_T} = \frac{1}{1 + \mu_0/\varkappa_T} = 1 - \frac{\mu_0}{\varkappa_T}. \tag{73}$$

In magnetic materials of permeability from 10 to 100

$$k^2(B) = (0.9-0.99)k^2(M), \tag{74}$$

i.e. differences between k may be less than 0.5 per cent and below 5 per cent.

Consideration of the final terms of relations (55) and (72) gives

$$\frac{k^2(B)}{k^2(M)} = \frac{1-E_H/E_B}{1-E_H/E_M} = \frac{(E_B-E_H)E_M}{E_B(E_M-E_H)}. \tag{75}$$

6. Definitions of the coefficient of magnetomechanical coupling in the nonrationalized cgsm unit system

In piezomagnetic equations in the nonrationalized magnetic cgs system of type B_c, the coefficient 4π occurs in one equation of each system, and the formulae relating the coupling coefficient k with the other piezomagnetic coefficients have the following form [6-12]

$$k^2 = \frac{d^2E_h}{4\pi\mu_T} = \frac{\Lambda^2 E_H}{4\pi\mu_T} = \frac{4\pi\mu_T g^2 E_B}{1+4\pi\mu_T g^2 E_B} = \frac{4\pi\mu_S h^2}{E_B}$$

$$= \frac{4\pi\mu_S\lambda^2}{E_B} = \frac{e^2}{e^2+4\pi\mu_S E_H} = 1-\frac{E_H}{E_B} = 1-\frac{\mu_S}{\mu_T}, \tag{76}$$

while the relation between the systems of type B and those of type M_0 is the following [6, 7]

$$k^2(B) = \frac{k^2(M_c)}{1+\mu_0/4\pi\mu_T}. \tag{77}$$

7. Conclusion

Most of piezomagnetic materials used so far are polycrystals for which the coefficient of magnetomechanical coupling k is defined for longitudinal and radial vibration, i.e. k_{33}, or for torsional vibration, k_{51}, e.g. [1, 2]. Coupling in a material or a piezomagnetic transducer is defined by the characteristic $k = f(H)$ for the primary magnetization curve and the hysteresis loop (Figs. 1, 2).

In the state of demagnetization and magnetic saturation, the magnetomechanical coupling vanishes ($E_H = E_B$, $\mu_T = \mu_S$, $k = 0$), and definition of characteristics of the coupling coefficient involves indirect determination of the state of magnetization of a material. The maximum values of k and the values of k in the remanence state are given in a short form (k_m, k_r). In good piezomagnetic materials $k > 0.15$ and, for example, in the metglasses of the Fe–Si–B alloys and in the rare earth alloys, k reaches now 0.85.

Coupling measurements show that saturation of magnetic material requires much larger fields than those in the standards of many renowned companies, which take, for example, 2,4 or 4 kA/m (30 or 50 Oe). The characteristics of k are also determined as a function of temperature, e.g. [1, 2, 4, 8, 9, 11, 14]. Synthetic

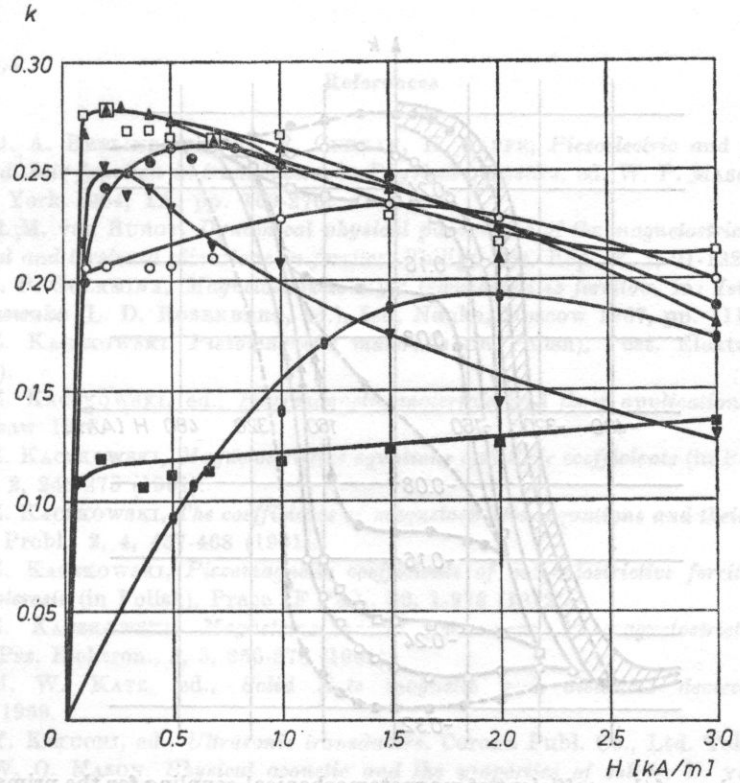


Fig. 1. Examples of the characteristics of the coefficients of magnetomechanical coupling of closen piezomagnetic transducers

n – unjoined transducer, l – soldered transducer, k – cemented transducer, alfer: (▲) k, (□) l; nickel: (o) n, (■) l; permendur (●) n, alcofer (▼) l; ferrite $Ni_{0.968}Co_{0.012}Mn_{0.02}Fe_2O_4$ (●); $f_r \approx$ 22-28 kHz

definitions and relations permit the comparison of the results obtained by authors using different definitions and systems of equations and units; or on the basis of knowledge of the values of the coefficient of magnetomechanical coupling, and moduli of elasticity and permeability, the determination of the other piezomagnetic coefficients occurring in piezomagnetic equations.

With a sufficient quality factor, the coefficient of magnetomechanical coupling can be defined from the mechanical resonance frequency f_r and the magnetomechanical antiresonance f_a, or from the ultrasonic velocities, proportional to these, for constant field intensity c_H or for constant induction c_B from the

following relation

$$k = \sqrt{1-\left(\frac{c_H}{c_B}\right)^2} \approx \sqrt{1-\left(\frac{f_r}{f_a}\right)^2} \approx \sqrt{2\left(1-\frac{f_r}{f_a}\right)}. \tag{78}$$

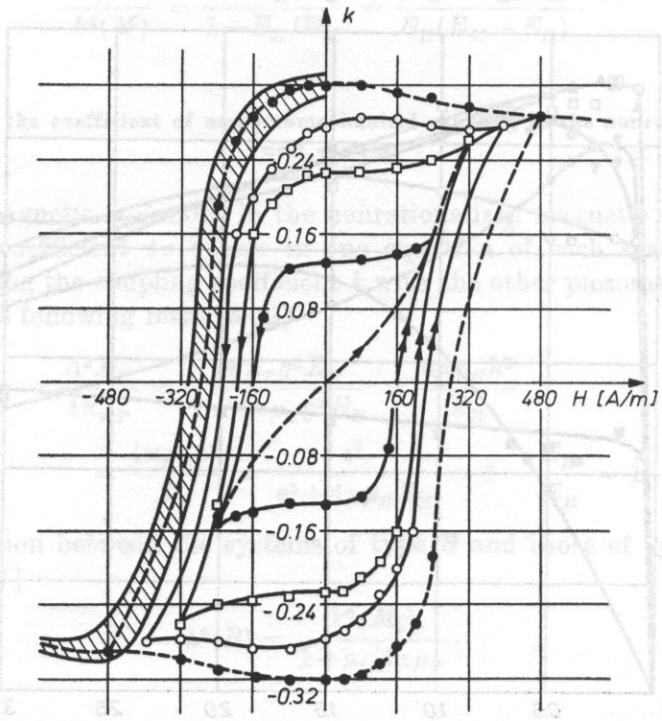


Fig. 2. The values of the coefficient of magnetomechanical coupling for the primary magnetization curve and the magnetic hysteresis loop of the ferrite $Ni_{0.953}Mn_{0.02}Co_{0.027}Fe_2O_4$ [8]
$H = 0.2$ A/m, $t = 20°C$, $f_r \approx 100$ kHz; $H_m = 240$ A/m ≈ 3 Oe (●); $H_m = 320$ A/m ≈ 4Oe (□); $H_m = 400$ A/m ≈ 5 Oe (o), $H_m = 4$ kA/m ≈ 50 Oe (●), $H_m = 0.8$-24 kA/m ≈ 10-300 Oe (- - -)

The validity of the proportional relation between the resonance frequencies and the velocities depends on their high mechanical quality factor Q_m of such a value that the product $k^2 Q_m \geqslant 10$. In addition to the mechanical quality factor Q_m and the magnetic quality factor, the coefficient k is one of the basic parameters defining the efficiency of a material or transducer. From these quantities the so called electromechanical efficiency η_{em} can, for example, be defined [1-3, 10-12, 14],

$$\eta_{em} = \frac{k^2 Q_m Q_\mu}{1 + k^2 Q_m Q_\mu}. \tag{79}$$

Thus, the materials with a high k are not always optimum, since their quality factor can be very low and, as a result, their efficiency will be worse, e.g. in ferrites and alfers.

The knowledge of the coupling coefficient and the quality factor permits piezomagnetic and piezoelectrical to be compared.

References

[1] D. A. BERLINCOURT, D. R. CURRAN, H. JAFFE, *Piezoelectric and piezomagnetic materials and their function in transducers*, in *Physical acoustics*, ed. W. P. MASON, Academic Press, New York 1964, **1A**, pp. 169-270.

[2] C. M. van BURGT, *Dynamical physical parameters of the magnetostrictive excitation of extensional and torsional vibrations in ferrites*, Philips Res. Rep., **8**, 2, 91-132 (1953).

[3] I. P. GOLAMINA, *Magnetostriktsonnyje izwuchateli iz ferritow*, in: *Istoczniki moszcznowo ultrazwuka* (L. D. ROSENBERG, ed.), Izd. Nauka, Moscow 1967, pp. 111-148.

[4] Z. KACZKOWSKI, *Piezomagnetic materials* (in Polish), Post. Elektroniki, **19**, 6, 45-59 (1974).

[5] Z. KACZKOWSKI, ed., *Piezomagnetic materials and their applications* (in Polish), PWN, Warsaw 1978.

[6] Z. KACZKOWSKI, *Magnetostrictive equations and their coefficients* (in Polish), Rozpr. elektrot., **7**, 2, 245-275 (1961).

[7] Z. KACZKOWSKI, *The coefficients of magnetostrictive equations and their relationship*, Proc. Vibr. Probl., **2**, 4, 457-468 (1961).

[8] Z. KACZKOWSKI, *Piezomagnetic coefficients of magnetostrictive ferrites and their magnetic hysteresis* (in Polish), Prace IF PAN, 39, 1-378 (1972).

[9] Z. KACZKOWSKI, *Magnetomechanical phenomena in magnetostrictive materials* (in Polish), Prz. Elektron., **2**, 3, 256-278 (1961).

[10] H. W. KATZ, ed., *Solid state magnetic and dielectric devices*, J. Wiley, New York 1959.

[11] Y. KIKUCHI, ed., *Ultrasonic transducers*, Corona Publ. Co., Ltd. Tokyo 1969.

[12] W. O. MASON, *Physical acoustic and the properties of solids*, D. van Nostrand, New York 1958.

[13] J. F. NYE, *Physical properties of crystals. Their representation by tensors and materials*, Clarendon Press, Oxford 1957.

[14] W. PAJEWSKI, Z. KACZKOWSKI, E. STOLARSKI, *Piezotronische Bauelemente*, in: *Handbuch der Elektronik*, Francis Verlag, München 1979, pp. 268-377.

[15] H. SUSSMANN, S. L. ERLICH, *Evolution on magnetostrictive properties of Hyperco*, J. Acoust. Soc. Am., **22**, 4, 499-506 (1950).

[16] S. P. TIMOSHENKO, J. N. GOODIER, *Theory of elasticity*, McGraw Hill, New York 1970.

Received on October 7, 1980; revised version on May 7, 1981.

ARCHIVES OF ACOUSTICS
6, 4, 401-408 (1981)

INVESTIGATIONS OF THE HYDRATION OF DEXTRAN USING AN ACOUSTIC METHOD*

ADAM JUSZKIEWICZ, JADWIGA POTACZEK

Institute of Chemistry, Jagiellonian University
(30-060 Kraków, ul. Karasia 3)

Measurements of the velocity of ultrasound in alcohol-water solutions of dextran with molecular weights 40 000 and 500 000 in the temperature range of 10-35°C have been made. The values of the hydration numbers have been determined and it has been shown that hydration is independent of the molecular weight of the dextran.

1. Introduction

According to the model accepted [1-6] the addition of ethyl alcohol to water causes an increase in the stabilization of the octahedral structure of water because the molecules of alcohol penetrate into the cages of the "open-work" structure of water. Spectroscopic infrared investigations [6-9] have shown that the stabilization of the structure of water by non-electrolyte molecules is connected with the increase of the stability of the hydrogen bonds between the water molecules. The observed order of water is due to the influence of non-polar groups of the non-electrolyte molecules of water; around the non-electrolyte molecules associated molecules of water called non-polar group solvates appear. This type of solvation is defined as second order solvation [10].

It is known that the dependence of the sound velocity and of the adiabatic compressibility on the alcohol concentration for alcohol-water solutions is parabolic. The maxima of the ultrasound velocity curves and the minima of the compressibility curves are dependent on temperature and they shift in the direction of lower concentrations of alcohol with increasing temperature. According to the previously presented considerations [6, 11] the point of maximum

* This paper was partly financed by the U.S. Agricultural Dept. under Public Law 480 agreement.

ultrasound velocity defines the concentration of alcohol at which all the cages of the undestroyed octahedral structure of water are occupied by alcohol molecules. The increase of the non-electrolyte concentration above the maximum value causes a break up of the primitive structure of water. The degree of order of the solution decreases and, as a result of this, the ultrasound velocity decreases. The addition of any solute to the alcohol-water mixtures causes a shift of the peak of the parabola in the direction of lower concentrations of alcohol. This is caused either by a break up of the water structure and the binding of the water molecules by the solute, or by the occupation by the solute molecules of part of the cages (holes) in the water structure. In both instances the proportion of empty cages occupied by alcohol molecules decreases, and the alcohol concentration at which the ultrasound velocity reaches a maximum is smaller. The difference between the concentration of the solute at the velocity maxima of the alcohol-water solution, and the concentration of the reference solution (alcohol-water) is caused by hydration of the solute. The amount of hydration water per number of grams of the solute W_x can be calculated from the following formula (1)

$$W_x = W_1 - \frac{W_0 A_1}{A_0}, \tag{1}$$

where A_0 and W_0 are the amounts of alcohol and water corresponding to the maximum for alcohol-water mixtures without a solute, A_1 and W_1 are amounts of alcohol and water at the maximum for alcohol-water solutions containing a certain amount of solute and W_x is the hydration in terms of percentage volume. In the previous paper [11] this method was used for the determination of the hydration of polyethylene glycols having molecular weights 400, 1500, 2000, 15 000 and 20 000. The optimum concentration of macromelecules at which this method gives repeatable results has been determined.

The acoustic and n.m.r. measurements of the longitudinal relaxation times showed that the hydration does not depend on the chain length of the polyethylene glycols tested.

This paper presents results of the measurements of the hydration numbers of dextran with molecular weights 40 000 and 500 000 over a range of concentration from 1.4 g to 17.1 g per 100 g of water in the temperature range 10-35°C.

2. Experimental arrangement

The dextran of molecular weight 40 000 used for the investigation was made by "Polfa" and the dextran of molecular weight 500 000 was made by Pharmacia Fine Chemicals *AB*.

Measurements of the sound velocity as a function of ethanol concentration were performed at 25°C for concentrations of 1.43, 2.68, 4.28, 5.71, 7.14, 8.57, 10.0, 11.43, 12.86, 14.29, 15.71 and 17.14 g per 100 g of water for each dextran.

Moreover, analogous measurements were carried out for 10 percent solutions of both fractions over the temperature range of 10-35°C. The measurements of the sound velocity were made with the system previously described [6, 11] based on the "sing-around" velocity meter. The results of the measurements are presented in Table 1 and in Figs. 1-4.

Table 1. Values of the hydration numbers

c [%]	Dextran 40 000		Dextran 500 000	
	n_h [cm³/g]	n_h $\left[\frac{\text{mH}_2\text{O}}{\text{m mon.}}\right]$	n_h [cm³/g]	n_h $\left[\frac{\text{mH}_2\text{O}}{\text{m mon.}}\right]$
1.43	0.760	6.8	0.676	6.1
2.68	0.511	4.6	0.511	4.6
4.28	0.340	3.1	0.340	3.1
5.71	0.383	3.4	0.340	3.1
7.14	0.360	3.2	0.321	2.9
8.57	0.370	3.3	0.333	3.0
10.00	0.370	3.3	0.333	3.0
11.43	0.319	2.9	0.287	2.6
12.86	0.340	3.1	0.311	2.8
14.29	0.333	3.0	0.300	2.7
15.71	0.325	2.9	0.279	2.5
17.14	0.325	2.9	0.279	2.5

3. Results and conclusions

In Table 1 the values of the hydration numbers for two dextrans are presented. It can be seen from these results that for both weights of dextran above a concentration of 4.28 g per 100 g of water the hydration numbers calculated on the basis of equation (1) are constant within experimental error with a tendency to become lower at the highest concentrations tested. At 25°C over the range of the medium concentrations tested the gram hydration n_g is equal, respectively, to: 0.366 cm³/g for dextran of molecular weight 40 000, and to 0.335 cm³/g for dextran of molecular weight 500 000. The contributions of the hydration n_i calculated from these data in terms of moles of H_2O per mole of monomer — ($C_6H_{10}O_5$) are equal to 3.3 for dextran of molecular weight 40 000 and to 3.0 for dextran of molecular weight 500 000.

Figs. 1 and 2 show the graph of ultrasound velocity as a function of ethanol content over the temperature range 5-35°C for water-alcohol solutions without dextran (Fig. 1) and for 10 percent solutions of dextran of molecular weight 40 000 (Fig. 2).

As can be seen in Figs. 1 and 2, the values of ultrasound velocity at the maximum points for water-alcohol solutions of dextran are higher than those

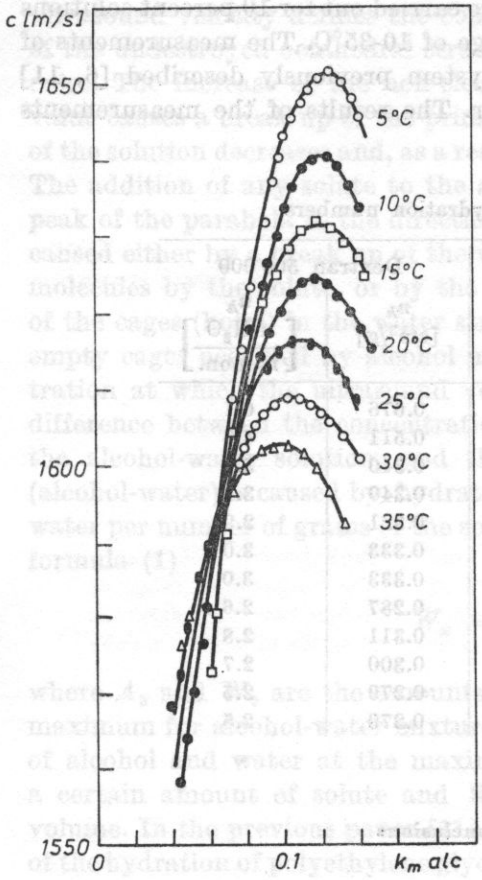


Fig. 1. Propagation velocity of ultrasonic waves as a function of ethanol concentration for a water-alcohol system in the temperature range 5-35°C

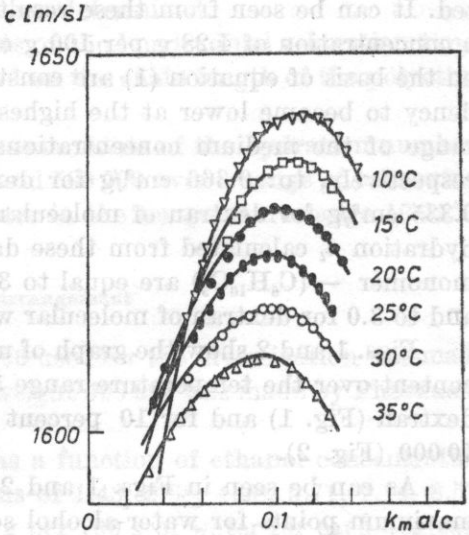


Fig. 2. Propagation velocity of ultrasonic waves as a function of ethanol concentration for dextran with molecular weight 40 000 in the temperature range 10-35°C

at the maximum point for the reference water-alcohol system. An increase in the ultrasound velocity is due to an increase in the order of the solution. In the solutions containing dextran this increase is caused by water molecules which are bound by hydrogen bonds to —OH groups of the dextran.

In Fig. 3 the temperature dependence of the number of moles of water per mole of ethyl alcohol at the maximum of ultrasound velocity for the water-

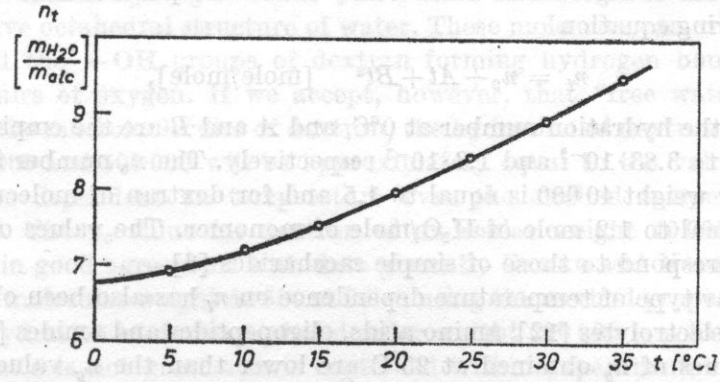


Fig. 3. Number of moles of water per mole of ethanol at the maximum of ultrasound velocity as a function of temperature for a water-alcohol system

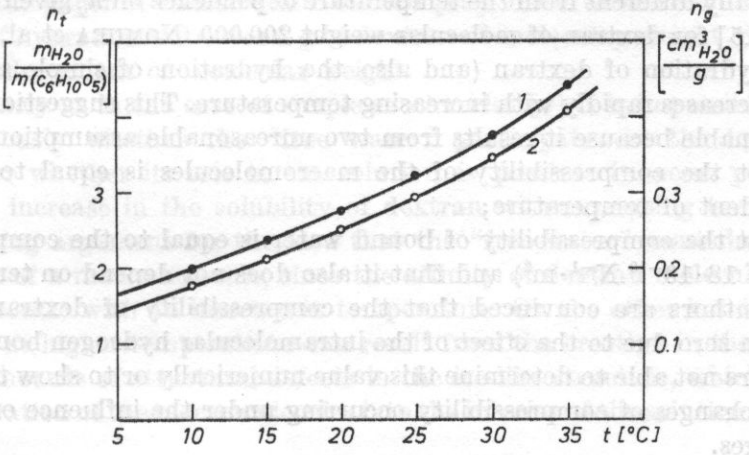


Fig. 4. Hydration numbers n_t [mole of H_2O/mole of mon.] and n_g [cm^3/g] as a function of temperature for dextran with molecular weight 500 000

$^1 - n_g$, $^2 - n_t$

alcohol mixture is presented. A shift of the maximum of the sound velocity in the direction of lower concentrations of alcohol with increasing temperature, i.e. an increase in the number of moles of water per mole of alcohol, is caused by the partial break up of the cages of the octahedral structure of the water. As

a result of this, the maximum of the ultrasound velocity is reached by the system at a lower concentration of alcohol.

In Fig. 4 the temperature dependence of the hydration numbers n_t [mole of H_2O/mole of mon.] and n_g [cm³/g] for dextran of molecular weight 500 000 is presented. Similar curves n_t and n_g were obtained for dextran of molecular weight 40 000.

The curves in Figs. 3 and 4 have very similar shapes and can be described by the following equation

$$n_t = n_0 + At + Bt^2 \quad \text{[mole/mole]}, \tag{2}$$

where n_0 is the hydration number at 0°C, and A and B are the empirical coefficients equal to $3.83 \cdot 10^{-2}$ and $1.3 \cdot 10^{-3}$ respectively. The n_0 number for dextran of molecular weight 40 000 is equal to 1.5 and for dextran of molecular weight 500 000 is equal to 1.2 mole of H_2O/mole of monomer. The values of n_t and n_0 obtained correspond to those of simple saccharides [6].

A similar type of temperature dependence on n_t has also been obtained for solutions of electrolytes [12], amino acids, oligopeptides and amides [13].

The values of n_g obtained at 25°C are lower than the n_g value presented by SHIIO et al. [14] for dextrin ($n_g = 0.4$ cm³/g) using a slightly different measuring method. The temperature dependence of the values of n_t and n_g for dextrans with molecular weights 40 000 and 500 000 presented in this paper is fundamentally different from the temperature dependence on n_g given by NOMURA et al. [15] for dextran of molecular weight 200 000. NOMURA et al. suggested that the hydration of dextran (and also the hydration of simple saccharides [16-18]) decreases rapidly with increasing temperature. This suggestion seems to be unreasonable because it results from two unreasonable assumptions:

1) that the compressibility of the macromolecules is equal to zero and is independent of temperature;

2) that the compressibility of bound water is equal to the compressibility of ice ($\beta = 18 \cdot 10^{-10}$ N$^{-1} \cdot$m²) and that it also does not depend on temperature.

The authors are convinced that the compressibility of dextran must be higher than zero due to the effect of the intramolecular hydrogen bonds. However, they are not able to determine this value numerically or to show the character of the changes of compressibility occurring under the influence of temperature changes.

SHIIO [14] stated the compressibility of dextrin macromolecules at 25°C to be equal to $9.2 \cdot 10^{-10}$ N$^{-1} \cdot$m², but this value was determined assuming that the compressibility of the hydration water at 25°C is the same as the compressibility of ice and that neither the compressibility of dextrin nor the compressibility of the hydration water changes with temperature.

It seems to be reasonable that the compressibility of the water in the hydration shells and the compressibility of the macromolecules should increase with increasing temperature. The changes of both compressibilities can give a told

effect markedly exceeding the variation of the compressibility in pure water. If the real numerical values of the compressibility of the hydration water and of the macromolecules were known, it could prove that the dependence of n_g on temperature has quite a different form.

The increase of the value of n_t with increasing temperature can be explained by an increase in the number of the molecules of "free water" with partially or completely broken hydrogen bonds which form as a result of the break up of the primitive octahedral structure of water. These molecules can orient themselves around the —OH groups of dextran forming hydrogen bonds with free electron pairs of oxygen. If we accept, however, that "free water" does not bind with the macromolecules of dextran, the hydration of dextran of molecular weight 40 000 and 500 000 will be approximately equal to the values of n_0 and will not be dependent on temperature over the studied temperature range of 10-35°C. The n_0 value for dextran of molecular weight 40 000 is equal to 1.5 and is in good agreement with data given by GEKKO and NOGUCHI [19] for dextran of molecular weight 44 500 at 20°C, using the method proposed by SHIIO.

The hydration of dextran of molecular weight 500 000 is a little lower but the difference is not great and is contained within experimental error. Apart from the influence of possible branching which can occur in the macromolecule of dextran of higher molecular weight, it can be concluded that the hydration of dextran does not depend on its molecular weight. The same conclusion can be found in the paper of GEKKO and NOGUCHI [19]. These authors showed that the amount of hydration water per 1 g of dextran of molecular weight greater than 2000 is independent of molecular weight.

Similarly as in the case of simple saccharides [6] it is impossible to decide simultaneously whether the "free water" participates in the hydration of dextran or whether its role in the solution is passive. It seems possible that the rapid increase in the solubility of dextran with increasing temperature is a convincing argument for the idea that the "free water" contributes to the hydration of a macromolecule, since the affinity of dextran molecules to water should increase with an increase in temperature. On the other hand, the better solubility at higher temperatures can result from the creation in the solution of greater amounts of structural defects in the ice-like structure, which facilitates the penetration of the macromolecules into the interior of the solution.

References

[1] H. S. FRANK. M. W. EVANS, *Free volume and entropy in condensed systems. III Entropy in binary liquid mistures: partial molal entropy in dilute solutions; structure and thermodynamics in aqueous electrolytes*, J. Chem. Phys., **13**, 507-532 (1945).

[2] G. G. MALENKOV, *Geomietriczeskoj aspekt jawlenija stabilizacji struktury wody molekulami nieelektrolitow*, Żur. Strukt. Chim., **7**, 331-336 (1966).

[3] M. J. BLANDAMER, *Acoustic properties, Water — a comprehensive treatise*, Plenum Press, New York 1973, vol. 2 pp. 495-528.

[4] F. FRANKS, *The solvent properties of water, Water — a comprehensive treatise*, Plenum Press, New York 1973, vol. 2, pp. 1-54.

[5] O. Ya. SAMOILOV, *K osnowam kineticzeskoj teorii gidrofobnoj gidratacii w rozbawlennych wodnych rostworach*, Zhur. Fiz. Chim., **52**, 1857-1862 (1978).

[6] A. JUSZKIEWICZ, *Ultrasound velocity measurements of hydration numbers of saccharides in alcohol-water solutions*, Archives of Acoustics, **6**, 3, 407-419 (1981).

[7] I. N. KOCHNIEW, A. I. KHALOIMOV, *Sostojanije wody w rostworach spirtow*, Żur. Strukt. Chim., **14**, 791-796 (1973).

[8] A. P. ZHUKOVSKY, M. V. DENGINA, *Spektroskopiczeskij mietod opriedielenija cziseł gidratacii niepolarnych grup w wodnych rostworach nieelektrolitow*, Żur. Strukt. Chim., **17**, 446-449 (1976).

[9] A. P. ZHUKOVSKY, M. V. DENGINA, *Spektroskopiczeskoje issledowanije struktury wodnych rostworow niekotorych nieelektrolitow*, Westn. Leningr. Uniw., **10**, 60-64 (1977).

[10] H. G. HERTZ, M. D. ZEIDLER, *Elementarvorgänge in der hydrathülle von ionen aus protonen- und deuteronenrelaxationszeiten*, Zeit. Elektrochem., **67**, 774-786 (1963).

[11] A. JUSZKIEWICZ, J. POTACZEK, *Investigations of the hydration of polyethyleneglycols using an acoustic method*, Archives of Acoustics, **12**, 3, 207-212 (1977).

[12] A. JUSZKIEWICZ, *Hydration numbers of electrolytes determined by an ultrasonic method* (in press).

[13] A. JUSZKIEWICZ, *Studies on the hydration of monocarbolic acids, amino acids and amides by the use of an ultrasonic method* (in press).

[14] H. SHIIO, T. OGAWA, H. YOSHIHASHII, *Measurements of the amount of bound water by ultrasonic interferometer*, J. Amer. Chem. Soc., **77**, 4, 4980-4982 (1955).

[15] H. NOMURA, S. YAMAGUCHI, Y. MIYAHARA, *Partial specific compressibility of dextran*, J. Appl. Polym. Sci., **8**, 2731-2734 (1964).

[16] A. W. PRYOR, R. ROSCOE, *The velocity and absorption of sound in aqueous sugar solutions*, Proc. Phys. Soc., **678**, 70–81 (1954).

[17] H. SHIIO, *Ultrasonic interferometer measurements of the amount of bound water*, Saccharides, J. Amer. Chem. Soc., **80**, 1, 70-73 (1958).

[18] S. GOTO, T. I. ISEMURA, *Studies of the hydration and the structure of water and their roles in protein structure*. IV. *The hydration of amino acids and oligopeptides*, Bull. Chem. Soc. Japan, **37**, 1697-1701 (1964).

[19] K. GEKKO, H. NOGUCHI, *Physicochemical studies of oligodextran. I. Molecular weight dependence of intrinsic viscosity, partial specific compressibility and hydrated water*, Biopolymers, **10**, 1513-1524 (1971).

Received on August 10, 1980; revised version on January 14, 1981.

L'identificatione della persona per mezzo della voce. Atti della tavola rotonda, Padova 14-15 Sept. 1978. F. E. Ferrero (ed.). ESA (Edizioni scientifiche associate) Roma.

Speech being an intra-species mode of communication of the biogenetic entity homo sapiens, it is only natural that ethnic language should be, and has been, for about 2 1/2 thousand years now, the object of scientific or quasi-scientific study. The spoken form of language contains, however, not only linguistic information — which until recently was the only object of language study — but also information about the speaker, which only became the object of scientific interest about 20 years ago, although the development of telephony and radiotelephony should have made the problem topical long before that. Personal voice characteristics are not irrelevant for psycholinguistics. Still, a serious motivation for wide-scale research did not arise before the forensic implications were realized. The human voice turned out of be not only a means of communication, but also a vehicle of crime, one of the commonest modes to which is blackmail by telephone. Some agencies of public security and state defence may also be concerned with voice recognition.

Scientific papers dealing with voice identification and verification appear sporadically in a few acoustical, communications, phonetic and criminological journals. The following monographs devoted to the subject should be mentioned: HECKER [2], KIRSTEIN [3], RAMISHVILI [4], TOSI [5] and BOLT et al. [1]. Though some of the work in this area is reported at such acoustic conferences as the International Congress on Acoustics or the International Conference on Acoustics, Speech and Signal Processing, yet the first large-scale meeting exclusively devoted to the subject was the round-table conference on "The Identification of the Speaker by his Voice" organized by the Italian phonetician Franco E. Ferrero (now director of the Institute of Glottology and Phonetics, University of Padua), under the auspices of the Italian Acoustical Society in the spring of 1979. The same year saw the appearance of the carefully edited proceedings of the conference, which are reviewed here. It was a national meeting, so almost all the participants were affiliated to Italian institutions, and represented the following areas: physics, phonetics, applied mathematics, electronics, electrical engineering, communications, computer science, linguistics, dialectology and psychology. Also present were representatives of the Defense Ministry.

The individual Sections included: (1) methodology and technology, (2) linguistic analysis, (3) recognition by temporal parameters, (4) recognition by spectral parameters, (5) identification based on the inspection of spectrograms and on auditory judgment, and (6) general problems of recognition.

The papers in Section 1 were mainly concerned with two problems, viz. the effect of disturbance and distortion of the signal on voice recognition rate and the alleged parallelism between identification by finger prints and by "voice prints", erroneously maintained by some American specialists. Particular note was taken of the invariability of the former compared with the substantial variability of the speech signal produced by a given voice, and of the relative ease of imposture in speech. Much attention was directed towards distortions and disturbances in telephone channels.

The leitmotif of Section 2 was the importance of dialect and linguistic-phonetic studies. There is probably no country interested in the forensic aspects of voice studies that lays so much emphasis on dialectal and linguistic-phonetic features as Italy. This accounts for an involvement of linguists in research and expertise that is stronger there than elsewhere.

The papers read in Section 3 pointed out the significance of such parameters as the duration of consonantal occlusion, relative phonation time and speech tempo for the procedures of grouping voices into classes and, in some cases, for the identification of the speaker.

Section 4 was devoted to tutorials of new techniques of acoustic parameter extraction, such as the Fast Fourier Transform and cepstral analysis. An application of this type of analysis is exemplified by forming a trajectory in the F_1 F_2 plane. The final decisions are taken either by the experimenter on the basis of the results of his analysis or – with the aid of an identifying algorithm – by a computer. Mathematical-statistical methods of multivariate analysis are frequently used.

Section 5 presented auditory and visual methods of voice recognition. The latter is based on the inspection of "visible speech" pictures (sonagrams). In the early sixties L. G. Kersta secured himself a sparkling though somewhat short-lived career in the USA promoting the sonagrams as fully reliable evidence in cases of voice identification, said to be as convincing as the finger prints (hence the term "voice print"). A large group of phoneticians, both from the USA and other countries launched a concentrated campaign against Kersta – a fact hardly noted in the proceedings – showing by scientific argumentation and counter-evidence that he was working under self-deception. It transpired with time (after a number of American acousticians cautiously joined the "anti-voice-print" front) that American courts of justice have passed several wrongful sentences on the basis of evidence given for the prosecution by Kersta or his disciples. Although, after a ten-year-long battle, Kersta had to suffer moral and financial defeat (his firm went bankrupt), American courts even now take recourse occasionally to the visual analysis of sonagrams. The problem seems at present to be scientifically settled. In this volume F. E. Ferrero and his associates once again prove thoroughly and comprehensively that sonagrams should not be used in court evidence and with the utmost caution in investigation, especially if the voice was transmitted over a telephone line. Ferrero supports his view by concrete evidence which he presents in another paper. Only a sketchy article is devoted to auditory recognition.

The reports included in Section 6 are general in character and present some legal aspects of voice recognition as well as some basic statistical methods which are, or can be, used in this area. However, such methods as statistical clustering or multiple scaling, which can also be applied, are not discussed.

The closing paper contains some proposals for the main lines of attack in the field under discussion, namely : (a) improvement of the methods of extracting such acoustical parameters from the speech signal as are most effective for characterizing individual voices, (b) development of statistical methods leading to correct identifications, (c) optimization of the decision-making procedures, and (d) normalization of the expertise.

The volume under review contains, on 305 pages, information which, though only coming from Italian sources, includes by virtue of its comprehensiveness, combined with a critical comparison of the methods described with those used elsewhere, valuable scientific and practical material. Research workers involved in the study of human voice and experts engaged in law-enforcing agencies should not fail to acquaint themselves with the results of the conference.

References

[1] R. H. Bolt et al., *On the theory and practice of voice identification*, National Academy of Sciences, Washington DC 1979.

[2] H. L. Hecker, *Speaker recognition. An interpretive survey of the literature*, American Speech and Hearing Association, Washington DC 1971.

[3] M. KIRSTEIN, *Akustische Untersuchungen zur automatischen Sprecherklassifikation*, H. Buske Verl. Hamburg 1971.

[4] G. S. RAMISHVILI, *The speech signal and the individuality of the voice* (in Russian), Izd. Mecniereba, Tbilisi 1976.

[5] O. TOSI, *Voice identification, theory amd legal applications*, Univ. Park Press, Baltimore 1979.

Wiktor Jassem (*Poznań*)