ARCHIVES OF ACOUSTICS
16, 2 195–205, (1991)

OUTLINE OF THE POLISH ACOUSTICIANS HISTORY
PART II (1970–1989)

L. FILIPCZYŃSKI

This study is the second part of a chronological complication of the most important events in the scientific life of Polish acousticians in the past 20 years, 1970–1989. The first part of the outline, covering the years before 1969, was published in the materials of 36th Open Seminar on Acoustics OSA'89, held in Szczyrk-Biła, September 1989 [1]. It was also published in Archives of Acoustics **15**, 1–2 (1990) [2]. The second part of the Outline was published in Polish in the materials of the 37 th Open Seminar on Acoustics held in Gdańsk, September 1990 [63].

Actually, the outline has been completed with new informations concerning the period which ended in 1969.

Events of major importance in Polish acoustics

1970

a) "Ultrasound in Biology and Medicine" – UBIOMED 1. The first international conference on physics and technology of ultrasound in applications in biology and medicine, organized in Jabłonna by Ultrasonic Department of Institute of Fundamental Technological Research, Polish Academy od Sciences [IFTR PAS] (head: prof. L. Filipczyński) and by Acoustics Committee PAS [3, 4]. In the following years, organization of the conference was taken over in succession by the GDR, the Czechoslovakia, Hungary and the USSR. In 1983 UBIOMED VI was again organized in Poland [5].

b) Open Seminar on Acoustics OSA-XVII, organized in Jaszowiec by Upper-Silesian Section of Polish Acoustic Society in cooperation with Acoustics Committee [6]. Annual continuation of these in 1970–1989.

c) Creation of Noise Control League. The first president – prof. H. Ryffert, and then T. Malinowski 1979, Witold Straszewicz 1983, J. Motylewski 1987.

d) Creation of Section of Acoustics of Molecular Interactions in Liquids, at Institute of Physics, Gdańsk University (head: prof. A. Śliwiński).

e) Environmental Laboratory of Noise and Vibration at Academy of Mining and Metallurgy in Kraków was organized (head: Z. ENGEL). Research on vibroacoustics, i.e. vibration and acoustic processes in machines, was started.

f) Piezomagnetics Lab. of Piezotronics Dept. at Institute of Electron Technology was transferred to Institute of Physics PAS. In 1984, in this Institute a Piezomagnetics Group of Magnetics Physics Dept. was created, headed by prof. Z. KACZKOWSKI.

g) Coordination Centre for Construction Acoustics, at the Council for Mutal Economic Aid, was organized with headquaarters in Bucharest. Prof. J. Sadowski was elected the first president of the Scientific Technological Board. Members of the board were prof. S. CZARNECKI and prof. Z. ENGEL, as well [56].

h) 2-nd International Conference on Noise Control organized in Warsaw by prof. S. CZARNECKI [10]. Subsequent conferences were organized in 1973 and 1975 [58, 60].

1971

a) Creation of Phoniatrics Lab. in the State College of Music in Warsaw (head: doc. dr med. Z. PAWŁOWSKI), transformed in 1974 into Dept. of Voice Pathophysiology and in 1980 into Phoniatrics Dept. The college was renamed to F. Chopin Academy of Music.

b) Intensification of research on semiconductor acoustics and solid state acoustooptics at Institute of Physics of Silesian Technical University (head: doc. A. OPILSKI), Department of Molecular Liquid Acoustics (head: doc. F. KUCZERA; mgr E. SOCZKIEWICZ – co-worker), Quantum Acoustics (head: doc. A. OPILSKI) and Solid State Acoustics (head: doc. Z. KLESZCZEWSKI) were created.

c) "Acoustics. Passive and Active Ultrasound Applications" Start of the first, 5-year research program PAN-20, worked out by Acoustics Committee and coordinated by IFTR PAS, prof. J. KACPROWSKI.

1972

a) The first Winter School "Molecular and Quantum Acoustics and Sonochemistry" organized in Zakopane by Institute of Physics, Silesian Technical University (head: prof. A. OPILSKI) and PTA. The conferences are held annually in Jaszowiec or Wisła [7]. Since 1979, annals, containing materials of the conferences, have been published by IFTR PAS [59].

b) Local circle Acoustics and Spectroscopy Lab. at Gdańsk University was created (head: dr J. Sułocki).

c) Publication of bibliographical index of papers on acoustics by Polish authors, for 1966–1971. The previous index convered years 1945–1965 [8].

d) The first, one-year, post-graduate course "Design of Noise Protection in Construction" organized in Institute of Construction Technology for employers of Design Bureau, Architecture Depts., Construction Supervision etc. (Head: prof. J. SADOWSKI). Subsequent courses took place in 1973, 1974 [9].

1973

a) The 2nd Congress of Polish Science. A synthethic report "Current condition and perspectives of development of scientific research on acoustics in Poland since 1985" prepared, based on 18 partial reports submitted by the acousticians community in Poland. The report includes Polish acoustic research centres, scope of their research and staff [10].

b) Start of publication of an English language quarterly "Archives of Acoustics". Chief editor: prof. S. CZARNECKI, then prof. I. MALECKI: 1983–1987 and doc. E. DANICKI: 1987.

c) The first Winter School "Noise and Vibroacoustic Hazard Control" organized in Jaszowiec by Institute of Physics, Silesian Technical University (head: prof. A. OPILSKI) and Upper-Silesian Branch of PTA. The school is annual.

d) Creation of Local Research Laboratory of Noise and Vibration head: doc. J. ZALEWSKI at Institute of Telecommunication and Acoustics Wrocław Technical University.

1974

a) Creation of Institute of Mechanics and Vibroacoustics at Academy of Mining and Metallurgy in Kraków (head: prof. Z. Engel, team: doc. J. ADAMCZYK, doc. S. BEDNACZ, prof. J. GIERGIEL, doc. dr S. KASPRZYK, prof. K. TOMASZEWSKI, prof. M. ZABAWA. In 1976 specialization in Machine Vibro-acoustics was started.

b) Prof. L. FILIPCZYŃSKI elected member of ICA for two terms till 1980 [11].

c) Start of publication of Scientific Transactions of the Research Centre of Radio and Television. The publication ceased in mid-eighties. Transactions comprised of papers on acoustics, electronics and electric circuit theory scientific editor: prof. S. MISZCZAK.

1975

Prof. L. FILIPCZYŃSKI elected vice-president of European Federation of Ultrasound in Medicine and Biology, for the term until 1979.

1976

a) "Acoustic Methods in Technology and Medicine". Start of research programme MR-I-24 for 1976–1980, worked out by Acoustics Committee and coordinated by prof. J. Ranachowski (IFTR PAS).

b) "Noise control" conference in Warsaw organized by prof. S. Czarnecki. Subseqent conferences were held in 1979, 1982, 1985 and 1988, under auspices of International Control Institute and were organized by prof. Z. Engel.

c) Organization of Polish Symposium on Vibration, Technology and Vibroacoustics in Kraków by prof. Z. Engel. Subsequent Symposia took place regularly in 1979, 1982, 1984, 1987. [12, 61, 62]

d) The first conference "Applications in Electronics and Electrical Engineering" took place in Jabłonna, organized by prof. J. Ranachowski [13, 14, 15]. Subsequent conferences were hwld annually. In 1980 the conference was called "Electric and Acoustic Methods of Inspection of Materials, Biological Structures and Mechanical Systems" [16, 17, 18].

The conferences are also a meeting-place for Polish acousticians and permit overview of annual achievements of CPBP 02.03 research programme "Acoustics in Technology, Medicine and Culture" [19].

e) "Environment protection against noise in Poland". Report prepared by Institute of Construction Technology and Institute of Environment Formation head: prof. J. Sadowski for European Committee [20, 21,22].

1978

a) FASE 78. The second International Congress of FASE in Warsaw (president – prof. S. Czarnecki), organized by IFTR PAS under auspices of Acoustic Committee PAS and Polish Acoustical Society (PTA). The Congress was opened by a report by prof. I. Malecki on quantum acoustics. During the congress, a meeting of FASE Council was held. Prof. F. François (France) was elected new president, prof. Malecki (Poland) – new vice-president and prof. F. Kolmer (Tchechoslovakia) – chancellor, all of them for the term 1979–1982 [23, 24].

b) The first Polish-French "Ultrasound colloque" organized in Paris by Ecole Pratique des Hautes Etudes, Groupement des Acousticiens de Langue Francaise and Polish Academy of Sciences [25, 26]. Subsequent colloques took place in Jabłonna and Paris, alternatively 1982, 1984, 1987, 1989 and were organized in cooperation with Acoustics Committee PAS and PTA [27, 28]. For the French part, the colloque was organized by prof. L. Pimonov.

c) "Research on development of ultrasound methods and apparatus for medical diagnostics". State Prize of the 2nd degree for the team of prof. L. Filipczyński, dr J. Etienne, dr J. Kopeć, dr G. Łypacewicz, dr A. Nowicki, dr T. Powałowski.

d) Musical Acoustics Laboratory of State Colege of Music in Gdańsk later: St. Moniuszko Academy of Music, Dept. of Composition and Music Theory was transformed into Section of Musical Acoustics, headed by doc. G. BUDZYŃSKI. The section organized courses on: Fundamentals of Acoustics and Electroacoustics, Technology of Experimental Music. Since 1981 the section has been headed by dr J. REGENT.

1980

The first Spring School "Acoustooptics and Applications" with foreign participants, organized by Institute of Experimental Physics, Gdańsk University head: prof. A. ŚLIWIŃSKI in cooperation with PTA and IFTR PAS. Subsequent schools were held in 1983, 1986, 1989 [29–35].

1981

"Acoustics in Technology and Medicine". Start of research MR-I-24 worked ot by Acoustics Committee and coordinated by prof. J. RANACHOWSKI, IFTR PAS.

b) Prof. A. ŚLIWIŃSKI elected a member of ICA, for two terms until 1987.

c) Prof. Z. JAGODZIŃSKI elected the president of Polish Acoustics Society. He held the position until 1987. [1]

1982

a) Jubilee of 70th anniversary of prof. I. MALECKI, nestor of Polish acoustics [36, 37].

b) Prof. I. MALECKI elected honorary member of FASE.

c) Institute, and since 1986 Research Centre of Health and Work Environment Protection in Mining, was created, headed by doc. A. LIPOWCZAN. The institute consisted of two departments. Research in Health Protection Dept. included vibration decrease diagnostics. In Dept. of Technical Acoustics, design of apparatus for vibroacoustic diagnostics of mining machinery, in particular: design of vibration transducers and corresponding instrumentation, was started.

d) "Psychoacoustics of Music". International Summer School organized in Jabłonna by Acoustic Committee PAS and F. Chopin Academy of Music in Warsaw head: prof. A. RAKOWSKI.

1984

a) "Noise and Vibration Hazard in Poland". An expertise prepared by a team of Acoustics Committee PAS, lead by prof. Z. ENGEL. The expertise was a matter of dispute at the State Council for Environment Protection and met some response in plans of environment protection [38].

b) "Hydroacoustics". The first symposium organized in Jastrzębia Góra by the Naval Military Academy, Hydroacoustics Dept. of Gdańsk Technical University and Polish Acoustical Society PTA. Subsequent symposia are organized annually at the coast [39–41].

c) "Development of methods and systems for speech analysis and recognition". State Prize of the second degree for the team: prof. W. JASSEM, prof. J. KACPROWSKI, doc. P. Łobacz, doc. dr med. W. TŁUCHOWSKI, dr R. GUBRYNOWICZ, dr H. KUBZDELA, mgr W. MIKIEL for 1971–1982 [42].

1985

a) "First Symposium on Sound Engineering" organized by doc. M. SANKIEWICZ of Dept. of Sound Engineering, Institute of Telecommunications, Gdańsk Technical University (head: doc. G. BUDZYŃSKI) [43]. Subsequent symposia were held biannually: 1987 in Warsaw organized by Dept. of Sound Direction (head: doc. H. CIOŁKOSZ and Dept. of Musical Acoustics (head: prof. A. Rakowski, F. Chopin Academy of Music [44]). 1989 in Kraków organized by Institute of Mechanics and Vibroacoustics, Academy of Mining and Metallurgy and Acoustics Committee PAS; organizer: J. ADAMCZYK [45].

b) Department of Acoustics, A. Mickiewicz University, Poznań, transformed into Institute of Acoustics (headed: doc. E. Ozimek).

c) Change of organization structure of Institute of Telecommunications and Acoustics, Wrocław Technical University (head: prof. W. MAJEWSKI and Dept. of Acoustics (head: doc. J. RENOWSKI) were created; Local Noise and Vibration Research Laboratory (head: doc. J. ZALEWSKI) continued its former research.

1986

a) Third Congress of Polish Science. Papers prepared, published in "Archiwum Akustyki" summarize an almost two-year discussion about most recent problems of acoustics development and specify forecast for Polish acoustics up to 2000 [46]. The declaration and Resolution of the Congress considering preferences in acoustics development and a list of Polish research plants doing in acoustics, their staff and programmes of research were published in "Archiwum Akustyki" as well [47].

b) Research programme CPBP 02.03 "Acoustics in Technology, Medicine and Culture". Its Applications in Device and Process Design, proposed for 1986–1990 by Acoustics Committee, was started, coordinated by prof. J. RANACHOWSKI (IFTR PAS).

c) First Winter School on Acoustic Intensometry in Karpacz 88 organized by Wrocław Branch of PTA and Institute of Telecommunications and Acoustics, Wrocław Technical University (head: dr H. IZOLCZUK). The Second School took place in 1988 in Karpacz. The school is heald biannually.

d) Signing of 3-year research cooperation agreement between PTA Polish Acoustics Society and La Societé Française d'Acoustique; default prolongation by subsequent 3-year period.

e) End of publication of "Archiwum Akustyki". Polish quarterly started in 1966. English language edition: "Archives of Acoustics" continued since 1987 chief editor is doc. E. DANICKI.

f) The third sub-programme "Enivironment Protection agains Noise and Vibration head: prof. J. SADOWSKI of the research programme CPBR II.4 was started.

g) The second sub-programme "Survey, estimation and reduction of noise and vibration hazards" of the research programme CPBR 11.1 "Human protection in work environment" was started coordinator Central Institute of Work Protection. Its aim is design of machines and tools with reduced noise and vibration level and methods of noise and vibration measurement as well as valuation of relevant national standards and guidelines for designers.

1987

"Prospects in Modern Acoustics Education and Development" International ICA-IUPAP Conference in Jastrzębia Góra organized by prof. A. ŚLIWIŃSKI.

b) "Acoustic Emission Method in Science and Technology". International school organized under auspices of CISM Int. Centre of Mechanical Science in Jabłonna by prof. I. MALECKI and prof. J. RANACHOWSKI.

c) Prof. A. ŚLIWIŃSKI elected the president of PTA

1988

a) The first map of Poland showing noise polution in 1986 was made by the team of prof. J. SADOWSKI.

b) "Protection Against Noise and Vibration" Report prepared by experts of Acoustics Committee PAS, Institute of Construction Technology, Institute of Environment Protection, Institute of Mechanics and Vibroacoustics, and Noise Control League for Dept. of Environment and Natural Resources Protection, as a part of natural resources programme till 2010, head of the team: prof. J. SADOWSKI.

1989

First Symposium on Electroacoustics organized in February 1989 in Czerniejew near Poznań by the Institute of Acoustics of the Adam Mickiewicz. University in Poznań in cooperation with ZWG "TONSIL" in Września.

Supplement to the period before 1989

1938

Witold JANKOWSKI, MDr, graduate of Faculty of Medicine, Lwów University published a paper "Powikłania wewnątrzczaszkowe zapalenia ucha środkowego", Pol. Gaz. Lek., **17**, 25, 519–524 (1938) [57].

1946

a) After emigration from Lwów to Zabrze in 1950, doc. dr med. T. CEYPEK graduate of Faculty of Medicine, Lwów University, organized and led Otolaryngology Dept. and Clinic ąt Silesian Medical Academy, Zabrze.

b) Doc dr med. W. JANKOWSKI arrived in Wrocław and organized Otolaryngology Clinics.

1951

Prof. B. ZAPIÓR started research in sonochemistry at Dept. of General Chemistry Jagiellonian University, Kraków. In 1961 the team was joined by dr A. JUSZKIEWICZ, who led the team after retirement of prof. B. ZAPIÓR since 1978.

1953

a) Mgr Z. KACZKOWSKI started research on magnetostrictive materials in Magnetic Materials Laboratory, Dept. of Electronics, IFTR PAS head: prof. J. GROSZKOWSKI.

b) Prof. T. CEYPEK commenced research in industrial audiology, and especially an injuries inmining and metallurgy, and cooperated with other Polish research centres.

c) Prof. W. JANKOWSKI became head of Dept. and Clinic of Otolaryngology, continued cooperation with other acoustic research centres in Poland.

1956

a) W. JASSEM, Ph.D., English philology, head of Phonography Dept. Poznań University, started work for IFTR PAS by organizing in Poznań Acoustic Phonetics Laboratory.

b) Mgr R. PŁOWIEC, Dept. of Vibration Research IFTR PAS, started research on problems of ultrasound measurement.

1965

Prof. H. RYFFERT elected the president of PTA. She held the position till 1981 [1].

1966

Institute of Electron Technology PAS was created. In 1966–1967 doc. Z. KACZKOWSKI was organizing Piezomagnetics Laboratory at Dept. of Piezoelectronics.

References

[1] *Materiały XXXVI Otwartego Seminarium z Akustyki OSA'* 89, Szczyrk-Biała 1989.

[2] Arch. Acoustics **15** 1–2 (1990).

[3] *Ultrasonics in biology and medicine*, L. Filipczyński [Ed.] PWN, Warszawa 1972.

[4] Arch. Akustyki **6**, 269–272 1971.

[5] Arch. Acoustics, **9**, 4–298 (1984).

[6] Arch. Akustyki **6**, 165–169 (1971).

[7] Arch. Akustyki, **9**, 445–446 (1974); **13**, 253–25 (1978); **12**, 275–276 (1977); **19**, 267–268 (1984).

[8] Arch. Akustyki, **2**, 3–68 (1967); 7, 3–77 (1972).

[9] *Działalność Instytutu Techniki Budowlanej w* r. 1973, Warszawa 1973, s. 88

[10] Arch. Akustyki **8**, 2 (1973).

[11] Nauka Polska **2**, 83–86 1976.

[12] Arch. Akustyki **15**, 115–122 (1980); **18**, 181–183 (1983); **10**, 317–465 (1985)

[13] *Zastosowanie polimerów w elektrotechnice i elektronice*, J. Ranachowski [Red.] Warszawa-Jabłonna, grudzień 1976 oraz grudzień 1977, IPPT PAN.

[14] *Special ceramics in electronics and electrical engineering* J. Ranachowski [Red.], Jabłonna, October 1978, IPPT PAN.

[15] *Wybrane metody badań cieczy i polimerów*, J. Ranachowski [Red.], Warszawa-Jabłonna, grudzień 1979, IPPT PAN.

[16] *Elektryczne i akustyczne metody badań dielektryków*, J. Ranachowski [Red.], Jabłonna, grudzień 1980, IPPT PAN.

[17] *Elektryczne i akustyczne metody badań materiałów*. J. Ranachowski [Red.] grudzień 1981, listopad 1982 oraz grudzień 1983, IPPT PAN.

[18] *Elektryczne i akustyczne metody badań materiałów i struktur biologicznych*, J. Ranachowski [Red.], Jabłonna grudzień 1984, IPPT PAN.

[19] *Akustyka w technice i medycynie. Przegląd wyników badań w problemie międzyresortowym MR* I24 *w latach* 1981–85, J. Ranachowski [Red.] IPPT PAN, 1985.

[20] *National approaches to community noise problem. Draft Final Report on Phase One. Activities of E. C. E. Task* Force *on Noise*. Economic Commission for Europe, Geneva, Switzerland. Task Force

Secretariat Office of Noise Abatement and Control. U.S. Environmental Protection Agency, January 1978.

[21] *Raport about the Research State and the output of Poland in the domain of Environmental Protection against Noise.* Building Research Institute ITB. Research Institute of Environmental Development (IKS) Warsaw 1976 (Praca zb. dostęp na w Bibliotece ITB).

[22] *9-th International Congress on Noise Abatement. Procceeding National Raport. Investigation progress at shaping domain and environment protection in Poland* (p. 527–529) Kozlekedesi Dokumentacios Vallalat. Budapest 1977.

[23] *Proc. of the Second Congress of the Federation of Acoustical Societes of Europe,* [Edts.] L. Filipczyński, Z. Pawłowski, H. Ryffert, J. Zieniuk IPPT PAN, Warsaw 1987, vol. I, II, III.

[24] Arch. Akustyki **14**, 185–194 (1979).

[25] Arch. Akustyki **9**, 243–245 (1974).

[26] *Colloque sur les ultra-sons.* M. L. Pimonow, M. R. D. Prudhome [Eds.]. École Pratique des Hautes Études -III e Section EPHE et le Centre Scientifique de l'Academie Polonaise des Sciences. Paris 1978, 1–37.

[27] *Colloque sur les ultra-sons et acoustique physique.* M. L. Pimonow, M. R. D. Prudhome [Eds.], École Practique des Hautes Études -III e Section EPHE et le Centre Scientifique de l'Academie Polonaise des Sciences Paris 1982, 1–174, oraz Paris 1987, 1–320.

[28] *Colloque sur les ultra-sons,* Instytut Podstawowych Problemów Techniki PAN, Ecole Practique des Hautes Études EPHE Warszawa 1980, 1–307.

[29] Arch. Akustyki **16**, 113–114 (1981); Ultrasonics **19**, 44–45 (1981).

[30] *Proc. First Spring School on Acustooptics and Applications,* Uniwersytet Gdański 1980. A. Śliwiński [Red.].

[31] *Proc. Second Spring School on Acoustooptics and Applications,* Uniwersytet Gdański 1983

[32] Arch. Acoustics **9**, 381 (1984); Ultrasonics, **22**, 45–46 (1984)

[33] *Proc. Third Spring School on Acoustooptics and Applications,* Uniwersytet Gdański 1986.

[34] Arch. Acoustics **11**, 413 (1986); J. Acoust. Soc. Am. **80**, 1548 (1986); Ultrasonics 25, 183 (1987).

[35] *4-th Spring School on Acoustooptics and Applications,* A. Śliwiński, P. Kwiek, A. Markiewicz [Eds.] World Scientific Singapore, New Yersey 1989.

[36] Nauka Polska 1–2, 107–110 (1983).

[37] *Problemy i metody współczesnej akustyki,* J. Ranachowski [Red.], PWN, Warszawa 1989, 7–23.

[38] *Raport „Zagrożenia hałasem i wibracjami w Polsce.* Rękopis opracowany pod kier. prof. Z. Engla dostępny w bibliotece IPPT PAN, Arch. Acoustics **10**, 199–210 (1985); Arch. Akustyki **20**, 90–109 (1985).

[39] *Materiały z I, II, III, IV Sympozjum z Hydroakustyki.* Wyższa Szkoła Marynarki Wojennej. Gdynia 1984, 1985, 1986, 1987.

[40] *Materiały z V i VI Sympozjum z Hydroakustyki.* Akademia Marynarki Wojennej, Gdynia 1988, 1989.

[41] Arch. Akustyki **21**, 153–154 (1986; **22**, 290 (1987).

[42] Arch. Akustyki **19**, 1–2 (1984).

[43] *I-e Sympozjum Inżynierii Dźwięku,* Gdańsk 1985, M. Sankiewicz [Red.]. 1986. Arch. Akustyki **22**, 154–155 (1987).

[44] *Sympozjum Reżyserii Dźwięku, Warszawa 1987. Zbiór tekstów referatów i materiałów, 23–25 kwietnia 1987* Akademia Muzyczna im. F. Chopina, Warszawa, Okólnik 2.

[45] *Inżynieria i reżyseria dźwięku.* J. Adamczyk [Red.]. AGH Kraków 1989

[46] Arch. Akustyki **20**, 329–334 (1985).

[47] Arch. Akustyki **21**, 3 (1986).

[48] *Prace I Szkoły Naukowej Intensometrii Akustycznej,* Instytut Telekomunikacji i Akustyki, Politechnika Wrocławska 1987.

[49] *Prace II Szkoły Naukowej Intensometrii Akustycznej,* Karpacz 6–9 czerwca 1988. Prace Naukowe II A Pol. Wrocławskiej Nr 65, Seria: Konferencji, nr 21

[50] *Prospects in modern acoustics — education and development,* A. Śliwiński, G. K. E. Budzyński [Eds.], World Scientific Singapore, New Jersey 1987.

[51] Acoustica **66**, 117 (1988); **68**, 76 (1989); Akusticzeskij Żurnał **5**, 954 (1988); Ultrasonics **26**, 47 (1988).
[52] Arch. Acoustics, **15**, 1–2 (1990).
[53] J. Sadowski, *Noise protection of environment by means of town-planning solution*, 13th International Congress on Acoustics, vol. 1. Invited lectures, Belgrad 1989, 107–118.
[54] J. Sadowski, Z. Engel, R. Kucharski, B. Lipowczan, B. Szudrowicz, *Narodowy program ochrony środowiska i zasobów naturalnych do 2010 r. Raport 7 nt. Ochrony środowiska przed hałasem i wibracjami.* Instytut Techniki Budowlanej, Warszawa 1987/88.
[55] J. Sadowski, *NPOS i ZN do 2010. Ochrona przed hałasem i wibracjami. Synteza 7*. Instytut Ochrony Środowiska, Warszawa 1988.
[56] B. Szudrowicz, J. Żuchowicz-Wodnikowska, *Prace Centrum Koordynacyjnego RWPG do Spraw Akustyki Budowlanej. Omówienie prac zrealizowanych w latach 1980–85 i plan pracy na lata 1986–91.* Instytut Techniki Budowlanej, Warszawa 1987.
[75] S. Iwankiewicz. *Prof. dr hab. Wiktor Jankowski*, Akademia Medyczna Wrocław 1975.
[58] Arch. Akustyki **9**, 365–369 (1974); **10**, 440–441 (1975); **12**, 187–193 (1977).
[59] *VII Zimowa Szkoła Akustyki Molekularnej*, Kwantowej i Sonochemii, Ustroń 1978, Akustyka Molekularna i Kwantowa t. I, Warszawa 1979; t. II Warszawa 1981; t. III Warszawa 1982; t. IV Warszawa 1983; t. V Warszawa 1984; t. VI Warszawa 1985; t. VII: Warszawa 1986; t. VIII Warszawa 1987; t. IX Warszawa 1988; t. X Warszawa 1989.
[60] Arch. Akustyki **17**, 209–210 (1982).
[61] *Proc. Noise Control*, [Ed.] Z. Engel Kraków 1982; vol. I, II, 1985, vol. I, II, 1988.
[62] *Sympozja Techniki Wibracyjnej*, Wyd. AGH Kraków, Wibroakustyka 3/4, 1979; Zeszyty Naukowe 1/2, 1979, 1982, 1984, 1987.
[63] *Materiały XXXVII Otwartego Seminarium z Akustyki SA '90*, Gdańsk 1990, 9–23.

ARCHIVES OF ACOUSTICS
16, 2, 207–228 (1991)

QUANTITATIVE ULTRASONOGRAPHY (DEVELOPMENT IN POLAND)

L. FILIPCZYŃSKI

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

Ultrasonography is based on the competent interpretation of ultrasonic images of the patients organs. Factors enabling a correct diagnosis are of qualitative character, mainly, because in the first place they depend on the experience of the specialist performing the examination.

However, a necessity of describing certain quantities with numbers appeared already in the early stages of development of ultrasonography. This need becomes more and more apparent and even essential in certain domains of ultrasonography where Doppler techniques are applied.

If 1958 would be accepted as the year of advent of clinical ultrasonography, then practically from that moment the following question was a live issue: if and when ultrasound can be harmful → especially to subtle fetal structures? This question gave origin to ultrasonic dosimetry, which was to answer how high intensities — how many W/cm² are introduced into examined structures and if it is permissible. This still remains a current and widely discussed problem.

Doppler techniques, strongly related at present with utrasonography, are the second domain requiring numerical parameters. The qualitative analysis of the morphology of the blood flow velocity curve was sufficient in the first stage of development of the Doppler technique. But before long quantitative indices were introduced and further developments were aimed at the determination of the blood's linear velocity in m/s, and then — volume velocity called also volume flow, measured in ml/min, most frequently.

The moment when microcomputer technique was introduced to ultrasonography, it became possible to determine linear dimensions of examined structures directly from ultrasonographic images, their surface, at certain additional assumptions concerning their volume, as well as the velocity of movable structures. It was also possible to introduce methods of spectral analysis in real time. This is of particular importance to cardiological ultrasonography. Quantitative determination of the blood pressure gradient from blood velocity measurement was also made possible.

And finally, the noninvasive determination of hemodynamic impendace (resistance) and vessel rigidity by means of associated ultrasonographic and Doppler techniques is the latest field developed at present.

This paper will discuss details of presented here methods and techniques, as well as prospects of their further development taking into account mainly the Polish achievements in this field.

Ultrasonography is based on the competent interpretation of ultrasonographic images of the patients internal organs. An old Chinese proverb says that one picture supplies more information than a thousand words. Factors which make a good diagnosis possible are mostly of qualitative character and are dependent on the experience of the specialist performing the examination.

Yet, the necessity of characterizing certain quantities with numbers appeared as early as the first period of development of ultrasonography. This necessity becomes more and more noticeable and even essential in certain ranges of ultrasonography, where Doppler techniques are applied.

The year 1958 can be accepted as the conventional date of birth of clinical ultrasonography, because it was then that the excellent British periodical "The Lancet" printed a paper by professor DONALD from the University Clinic in Glasgow, doctor VICAR and enginner BROWN from the Kelwin Hughes Company, containing for the first time images obtained with ultrasound of the uterus, cysts of the ovaries, ascites and neoplasmas in a pregnant woman [3]. In order to evaluate the progress made from that time, you only need to look at present obtained from computer-aided ultrasonographs ultrasonograms of the fetal head, which might be called portraits from profile or en face.

We owe this giant progress, happening during the life time of one generation, to the incredible development and pioneer research of acousticians, electronic engineers and computer scientists, supported by the research of clinicians applying the method of successive step by step development.

However, from the first moment the question arouse: may ultrasound be harmful – to subtle fetal structure especially and when can, this happen? Can effects of ultrasonic activity turn out to be as harmful as ionizing radiation? What is the permissible dose? These questions have accompanied all initial research within ultrasonography and they still remain the subject of international interest.

The first ultrasonograms in Poland were obtained in 1966 [14] with a UG-1 ultrasonograph (Figs. 1, 2) constructed by our team in the Institute of Fundamental Technological Research (IFTR) of the Polish Academy of Sciences. We obtained first ultrasonograms of pregnant women (Figs. 3, 4). As far back as then we developed two original methods of measuring the ultrasonic intensities radiated into the patients abdominal cavity (Fig. 5), achieving results of 2 mV/cm^2 SATA equivalent to about 5 mV/cm^2 SPTA [16]. Fig. 6 presents permissible doses of ultrasound intensity according to the American Institute of Ultrasound in Medicine AIUM [1]. Doses applied in our first and further research should be located near the lowest dashed line. Hence, they were about 10 times lower than permissible values.

The problem of permissible ultrasonic doses has acquired new dimensions at present, because of new additional safety criteria imposed on ultrasonic examinations including temperature effects. The last conference organized at the Illinois University in Urbana in June 1990 was dedicated to this issue. The AIUM-Institute recomends the criterion of maximal permissible increase of the fetus'es temperature equal to 1°C over the physiological temperature [1]. Examinations of seven species of mammals

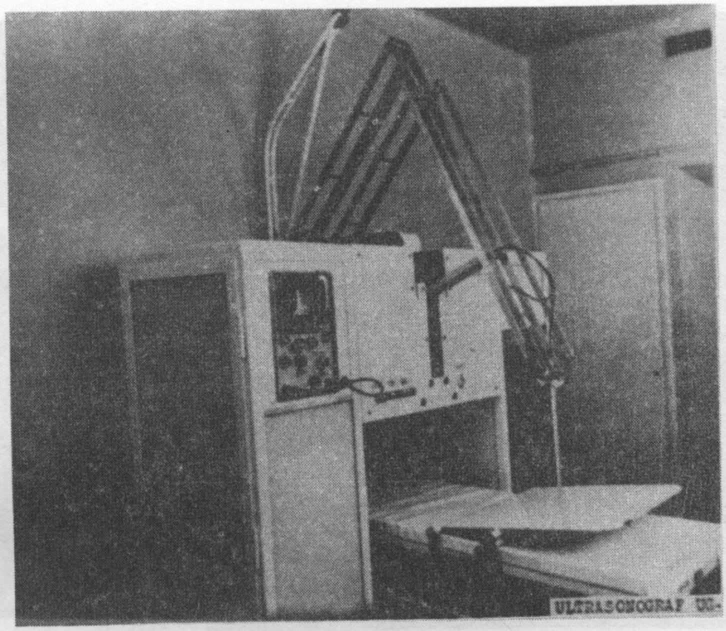


FIG.1. The original ultrasonograph UG–1. The ultrasonic probe frequency 2 MHz was coupled mechanically by means of a pantograph with the movable cathode ray tube [14] (1966).

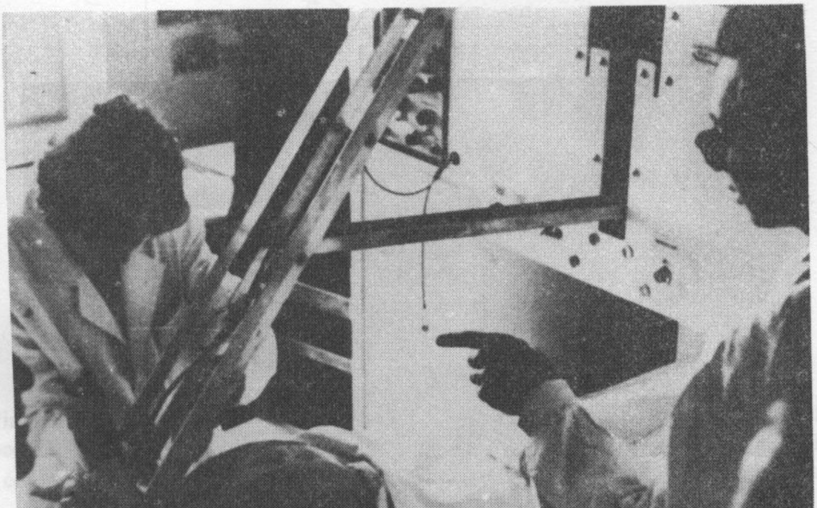


FIG. 2. Ultrasonograph UG–1 during examination of a pregnant woman. Left – dr G. Łypacewicz, right – dr J. Etienne [14] (1966).

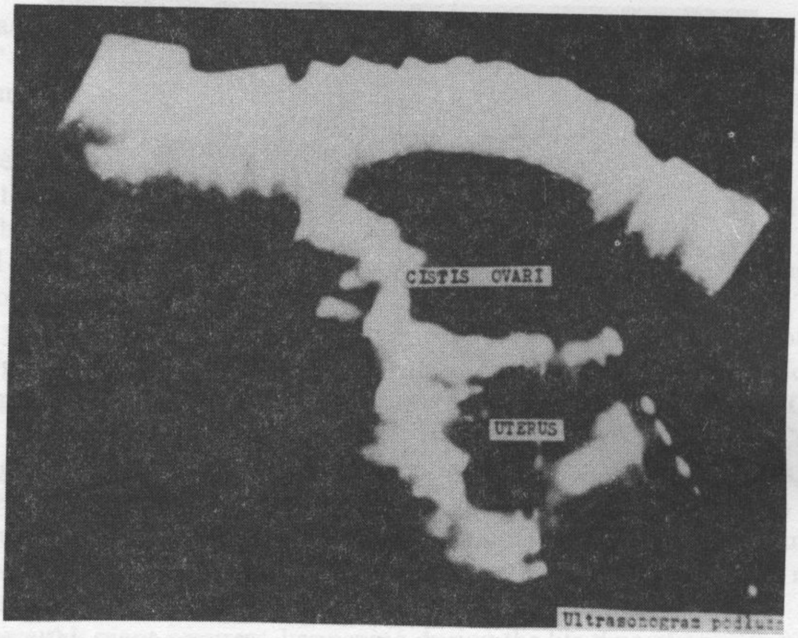


FIG. 3. Longitudinal ultrasonogram of the woman with visible cistis ovari and uterus [14] (1966).

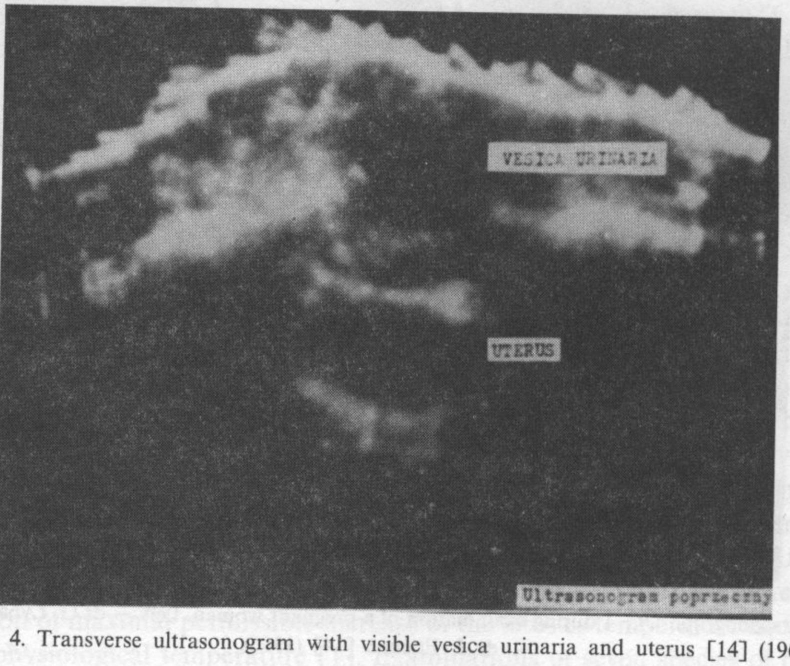


FIG. 4. Transverse ultrasonogram with visible vesica urinaria and uterus [14] (1966).

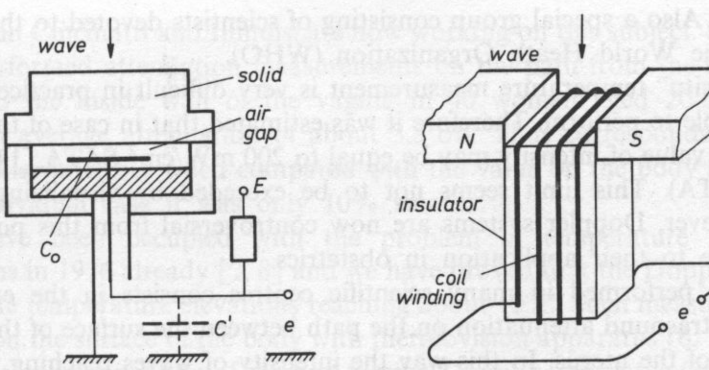


FIG. 5. Principle of the capacitance and electrodynamic absolute measurement methods of ultrasound intensity [14], [16], (1966).

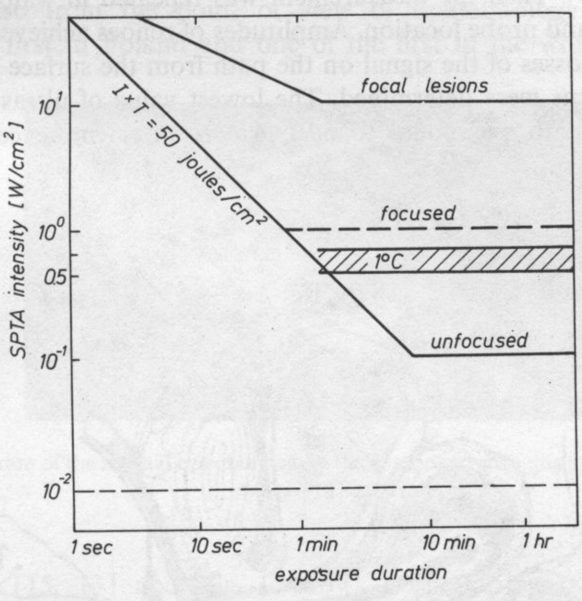


FIG. 6. Minimum SPTA (Space Peak Time Averaged) intensities required for ultrasonic bioeffects based on the specified AIUM (American Institute of Ultrasound in Medicine) recommendation. The hatched area corresponds to the lowest intensity levels which may cause a temperature rise of 1 C.

and anamnesis with mothers who had fever during their pregnancy indicate explicitly that a temperature elevation exceeding 2°C is very frequently the reason for teratism due to dysplasia of the embryo, since hyperthermia destructively influences the development of its nervous system.

This problem is of lively interest to scientists investigating the effect of ultrasound on living tissues, clinicians applying ultrasonic devices and producers of these devices. Also a special group consisting of scientists devoted to this issue was created by the World Heath Organization (WHO).

The "in situ" temperature measurement is very difficult in practice or may be even impossible to perform. Therefore it was estimated that in case of this criterion, the threshold value of intensity may be equal to 200 mW/cm² SATA [1] (about 500 mW/cm² SPTA). This limit seems not to be exceeded by visualizing ultrasonic systems, however, Doppler systems are now controversial from this point of view with reference to their application in obstetrics.

Research performed in many scientific centres consists in the experimental research of ultrasound attenuation on the path between the surface of the body and internal wall of the uterus. In this way the intensity of waves reaching the embryo can be determined and the temperature increase may be estimated [25].

We have taken up this subject concerning early pregnancy [4] as the first already in 1976. Our method consisted in the insertion of reflectors in the form of stainless steel balls mounted on a rod into the uterus of a patient immediately before legal abortion (Fig. 7). Then the measurement was repeated in water for the same geometry of the ball and probe location. Amplitudes of echoes achieved in both cases were compared and losses of the signal on the path from the surface of the body to the inside of the uterus were determined. The lowest value of ultrasound intensity attenuation on this path was equal to 6dB. American scientific centers, such as

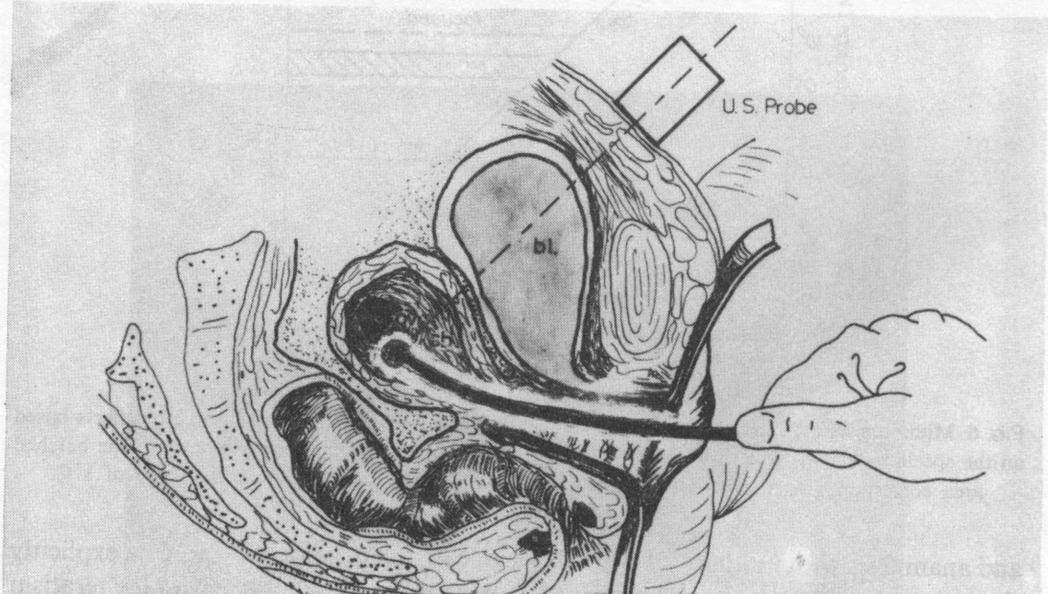


FIG. 7. Principle of signal loss measurements in pregnant women [4] (1976).

universities in Cincinatti and Illinois, are now working on this subject. Recently prof. O'BRIEN performed attenuation measurements on the path from the surface of the abdomen to the inside wall of the vagina in 40 women aged 20–30 years and obtained an average attenuation of about 3.5 dB. Thus corresponds to an about 2 times smaller intensity when compared with the value on the body surface. In an extreme individual case it was only 10% smaller.

We have been occupied with the problem of temperature in diagnostic examinations in 1976 already [7, 8] and we have proved that the Doppler apparatus may generate temperature elevations reaching about 10°C. Such measurements were performed on the surface of the body with thermovision apparatus [6, 10]. It was an extreme case concerning one of the first foreign Doppler apparatus for examination of the peripheral circulatory system.

We have dedicated several papers concerning diagnostic devices, as well as lithotripsy lately, [9, 11, 23, 25] to the problem of heat generation.

We have continued work on the visualization and permissible doses of ultrasounds also from the point of view of ophtalmology since 1967. Figure 8 presents the first in Poland and one of the first in the world ultrasonogram of

FIG. 8. Ultrasonogram of the normal eye obtained by the compound scanning system with frequency of 10 MHz [15] (1967).

a normal eye [15, 13] and Fig. 9 shows a special ophtalmic ultrasonograph constructed by us at that time. Performed by us quantitative investigations of the influence of ultrasound intensity on eye tissues of a rabbit proved that first signs of lesions in the eyes retina were observed at intensities of approximately 200 W/cm^2 (SATP) [21]. On the basis of these results the intensity of ultrasonographs produced in Poland according to our conception was limited [18].

After 20 years we still perform intensity measurements of various types of foreign and Polish ultrasonographs in our laboratory using for this purpose membrane PVDF hydrophones. As for ophtalmology, we should mention the quantitative determination of the distance between various structures in the eye the so-called eye biometry with special ultrasonographs with presentation A. This method is nowa-

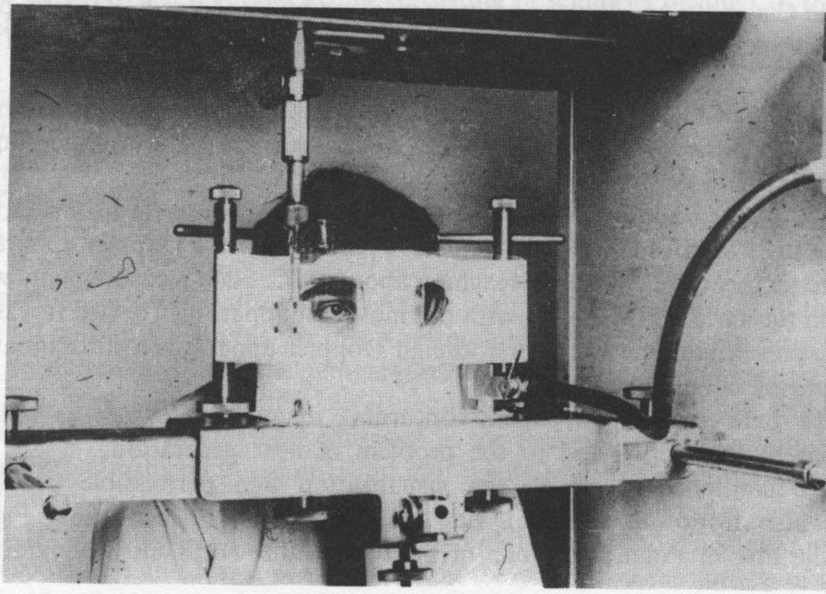


FIG. 9. The original compound scanning ultrasonograph for visualization of the eye [15] (1967).

days also applied to determine the parameters of plastic lenses implanted in the eye during cataract operations.

Noninvasive Doppler technique, strongly associated with ultrasonography at present, is the next domain which requires numerical parameters. In the first stage of development of this technique, the qualitative morphological analysis of the blood velocity curve was sufficient. Before long quantitative indices were introduced and further development was aimed at the determination of the linear velocity of blood in m/s, and then volume velocity, called volume flow, measured most frequently in ml/min. It is impossible to localize an examined vessel with the simplest continuous wave technique, whereas by means of the pulse method not only vessels can be localized, but also blood velocity and its distribution inside a vessel can be measured.

We have made first quantitative measurements of volume flow of blood in the carotid arteries in humans in 1975 [2], using several various ultrasonic methods at the same time [34, 46, 17] and obtaining the average value of 0.53 l/min (Fig. 10). Several techniques based on designed by us instruments have been applied in this research; visualization (Fig. 11) continuous wave Doppler technique, spectral analysis, pulse Doppler with double switchable beams. This last technique made it possible to eliminate the generally unknown inclination of the ultrasonic beam with respect to the direction of the blood flow. It was used by us to measure profiles of blood velocity distribution in the artery during the full cycle of the heart action. The information on the inclination angle and the geometry of the vessel can be nowadays read directly off the screen when the Doppler method is incorporated into the ultrasonographic equipment.

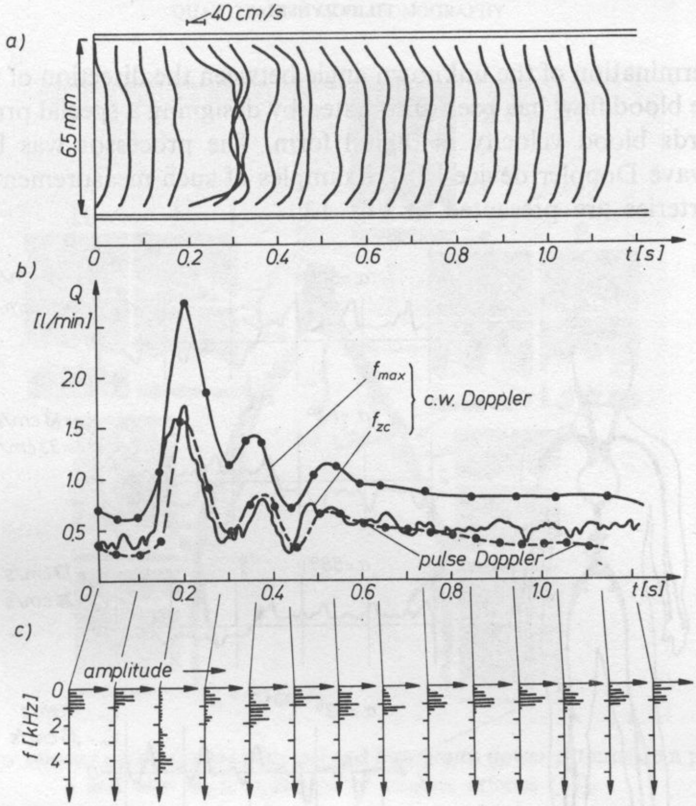


FIG. 10. Results of blood flow measurements in the carotid artery during one cycle of the heart action carried out by means of the pulse method (a), continuous wave method (b), and spectral analysis (c) [2] (1976).

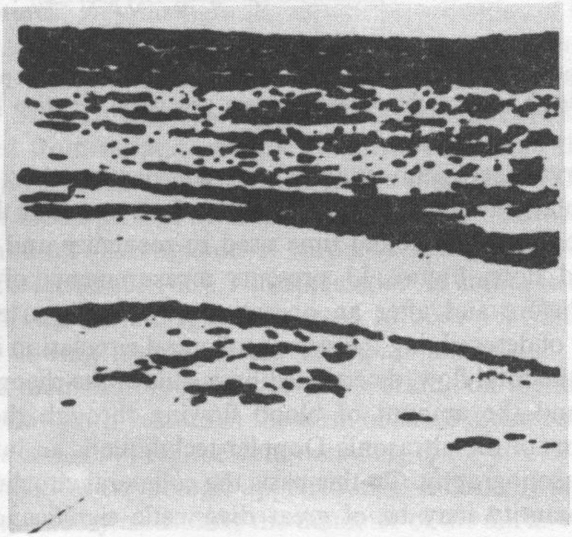


FIG. 11. The image of the carotid artery where measurements of the blood flow velocity and its distribution were carried out [2] (1976).

The determination of the unknown angle between the direction of the ultrasonic beam and the blood flow has been automated by designing a special processor which directly records blood velocity in digital form. The processor was built into the continuous wave Doppler device [45]. Examples of such measurements carried out in various arteries are presented in Fig. 12.

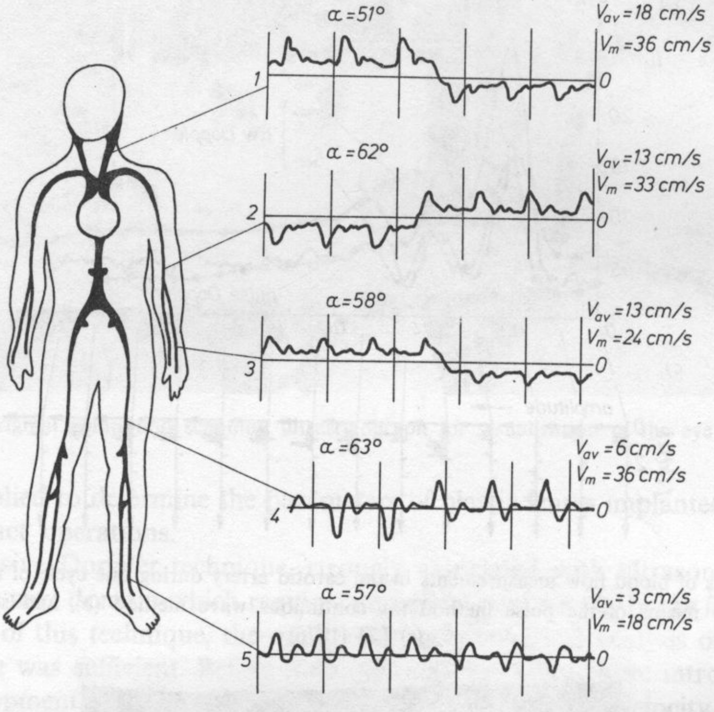


FIG. 12. Quantitative blood velocity examinations by means of an automatic Doppler system eliminating the angle dependence between the ultrasonic beam and blond direction [17, 45] (1978).

The time interval histogram analysis was the next method which made the evaluation of vascular flow disturbances possible. It is a certain kind of simplified spectral analysis performed in actual time used to recognize and evaluate disturbances of the blood flow. Figure 13 presents measurements of blood flow and histograms made before and after an operation of stenotic arteries [54].

The possibility of determining the size of collateral circulation in the thigh as the difference between the total flow, determined by means of the rheographic method in a thigh segment, and the amount of blood flowing through the femoral artery, determined by means of the ultrasonic Doppler technique is an interesting example of quantitative ultrasonography. On this basis the collateral circulation index can be determined; this quantity may be of great diagnostic significance [55, 37].

From the moment when microcomputer techniques were introduced to ultrasonography it became possible to determine linear dimensions of examined struc-

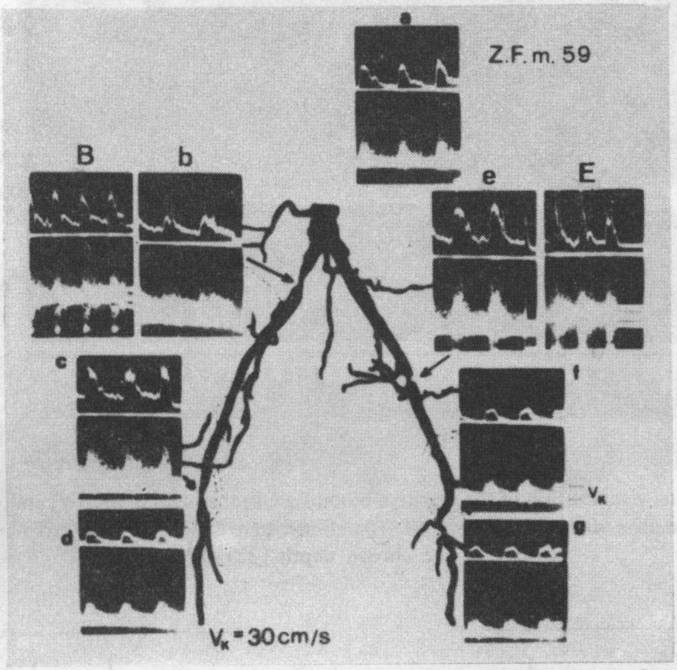


FIG. 13. Blood flow velocity curves (upper patterns) and histograms (lower patterns) in a patient before b, e and after B, E operation of stenotic arteries [53].

tures directly from ultrasonographic images, their surface, their volume at certain additional assumptions, grey histograms, as well as velocities of moving structures. This is of particular importance in cardiology. The visualization of flowing blood became possible by connecting the ultrasonograph with the Doppler technique and by developing the SEC Doppler technique (Stationary Echoes Cancellation) [39], which was adapted from radar technique. Figure 14 presents the visualization of blood flow in the jugular vein and the aorta at the same time [22], while in Fig. 15 we have records of blood flow in several heart structures obtained by means of this method. The method for investigation of coronary graft patency was developed on the basis of the SEC technique [40]. The application of colour is the further stage of development of this technique. It can be used to code the direction or even the velocity of blood flowing in big vessels, the heart above all.

At the present moment foreign centres are working on the idea of denoting the direction of blood flow with vectors in the form of small arrows additionally printed on the colour. Also a method for the determination of the stroke volume in liters of blood was worked out for cardiologic application. It is done by measuring the blood velocity in the aorta and by measuring the diameter of the aorta. An original method for detecting innocent vibratory heart murmurs in children was elaborated on the basis of the cepstrum analysis [38].

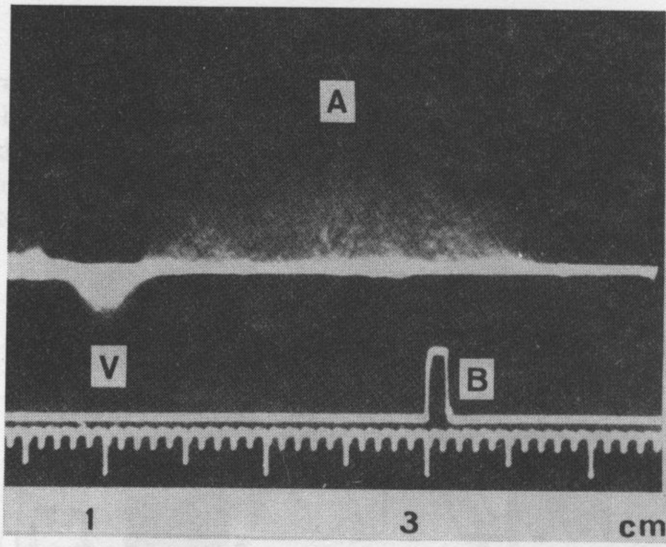


FIG. 14. Simultaneous visualization of the negative blood flow in the jugular vein (V) and the positive one in the aorta (A) by means of the SEC technique. The electric gate (B) enables measurements of the blood velocity at the chosen depth [22] (1982).

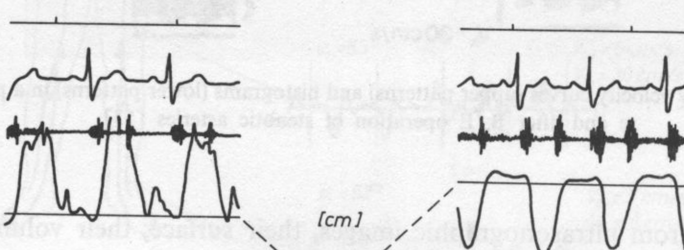


FIG. 15. The succeding records of time markers, ECG (a) FONO (B) and of Doppler blood velocities (c) carried out at 6 points at different depth in a childs heart by the UDP-30-SEC apparatus [22] (1982).

The Doppler transcranial technique for measurements of blood flows in cerebral arteries is an interesting method. Qualitative data achieved from the analysis of the blood veocity curve in terms of time are supplemented with spectral analysis in real time. However, this method does not result in quantitative data, yet it makes it possible to estimate disturbances of blood flowing in examined arteries. An example of such an examination is shown in Fig. 16.

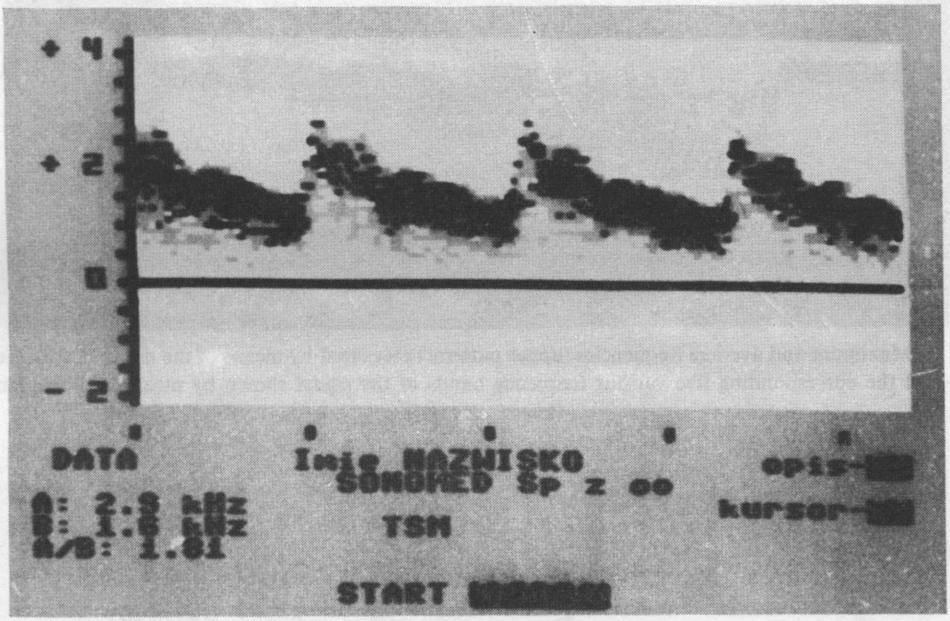


FIG. 16. Blood velocity measurements in the middle brain artery in terms of time carried out by means of the transcranial Doppler device IPPT PAN-Sonomed, Warsaw

The analysis of blood flow disturbances due to vascular pathologies is the subject of many papers, including Polish ones [41]. First of all, a system and device for bidirectional measurements of blood flow was developed to make correct interpretations of flows in disturbed areas possible [44]. Further research concentrated on methods for evaluation of the degree of disturbance by means of the analysis of Doppler signals in the domain of time [42] (Fig. 17). They were found helpful in determining methods of maximum blood flow velocity measurements using time interval histograms.

The determination of pressure gradients by means of blood velocity measurements is an interesting example of quantitative application of ultrasonography. This technique is based on the simplified Bernoulli's theorem [29, 37].

Another method of determining pressure in vessels was elaborated in our Department in 1985. This method consists in tracing movements of both vessel walls and measuring their displacement by means of the echo method in presentation A

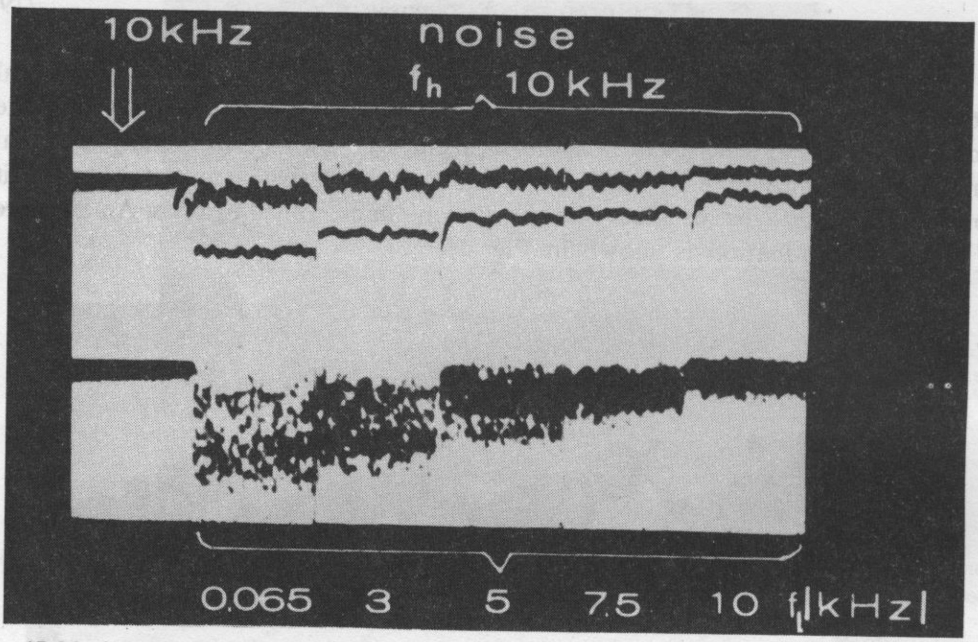


FIG. 17. Maximum and average frequencies (upper patterns) measured by means of the developed method [42] and the corresponding five various frequency bands of the signal shown by means of histograms (lower patterns).

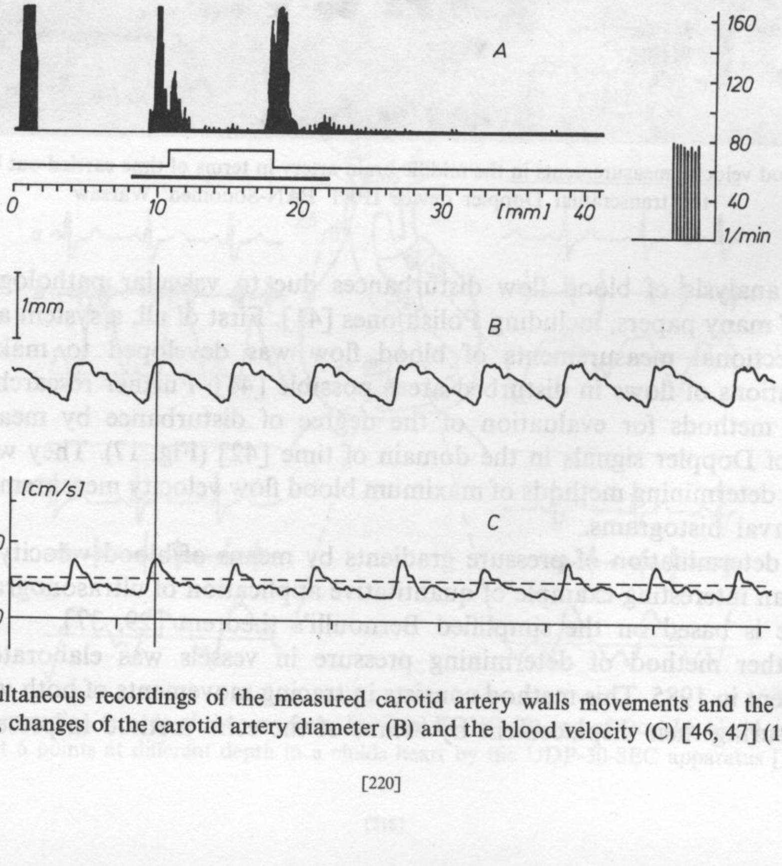


FIG. 18. Simultaneous recordings of the measured carotid artery walls movements and the heart beat frequency (A), changes of the carotid artery diameter (B) and the blood velocity (C) [46, 47] (1981, 1988).

[49, 51]. On the basis of a wide experimental material a relationship between the change in the vessels cross-section and internal pressure can be found [50]. The volume flow in ml/min is determined by measuring blood flow velocity with the Doppler method and by measuring the vessels cross-section (Fig. 18) On the basis of blood pressure and volume flow changes determined at the same time, the microprocessor calculated the hemodynamic impedance of the vascular system [48, 47].

Figure 19 presents the hemodynamic impedance of the carotid artery, obtained with this method. The whole calculation procedure is automatic and it is carried out by a special computer while the measurement is made with a duplex type special

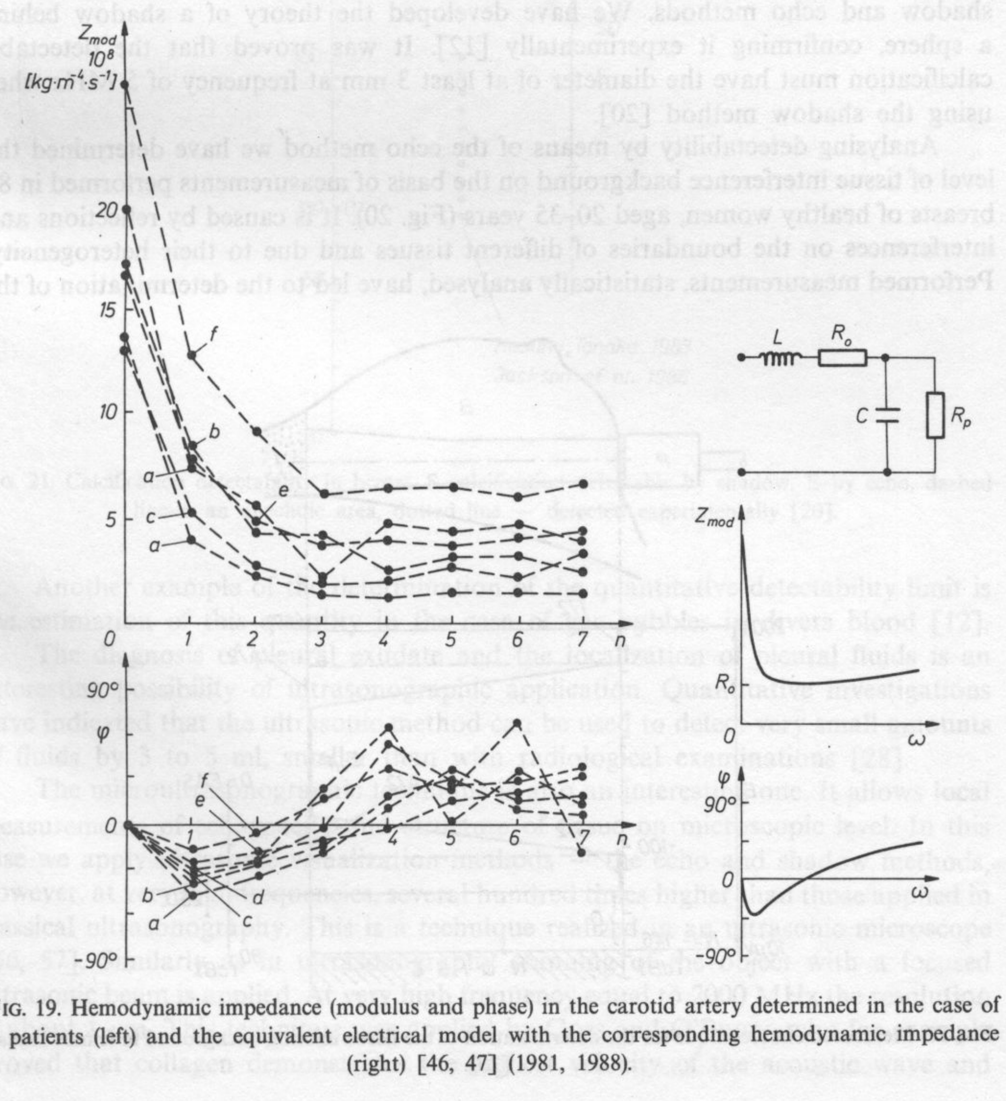


FIG. 19. Hemodynamic impedance (modulus and phase) in the carotid artery determined in the case of 6 patients (left) and the equivalent electrical model with the corresponding hemodynamic impedance (right) [46, 47] (1981, 1988).

head. This method can be used not only for the noninvasive determination of the hemodynamic impedance, but also for the determination of the vessels rigidity by means of joined ultrasonographic and Doppler technique.

The determination of a quantitative limit of pathologic structure detectability with ultrasonography is a very significant problem. The general case of this problem is very complicated because of various shapes, acoustic properties and structures of detected tissular systems. However, it is worth to make an atempt of answering this question for particular cases.

We have chosen calcifications occurring in the early stage of the neoplastic disease in breasts. This issue was analysed on two models, namely calcifications in the form of elastic and rigid spheres. Two detection methods were considered, shadow and echo methods. We have developed the theory of a shadow behind a sphere, confirming it experimentally [12]. It was proved that the detectable calcification must have the diameter of at least 3 mm at frequency of 5 MHz when using the shadow method [20].

Analysing detectability by means of the echo method we have determined the level of tissue interference background on the basis of measurements performed in 82 breasts of healthy women, aged 20–35 years (Fig. 20). It is caused by reflections and interferences on the boundaries of different tissues and due to their heterogeneity. Performed measurements, statistically analysed, have led to the determination of the

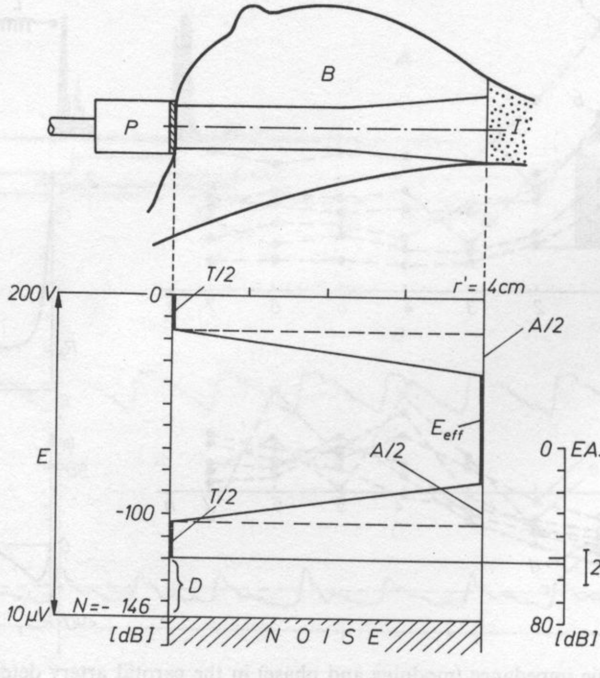


FIG. 20. Measurement technique for the determination of the tissue infering background in female breasts [20].

minimum diameter of detected calcifications, equal to 0.4 mm [20, 52]. The possibility of distinguishing echoes of calcifications increases during the detection of calcifications on the background of the neoplasms shadow.

If the tissue interference background is neglected, then detectability increases by one order of magnitude, as marked with a dashed line in Fig. 21.

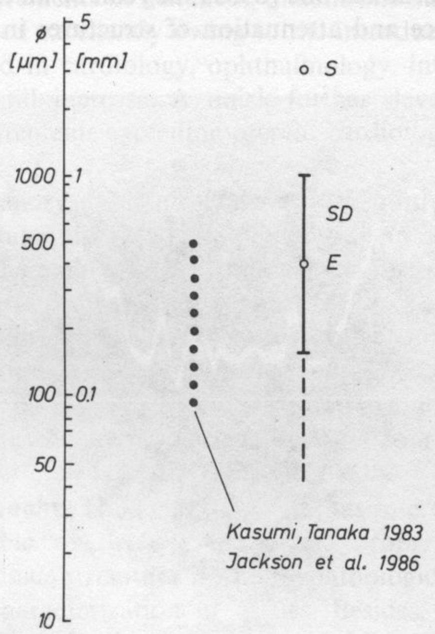


FIG. 21. Calcification detectability in breast. S–calcification detectable by shadow, E–by echo, dashed line-in an anechoic area, dotted line – detected experimentally [20].

Another example of the determination of the quantitative detectability limit is the estimation of this quantity in the case of gas bubbles in divers blood [12].

The diagnosis of pleural exudate and the localization of pleural fluids is an interesting possibility of ultrasonographic application. Quantitative investigations have indicated that the ultrasonic method can be used to detect very small amounts of fluids by 3 to 5 ml, smaller than with radiological examinations [28].

The microultrasonographic technique is also an interesting one. It allows local measurements of echogeneity and structure of tissue on microscopic level. In this case we apply ultrasonic visualization methods – the echo and shadow methods, however, at very high frequencies, several hundred times higher than those applied in classical ultrasonography. This is a technique realized in an ultrasonic microscope [56, 57]. Similarly as in ultrasonography scanning of the object with a focused ultrasonic beam is applied. At very high frequency equal to 2000 MHz the resolution is about 1 µm. This technique was applied by GOSS and O'BRIEN, who for example proved that collagen demonstrates the highest velocity of the acoustic wave and therefore highest impedance and, what follows, the highest echogeneity among soft tissular structures [27].

The possibility of investigations of cells in vivo, e.g. neoplastic cells, of determining acoustic parameters of these objects and of observing their dynamics are the important advantages of an acoustic microscope. An in vivo image with a signal distribution along a determined line [33, 31, 32] can be an example. It can be used to determine the impedance and attenuation of structures in observed cells (Fig. 22).

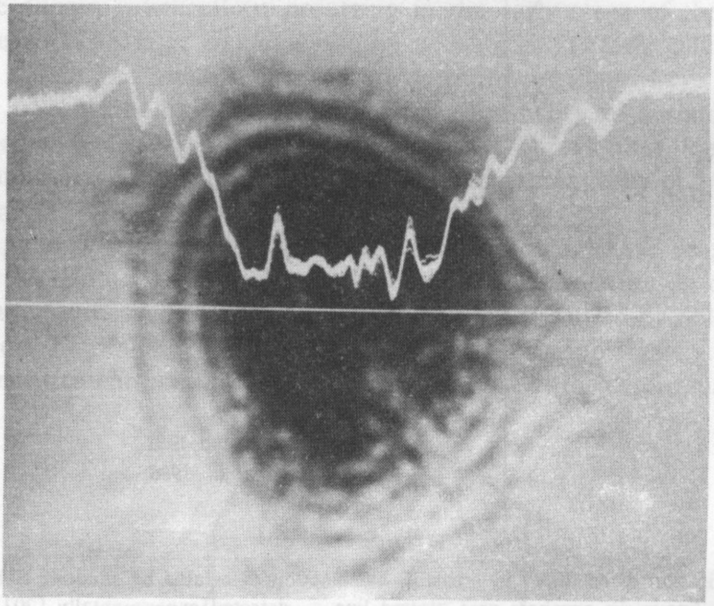


FIG. 22. Distribution of the acoustic signal along the living cell obtained by means of the acoustical microscope with the frequency of 2GHz enabling the determination of the local impedance, attenuation and thickness [33, 32]

Conclusions

To finish this review it seems useful to consider the following five trends of quantitative ultrasonography development:

1. Tissue characterization. Research is aimed at the determination of types of pathologies of detected structures. It is a very complicated problem. Information about these structures which should be obtained in vivo depends on many parameters, such as: frequency, attenuation, scattering, geometry, spectral parameters of tissues, parameters of ultrasonic beams penetrating the tissues etc. Thus, the work performed within this topic during the last 20 years has not given satisfactory results. Several leading centres have developed certain conceptions, but they are put to clinical use only within those centres for their own use [53]. Clinical investigations on a large scale, which would allow the acceptation or rejection of these conceptions, can not be carried out due to a lack of highly advanced and availabe apparatus. Producers of ultrasonographs have yet not decided to take the risk of investing in this type of equipment.

2. Biometry. Measurements of distances, dimensions and velocities of examined structures have developed excellently owing to the introduction of microcomputers. They are widely applied in cardiology, ophthalmology, internal medicine etc. and their application will still increase. A quick further development of blood flow visualization and measurements exceeding present cardiological applications should be expected.

3. Ultrasonic dosimetry. This problem is constantly current' especially in obstetrics, since it was stated that ultrasound may have an harmfull thermal effect on the nervous system of the embryo. Several groups of scientists abroad are working on this problem right now. The creation of a special committee for the investigation of this problem in the World Health Organization indicates its importance. It seems an optimal solution to design a special dosimetric device which would measure the dose accepted by every patient during an ultrasonic examination. On the basis of such an information it would be possible to carry out credible epidemiologic research in this field.

4. Microultrasonography. This new technique has interesting perspectives. First of all it makes possible to determine acoustic properties and, consequently, echogeneity of various tissue structures including pathological ones. This could be of great importance for characterization of tissues. Besides, this new field can find application in examinations "in vivo" of tissues and cells, their dynamics etc., what is impossible with such techniques as optic or electron microscopy.

5. Vascular ultrasonography. It seems that this specialization will develop quickly and will be quickly introduced to clinical ultrasonography. There are already possibilities of measurements of movements of vascular walls of blood flow determination, measurements of blood pressure dynamics as well as determination of hemodynamic impendance and even the rigidity of vessels. This absolutely new information can be achieved with fully noninvasive ultrasonic method "in vivo".

Moreover, further progress in microelectronics and in ultrasonic technology undoubtedly will result in the development and price decrease of the existing ultrasonic instrumentation and in the construction of new models, which will enable the introduction of new ideas of quantitative ultrasonography to clinical practice.

References

[1] *AIUM. Bioeffects considerations for the safety of diagnostic ultrasound*, J. Ultrasound in Medicine, 7, 9 (1988) Suplement S. 3 and S. 35.

[2] K. Borodziński, L. Filipczyński, A. Nowicki, T. Powałowski, *Quantitative transcutaneous*

measurements of blood flow in carotid artery by means of pulse and continuous wave Doppler methods, Ultrasound in Medicine and Biology, **2**, 3, 189–193 (1976).

[3] I. Donald, Mac Vicar, T. G. Brown, *Investigation of abdominal masses by pulsed ultrasound*, The Lancet 1168–1195 (1958).

[4] J. Etienne, L. Filipczyński, A. Firek, J. Groniowski, J. Kretowicz, J. Sałkowski, *Intensity determination of ultrasonic focused beams used in ultrasonography in the case of gravid uterus*, Ultrasound in Medicine and Biology, **2**, 119–122 (1976).

[5] L. Filipczyński [Ed.] *Ultrasonics in biology and medicine*, PWN, Warsaw 1972.

[6] L. Filipczyński, *Measurement of the temperature increases generated in soft tissue by means of ultrasonic diagnostic Doppler equipment*, Ultrasound in Medicine and Biology, **4**, 2, 151–155 (1978).

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

[8] L. Filipczyński, *Thermal effects in soft tissue developed under the action of ultrasonic fields of long duration*, Archives of Acoustics **2**, 4, 297–303 (1977).

[9] L. Filipczyński, *Absolute measurement methods and possibilities of thermographic estimation of the ultrasonic dose in vivo*, In: Investigate ultrasonology Technical Advances, R. C. Hill, Alvisi [Edts.] Pitman Medical, Kent 1980.

[10] L. Filipczyński, *Experiments and estimation of the thermal effect generated under the action of ultrasonic waves in soft tissues by means of the Doppler diagnostic instrumentation* (in Polish) Archiwum Akustyki **13**, 3, 215–222 (1978).

[11] L. Filipczyński, *Absorption of longitudinal and sheare waves and generation of heat in soft tissues*, Ultrasound in Medicine and Biology, **12**, 3, 223–228 (1986).

[12] L. Filipczyński, *Diffusion ultrasonore dans le sang et detection des bulles de gaz*. Innovation et Technologie en Biologie et Medicine, **5**, 1, 6–7 (1984).

[13] L. Filipczyński, *Compound and rapid scan ultrasonic imaging of eye structures*, Ophtalmic Ultrasound. Proc. Ultrasonography in Ophthalmology [Ed.] R. Sarin Mosby. St. Luis 1969, 207–212.

[14] L. Filipczyński, J. Etienne, G. Łypacewicz, *Visualizing the inside of the abdomen by means of ultrasonics*, Proc. of Vibration Problems, Warsaw, 7, 3, 211–220 (1966).

[15] L. Filipczyński, J. Etienne, G. Łypacewicz, J. Sałkowski, *Visualizing internal structures of the eye by means of ultrasonics*, ditto **8**, 4, 357–368 (1967).

[16] L. Filipczyński, J. Groniowski, *Visualization of the inside of the abdomen by means of ultrasonics and two methods for measuring ultrasonic doses.* Digest of the 7-th International Conference on Medical and Biological Engineering, Stockholm 1987, 320.

[17] L. Filipczyński, R. Herczyński, A. Nowicki, T. Powałowski, *Hemodynamics and ultrasonic Doppler measurement methods* (in Polish) PWN Warsaw 1980.

[18] L. Filipczyński, D. Komitowski, G. Łypacewicz, B. Rościszewska, J. Sałkowski, *Ultrasonograph UO-3 for eye vizualization and its clinical application* (in Polish) Archiwum Akustyki **8**, 3, 265–282 (1973).

[19] L. Filipczyński, T. Kujawska, T. Waszczuk, *The shadow behind a sphere immersed in water-measured, estimated and computed*, Trans. of IEEE Ultrasonics, Ferroelectrics and Frequency Control **38**, 1, 35–39 (1991).

[20] L. Filipczyński, T. Kujawska, G. Łypacewicz, *Evaluation of calcification detectability in female breasts*, Archives of Acoustics **15**, 1–2, 107–119 (1990).

[21] L. Filipczyński, G. Łypacewicz, D. Komitowski, B. Rościszewska, *Ultrasonic intensity levels in imaging of eye structures and their influence on eye tissues*, Proc. SIDUO IV [Ed.] J. Poujol Paris 1971, 17–23.

[22] L. Filipczyński, A. Nowicki, A. Chrościcki, *Application of the stationary echo cancellation technique SEC in ultrasonic Doppler mesurements of blood flow in childrens hearts*, Archives of Acoustics, **7**, 3–4, 247–252 (1982).

[23] L. Filipczyński, M. Piechocki, *Estimation of the temperature increase in the focus of the lithotripter*

for the case of high rate administration, Ultrasound in Medicine and Biology, **16**, 2, 149, 727 (1990).

[24] L. Filipczyński, I. Roszkowski, [Edts.], *Ultrasonic diagnostics in obstetrics and in ginecology* (in Polish) PZWL Warsaw 1977.

[25] L. Filipczyński, J. Wójcik, *Estimation of transient temperature elevation in lithotripsy and in ultrasonography*. Ultrasound in Medicine and Biology 1991 (in print).

[26] L. Filipczyński, J. Zieniuk [Eds.], *Proceeding of the Second Congress of the Federation of Acoustica Societes of Europe*, Vol. II, Warsaw 1978.

[27] S. A. Goss, W. D. Jr. O'Brien, *Direct ultrasonic velocity measurements of mammalian colagen threads*, J. Acoust. Soc. Am. **65**, 2, (1979).

[28] J. Gryminski, P. Krakówka, G. Łypacewicz, *The diagnosis of pleural effusion by ultrasonic and radiological techniques*, Chest **70**, 33–37 (1976).

[29] Holen et al., *Determination of pressure gradient in mitral stenosis with noninvasive ultrasound Doppler technique*, Acta Med. Scand., **199**, 455–460 (1976).

[30] W. Jakubowski [Ed.], *Ultrasonic diagnostics* (in Polish), PZWL Warsaw 1989.

[31] J. Litniewski, *An acoustic microscope in measurements of mechanical properties of surface layers — V(z)*, Archives of Acoustics **11**, 3, 299–312 (1986).

[32] J. Litniewski, *On the possibility of the visualization of the velocity distribution in biological samples using SAM*, IEEE Transactions of Ultrasonics, Ferroelectrics and Frequency Controll, **36**, 1, 134–135 (1989).

[33] J. Litniewski, J. Bereiter-Hahn, *Measurements of cells in culture by scanning acoustic microscope*, Journal of Microscopy, **158**, 1, 95–107 (1990).

[34] A. Nowicki, *Ultrasonic pulse Doppler method in blood flow measurement*, Archives of Acoustics, **2**, 4, 305–323 (1977).

[35] A. Nowicki, *Simplified automatic measurements of blood flow by the ultrasonic pulse Doppler method*, Archives of Acoustics, **4**, 359–366 (1979).

[36] A. Nowicki, P. Karłowicz, M. Piechocki, W. Secomski, *Method for the measurement of the maximum Doppler frequency*, Ultrasound in Medicine and Biology, **11**, 3 479–486 (1985).

[37] A. Nowicki, *Doppler echography* (in Polish) PWN, Warsaw 1989.

[38] A. Nowicki, A. Marciniak, *Detection of wall vibrations by means of cepstrum analysis*, Ultrasonic Imaging. **11**, 273–282 (1989).

[39] A. Nowicki, J. M. Reid, *An infinite gate pulse Doppler*, Ultrasound in Medicine and Biology, **7**, 41–50 (1981).

[40] A. Nowicki, J. Reid, *An ultrasonic method for the investigation of coronary graft patency*, Archives of Acoustics **8**, 3, 223–234 (1983).

[41] A. Nowicki et al., *Attempts at the determination of the arterial blood flow disturbance intensity* (in Polish), Polski Przegląd Chirurgiczny **55**, 2, 105–108 (1983).

[42] M. Piechocki, *Application of a phase loop system to analysis of a Doppler signal in ultrasonic systems*, Archives of Acoustics, **10**, 3, 273–280 (1985).

[43] M. Piechocki, *Flow perturbation measurements by numerical analysis of the power spectrum of a Doppler signal*, Archives of Acoustics, **10**, 3, 281–302 (1985).

[44] M. Piechocki, *Bi-directional continuous wave Doppler flowmeter and its application in investigations of disturbed blood flows*, Archives of Acoustics, **11**, 1, 75–88 (1986).

[45] T. Powałowski, *Real time automatic transcutaneous determination of blood velocity by means of c.w. Doppler methods eliminating angle dependence*, Proc. Second Congress of the Federation of Acoustic Soc. of Europe, [Eds.] L. Filipczyński, J. Zieniuk, Warsaw 1978 vol. II, 157–160.

[46] T. Powałowski, *An ultrasonic c. w. Doppler method of measurement of the blood flow velocity*, Archives of Acoustics, **6**, 3, 257–306 (1981).

[47] T. Powałowski, *Ultrasonic system for noninvasive measurement of hemodynamic parameters of human arterial-vascular system*, Archives of Acoustics, **13**, 1–2, 89–108 (1988).

[48] T. Powałowski, *A noninvasive ultrasonic method for vascular input impedance determination applied in diagnosis of the carotid arteries*, Archives of Acoustics, **14**, 3–4, 293–312 (1989).

[49] T. Powałowski, B. Peńsko, *Noninvasive ultrasonic method of pressure and flow measurements for estimation of hemodynamical properties of cerebrovascular system.*, Archives of Acoustics, **10**, 3, 303–314 (1985).

[50] T. Powałowski, B. Peńsko, *A noninvasive ultrasonic method for the elasticity evaluation of the carotid arteries and its application in the diagnosis of the cerebro-vascular system*, Archives of Acoustics, **13**, 1–2–109–126 (1988).

[51] T. Powałowski, B. Peńsko, Z. Trawiński, L. Filipczyński, *An ultrasonic transcutanepis method of simultaneous evaluation of blood pressure, flow rate and pulse wave velocity in the carotid artery*, Abstracts of Satellite Symposium of the XIIIth World Congress of Neurology, Aachen, M, 68 1985.

[52] *Proceeding of the Conf. Ultrasound in Biology and Medicine UBIOMED VI*, Archives of Acoustics **9**, 1–2, 1–298 (1984).

[53] D. E. Robinson, R. W. Gill, G. Kossoff, *Quantitative Sonography*, Ultrasound in Medicine and Biology, **12**, 7, 555–565 (1986).

[54] J. Wesołowski, A. Wycech, A. Nowicki, H. Rykowski, *The usefulness of Doppler time interval histogram analysis in arterial occlusion*, International Angiology, **7**, 4, 309–311 (1988).

[55] J. Wesołowski, A. Nowicki, B. Topolska, G. Pawlicki, I. Pałko, L. Filipczyński, H. Rykowski, *Attempts at quantitative determination of the collateral blood flow in normal and ischaemic extremities by means of the ultrasonic and rheographic methods* (in Polish). Polski Przegląd Chirurgiczny **53**, 2, 129–135 (1981).

[56] J. K. Zieniuk, *Mikroskop ultradźwiękowy: Technika. Zastosowania.* Materiały XXV OSA, Białowieża 1988 V. I p. 66–80.

[57] J. K. Zieniuk, A. Latuszek, *Ultrasonic pin scanning microscope a new approach to ultrasonic microscopy*, Ultrasonic Symp. 1986, 1037–1039.

Received October 25, 1990

ARCHIVES OF ACOUSTICS
16, 2, 229–235 (1991)

TRANSFER IMPEDANCE OF A THREE-LAYER VISCOELASTIC ROD

W. BANDERA

University of Gdańsk, Laboratory of Applied Acoustics and Spectroscopy
(80-952 Gdańsk ul. Wita Stwosza 57)

The present work is an attempt to apply the transfer impedance method to the investigation of systems composed of several layers of a material with known viscoelastic properties. A multilayer rod excited to longitudinal vibrations is examined theoretically. The end of the rod is stiffly connected with an optional mass. We assume that the component layers are homogeneous whereas the specific wave impedance undergoes jump-like changes at the border of the layers. In order to find the expression for the transfer impedance of the rod considered, the electromechanical analogies are used. The formulae given in the work allow to find the transfer impedance modulus which well describes the ability of a given structure to energy transmission. So, knowing the sound velocity c_n and loss factor η_n for a given layer of the material, one can determine the components of propagation constant α_α and β_n which subsequently allow to find the modulus of the transfer impedance for an optional frequency, at the fulfilled condition that the length of the longitudinal wave in the rod is much bigger than the lateral dimension of the rod. To illustrate the usefulness of the introduced formulae, a number of numerical investigations for vibroisolating materials are made. The influence of the mutual configuration of layers and their properties as components on the transfer impedance of the composed rod-like specimen is discussed.

Praca stanowi próbę zastosowania impedancji przenoszenia do badania układów złożonych z kilku warstw materiału o znanych własnościach lepkosprężystych. Rozpatrzono teoretycznie przypadek pręta lepkosprężystego złożonego z trzech różnych warstw i pobudzanego do drgań harmonicznych. Założono, że warstwy składowe są jednorodne, przy czym na granicy warstw impedancja falowa zmienia się skokowo. Wyprowadzono wzory na część rzeczywistą i urojoną impedancji przenoszenia pręta połączonego sztywno z dowolną masą oraz podano przykładowe wyniki obliczeń numerycznych, pokazujących wpływ konfiguracji warstw na charakter zmian modułu impedancji przenoszenia rozpatrywanego pręta.

1. Introduction

Designing of sandwich type partitions of given insulating properties requires determination of acoustic impedance not only for separate layered components but also for complex structures as their effective values. There are papers related to evaluation of transmission and reflection coefficients of layered media [4, 5]. In [8] a partition of the impedance gradient distribution across its thickness was suggested. Such a system may be constructed as consisting of many layers differing among each other of a small jump of impedance. Construction of layered elements requires measurements of impedance variations during joining layers together. Works [6, 1, 2, 3] have shown that when applying measurements of mechanical impedance (input or transfer ones), it is easy to determine the viscoelastic properties of rubber-like materials. The present work is an attempt of applying the transfer impedance method to the investigation of systems composed of several layers of a material with known viscoelastic properties.

In this paper the general equation of transfer impedance of a three-layer viscoelastic rod, stiffly connected with an optional mass, has been derived and the influence of the mutual configuration of layers and their properties as components on the transfer impedance modulus of the composed rod-like specimen have been presented.

2. Theory

Let us consider the situation shown in Fig. 1 which presents a theoretical model of a three-layer rod excited to longitudinal vibrations at point $x = 0$. The end of the

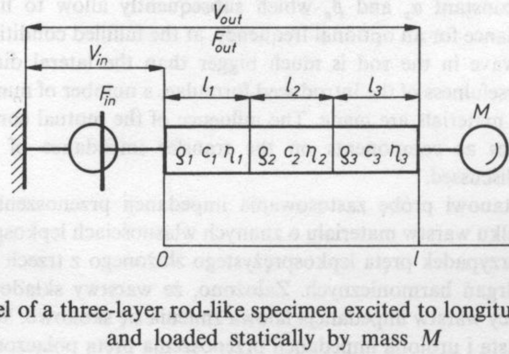


FIG. 1. Theoretical model of a three-layer rod-like specimen excited to longitudinal vibrations by a shaker and loaded statically by mass M

rod is stiffly connected with as optional mass M. It is assumed that the component layers are homogeneous whereas the specific wave impedance undergoes jump-like changes at the border of the layers. The considered rod is, thus, a combination of three rods of different wave impedances Z_{01}, Z_{02} and Z_{03}. Besides it is assumed that the cross-section is constant along the rod and its lateral dimension is much smaller than the length of the longitudinal wave propagating in the rod. In order to find the expression for the transfer impedance of the rod considered, one makes use of electromechanical analogies. The analogue of the rod presented in Fig. 1 is the wave-guide loaded by inductivity which can be regarded as a chain connection of three mechanical four-poles [9]. The connection between the force and vibration velocity at the ends of the rod can be thus expressed by the following matrix equation:

$$\begin{bmatrix} F_{\text{in}} \\ V_{\text{in}} \end{bmatrix} = \begin{bmatrix} \operatorname{ch}\gamma_1 l_1 & Z_{01}\operatorname{sh}\gamma_1 l_1 \\ Z_{01}^{-1}\operatorname{sh}\gamma_1 l_1 & \operatorname{ch}\gamma_1 l_1 \end{bmatrix} \begin{bmatrix} \operatorname{ch}\gamma_2 l_2 & Z_{02}\operatorname{sh}\gamma_2 l_2 \\ Z_{02}^{-1}\operatorname{sh}\gamma_2 l_2 & \operatorname{ch}\gamma_2 l_2 \end{bmatrix} \begin{bmatrix} \operatorname{ch}\gamma_3 l_3 & Z_{03}\operatorname{sh}\gamma_3 l_3 \\ Z_{03}^{-1}\operatorname{sh}\gamma_3 l_3 & \operatorname{ch}\gamma_3 l_3 \end{bmatrix} \begin{bmatrix} F_{\text{out}} \\ V_{\text{out}} \end{bmatrix} \tag{1}$$

where F_{in}, V_{in} and F_{out}, V_{out} are the complex forces and complex vibration velocities at input and output of the specimen, Z_0 is the wave impedance, l is the component rod length, γ is the complex propagation constant, $\gamma = \alpha + i\beta$ where α and β are the attenuation constant and phase constant, respectively. These latter can be expressed in terms of the material parameters [7] as

$$\beta = \beta_0\sqrt{D+1}/\sqrt{2D}, \qquad \alpha = \beta_0\sqrt{D-1}/\sqrt{2D}, \tag{2, 3}$$

$$\beta_0 = 2\pi f(\varrho/E_d)^{1/2}, \qquad D = \sqrt{1+\eta^2}, \tag{4, 5}$$

in which E_d is the dynamic Young's modulus, η is the loss factor, ϱ is the density and f is the frequency. The chain of three four-poles presented in Fig. 2 can be reduced to

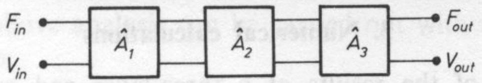


FIG. 2. Three-layer rod-like specimen regarded as a chain connection of three mechanical four-poles

an equivalent four-pole whose chain matrix is equal to the product of the chain matrixes of the component fourpoles [9], and equation (1) can be put down as

$$\begin{bmatrix} F_{\text{in}} \\ V_{\text{in}} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} F_{\text{out}} \\ V_{\text{out}} \end{bmatrix} \tag{6}$$

from which one obtains the expression for force F_{in}

$$F_{\text{in}} = AF_{\text{out}} + BV_{\text{out}} \tag{7}$$

Substituting $F_{\text{out}} = i\omega M V_{\text{out}}$ for (7) one gets

$$F_{\text{in}} = (i\omega MA + B)V_{\text{out}} \tag{8}$$

and finally the complex transfer impedance of the three-layer rod is found from the ratio of the exciting force F_{in} to the vibration velocity V_{out}

$$Z_t = \frac{F_{\text{in}}}{V_{\text{out}}} = i\omega MA + B \tag{9}$$

where elements A and B of the equivalent matrix (6) are expressed by the viscoelastic parameters of each rod layer as follows:

$$A = \mathrm{ch}\gamma_1 l_1 \mathrm{ch}\gamma_2 l_2 \mathrm{ch}\gamma_3 l_3 + Z_{01}\,\mathrm{sh}\gamma_1 l_1 Z_{02}^{-1}\,\mathrm{sh}\gamma_2 l_2 \cdot \mathrm{ch}\gamma_3 l_3 + \\ + \mathrm{ch}\gamma_1 l_1 \cdot Z_{02}\,\mathrm{sh}\gamma_2 l_2 Z_{03}^{-1}\,\mathrm{sh}\gamma_3 l_3 + Z_{01}\,\mathrm{sh}\gamma_1 l_1 \cdot \mathrm{ch}\gamma_2 l_2 \cdot Z_{03}^{-1}\,\mathrm{sh}\gamma_3 l_3 \quad (10)$$

$$B = \mathrm{ch}\gamma_1 l_1 \mathrm{ch}\gamma_2 l_2 \cdot Z_{03}\mathrm{sh}\gamma_3 l_3 + Z_{01}\,\mathrm{sh}\gamma_1 l_1 \cdot Z_{02}^{-1}\,\mathrm{sh}\gamma_2 l_2 \cdot Z_{03}\mathrm{sh}\gamma_3 l_3 + \\ + \mathrm{ch}\gamma_1 l_1 \cdot Z_{02}\,\mathrm{sh}\gamma_2 l_2 \mathrm{ch}\gamma_3 l_3 + Z_{01} \cdot \mathrm{sh}\gamma_1 l_1 \cdot \mathrm{ch}\gamma_2 l_2 \cdot \mathrm{ch}\gamma_3 l_3 \quad (11)$$

The real and imaginary parts of the complex transfer impedance written in an explicit form are given in the Appendix.

Formulas (9), (10) and (11) allow to find real and imaginary part of the transfer impedance for a given frequency if only one knows viscoelastic properties, i.e. sound velocity c and loss factor η for each layer of the material. The calculations can be carried out at optional boundary conditions accepting different values of loading mass M. Finally the transfer impedance modulus which describes well the ability of a structure to energy transmission is calculated. Putting in calculations $l_3 = 0$ one can proceed in a simple way to a two-layer system or when accepting $l_3 = 0$ and $l_2 = 0$ one obtains expressions for transfer impedance for a homogeneous rod, given earlier in [2].

3. Numerical calculations

Some examples of the results of a three-layer rod calculations versus the sequence of the layers are shown in Fig. 3. The calculations were made in the

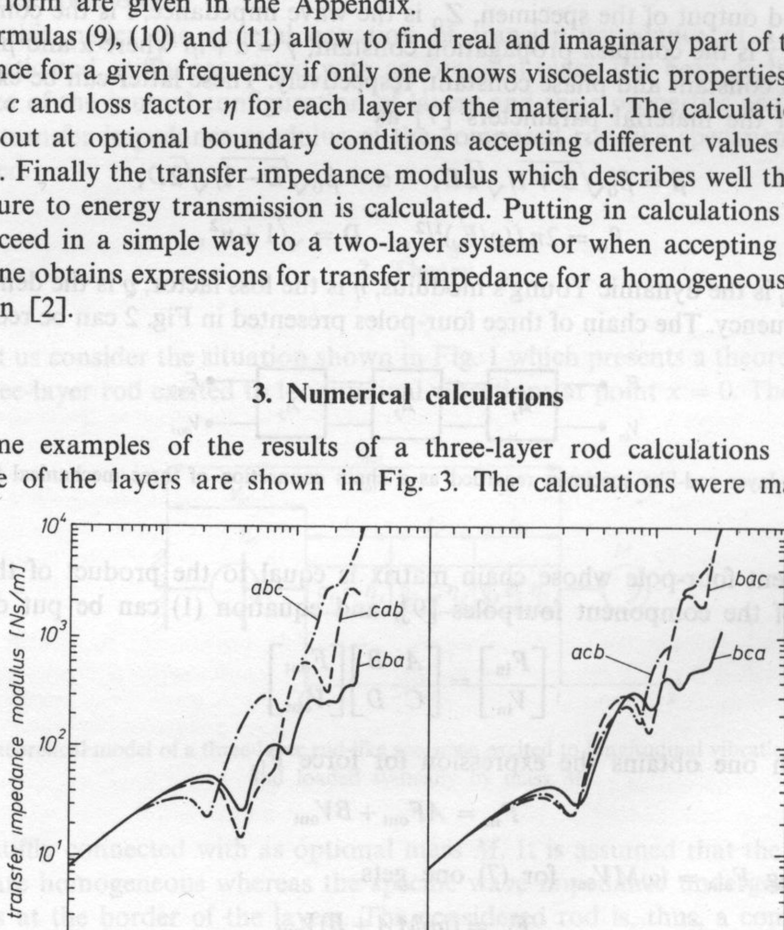


FIG. 3. Transfer impedance modulus for the different configuration of layers; a–Elastomer Z-22, b–Elastomer Z-7, c–Polyurethane Syntactic Foam. Calculations were made for $M/m = 1$, where M is the loading mass and m is the total mass of the three-layer specimen

frequency range from 0 Hz up to 3000 Hz. The following values have been accepted for the calculations: $c_1 = 60$ m/s, $\eta_1 = 0.35$, $\varrho_1 = 936$ kg/m^3, $c_2 = 95$ m/s, $\eta_2 = 0.14$, $\varrho_2 = 1129$ kg/m^3, $c_3 = 270$ m/s, $\eta_3 = 0.13$, $\varrho_3 = 1313$ kg/m^3 appearing for real rubber-like materials i.e. Polyurethane Syntactic Foam, Elastomer Z-7 and Elastomer Z-22, respectively. Six different configurations abc, cab, cba, acb, bac and bca were numerically examined. As seen from the figure, the character of the changes of transfer impedance modulus of the regarded layer system is similar to the configuration of layers abc and acb as well as bca and cba. The influence of the mutual configuration of layers on the transfer impedance modulus is evident.

4. Conclusions

The introduced formulas enable a quick numerical analysis of transfer impedance modulus for an optional two- or three-layer system with required viscoelastic properties. The method enables to control how the transfer impedance modulus is varied depending on the sequence of layers and their viscoelastic properties. In the author's opinion, the presented method of numerical analysis of transfer impedance should be helpful in the estimation of the properties of viscoelastic layer systems and may enable their right selection for the vibration minimization. It is worth emphasizing that the above analysis can be carried out with the help of relatively simple microcomputers.

APPENDIX

In order to write down expression for real and imaginary parts of the complex transfer impedance in an explicit form, elements A and B (see eq. (10) and (11)) of matrix (6) should be expressed by the viscoelastic parameters of each rod layer. To do this, the following substitutions can be introduced:

$$Z_{\text{on}} = x_n + iy_n, \quad x_n = \frac{S\omega\varrho_n\beta_n}{\alpha_n^2+\beta_n^2}, \quad y_n = \frac{S\omega\varrho_n\alpha_n}{\alpha_n^2+\beta_n^2} \tag{A1}$$

$$Z_{\text{on}}^{-1} = u_n - iw_n, \quad u_n = \frac{\beta_n}{\omega\varrho_n S}, \quad w_n = \frac{\alpha_n}{\omega\varrho_n S}, \quad E_n = \operatorname{ch}\alpha_n l_n \cos\beta_n l_n \tag{A2}$$

$$F_n = \operatorname{sh}\alpha_n l_n \sin\beta_n l_n, \quad H_n = \operatorname{sh}\alpha_n l_n \cos\beta_n l_n, \quad K_n = \operatorname{ch}\alpha_n l_n \sin\beta_n l_n \tag{A3}$$

where Z_{on} is the wave impedance, α_n is the attenuation constant, β_n is the phase constant, ϱ_n is the material density, S is the cross-sectional area of the rod, l_n is the component rod length, n is the number of successive layer ($n = 1, 2, 3$), $i = \sqrt{-1}$ is the imaginary unit. The attenuation constant α_n and the phase constant β_n can be expressed in terms of dynamic characteristics as

$$\beta_n = \frac{\omega}{c_n}, \quad \alpha_n = \beta_n \frac{\sqrt{D_n - 1}}{\sqrt{D_n + 1}}, \quad D_n = \sqrt{1 + \eta_n^2} \tag{A4}$$

where c_n is the sound velocity and η_n is the loss factor.

Using the above given substitutions, after the calculations according to (10), (11), (9) and by applying several mathematical manipulations and transformations one obtains the final form of the expression for the real and imaginary part of transfer impedance of three-layer rod in the following form:

$$\begin{aligned}
Re(Z_t) = \omega M[&E_3(E_1F_2+F_1E_2)+F_3(E_1E_2-F_1F_2)+\\
&+(H_1K_2+K_1H_2)\cdot(x_1w_2F_3+x_1u_2E_3+y_1w_2E_3-y_1u_2F_3)+\\
&+(H_1H_2-K_1K_2)(x_1u_2F_3+y_1u_2E_3+y_1w_2F_3-x_1w_2E_3)+\\
&+(E_1K_2+F_1H_2)(x_2u_3H_3+x_2w_3K_3+y_2w_3H_3-y_2u_3K_3)+\\
&+(E_1H_2-F_1K_2)(x_2u_3K_3+y_2u_3H_3+y_2w_3K_3-x_2w_3H_3)+\\
&+(H_1F_2+K_1E_2)(x_1u_3H_3+x_1w_3K_3+y_1w_3H_3-y_1u_3K_3)+\\
&+(H_1E_2-K_1F_2)(x_1u_3K_3+y_1u_3H_3+y_1w_3K_3-x_1w_3H_3)]+\\
&+E_1E_2(x_3H_3-y_3K_3)-E_1F_2(y_3H_3+x_3K_3)-F_1E_2(y_3H_3+x_3K_3)-\\
&-F_1F_2(x_3H_3-y_3K_3)+E_1E_3(x_2H_2-y_2K_2)-E_1F_3(x_2K_2+y_2H_2)-\\
&-F_1E_3(x_2K_2+y_2H_2)-F_1F_3(x_2H_2-y_2K_2)+E_2E_3(x_1H_1-y_1K_1)+\\
&+F_2F_3(y_1K_1-x_1H_1)-E_2F_3(x_1K_1+y_1H_1)-\\
&-F_2E_3(x_1K_1+y_1H_1)+(x_3H_3-y_3K_3)(x_1u_2H_1H_2+\\
&+x_1w_2H_1K_2)+(x_3K_3+y_3H_3)(x_1w_2H_1H_2-x_1u_2H_1K_2)+\\
&+(x_3H_3-y_3K_3)(x_1w_2K_1H_2-x_1u_2K_1K_2)-(x_3K_3+y_3H_3)\cdot\\
&\cdot(x_1K_1u_2H_2+x_1K_1w_2K_2)-(x_3K_3+y_3H_3)(y_1H_1u_2H_2+\\
&+y_1H_1w_2K_2)+(x_3H_3-y_3K_3)(y_1H_1w_2H_2-y_1H_1u_2K_2)+\\
&+(x_3K_3+y_3H_3)(y_1K_1u_2K_2-y_1K_1w_2H_2)-\\
&-(x_3H_3-y_3K_3)(y_1K_1u_2H_2+y_1K_1w_2K_2).
\end{aligned} \tag{A5}$$

$$\begin{aligned}
\mathrm{Im}(Z_t) = \omega M[&E_3(E_1E_2-F_1F_2)-F_3(E_1E_2-F_1F_2)+(H_1H_2-\\
&-K_1K_2)(x_1u_2E_3+x_1w_2F_3+y_1w_2E_3-y_1u_2F_3)+(H_1K_2+\\
&+K_1H_2)(x_1w_2E_3-x_1u_2F_3-y_1u_2E_3-y_1w_2F_3)+(E_1H_2-\\
&-F_1K_2)(x_2u_3H_3+x_2w_3H_3+y_2w_3H_3-y_2u_3K_3)+(E_1K_2+
\end{aligned} \tag{A.6}$$

$$+F_1H_2)(x_2w_3H_3-x_2u_3K_3-y_2u_3H_3-y_2w_3K_3)+(H_1E_2- \tag{A6}$$
$$-K_1F_2)(x_1u_3H_3+x_1w_3K_3+y_1w_3H_3-y_1u_3K_3)+(H_1F_2+ \quad \text{[cont.]}$$
$$+K_1E_2)\cdot(x_1w_3H_3-x_1u_3K_3-y_1u_3H_3-y_1w_3K_3)]+$$
$$+E_1E_2(y_3H_3+x_3K_3)+E_1F_2(x_3H_3-y_3K_3)+F_1E_2(x_3H_3-$$
$$-y_3K_3)-F_1F_2(y_3H_3+x_3K_3)+E_1E_3(x_2K_2+y_2H_2)+$$
$$+E_1F_3(x_2H_2-y_2K_2)+F_1E_3(x_2H_2-y_2K_2)-F_1F_3(x_2K_2+$$
$$+y_2H_2)+E_2F_3(x_1H_1-y_1K_1)+F_2E_3(x_1H_1-y_1K_1)+$$
$$+E_2E_3(x_1K_1+y_1H_1)-F_2F_3(x_1K_1+y_1H_1)+(x_3K_3+$$
$$+y_3H_3)(x_1H_1u_2H_2+x_1H_1w_2K_2)+(x_3H_3-y_3K_3)(x_1H_1u_2K_2-$$
$$-x_1H_1w_2H_2)+(x_3H_3-y_3K_3)(x_1K_1u_2H_2+x_1K_1w_2K_2)+$$
$$+(x_3K_3+y_3H_3)(x_1K_1w_2H_2-x_1K_1u_2K_2)+(x_3H_3-$$
$$-y_3K_3)(y_1H_1u_2H_2+y_1H_1w_2K_2)+(x_3K_3+y_3H_3)(y_1H_1w_2H_2-$$
$$-y_1H_1u_2K_2)+(x_3H_3-y_3K_3)(y_1K_1w_2H_2-y_1K_1u_2K_2)-$$
$$-(x_3K_3+y_3H_3)(y_1K_1u_2H_2+y_1K_1w_2K_2).$$

The transfer impedance modulus is calculated from well-known formula

$$|Z_t| = [Re^2(Z_t)+I_m^2(Z_t)]^{1/2}. \tag{A7}$$

References

[1] W. Bandera, *Mechanical impedance method for measurement of dynamic properties of high loss and low loss materials*, Proceedings of the 3rd European Conference on Nondestructive Testing, Florence 1984, Vol. 4, pp. 259–268.

[2] W. Bandera, *Application of the transfer mechanical impedance method for investigation of viscoelastic materials utilised in noise and vibration protection*, Proceeding of the 5th FASE Symposium, Thessaloniki 1985, pp. 231–234.

[3] W. Bandera, *Sposób i urządzenie do wyznaczania dynamicznego modułu Young'a oraz współczynnika strat gumy i materiałów gumopodobnych*, Patent PRL Nr 240897, 1983.

[4] P. D. Jackins, G. C. Gaunaurd, *Resonance scattering of sound waves by sets of flat inhomogeneous rubber layers*, Proceedings of the 11-th ICA Congres, Paris 1983, Vol. 1, pp. 95–98.

[5] W. M. Madigosky, R. Fiorito, H. Uberall, *Resonance theory for layered media*, Proceedings of the 11-th ICA Congres, Paris, 1983, Vol. 1, pp, 99–102.

[6] G. S. Rosin, *Measurement of dynamic properties of acoustic materials*, Moscow, Izdatelstvo Literatury po Stroitelstvu, 1972 (in Russian).

[7] I. C. Snowdon, *Vibration and shock in damped mechanical systems*, New York: John Wiley and Sons, 1968.

[8] A. Śliwiński, *On sound transmission through insulating walls*, Proceedings of the 5-th FASE Symposium, Thessaloniki 1985, pp. 147–150.

[9] R. Wyrzykowski, A. Puch, J. Snakowski, *Filtry akustyczne*, Rzeszów, WSP, 1972.

Received September 12, 1989

ARCHIVES OF ACOUSTICS
16, 2, 237–247 (1991)

IMAGE SIMILARITY FUNCTIONS IN NON-PARAMETRIC ALGORITHMS OF VOICE IDENTIFICATION

CZ. BASZTURA AND J. ZUK

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

This paper is dedicated to the question of the choice of a function of similarity between images in non-parametric alogorithms of voice recognition. The usefulness of 10 similarity functions (8 distances and 2 nearness'es) in three non-parametric identification algorithms — *NN* (nearest neighbour), *k–NN* (*k*–nearest neighbours) and *NM* (nearest mean) — was investigated for three sets of parameters (1 natural and 2 normalized). Results obtained for a population of speakers from a closed set with size $M = 20$ (after 10 repetitions of the learning and test sequences) have proved that the Camberr distance function prevails in all types of parameters and algorithms. Other functions ensure a differentiated discrimination force strongly dependent on the algorithm and form of parameters. Limited usefulness of the square of Mahalonobis distance in comparison to other similarity functions was proved, as well as generally worse results for the *NM* algorithm.

Praca jest poświęcona problemowi doboru funkcji podobieństwa pomiędzy obrazami w nieparametrycznych algorytmach rozpoznawania głosów. Dla trzech zespołów parametrów (1 naturalnego i 2 znormalizowanych), pochodzących z ekstrakcji sygnału mowy hasła kluczowego, zbadano przydatność 10 funkcji podobieństwa (8 odległości i 2 bliskości) w trzech nieparametrycznych algorytmach identyfikacji: *NN* (najbliższy sąsiad), *k-NN* (k–najbliższych sąsiadów) oraz *NM* (najbliższa średnia). Uzyskane wyniki dla populacji mówców zbioru zamkniętego o liczebności $M = 20$ (po 10 powtórzeń ciągu uczącego się i testowego), wykazały zdecydowaną przewagę funkcji odległości Camberra we wszystkich rodzajach parametrów i algorytmów. Pozostałe funkcje zapewniają zróżnicowaną siłę dyskryminacyjną zależną mocno od algorytmu i postaci parametrów. Wykazano słabą przydatność kwadratu odległości Mahalanobisa w porównaniu z innymi funkcjami podobieństwa oraz ogólnie gorsze wyniki dla algorytmu *NM*.

1. Introduction

Computer recognition of voices includes several partial procedures which can be divided into three basic blocks:

a) source
b) measurement block
c) classification block

The sender of the signal (speaker) and the set of phenomena and conditions related with sending and registration of the speech signal is the source. The measurement block includes processing and analysis procedures of the input signal $u(t)$. This signal is the speaker's voice representation in acoustic images **x**. Quantity **x** usually denotes vectors from the space of parameters R^P (P – dimension of the space).

The classification block is a set of procedures or a procedure converting the input vector information **x** into a scalar m from the space of classes and decisions. Quantity m is the indicator of voices among which the system included the recognised signal. In parametric recognition algorithms with full (or estimated) probabilistic information the problem of finding the function of similarity of recognised voice's image and the standard is included in the classification algorithm [1]. "Voice standards" in the classical Bayes algorithm [1] are contained in a multidimensional distribution of conditional probability $p(x|m)$ (x–vector of individual's parameters, m–class speaker's number). The decision criterion is based on the minimum of average risk, which includes the loss matrix and the probability of appearance of images from the given class and of course the $p(x|m)$ distribution [1].

There is always a definite correlation between the space of parameters and accepted functions of similarity of the recognised image and standards in non-parametric recognition algorithms [1, 4, 5]. Frequently the simplification of the classification procedure in non-parametric agorithms leads to worse recognition results, because a similarity functions inadequate to the space of parameters is applied. The significance of this problem with regard to automatic speech recognition is among others confirned by TADEUSIEWICZ's paper [5] which present the usability evaluation of the similarity function (in the form of distance measures) in the recognition of vowel in the Polish language; and by the paper by Zalewski [6] who analysed the effectiveness of distance measures in the recognition of speakers with the application of linear predictive coding (LPC). Also BASZTURA [3] tired to check 8 chosen similarity functions as indices for speech transmission quality estimation. Because of frequent use of computer voice recognition in classification procedures it seems advisable to investigate a group of chosen similarity functions with regards to their effectiveness, using homogeneous experimental material.

2. Methods

To achieve a clear evaluation of the influence of investigated similarity functions on results of voice identification, the comparative procedure has to be free of all types of variability which influence the evaluation. At the same time it is advisable and necessary to check the "behaviour" of the similarity function in definite non-parametric classification algorithms. Considering this, the folowing assumption concerning methods made to systemize further experiments:

a) Voices of 20 speakers (men) aged 20–35 in a so-called closed set (i.e. recognised speakers will be included among the set of speakers in the learning sequences) were accepted as phonetic material.

b) In order to eliminate the influence of information which is not individual (linguistic and sociolinguistic) on identification results, a short-term analysis model with a fixed key-word for all statements was chosen. The maxim „Jutro będzie ładny dzień" ("tomorrow will be a fine day") was chosen as the key-word. It was used previously in paper [3] among others. The test series wās TS recorded 7 days after the recoding of the learning sequences.

c) Vectors with components x_p (V_p) which are numbers corresponding with the number of time intervals between zero-crossings of the speech signal were applied as individual parameters forming images of statements x and standards V [2]. Components x_p are calculated from:

$$x_p = x(t_{p-1}, t_p) = \begin{cases} x(t_{p-1}, t_p)+1 & \text{for} \quad t_j \in (t_{p-1}, t_p\rangle \\ x(t_{p-1}, t_p) & \text{for} \quad t_j \notin (t_{p-1}, t_p\rangle \end{cases} \tag{2.1}$$

where: t_p — boundary values of so-called time channes; $p = 1, 2, 3, \ldots, P$, P – number of time channels. It was accepted that $P = 7$, while t_p was chosen in accordance with the exponential division [2] from range $t \in (0.2$ ms–6.2 ms).

d) It was accepted that the usability of the similarity function will be evaluated for three most frequently applied heuristic classification algorithms, i.e. *NN* (Nearest Neighbour) *k–NN* (*k*–Nearest Neighbours) and *NM* (Nearest Mean). These algorithms have the following form:

NN algorithm

Image x belongs to class (voice) m, i.e. $x \to m$ if

$$FP(x, V_{m,i}) < FP(x, V_{l,i}) \tag{2.2}$$

where: $m = 1, 2, \ldots, M$; M – number of classes (voices), $- = 1, 2, \ldots, m-1, m+1, \ldots, M$, $V_{m,i} = X_{m,i}$–image of speaker's voice $i = 1, 2, \ldots, I_m$, I_m – number of repetitions of the statement in the learning sequences.

k–NN Algorithm

Functions of similarity (let us accept these as distances) between image x and all images in the learning series $x_{m,i}$ are calculated and ordered according to increasing order (decreasing order for nearness functions). Then first k distance values are considered and it is determined how many of them correspond with individual classes. If among k minimal distances there is $k_1, k_2, \ldots, k_m$ which belong to first, second, ... etc. class respectively, then values k_m are accepted as new similarity functions. Image x belongs to class m, i.e. $x \to m$ if

$$k_m > k_i \qquad l = 1, 2, 3, \ldots, m-1, m+1, \ldots, M \tag{2.3}$$

Value k is chosen in suitable proportion to the length of the learning sequences LS.

NM Algorithm

Most frequently the mean vector the voice (class) standard in the NM algorithm. The decision rule of the algorithm is as follows:

$$x \to m \quad \text{if} \quad FP(x, V_m) < FP(x, V_l) \tag{2.4}$$

where

$$V_m = x_m = \frac{1}{I_m} \sum_{i=1}^{I_m} x_{m,i} \tag{2.5}$$

m, l – as in expressions (2.2) and (2.3).

e) A set of ten similarity functions was chosen for investigation from among known similarity functions. Eight of them are distance functions, also called distance measures, while two are nearness functions. These functions corresponded with given below relationships (between x as the recognised image and V as the standard image). The first group of similarity functions can be noted with Minkowski's dependence:

$$d^{\mathrm{MIN}}(x, V) = \Big[\sum_{p=1}^{P} |x_p - V_p|^r\Big]^{1/r}; \; r \geq 1 \tag{2.6}$$

where: $p = 1, 2, \ldots, P$, x_p – p-th – element of vector x, V_p – p-th – element of vector V. For $r = 1$ d^{MIN} is known as Hamming's distance or street distance [6]. For $r = 2$ it is the Euclides metric. Two Minkowski's distances were additionally accepted for investigation, mamely d^{MIN3} ($r = 3$) and d^{MIN5} ($r = 5$).

Other similarity functions are as follows:

CHI–square distance

$$d^{\mathrm{CHI}}(x, V) = \sum_{p=1}^{P} \frac{1}{x_p + V_p} \left[\frac{X_p}{\sum_{p=1}^{P} x_p} - \frac{V_p}{\sum_{p=1}^{P} V_p} \right] \tag{2.7}$$

Czebyszew's distance

$$d^{\mathrm{CZE}}(x, V) = \max_p (x_p - V_p) \tag{2.8}$$

Camberr's distance

$$d^{\mathrm{CAM}}(x, V) = \sum_{p=1}^{P} \frac{x_p - V_p}{x_p + V_p} \tag{2.9}$$

square of Mahalanobis'es distance

$$d^{\mathrm{MAH}}(x, V) = (x-V)^{Tr} C^{-1}(x-V) \tag{2.10}$$

where C–mean covariance matrix (intraclass scatterings [1])

directional cos nearness function

$$b^{\cos}(x, V) = \frac{x \cdot V^{\mathrm{Tr}}}{|x||V|} \tag{2.11}$$

and

Tanimoto's nearness function

$$b^{\mathrm{TAN}}(x, V) = \frac{x \cdot V^{Tr}}{x \cdot x^{Tr} + V \cdot V^{Tr} - x \cdot V^{Tr}} \tag{2.12}$$

3. Identification experiment

An experiment of voice identification was carried out in compliance with paragraph 2. It was aimed at the determination of numerical relations and dependences for 3 identification algorithms and 10 similarity functions.

It was accepted that the learning series LS will consist of 200 statements of 20 speakers (20×10 repetitions), while the test series will also consist of 10 statements of every speaker. Numerical methods were used for parameter extraction ($P = 7$). Statements were recorded on professional equipment in a quiet room. The band of the signal was limited to the 75 − 4500 Hz range. The sampling frequency of the a/d converter was equal to $f_{pr} = 10\,000$ samples/s and the dynamics were described by a 10 bit word.

All experiments were repeated for all three forms in order to analyse the influence of the form of sets of parameters on the effectiveness of identification. The first set ($ZP1$) is a set of measurement parameters (Table 1). The second ($ZP2$) is a set of parameters with components normalized with respect to the value of their variability range (Table 2) (expression 2.1)).

Let

$$\Delta x_p^{sr} = \frac{1}{M} \sum_{m=1}^{M} (x_{m,p}^{\max} - x_{m,p}^{\min}) \tag{3.1}$$

be the mean variability range of the p − element, where

$$x_{m,p}^{\max} = \max_i \{x_{m,i,p}\} \tag{3.2}$$

and

$$x_{m,p}^{\min} = \min_i \{x_{m,i,p}\} \tag{3.3}$$

Table 1. Set of parameters–$ZP1$ (not normalized). An example of 20 repetitions for speaker 1.

Series	Repetition	Parameter no						
		1	2	3	4	5	6	7
LS	1	68	98	77	104	133	261	59
	2	75	110	76	92	127	213	63
	3	105	127	101	103	175	236	68
	4	65	87	75	97	127	267	73
	5	117	113	95	78	136	161	27
	6	121	216	144	88	120	200	42
	7	139	187	153	114	116	180	36
	8	127	195	146	105	117	175	39
	9	79	106	80	97	128	225	62
	10	126	182	137	95	154	245	46
TS	1	120	117	83	81	126	183	32
	2	109	140	116	130	187	226	101
	3	99	130	126	117	178	225	96
	4	87	125	91	97	174	250	56
	5	108	129	88	109	156	199	54
	6	96	136	94	110	154	201	53
	7	116	170	134	102	52	259	41
	8	122	201	137	77	106	185	37
	9	78	117	84	97	93	158	37
	10	84	120	69	103	95	155	45

We determine the maximal mean variation range

$$\Delta x^{sr\max} = \max_p \{\Delta x_p^{sr}\} \tag{3.4}$$

The regraduated p element is calculated from

$$x_p^{(ZP2)} = x_p \frac{\Delta x^{sr\max}}{\Delta x^{sr}} \tag{3.5}$$

The third set ($ZP3$) is a set of parameter with components normalized with respect to the variations range of their variances (Table 2)

$$x_p^{(ZP3)} = x_p \frac{\delta^{\max}}{\delta_p^{\max}} \tag{3.6}$$

where

$$\delta^{\max} = \max_p \{\delta_p\} \tag{3.7}$$

Table 2. Set of parameters–$ZP2$ normalized with respect to maximal range of parameter's variability. An example of 20 repetitions for speaker 2

Series	Repetition	Parameter no 1	2	3	4	5	6	7
LS	1	135	98	101	297	791	1178	630
	2	149	110	100	263	755	961	673
	3	209	127	133	294	1041	1065	727
	4	129	87	99	277	755	1205	289
	5	233	113	125	223	809	727	289
	6	241	216	189	251	714	902	449
	7	277	187	201	325	690	812	385
	8	253	195	192	300	696	790	417
	9	157	106	105	277	761	1015	663
	10	251	182	180	271	916	1106	492
TS	1	239	117	109	231	749	826	342
	2	217	140	153	371	1112	1020	1079
	3	197	130	166	334	1059	1015	1026
	4	173	125	120	277	1035	1128	598
	5	215	129	116	311	928	898	577
	6	191	136	124	314	916	907	566
	7	231	170	176	291	309	1169	438
	8	243	201	180	220	630	835	395
	9	155	117	110	277	553	713	395
	10	167	120	91	294	565	699	481

and

$$\delta_p = \frac{1}{M} \sum_{m=1}^{M} \delta_{m,p} \tag{3.8}$$

while $\delta_{m,p}$ – variance of p parameter of m speaker calculated on the basis of learning series *LS*.

4. Analysis of results and conclusions

The series of carried out identification experiments led to definite comparisons and analysis aimed at the usability evaluation of individual similarity functions in investigated non-parametric identification algorithms. Two additional sets of parameters ($ZP2$ and $ZP3$) resulting from normalizing transformations improved the results in terms of static likelihood. The following conclusions can be drawn from the set of results presented in Tables 4, 5, 6 and 7:

Table 3. Set of parameters–$ZP3$ normalized with respect to the variability range of parameters' variances. An example of 20 repetitions for speaker 3

Series	Repetition	Parameter no						
		1	2	3	4	5	6	7
LS	1	122	98	97	274	732	1094	597
	2	134	110	96	242	699	892	637
	3	188	127	127	271	963	989	688
	4	116	87	95	255	699	1119	738
	5	209	113	120	205	748	675	273
	6	216	216	182	232	660	838	425
	7	249	187	193	300	638	754	364
	8	227	195	184	276	644	733	395
	9	141	106	101	255	704	943	627
	10	225	182	173	250	847	1027	465
TS	1	215	117	105	213	693	767	324
	2	195	140	146	342	1029	947	1022
	3	177	130	159	308	979	943	971
	4	156	125	115	255	957	1047	566
	5	193	129	111	287	858	834	546
	6	172	136	119	290	847	842	536
	7	208	170	169	269	286	1085	415
	8	218	201	173	203	583	775	374
	9	140	117	106	255	512	662	374
	10	150	120	87	271	523	649	455

Table 4. Results of voice identification in % for sets of parameters–$ZP1$.

Algorithm	Measure	1	2	3	4	5	6	7	8	9	10
		d^{HAM}	d^{EUK}	d^{MIN3}	d^{MIN5}	d^{CHI}	d^{CZE}	d^{CAM}	d^{MAH}	d^{COS}	d^{TAN}
NN	S_r	91.5	90.5	90.5	88.5	92.0	87.0	95.5	91.0	93.0	91.5
	δ	12.3	12.8	12.8	15.0	15.8	17.5	12.8	14.1	10.8	12.7
	$S_{r\max}$	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$S_{r\min}$	60.0	60.0	60.0	50.0	40.0	40.0	50.0	60.0	60.0	60.0
k-NN	S_r	90.5	89.5	87.0	86.5	89.5	83.5	94.5	87.5	89.0	89.0
	δ	14.3	14.7	15.9	18.4	23.3	21.6	11.8	14.5	12.1	14.8
	$S_{r\max}$	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$S_{r\min}$	50.0	50.0	50.0	30.0	10.0	20.0	50.0	50.0	60.0	50.0
NM	S_r	88.0	84.5	81.0	81.0	89.5	81.0	96.0	80.5	87.0	85.0
	δ	19.0	21.4	22.2	24.0	23.1	24.3	12.7	24.8	15.9	21.4
	$S_{r\max}$	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$S_{r\min}$	30.0	30.0	30.0	30.0	20.0	30.0	50.0	10.0	40.0	30.0

Table 5. Results of voice identification in % for sets of parameters–$ZP2$

Algorithm	Measure	1	2	3	4	5	6	7	8	9	10
		d^{HAM}	d^{EUK}	d^{MIN3}	d^{MIN5}	d^{CHI}	d^{CZE}	d^{CAM}	d^{MAH}	d^{COS}	d^{TAN}
NN	s_r	95.0	97.0	97.0	95.5	94.0	93.5	95.5	91.0	93.0	97.0
	δ	14.0	9.2	9.2	10.0	14.3	13.1	12.8	14.1	14.5	9.8
	$s_{r\max}$	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$s_{r\min}$	50.0	70.0	70.0	70.0	40.0	50.0	50.0	60.0	40.0	60.0
k-NN	s_r	95.0	93.5	92.5	92.0	90.5	90.5	94.5	87.5	95.0	95.0
	δ	12.4	15.7	15.5	15.8	21.6	17.6	11.9	14.5	14.0	11.9
	$s_{r\max}$	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$s_{r\min}$	60.0	40.0	40.0	40.0	20.0	30.0	50.0	50.0	40.0	50.0
NM	s_r	94.5	95.0	95.0	94.5	90.5	93.0	96.0	80.5	93.0	95.5
	δ	12.3	11.5	11.0	11.0	23.1	10.6	12.7	24.8	15.9	8.9
	$s_{r\max}$	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$s_{r\min}$	60.0	60.0	60.0	60.0	20.0	60.0	50.0	10.0	40.0	70.0

Table 6. Results of voice identification in % for sets of parameters–$ZP3$

Algorithm	Measure	1	2	3	4	5	6	7	8	9	10
		d^{HAM}	d^{EUK}	d^{MIN3}	d^{MIN5}	d^{CHI}	d^{CZE}	d^{CAM}	d^{MAH}	d^{COS}	d^{TAN}
NN	s_r	95.5	97.0	97.0	95.5	93.5	94.5	95.5	91.0	93.0	96.0
	δ	12.3	9.2	9.2	9.4	15.0	11.9	12.8	14.1	14.2	11.9
	$s_{r\max}$	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$s_{r\min}$	60.0	70.0	70.0	70.0	40.0	50.0	50.0	60.0	40.0	50.0
k-NN	s_r	95.0	94.5	94.5	92.5	90.5	91.5	94.5	87.5	95.0	95.0
	δ	12.4	12.8	12.8	14.8	21.6	17.6	11.9	14.5	14.0	11.9
	$s_{r\max}$	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$s_{r\min}$	60.0	60.0	70.0	60.0	20.0	60.0	50.0	10.0	60.0	70.0
NM	s_r	94.5	95.0	95.0	94.5	90.5	93.0	96.0	80.5	93.0	95.5
	δ	12.3	11.5	11.0	11.0	23.1	10.6	12.7	24.8	15.9	8.9
	$s_{r\max}$	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	$s_{r\min}$	60.0	60.0	60.0	60.0	20.0	60.0	50.0	10.0	40.0	70.0

1) Positively best average results of correct identification for sets of natural parameters (not normalized) $ZP1$ come from Camberr's distance function (enc. 9). This is due to a somewhat normalizing form of this function. Differences with respect to other similarity functions are smallest for the NN algorithm (2.5%), and greatest for NM (6.5%). Tests of significance performed for differences of results between d^{CAM} and $b^{\cos}$, d^{CHI}, which give closest average results of correct identification (see Table 4), have indicated the significance of these differences on significance lovel $\alpha = 0.05$.

Tabele 7. Parameters of the voice identification experiment arranged according to decreasing values of correct decisions

N°	Set of parameters	FP	Algorithm	s_r [%]	σ [%]	$s_{r\max}$ [%]	$s_{r\max}$ [%]
1	*ZP*2	Euk	NN	97.0	9.2	100	70
2	*ZP*3	Euk	NN	97.0	9.2	100	70
3	*ZP*2	Min3	NN	97.0	9.2	100	70
4	*ZP*3	Min3	NN	97.0	9.2	100	70
5	*ZP*2	Tan	NN	97.0	9.8	100	60
6	*ZP*3	Tan	NN	96.0	11.9	100	50
7	*ZP*1	Cam	NM	96.0	12.7	100	50
8	*ZP*2	Cam	NM	96.0	12.7	100	50
9	*ZP*3	Cam	NM	96.0	12.7	100	50
10	*ZP*2	Tan	NM	95.5	8.9	100	70
11	*ZP*3	Tan	NM	95.5	8.9	100	70
12	*ZP*3	Min3	NM	95.5	9.4	100	70
13	*ZP*3	Min5	NN	95.5	9.4	100	70
14	*ZP*2	Min5	NN	95.5	10.0	100	70
15	*ZP*1	Cam	NN	95.5	10.0	100	50
16	*ZP*3	Ham	NN	95.5	12.3	100	60
17	*ZP*2	Cam	NN	95.5	12.8	100	50
18	*ZP*3	Cam	NN	95.5	12.8	100	50

2) For sets of parameters *ZP*1 all other similarity functions gave best identification results for the *NN* algorithm and worst for the *NM* algorithm. This is also confirmed by so-called minimal probabilities of correct identification $s_{r\min}$ for individual speakers. Their values decreased to 30, 20 and even 10% for Mahalanobis'es distance (Table 4).

3) The normalization of sets of parameters *ZP*2 and *ZP*3 resulted in an increase of voice identification correctness by several percent on the average for all algorithms except for Camberr's distance function (see point 1) which had exactly the same effectiveness as for *ZP*1.

4) Greatest differentiation of effectiveness occured for individual similarity functions in case of normalized parameters. Distance functions such as d^{EUK}, d^{MIN3} and neamess function b^{TAN} were distinguished, and the *NN* algorithm was distinguished as for *ZP*1.

5) Positively worst results (for the *NM* algorithm especially) were achieved with the square of Mahalanobis'es distance (enc. 10). This conclusion confirms results and

conclusions presented in TADEUSIEWICZ'es paper [5]; namely, that in certain cases better results can be reached with less complex similarity functions.

To recapitulate we can accept a general conclusion that it is advisable to use Camberr's distance function for natural parameters (directly from measurements). While it is sufficient to use Euklides'es distance function or Tanimoto's nearness function when parameters are normalized. The application of the square of Mahalanobis'es distance is not recommended, for short learning series especially. As to the evaluation of algorithms, the nearest mean *NM* algorithm achieves the positively lowest general rating. It is understandable that presented results can not (this concerns exact numerical values) be transferred directly for experiments with sets of parameters with different dimensions and structure. This finds confirmation in the differentiations achieved for *ZP*2 and *ZP*3. In order to achieve exact numerical values some experiments out of these presented above should be repeated at random at least.

References

[1] Cz. BASZTURA, *Sources, signals and acoustic images* (in Polish), WKiŁ, Warszawa 1988.

[2] Cz. BASZTURA, J. JURKIEWICZ, *Analysis of zero-crossings of a speech signal in a short-term model of automatic speaker identification* (in Polish) Arch. Akustyki **13**, 3, 203–214 (1978).

[3] Cz. BASZTURA, *Similarity functions of acoustic images as indicators of objective evaluation of speech quality transmission* (in Polish) Arch. Akustyki **22**, 3, 217–233 (1987).

[4] A. J. GRAY, J. D. MARKEL, *Distance measures for speech processing*, IEEE ASSP–24, **5**, 380–391 (1976).

[5] R. TADEUSIEWICZ, *Usability evaluation of chosen metric in minimal distance methods of speach recognition* (in Polish) Arch. Akustyki, **18**, 3, 275–284 (1983).

[6] J. ZALEWSKI, *A comparison of the effectiveness of some distances measures in speaker recongnition experiments.* Paper on the Speaker Recognition Working Group on the X-th International Congress of Phonetic Sciences, Utrecht 1983.

Received June 6, 1989

ARCHIVES OF ACOUSTICS
16, 2, 249–267 (1991)

FULLWAVE THEORY OF $\Delta v/v$ SAW WAVEGUIDES AND COUPLERS

D. GAFKA

Institute of Fundamental Technological Research Polish Academy of Sciences
(00-049 Warsaw Świętokrzyska 21)

This paper considers propagation of a surface acoustic wave (SAW) along a multiperiodic system of electrodes (note: a multiperiodic system is a system with a group of several equidistant electrodes with equal width occuring periodically with a certain period) distributed on the surface of a piezoelectric half-space. The boundary problem with homogeneous mechanical and mixed electric boundary conditions was solved on the basis of properties of the effective surface permittivity and Floquet's theorem for periodic structures. New functions satisfying adequate conditions in an assumed multiperiodic system were formulated. A dispersion relation for the velocity of a SAW guided along electrodes was derived.

The presented theory was applied in numerical analysis of SAW velocity dispersion in a $\Delta v/v$ single and two-electrode waveguide, assuming an adequately long repetition period of groups of electrodes. Velocities and field decay distributions for two modes – symmetric and asymmetric are given for a two-electrode waveguide. The coupling coefficient between both electrodes of a coupler constructed on the basis of such a waveguide was calculated.

1. Introduction

Most modern SAW devices apply a wide beam of surface waves. The application of such a beam involves certain undesirable effects: beam spreading which accompanies its propagation, difficulties with changes of direction of propagation and inefficient use of the piezoelectric substrates surface. The application of waveguides eliminates all these problems, because a previously excited SAW can be guided. However, there are certain difficulties with the general application of such solutions in SAW technology: high losses and ineffective excitation (small aperture of waveguides) [10, 20, 21, 24]. Nevertheless they are used mainly in constructions of:
– long delay lines, storing analog or digital signals [1].
– convolvers, performing nonlinear operations on signals e.g. Fourier transformation [11, 19],
– monolithic amplifiers on a semiconductor substrate [7, 8, 10]
– filters with high quality factor (a pair of coupled waveguides) [23].

* This work was partly supported from CPBP 01.08 and CPBP 02.02

There are many methods of guiding a wave within a certain separated region. The guiding region has lower velocity of wave propagation in relation to the rest of the surface [6, 20, 22, 24]. Deposition of a thin conducting film on the surface of the piezoelectric substrate is one of the simplest methods in the SAW technology. A local electric field shorting constitutes a SAW guide, because as we know [10, 18] SAW velocity under a shorted surface, V_0, is lower than SAW velocity on a free (adjoining vacuum) surface V_v.

These types of guides (in a certain specific configuration) called piezoelectric waveguides or $\Delta v/v$ waveguides, will be the subject of further fullwave analysis. The method of analysis applied in this paper was previously used in investigations of perpendicular SAW propagation in relation to the system of electrodes [2, 4, 5, 10, 12] and in the analysis of the guidance of a wave in a periodic electrode system (including the transition to the "rare" system, i.e. single $\Delta v/v$ waveguide) [10]. This paper is a generalization of the theory presented in [10, Chapter 3] for the case of a wave guided in a multiperiodic electrode system, i.e., system of electrodes where not one but a group of N electrodes repeats periodically (Fig. 1). In such a case the transition to the "rare" system leads to a desciption of a finite number of coupled waveguides. A two-electrode guide is discussed in detail.

The following part of this paper presents the boundary problem, simplifying assumptions and changes in the formulation of the problem suitable for the accepted method of analysis [2, 4, 5, 10]. The construction of the problem's solution on the basis of new special functions, described in the Appendix, is given in Section 3. Section 4 presents results of numerical analysis of a one- and two-electrode SAW guide.

2. Formulation of the problem

We are looking for the field of a surface acoustic wave propagating along a multiperiodical system of metal electrodes with width w, distributed on the surface of a piezoelectric half-space. A group of a definite number (N) of electrodes repeats periodically with period Λ. The distance between electrodes within a group is equal to $2p$. The structure is unlimited. The cross section of the analysed system is shown in Fig. 1. The structure is homogeneous in the direction perpendicular to the cross section (x_1).

From the mathematical point of view the problem can be reduced to the solution of a set of partial differential equations with homogeneous mechanical and

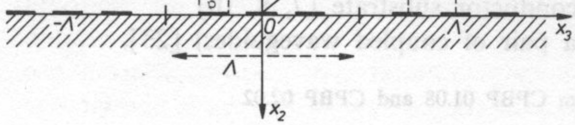


FIG. 1. Analysed system of electrodes (here: $N = 3$)

mixed electrical boundary conditions. The set of differential equations describing coupled mechanical-electrical wave processes in a linear and homogeneous medium has the following form for a case of electrostatic approximation [18] and for a solution harmonic in time:

$$\begin{cases} c_{ijkl}U_{k,jl}+e_{kij}\phi_{,jk}+\omega^2 dU_i=0, & i,j,k,l=1,2,3 \\ e_{ikl}U_{k,li}-\varepsilon_{ij}\phi_{,ji}=0, & d \text{ — mass density of the substrate} \end{cases} \tag{1}$$

where the vector of unknown quantities $\mathbf{U}=[U_1, U_2, U_3, \phi]$ consists of components of the displacement vector $\mathbf{U}$ and electric potential ($\mathbf{E}=-\nabla\phi$), c is the elastic tensor of the substrate e is the tensor of piezoelectric coefficients and ε is the permittivity tensor. The set of equations (1) has to be satisfied in the top ($x_2<0$) and bottom ($x_2>0$) half-space separately. Boundary conditions on the boundary of media ($x_2=0$) are described by the following equations:

a) mechanical

$$T_{2j}=c^{2jkl}U_{k,l}+e_{k2j}\phi_{,k}=0 \quad \text{for } x_2=0 \tag{2}$$

b) electrical

$$\phi=0 \quad \text{on the surface of electrodes}$$

$$\Delta D_2=D_2^+-D_2^-=0 \quad \text{between electrodes} \tag{3}$$

The following additional idealizing assumptions were accepted:

– the metallization is a perfect conductor
– the thickness of the metallization is infinitely small
– metal electrodes do not load the surface of the piezoelectric otherwise condition (2) would not be fullfilled.

The problem in such a form is a three-dimensional problem. The method proposed in [4, 5] and developed in [2, 10, 12, 13, 14] was expanded in order to solve this problem. The mentioned method allows the problem to be reduced to a pseudoelectrostatic boundary problem on the boundary surface $x_2=0$ and its algebraization.

The expansion consist in the introduction of the effective surface permittivity [10, 18] which implicite contains the fulfilment of wave mechanical properties of SAW (1) and of the mechanical boundary condition (2). Therefore, the set of mixed electrical boundary conditions on the surface remains to be included.

The effective surface permittivity in the problem under consideration sufficiently characterizes the substrate. It is determined by the $\Delta D_\perp/E_{||}$ relation where $\Delta D_\perp$ is the electric charge density on the substrate surface (it can be the charge on electrodes located on the surface), equal to the difference between the component of the electric displacement vector perpendicular to the surface $x_2=0$ considered from the substrate side ($x_2=0^+$) and from the vacuum side ($x_2=0^-$); $E_{||}$ is the electric strength field on the substrate surface. In further text $\Delta D_\perp$ and $E_\parallel$ denote complex amplitudes of equivalent harmonic waves:

$$\exp j(\omega t - k\cdot x) \tag{4}$$

where ω – angular wave frequency assumed in the paper as definite; factor $\exp(j\omega t)$ will be neglected further on, k – wave vector which assumes values from k_v on the free surface of the piezoelectric adjoining vacuum to k_0 on the metallized surface. In the considered case we assume that k_0 and k_v are real numbers and $k_0 \geqslant k_v$, $k_0 = \omega/V_0$, $k = \omega/V_v$, where V_0 and V_v are corresponding SAW velocities and the generally used in the SAW theory $\Delta v/v$ factor is defined as $\Delta v/v = (V_v - V_0)/V_v$.

In an approximation resulting from the neglect of acoustic bulk waves the relation between amplitudes $\Delta D_\perp$ and $E_\parallel$ of harmonic waves in form (4) is expressed by [10, 18] (for $k > 0$)

$$\varepsilon(k) = -j\frac{\Delta D_\perp}{E_\parallel} = \varepsilon_0 \varepsilon_{ef}\frac{k^2 - k_v^2}{k^2 - k_0^2}, \tag{5}$$

where $k = |\mathbf{k}| = \sqrt{k_1^2 + k_3^2}$, ε_0 – permittivity of the vacuum.

As we know, a wave propagating in a periodicaly non-homogenous system can be described with the sum of harmonic components related to the period of the system, Λ. Consistently, in accordance to Floquet's theorem electric field $E_\parallel$ and surface charge density $\Delta D_\perp$ on a the piezoelectric surface $x_1 x_3$ with a periodic electrode system have the folowing form

$$\mathbf{E}_\parallel = \sum_{n=-\infty}^{\infty} \mathbf{E}_n \exp(-j(\mathbf{k} + n\mathbf{K})\cdot\mathbf{x}) \tag{6a}$$

$$\Delta D_\perp = \sum_{n=-\infty}^{\infty} D_n \exp(-j(\mathbf{k} + n\mathbf{K})\cdot\mathbf{x}) \tag{6b}$$

where $\mathbf{x} = [x_1, x_3]$, $\mathbf{K}$ – wave vector for considered system of electrodes with one component $\mathbf{K} = [0, K]$, $K = 2\pi/\Lambda$; and $\mathbf{k} = [k_1, k_3]$ is the wave vector of SAW propagation, e.g. of guiding SAW along electrodes $k_3 = 0$ (Fig. 1) in the considered case.

It can be easily checked that even when a wave propagates along electrodes, spatial harmonics into which the wave was decomposed have generally slant directions to the electrode system what, is expressed by the following

$$\mathbf{E}_\parallel = \sum_{n=-\infty}^{\infty} E_n \mathbf{e}_n \exp(-j\varkappa_n \xi_n)$$

$$\Delta D_\perp = \sum_{n=-\infty}^{\infty} D_n \exp(-j\varkappa_n \xi_n), \tag{7}$$

where $\varkappa_n = |\mathbf{k} + n\mathbf{K}| = \sqrt{k_1^2 + (nK)^2}$, ξ_n coordinate measured along the wave vector

$\varkappa_n$, in such a manner that

$$\varkappa_n \xi_n = (\mathbf{k} + n\mathbf{K}) \cdot \mathbf{x} \quad \text{(see Fig. 1)} \tag{8}$$

and $\mathbf{e}_n$ is a versor in the direction of the wave vector $\mathbf{e}_n = (\mathbf{k} + n\mathbf{K})/\varkappa_n$. Equivalent components of the potential have the following form:

$$\phi_n \exp(-j\varkappa_n \xi_n), \quad \text{where} \quad \phi_n = -jE_n/\varkappa_n, \tag{9}$$

whereas amplitudes E_n and D_n are related to each other when the effective surface permittivity is introduced for every n separately, as follows

$$D_n = j\varepsilon_n E_n, \tag{10}$$

where

$$\varepsilon_n = \varepsilon(\varkappa_n). \tag{11}$$

The boundary problem presented at the beginning can now be formulated as follows:

find amplitudes of the electric field E_n and wave number k_1 for given frequency ω, so that:

– electric field $E_{\parallel}$ is zero on electrodes

$$\mathbf{E} = 0 \text{ on electrodes} \tag{12}$$

– distribution of surface charge density $\Delta D_{\perp}$ vanishes between electrodes

$$\Delta D_{\perp} = 0 \text{ between electrodes.} \tag{13}$$

3. Dispersion equations

We obtain the following conditions from (6a) for component e_3 and from (6b)

$$\bar{\mathbf{E}} = \sum_{n=-\infty}^{\infty} \frac{nK}{\varkappa_n} \cdot E_n e^{-jnKx_3} = 0 \quad \text{on electrodes}$$

$$\bar{\mathbf{D}} = \sum_{n=-\infty}^{\infty} \varepsilon_n E_n e^{-jnKx_3} = 0 \quad \text{between electrodes} \tag{14}$$

We should notice that $\varepsilon_n \to \varepsilon_\infty = \varepsilon_0 \varepsilon_{ef}$ for $|n| \to \infty$ what more $\varepsilon_n \cong \varepsilon_\infty$ is fulfilled beginning from even small n for wave numbers K not too small in comparison to $k_{0,v}$. Similarly, $nK/\varkappa_n \to \pm 1$ for $n \to \pm\infty$.

Conditions (14) ensure vanishing on the electrodes of one component of the electric field-component E_3 perpendicular to electrodes. These conditions should be suplemented with an additional relationship, namely vanishing condition for the second component of the field along the electrodes in an arbitrary point of the electrodes. It is convenient to relate this condition to the axis of electrodes [10]. In accordance with (6a) we obtain

$$\sum_{n=-\infty}^{\infty} \frac{k_1}{\varkappa_n} \cdot E_n e^{-jnK(2l-N-1)p} = 0, \quad l = 1, 2, \ldots, N. \tag{15}$$

There are as many conditions as there are electrodes in a group.

The set of equations (14) will be solved jointly in the first place, and afterwards condition (15) will be satisfied.

As we know [2, 3], the set of two functions

$$g(\theta) = \sum_{n=-\infty}^{\infty} S_n P_n(\cos\Delta) e^{-jn\theta}$$
$$f(\theta) = \sum_{n=-\infty}^{\infty} P_n(\cos\Delta) e^{-jn\theta} \tag{16}$$

where $\Delta = Kw/2$, $\theta = Kx_3$, $S_n = n/|n|$ is the function of sign n and P_n are Legendre's polynomials, satisfies the following conditions

$$g(\theta) = \begin{cases} 0, & 0 \le |\theta| \le \Delta \\ -j\dfrac{\theta}{|\theta|} 2^{1/2} e^{j\theta/2} (\cos\Delta - \cos\theta)^{-1/2}, & \Delta < |\theta| \le \pi \end{cases}$$
$$f(\theta) = \begin{cases} 2^{1/2} e^{j\theta/2} (\cos\theta - \cos\Delta)^{-1/2}, & 0 \le |\theta| \le \Delta \\ 0, & \Delta < |\theta| \le \pi \end{cases} \tag{17}$$

Indeed, functions $g(\theta)$ and $f(\theta)$ are solutions to a corresponding electrostatic problem for a dielectric (i.e. for $k_0 = k_v$ in expression (5)). However, we have $k_0 \cong k_v$ for piezoelectrics and expression (5) differs significantly from a similar expression for a dielectric in a narrow range of wave numbers $k \in [k_v, k_0]$. This has been noticed and utilized in [2, 4, 5, 10] to construct a solution to the boundary problem (12), (13) for a piezoelectric substrate ($k_0 \neq k_v$) with a periodic system of electrodes.

Following a similar procedure as in [2, 12, 13, 14], functions which are analogical to (16) solutions to an electrostatic problem for the considered system of groups of electrodes from Fig. 1 are introduced

$$G_N(\theta; \alpha) = \sum_{n=-\infty}^{\infty} S_n X_n^N e^{-jn\theta}$$
$$F_N(\theta; \alpha) = \sum_{n=-\infty}^{\infty} X_n^N e^{-jn\theta}, \tag{18}$$

where parameter related to the distance between electrodes in a group $\alpha = Kp$, while conditions

$$G_N(\theta;\alpha)=0 \quad \text{on electrodes}$$
$$F_N(\theta;\alpha)=0 \quad \text{between electrodes} \tag{19}$$

are satisfied. Forms and properties of these functions are given in the Appendix.

Applying method [2, 10] quantities $\bar{E}$ and $\bar{D}$ were expressed by functions X_m as follows

$$\bar{E}=\sum_{m=-\infty}^{\infty}\alpha_m G_N(Kx_3;Kp)e^{-jmKx_3}$$
$$\bar{D}=j\sum_{m=-\infty}^{\infty}\beta_m F_N(Kx_3;Kp)e^{-jmKx_3} \tag{20}$$

In this case boundary conditions (14) are reduced to:

$$\frac{nK}{\varkappa_n}E_n=\sum_{m=M_1}^{M_2}\alpha_m S_{n-m}X_{n-m}^N$$
$$\varepsilon_n E_n=\sum_{m=M_1}^{M_2}\beta_m X_{n-m}^N, \tag{21}$$

where infinite summation according to m is substituted with summation with in the range $M_1\le m\le M_2$. This corresponds to the cut-off of an equivalent infinite set of equations, what in this case is possible [2, 10, 14], because $nK/\varkappa_n$ as well as ε_n quickly achieve constant values for $n\to\pm\infty$. α_m are determined from a comparison between left and right sides of (21) for $N_1\le n\le N_2$ (we will discuss the choice of such limits below; but it is generally known that all terms for which $nK/\varkappa_n$ as well as ε_n differ significantly from their fixed values for $n\to\pm\infty$have to be taken into account). Expressions (21) should have the form of a compatible set of equations for $n<N_1$ and $n>N_2$ for α_m calculated as described above. This determinates the choice of M_1 and M_2, depending on N_1 and N_2. Additionally [2, 10, 14] the compatibility condition of this set of equations imposes the method of selecting N_1 and N_2 namely for $n<N_1$ and $n>N_2$ terms on the left side in equation (21) have to have practically constant values (approximately equal to values for $|n|\to\infty$).

We should notice that the number of variables α_m within range $M_1\le m\le M_2$ has to exceed the number N_2-N_1+1 (i.e. number of equations resulting from a comparison between the right and left side of (21), because N more equations (15) resulting from field vanishing along the axis of electrodes remain to be satisfied.

Let us recapitulate. A comparison between the pair of equations (21) for $N_1\le n\le N_2$ results in the following expression

$$\sum_{m=M_1}^{M_2}\left(\frac{nK}{\varkappa_n}\cdot\varepsilon_{ef}-S_{n-m}\varepsilon_n\right)\alpha_m X_{n-m}^N=0, \tag{22}$$

which has to be fulfilled for every $n \in [N_1, N_2]$, $N_1 < 0$, $N_2 > 0$.

Additionally from (15) and (21) we obtain

$$\sum_{m=M_2}^{M_2} \alpha_m \left(\sum_{n \neq 0} \frac{S_{n-m} X_{n-m}^N}{nK} e^{-jnK(2l-N-1)p} + \frac{\varepsilon_{ef}}{\varepsilon_0} X_{-m}^N \right) = 0, \tag{23}$$

where $l \in [1, N]$.

M_1 and M_2 have to be chosen after N_1 and N_2 are determined. They are chosen as shown below in order to secure an equal number of unknown quantities and equations

$$M_1 = N_1 - \mathrm{int}(N/2) \quad (N_1 < 0)$$
$$M_2 = N_2 + \mathrm{int}((N+1)/2). \tag{24}$$

The total number of equations and unknowns is equal to $N_2 - N_1 + N + 1$ in such a case. The condition for the existance of a solution to the set of equations (22) and (23), i.e. when the equivalent characteristic determinant $\Delta(k_1)$ is equal to zero, gives the sought for dispersion relation

$$\Delta(k) = 0 \tag{25}$$

where k is the value of wave number k_1 which determines the guiding velocity of SAW in the waveguide under consideration

$$V = \omega/k. \tag{26}$$

4. Numerical results

This Section presents numerical results of the analysis of SAW guidance along a system of electrodes deposited on a $LiNbO_3$ YZ. A propagation constant k was sought which would fulfill the following condition

$$k_v \le k \le k_0 \tag{27}$$

There are many solutions for the wave number in a periodic system, but it was accepted that k is a solution of an equivalent dispersion equation (25) from Brillouin's I zone [10]. K is large enough (Λ small) in numerically analysed structures that there is only one n which fulfills condition (27).

Presented below results have been divided into two groups according to the size of the parameters:

a) the distance between groups of electrodes greatly exceeds the distance between electrodes with in a group and the width of electrodes

$$\Lambda \gg w, p$$

b) mentioned above distances are comparable

$$\Lambda \cong w, p$$

(a) type cases have an easy physical interpretation and practical application. Condition $\Lambda \gg w, p$ means that approximately there is no interaction between electrodes belonging to neighbouring groups of electrodes. Therefore, results obtained for such parameters can be accepted as propagation constants of waveguides with SAW ($\Delta v/v$ waveguide for $N = 1$, pair of coupled waveguides – SAW waveguide coupler for $N = 2$). It should be noted that a method of the fullwave analysis of an isolated waveguide with SAW does not exist; whereas when we assume $\Lambda \gg w, p$, then such an analysis is possible with certain approximation within the framework of the theory of periodic systems.

(b) type cases do not have a simple physical interpretation, because it is a problem of guiding a wave along a coupled infinite number of groups of electrodes. However, chosen relationships have been described here in order to make this elaboration complete.

a) SAW waveguides

The case of a one-electrode insulated approximately waveguide has been already analysed in papers [9, 10, 17, 23, 24]. A system of two coupled electrodes is considered below. The width of electrodes was accepted at $w = 0.015$ mm and repetition period of the pair of electrodes was accepted at $\Lambda = 3$ mm, considered as sufficiently large. Figure 2 presents the $1/V\,(w)$ dependence (V wave propagation velocity, $V = \omega/k$) in enlarged scale for $\Lambda = 3$ and 6 mm, and $N = 1$ in order to ilustrate the influence of neighbouring groups of electrodes.

When the distance between neighbouring electrodes is doubled the velocity of a wave propagating in the structure changes by about 0.09%. The distance of 3 mm between electrodes with width equal to $w = 0.015$ mm was considered as sufficiently

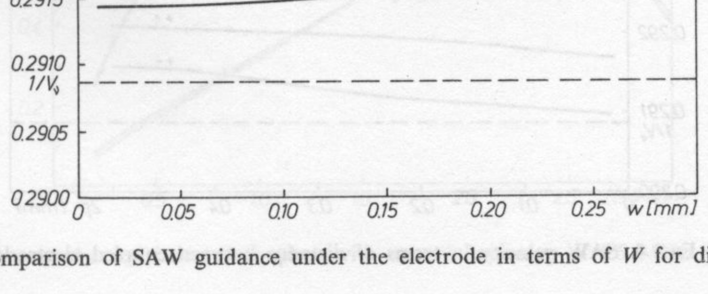


FIG. 2. Comparison of SAW guidance under the electrode in terms of W for different Λ.

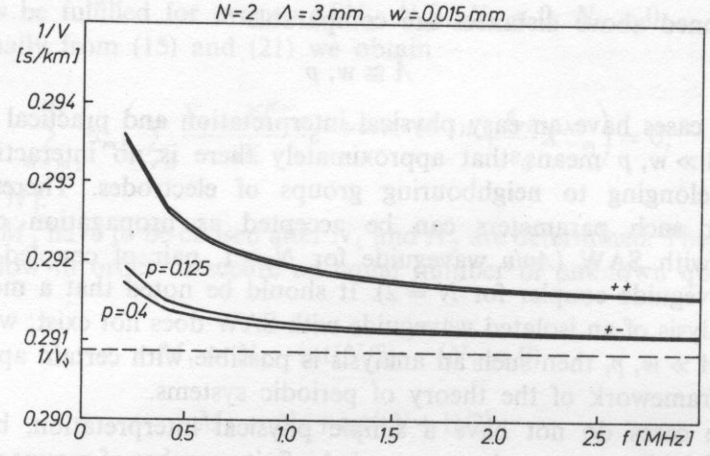


FIG. 3. SAW velocity dispersion in a two-electrode guide for two modes

to satisfy the condition of insulation of the waveguide from neighbouring electrodes.

In the case of a two-electrode waveguide, there are two, orthogonal with respect to each other modes with different SAW propagation velocities. Figure 3 presents dispersion characteristics $1/V(f)$ for both modes even, symmetric $++$ and odd, antisymmetric $+-$ for $p = 0.04$ and 0.125 mm. Figure 4 illustrates the influence of the distance between electrodes $2p$ on $1/V$ changes for $f = 1$ MHz. The odd mode for $2p$ close to w should be cutt-off just as the first higher mode with odd charge distribution in a one-electrode waveguide [17]. The curve in Fig. 4 only approaches the cut-off value $1/V_v$. In order to closely investigate the characteristic near the cut-off, the complex value of the SAW wave number with damping should be permitted.

Taking advantage of the difference of wave velocity a SAW coupler can be

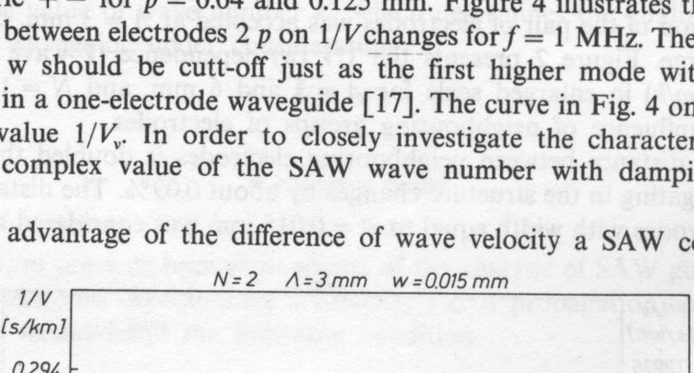


FIG. 4. SAW velocity in terms of distance between coupled electrodes

constructed, because of the existance of two modes with different velocities in a two-electrode waveguide. Power transmission from one port to the other of an ideal coupler is expressed by:

$$\tau = \sin^2\left(\frac{1}{2}\cdot(k_e - k_o)Z\right) \tag{28}$$

where τ — power transmission factor, k_e — value of even mode wave vector, k_o — value of odd mode wave vector, l — length of coupler.

Figure 5 presents the required length of couplers 3 dB and 10 dB in terms of

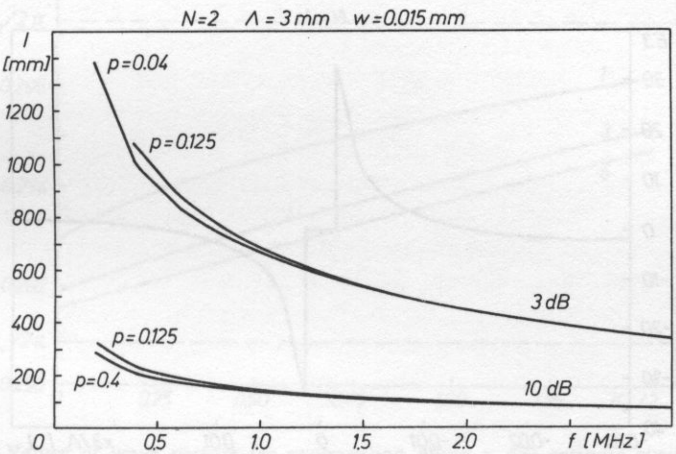


FIG. 5. Required lengths of SAW couplers — 3 dB and 10 dB for various frequencies

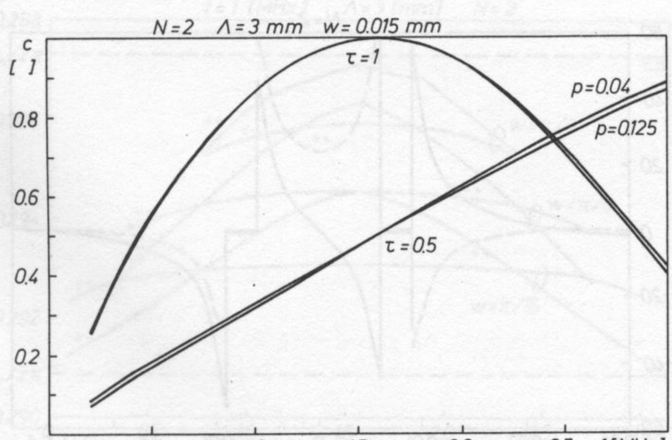


FIG. 6. Frequency characteristics of couplers with central frequency of 1.6 MHz ($\tau = 0.5,\ 1$)

frequency for $p = 0.04$ and $p = 0.125$ mm. For comparison the wavelength of a SAW is equal to about 3.4 mm for $f = 1$ MHz. A frequency coupling characteristic of couplers designed for frequency $f = 1.6$ MHz are shown in Fig. 6 for $\tau = 1$ full power transmission from one port to the other and $\tau = 0.5$ (3 dB coupler).

To illustrate the decay rate of the electric field around the electrodes, E_3 in terms of x_3/Λ was plotted for the following systems:

Fig. 7 $N = 1$, $w = 0.015$ mm, $\Lambda = 3$ mm

Fig. 8 $N = 2$, $w = 0.015$ mm, $p = 0.04$ mm

$\Lambda = 3$ mm both modes

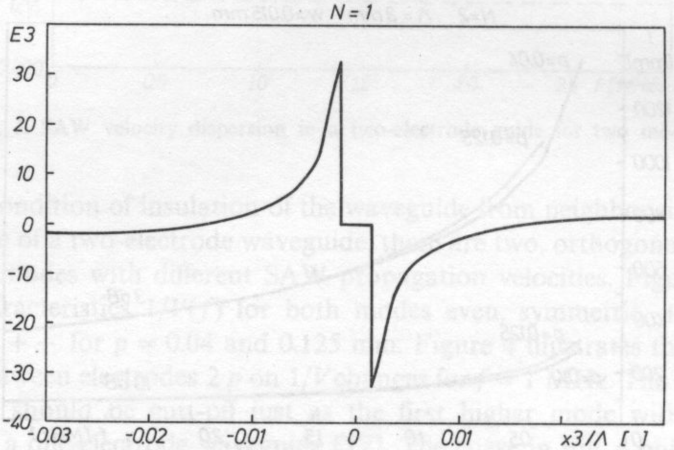


FIG. 7. Electric field decay of SAW guided under the electrode

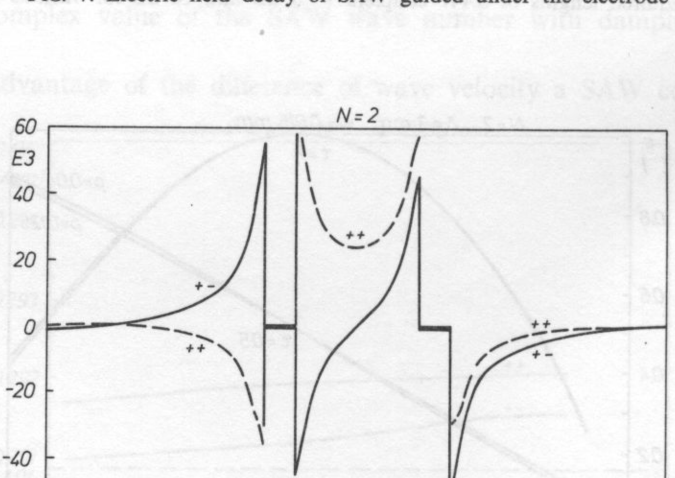


FIG. 8. Electric field decay of SAW guided under coupled electrodes

b) SAW propagation along a multiperiodic electrode system

Two more dependences are presented to make the elaboration complete:

Fig. 9 $k/2\pi$ $(Kw/2)$ for $N = 1$, $f = 1$ MHz, $\Lambda = 1$, 3, 6 mm

Fig. 10 $k/2\pi$ (Kp) for $N = 2$, $f = 1$ MHz, $\Lambda = 3$ mm, $w = \Lambda/4$, $\Lambda/8$

When the metallized surface is increased the value of the wave vector approximates k_0. Curves in Fig. 10 are symmetrical in respect to $Kp = \pi/2$, because

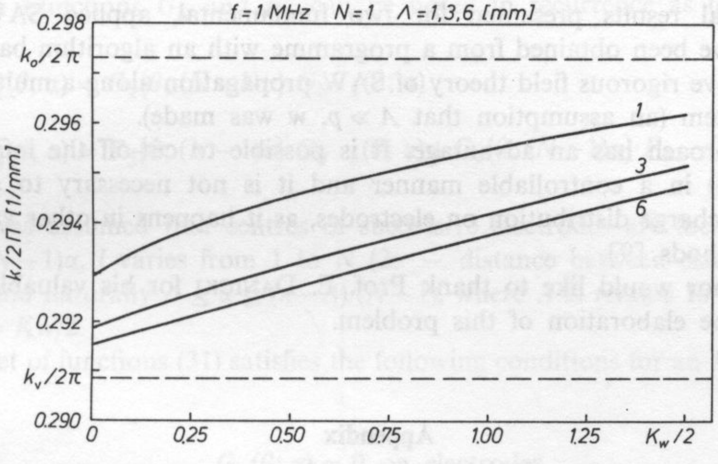


FIG. 9. Values of wave vectors for propagation along a one-periodic structure

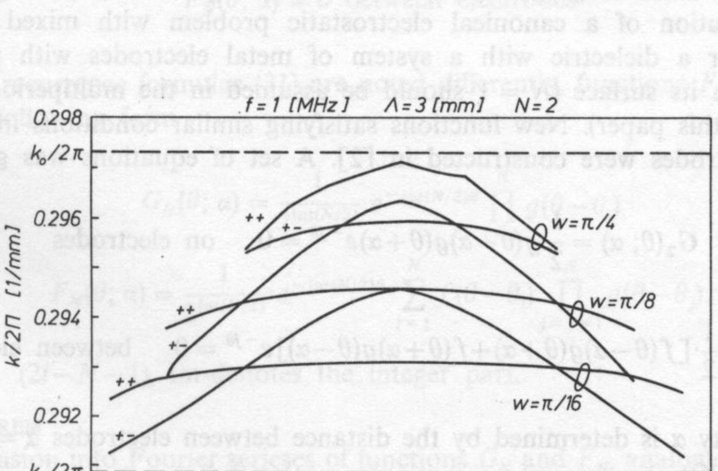


FIG. 10. Values of wave vectors for propagation along a two-periodic structure

for greater p electrodes interact with outer electrodes from the neighbouring group, i.e. $k(Kp) = k(\pi - Kp)$. As it can be seen for a case of SAW propagation along a multiperiodic electrode system there is a strong interaction between all electrodes for such parameters. There are p values for which velocities of both types i.e.: $++$ $++$ $++$ etc. and $+-$ $+-$ $+-$ etc. are equal.

5. Final remarks

Numerical results presented for two fundamental, applied SAW guiding structures have been obtained from a programme with an algorithm based on the discussed above rigorous field theory of SAW propagation along a multiperiodical electrode system (an assumption that $\Lambda \gg p$, w was made).

This approach has an advantage. It is possible to cut-off the infinite set of equations (14) in a controllable manner and it is not necessary to assume an approximate charge distribution on electrodes, as it happens in other variation or Galerkin methods [9].

The author would like to thank Prof. E. DANICKI for his valuable remarks. concerning the elaboration of this problem.

Appendix

The set of functions $g(\theta)$ and $f(\theta)$, described with equations (16) and given in [3, 16] is a solution of a canonical electrostatic problem with mixed boundary conditions for a dielectric with a system of metal electrodes with period 2π, distributed on its surface ($N = 1$ should be assumed in the multiperiodic system presented in this paper). New functions satisfying similar conditions in a case of separate electrodes were constructed in [2]. A set of equations was given there

$$G_2(\theta;\alpha) = \frac{1}{2}\cdot g(\theta-\alpha)g(\theta+\alpha)e^{-j\theta} = 0 \quad \text{on electrodes}$$

$$F_2(\theta;\alpha) = \frac{1}{2}\cdot[f(\theta-\alpha)g(\theta+\alpha)+f(\theta+\alpha)g(\theta-\alpha)]e^{-j\theta} = 0 \quad \text{between electrodes} \tag{29}$$

where quantity α is determined by the distance between electrodes $\alpha = Kp$, while variable $\theta = Kx_3$.

Their expansions into Fourier series in slightly changed notation are as follows:

$$G_2(\theta;\alpha) = \sum_{n=-\infty}^{\infty} X_n^2 S_n e^{-jn\theta} \qquad F_2(\theta;\alpha) = \sum_{n=-\infty}^{\infty} X_n^2 e^{-jn\theta}, \tag{30}$$

where

$$X_0^2 = 0; \quad X_n^2 = \sum_{m=0}^{n-1} P_m(\cos\Delta)\cdot P_{n-m-1}(\cos\Delta)\cdot\cos(2m-n+1)\alpha$$

$$X_{-n}^2 = -X_n^2, \quad \pm\alpha = \pm Kp \text{ are centres of electrodes.}$$

Taking advantage of results of (29) and (30) a generalized set of functions was constructed below. It fulfills analogic mixed boundary conditions in a system with a group of an arbitrary number (N) of electrodes repeated with a 2π period (in variable θ). Functions G_N and F_N can be noted in recurrence as follows

$$G_N(\theta;\alpha) = G_2(\theta;(N-1)\alpha)\cdot G_{N-2}(\theta;\alpha)$$
$$F_N(\theta;\alpha) = F_2(\theta;(N-1)\alpha)\cdot G_{N-2}(\theta;\alpha) + G_2(\theta;(N-1)\alpha)\cdot F_{N-2}(\theta;\alpha), \tag{31}$$

where it was assumed that centres of successive electrodes are located at points $\theta_l = (2l-N-1)\alpha$, l varies from 1 to N (2α — distance between electrodes within a group) and naturally $\Delta \le \alpha \le (\pi-\Delta)/(N-1)$, where Δ is related to the electrodes width $\Delta = Kw/2$.

The set of functions (31) satisfies the following conditions for an N — electrode group:

$$G_N(\theta;\alpha) = 0 \text{ on electrodes}$$
$$F_N(\theta;\alpha) = 0 \text{ between electrodes} \tag{32}$$

After recurrence formulas (31) are noted differently, functions F_N and G_N will have the following form

$$G_N(\theta;\alpha) = \frac{1}{2^{\mathrm{int}(N/2)}}\cdot e^{-\mathrm{int}(N/2)\theta}\prod_{l=1}^{N} g(\theta-\theta_l)$$
$$F_N(\theta;\alpha) = \frac{1}{2^{\mathrm{int}(N/2)}}\cdot e^{-\mathrm{int}(N/2)\theta}\cdot\sum_{l=1}^{N} f(\theta-\theta_l)\prod_{j=1\neq l}^{N} g(\theta-\theta_j), \tag{33}$$

where $\theta_i = (2i-N-1)$, int-denotes the integer part.

THEOREM

Expansion into Fourier serieses of functions G_N and F_N, analogical to (16) and (30), are as follows

$$G_N(\theta;\alpha) = \sum_{n=-\infty}^{\infty} S_n X_n^N e^{-jn\theta} \qquad F_N(\theta;\alpha) = \sum_{n=-\infty}^{\infty} X_n^N e^{-jn\theta}, \tag{34}$$

while for $n \geq 0$

$$\begin{cases} X_n^1 = P_n(\cos\Delta) \\ X_n^2 = \sum_{m=0}^{n-1} P_m P_{n-m-1} \cos[(2m-n+1)\alpha];\ X_0^2 = 0 \\ \dots\dots\dots\dots\dots\dots\dots\dots \\ X_n^N = 2\sum_{m=1}^{n} X^2_{m|(N-1)\alpha} X_{n-m}^{N-2} \end{cases} \tag{35}$$

where $X_m^2|_\beta$ means that X_m^2 are expansion coefficients of function $F_2(\theta;\beta)$

For $n \leq -1$ the following symmetry relations are fulfilled

$$\begin{aligned} X_{-n-1}^N &= X_n^N \quad \text{for odd } N \\ X_{-n}^N &= -X_n^N \quad \text{for even } N \end{aligned} \tag{36}$$

Relationships (35) and (36) require an explanation. Below they are proved on the basis of principles of mathematical induction.

PROOF.
1) For $N = 1$, (35) is trivial
2) For $N = 2$, (35) was proved in [2], different notation.
3) Assuming that the expansion for $N = 2$ (30) and for $N-2$ is given

$$G_{N-2}(\theta;\alpha) = \sum_{n=-\infty}^{\infty} X_n^{N-2} S_n e^{-jn\theta} \quad F_{N-2}(\theta;\alpha) = \sum_{n=-\infty}^{\infty} X_n^{N-2} e^{-jn\theta} \tag{37}$$

while X_n^{N-2} fulfills relationships (35–36).
4) It should be proved that coefficients X_n^N in expansions (34) fulfill the last equation (35) and relation (36). Taking advantage of (31) and substituting (30) and (37), for G_N as first, we obtain

$$G_N = \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} S_n S_m X_m^2 X_n^{N-2} e^{-j(n+m)\theta} = \sum_{n=-\infty}^{\infty}\Big[\sum_{m=-\infty}^{\infty} S_{n-m} S_m X_m^2 X_{n-m}^{N-2}\Big] e^{-jn\theta} \tag{38}$$

Comparing (34) with (38) we have

$$S_n X_n^N = \sum_{m=-\infty}^{\infty} S_{n-m} S_m X_m^2 X_{n-m}^{N-2}, \tag{39}$$

what for different n can be noted as

$$\begin{aligned} S_n X_n^N &= -C_n + 2\sum_{m=0}^{n} X_m^2 X_{n-m}^{N-2} && \text{for} \quad n \geq 0 \\ S_n X_n^N &= -C_n && \text{for} \quad n = -1 \\ S_n X_n^N &= -C_n + 2\sum_{m=n+1}^{-1} X_m^2 X_{n-m}^{N-2}, && \text{for} \quad n \leq -2 \end{aligned} \tag{40}$$

where

$$C_n = \sum_{m=-\infty}^{\infty} X_m^2 X_{n-m}^{N-2} \equiv 0, \tag{41}$$

are coefficients of Fourier series expansion of the function

$$F_2 F_{N-2} = \sum_{n=-\infty}^{\infty} \Big[\sum_{m=-\infty}^{\infty} X_m^2 X_{n-m}^{N-2} \Big] e^{-jn\theta}, \tag{42}$$

equal identically to zero [2].

When we take advantage of equation $X_0^2 = 0$ in (40), we have

$$\begin{aligned} S_n X_n^N &= 2 \sum_{m=1}^{n} X_m^2 X_{n-m}^{N-2} \quad \text{for} \quad n \geq 1 \\ S_n X_n^N &= 0 \quad \text{for} \quad n = 0, -1 \\ S_n X_n^N &= 2 \sum_{m=n+1}^{-1} X_m^2 X_{n-m}^{N-2} \quad \text{for} \quad n \leq -2. \end{aligned} \tag{43}$$

Similarly for F_N, from (31), (30) and (37) we have

$$\begin{aligned} F_N &= \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} (S_m X_m^{N-2} X_n^2 + S_m X_n^{N-2} X_m^2) e^{-j(n+m)\theta} = \\ &= \sum_{n=-\infty}^{\infty} \Big[\sum_{m=-\infty}^{\infty} S_m (X_m^{N-2} X_{n-m}^2 + X_{n-m}^{N-2} X_m^2) \Big] e^{-jn\theta}. \end{aligned} \tag{44}$$

Comparing (44) with (34)

$$X_n^N = \sum_{m=\infty}^{\infty} S_m (X_m^{N-2} X_{n-m}^2 + X_{n-m}^{N-2} X_m^2), \tag{45}$$

what for different ranges of n results in:
for $n \geq 0$

$$\begin{aligned} X_n^N &= \sum_{m=0}^{\infty} (X_m^{N-2} X_{n-m}^2 + X_{n-m}^{N-2} X_m^2) - \sum_{m=-\infty}^{-1} (X_m^{N-2} X_{n-m}^2 + X_{n-m}^{N-2} X_m^2) = \\ &= \sum_{m=0}^{\infty} (X_m^{N-2} X_{n-m}^2 + X_{n-m}^{N-2} X_m^2) - \sum_{m=n+1}^{\infty} (X_m^{N-2} X_{n-m}^2 + X_{n-m}^{N-2} X_m^2) = \\ &= \sum_{m=0}^{n} (X_m^{N-2} X_{n-m}^2 + X_{n-m}^{N-2} X_m^2) = 2S_n \sum_{m=0}^{n} X_m^{N-2} X_{n-m}^2 = 2S_n \sum_{m=0}^{n} X_m^2 X_{n-m}^{N-2} \end{aligned}$$

for $n = -1$

$$X_{-1}^N = \sum_{m=-\infty}^{\infty} S_m (X_m^{N-2} X_{-1-m}^2 + X_{-1-m}^{N-2} X_m^2) = \sum_{m=-\infty}^{\infty} X_m^2 X_{-1-m}^{N-2} (S_m + S_{-m-1}) \equiv 0$$

for $n \leq -2$

$$X_n^N = \sum_{m=0}^{\infty} (X_m^{N-2} X_{n-m}^2 + X_{n-m}^{N-2} X_m^2) - \sum_{m=-\infty}^{-1} (X_m^{N-2} X_{n-m}^2 + X_{n-m}^{N-2} X_m^2) =$$

$$= \sum_{m=-\infty}^{n} (X_m^{N-2} X_{n-m}^2 + X_{n-m}^{N-2} X_m^2) - \sum_{m=-\infty}^{-1} (X_m^{N-2} X_{n-m}^2 + X_{n-m}^{N-2} X_m^2) =$$

$$= -\sum_{m=n+1}^{-1} (X_m^{N-2} X_{n-m}^2 + X_{n-m}^{N-2} X_m^2) = 2S_n \sum_{m=n+1}^{-1} X_m^2 X_{n-m}^{N-2},$$

or finally

$$X_n^N = 2S_n \sum_{m=0}^{n} X_m^2 X_{n-m}^{N-2} \quad \text{for} \quad n \geq 1,$$

$$X_n^N = 0 \quad \text{for} \quad n = -1, 0 \tag{46}$$

$$X_n^N = 2S_n \sum_{m=n+1}^{-1} X_m^2 X_{n-m}^{N-2} \quad \text{for} \quad n \leq -2$$

Obtained relations (43), and (46), prove that expansions (34) with coefficients X_n^N defined by (35) are correct. Symmetry relations (36) remain to be proved: If $n \leq -2(k = -n)$, then from (46) we have

$$X_{-k}^N = 2S_{-k} \sum_{m=-k+1}^{-1} X_m^2 X_{-k-m}^{N-2}, \tag{47}$$

when the index is changed, $m = p-k$, we have

$$X_{-k}^N = -2 \sum_{p=1}^{k-1} X_{p-k}^2 X_{-p}^{N-2} = -2 \sum_{p=1}^{k} X_{p-k}^2 X_{-p}^{N-2}. \tag{48}$$

For even N ($X_p^{N-2} = -X_p^{N-2}$ from assumption (3)) we have

$$X_{-k}^N = 2 \sum_{p=1}^{k} X_{-(k-p)}^2 X_p^{N-2} = -2 \sum_{p=1}^{k} X_{k-p}^2 X_p^{N-2}. \tag{49}$$

A comparison between this expression and first equation (46) results in:

$$X_{-n}^N = -X_n^N. \tag{50}$$

For edd N ($X_{-p}^{N-2} = X_{p-1}^{N-2}$) we have

$$X_{-k}^N = 2 \sum_{p=1}^{k} X_{k-p}^2 X_{-p}^{N-2} = 2 \sum_{p=1}^{k} X_{k-p}^2 X_{p-1}^{N-2} = 2 \sum_{m=0}^{k-1} X_{k-m-1}^2 X_m^{N-2} = X_{k-1}^N,$$

and by analogy

$$X_n^N = X_{-n-1}^N \tag{51}$$

what brings to the end the proof for the correctness of expansions (34) with (35) and (36).

References

[1] P. L. Adams, H. J. Shaw, *Surface waveguide for long time delay line*, 1972 Ultras. Symp., Boston, MA, paper M-S.

[2] T. Aoki, K. A. Ingebrigtsen, *Acoustic surface waves in split strip periodic matal gratings on a piezoelectric surface*, IEEE Trans. SU, **24**, 3, 179–193 (1977).

[3] H. Bateman, A. Erdelyi, *Higher transcendental functions*, Vol. 1, New York 1953.

[4] K. Blotekjar, K. A. Ingebrigtsen, H. Skeie, *A method for analyzing waves in structures consisting of metal strips on dispersive media*, IEEE Trans. ED, **20**, 12, 1133–1138 (1973).

[5] K. Blotekjar, K. A. Ingebrigtsen, H. Skeie, *Acoustic surface waves in piezoelectric materials with periodic metal strip on the surface*, IEEE Trans. ED, **20**, 12, 1139–1146 (1977).

[6] L. A. Coldren, *Rayleigh wave guidance using anizotropic topographic structures*. APPL. Phys. Lett., **25**, 7, 367–370 (1974).

[7] A. Coldren and G. S. Kino, *CW monolitic acoustic surface wave amplifier incorporated in a $\Delta v/v$ waveguide*, Appl. Phys. Lett., **23**, 3, 117–118, (1983).

[8] E. Danicki, *Amplification of ultrasonic surface waves in a piezoelectric-semiconductor system*, Biul. PAS, **18**, 5, 47–51, (1970).

[9] E. Danicki, *Dispersion of velocity in piezoelectric acoustic waveguides*, J. Tech. Phys., **10**, 137–140,(1975).

[10] E. Danicki, *Propagation and generation of SAW, Biul.* WAT 11 (1976), suplement in Polish see also E. Danicki, *Theory of surface acoustic wave slant propagation in the periodic electrode system*, J. Tech. Phys., **19**, 1, 69–77 (1978).

[11] E. Danicki. *Acoustic surface waves in electronics*, Elektronika, **19**, 2, 52–60 (1978), (in Polish).

[12] E. Danicki, D. Gafka, *SAW dispersion in multiperiodic system of electrodes*, Akustyka Molekularna i Kwantowa, **8**, 83–87 (1988), (in Polish).

[13] E. Danicki, D. Gafka, *Slant propagation of SAW in split strip periodic system of electrodes*, Proc. of OSA'88, **1**, 255–260, Białowieża 1988, (in Polish)

[14] E. Danicki, D. Gafka, *Propagation, generation and detection of SAW in a multi-periodic system of metal strips on a piezoelectric substrate*, J. Acoust. Soc. Am. **89**, 1, 27–38 (1991)

[15] P. A. Dufilie, *Low cost SAW convolver with high spurious suppression*, 1980 Ultrasonic Symp. Proc., 43–47.

[16] H. Engan, *Excitation of elastic surface waves by spatial harmonics of interdigital transducer*, IEEE Trans. ED, **16**, 12, 1014–1017 (1969).

[17] H. Engan, K. A. Ingebrigtsen and A. Ronnenkleiv, *Characterization of SAW metal strip waveguides*, 1980 Ultras. Symp. Proc., 77–81.

[18] K. A. Ingebrigtsen, *Surface waves in piezoelectrics*, J. Appl. Phys., **40**, 6, 2681–2686 (1969).

[19] M. Luukkala and J. Surakka, *Acoustic convolution and correlation and the associated nonlinearity parameters in $LiNbO_3$*, J. Appl. Phys., **43**, 6, 2510–2518, (1972).

[20] A. A. Oliner, *Waveguides for acoustic surface waves: a review*, Proc. IEEE, **64**, 5, 615–627 (1976).

[21] A. A. Oliner (ed.), *Acoustic surface waves*, Springer-Verlag, New York 1978, 331.

[22] H. F. Tiersten, *Elasic surface waves guided by thin films*, J. Appl. Phys., **40**, 770–789 (1969).

[23] H. F. Tiersten, R. C. Smythe, *Guided surface wave filter*, 1975 Ultras. Symp., Los Angeles, CA, paper J–6.

[24] R. V. Schmidt, L. A. Coldren, *Thin film acoustic surface waveguides on anizotropic media*, IEEE Trans. SU, **22**, 2, 115–122 (1975).

Received May 18, 1989

ARCHIVES OF ACOUSTICS
16, 2, 269–282 (1991)

OBJECTIVE CHARACTERISTIC PARAMETERS OF LOUDSPEAKERS IN THE ASPECT OF PSYCHOACOUSTIC DATA

E. HOJAN

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

The paper presents the proposal of an objective evaluation of loudspeakers. Certain subjective properties of the hearing organ should be taken into consideration. The aim of these loudspeaker investigations is to correlate results of objective investigations with a subjective evaluation of the sound emitted by loudspeakers.

1. Introduction

The correlation of objective and subjective evaluations of a sound propagating in a given room or open space is one of the basic problems faced by audio-acousticians [3, 10, 14, 15].

When an electric transducer, e.g., a loudspeaker, acts as a sound source, the physical parameters of the loudspeaker can affect the subjective evaluation of the sound. The question arises about the relation between the parameters and the evaluation. Transfer response and harmonic distortions are the most frequently determined physical parameters of a loudspeaker.

The method of subjective evaluation of a sound transmitted by a loudspeaker used to date lead to their one-dimensional arrangement with respect to, for example, the degree of fidelity of sound reproduction understood as, for example, global evaluation or with respect to selected attributes of sound sensation [14].

A critical analysis of methods of subjective evaluation of loudspeakers indicates that such factors as:

– the reference sound source, the number of loudspeakers tested, the test material, the acoustic properties of the listening room, the selection and number of subjects, the tasks assigned to the subjects,

– the aim of the investigations, e.g., the arrangement of loudspeakers with respect to the selected attribute of the sensation perceived, the selection of a loudspeaker which reproduces the sound signal with the highest fidelity,

significantly affect the results of the investigations.

In reality, a subject is guided by more than one criterion of evaluation – also in the case of a global subjective evaluation of a sound emitted by loudspeakers.

The multidimensional character of the sound perception space was the basis of an attempt to apply the multidimensional scaling technique (MDS) to investigate this space [3].

In the subjective evaluation of sounds by the multidimensional scaling technique, the number of the dimensions of the perception space indicates how many independent criteria are used by the subject during this evaluation.

The measurements of objective parameters of loudspeakers provide data which help arrange loudspeakers with respect to a given parameter and then compare this arrangement with arrangements which are the result of the subjective evaluation along another dimension of perception [6, 7, 8, 9].

The results obtained indicate, however, that the correlation between these arrangements is still not satisfactory.

2. Aim

The aim of the investigations was to determine the relation between the subjective evaluation of a sound and modified objective parameters of loudspeakers. We employed the technique of multidimensional scaling for the subjective evaluation, whereas in the objective evaluation excitation with a stationary signal as well as the impulse technique were used.

It would seem that this domain has been fully investigated and only a further development of technology can increase the precision and rate at which objective parameters of a loudspeaker are determined [6, 11].

Our working hypothesis was that if the results of psychoacoustic investigations are used to evaluate objectively the parameters of a loudspeaker, the results of subjective evaluations can further approach objective investigations.

Therefore, through the weighing of results of evaluations of objective parameters, given that the function of the weight is the result of psychoacoustic investigations, it is possible to determine the parameters of a loudspeaker which incorporate some kind of effects, namely subjective evaluations.

3. Objective parameters of the loudspeaker

When selecting loudspeaker parameters, the author's experience was helpful in investigating the perception of changes in transient sounds [4], in checking loudspeakers by means of impulse methods and in studying the perception of the deformation of impulses in the acoustic field of the loudspeaker [5].

It follows from this experience that the parameters connected with the loudspeaker response to the excitation by the Tone Burst sound are a necessary complementation of parameters which directly describe the transmission properties of the loudspeaker in the steady state. This is so as it is necessary to treat the loudspeaker as a nonminimal-phase system, especially in the high frequency range.

The justification of the selection and detailed definitions of objective parameters of the loudspeaker have been quoted elsewhere [1, 4, 6]; those which have been used in the present investigations are discussed below.

3.1. Transfer response

This response was determined by excitation using a sinusoidal signal with retunable frequency. Next, the width of the transfer band Δf was determined. Maximal values of deviations from the value of the mean level indicate the irregularity of ΔL response in the response band.

3.2. Transient response

The loudspeaker transient response, i.e., its transfer response for a given time moment, after the disconnection of the excitation Tone Burst signal, was recorded by the measurement system described in detail elsewhere [6]. The transient response index D was adopted as the measure of the deviation of the loudspeaker transient response from its transfer response in the steady state (for a given time τ after the disconnection of the excitation signal) where

$$D = \frac{1}{n}\sum_{1}^{n}(L_{ssi} - L_{\tau i}) \quad [\mathrm{dB}]$$

L_{ssi} – the value of the acoustic pressure level of a steady state signal at its i-th frequency. $L_{\tau i}$ – the value of the acoustic pressure level of the signal final transient at a given time after the disconnection of the excitation signal with the i-th frequency, n – the overall number of frequency components.

Additionaly, the quantities D_r which define the difference between the maxima and minimal values of the index D for successive times τ were determined.

3.3. Duration of the initial and final transients

The duration of transients connected with the signal growth "t_n" – the initial transient, and with the signal reverberation "t_z" – the final transient, is one of the basic measures of transient distortions introduced by the loudspeaker when it is excited by an impulse.

Taking into account the character of transients of loudspeakers excited by tone

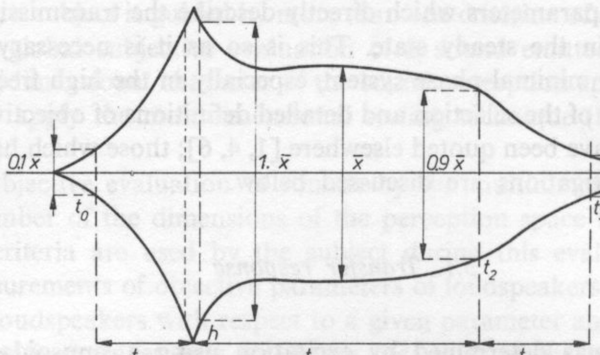


FIG. 1. Time response of the loudspeaker to impulse excitation
x – mean value curve at the steady time segment
t_{01} – duration of the initial transient
t_{12} – duration of the steady state
t_{23} – duration of the final transient

impulse, we have defined the durations of the initial and final transients as time segments $t_n = t_{01}$ and $t_z = t_{23}$ (Fig. 1).

The duration of the transients can be expressed directly in time units (ms). It is also possible to take into account a specific value of the frequency of the signal excited by means of the number of vibration periods (T). Hence, for $f = 2000$ Hz ($T = 0.5$ ms) the duration of the 2 ms transient is equal to 4 in absolute units.

3.4. *Energy, coefficient*

The energy ratio at time t_{01}, t_{23} reffered to the steady state energy at time segment t_{12} was adopted as the measure of energy of transient during the signal growth H_n and its decay H_z.

4. Selected results of psychoacoustic investigations

At was mentioned above, the loudspeaker parameter discussed in Section 3 were selected from a greater number of parameters. The selection criterion followed from the degree of their correlation with the subjective evaluation of the sound, defined through the arrangement of loudspeakers using the method of multidimensional scaling.

The results obtained were not satisfactory for the author ($\varrho < 0.80$). It was decided to modify the objective parameters through their appropriate weighing resulting from the data obtained from psychoacoustic investigations.

When defining the weight of particular parameters, it was decided to take into account the following facts known from the theory of hearing [1, 2, 12]:

– the width of the critical band up to the band mid frequency $f_m = 500$ Hz is equal 100 Hz and for higher frequencies it is equal to 17% of the value of the mid frequency,

– the resolution capability of the hearing organ in the frequency domain changes in accordance with the curve shown in Fig. 2,

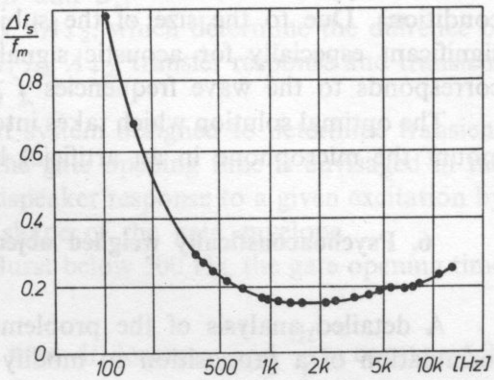


FIG. 2. The curve of changes in the resolution capability of the hearing organ in the frequency function [12]

– the masking effects affect the overall loudness of the sound, tonal balancing and hence the sound timbre. Sound loudness can be determined for example on the basis of, what are known as Zwicker nomograms. When determining sound loudness, its spectral composition and the effects of the mutual masking of components are taken into consideration,

– the hearing organ values are an acoustic signal in the frequency function in the band of up to 500 Hz according to the linear scale and above according to the logarithmic scale,

– the hearing organ is capable of recording changes in the values of the signal acoustic pressure; this capability is greater in the case of pressure increase than decrease.

5. Conditions of the investigations of loudspeaker objective parameters

When discussing problems connected with the determination of loudspeaker objective parameters, other problems which can affect the final result should also be taken into consideration. The latter include the determination of the conditions of objective measurements.

The first problem is where, in what room, the objective measurements of loudspeaker should take place. As has been indicated by the investigations conducted by the present author, among others, a correlation of the objective evaluations with parameters of a loudspeaker for the steady state occurs only when these parameters have been determined in the same room in which subjective evaluation takes place.

When loudspeaker parameters were related to its work in the transient state, they correlated with the subjective evaluation also in the case when the parameters were determined in an anechoic chamber.

The second methodological problem connected with objective investigations is that of the approximation of diffraction around the measurement microphone to real conditions. Due to the size of the subject's head (Ø $\simeq$ 17 cm), these effects are significant especially for acoustic signals whose wave length $\lambda \leq 17$ cm, which corresponds to the wave frequencies $f \geq 2000$ Hz.

The optimal solution which takes into account the diffraction effects seems to be mount the microphone in an artificial head.

6. Psychoacoustically weighed objective parameters of the loudspeaker

A detailed analysis of the problems discussed above was the basis for the formulation of a proposition to modify the definitions and the methodology of measurement of selected parameters of the loudspeaker.

The propositions were presented with reference to each parameter.

6.1. Steady state

6.1.1. Transfer response. The measurement of this response was made in a listening room by exciting the loudspeaker with a sinusoidal signal. The microphone was mounted in an artificial head, and its output signal was recorded by the *A/C* transducer.

The following are adopted as parameters of the transfer response:

– transfer band Δf [Hz],

– irregularities of the response ΔL [dB] determined in the entire transmission band $(f_d \div f_g)$ and in bands $(f_d \div 500)$ Hz, $(500 \div f_g)$ Hz,

where:

– f_d, f_g – the bottom and top frequency of the loudspeaker transfer band $(\Delta f = f_g - f_d)$,

– loudness N in sones and the level of audible intensity L_n in phones, determined on the basis of Zwicker nomograms (transfer response is treated as a spectrum of the acoustic signal registered at the loudspeaker output),

– area A_{+2} [dB], exceeding the mean level of the transfer response by values greater than +2 dB in the entire transfer band and in ranges $(f_d \div 500)$ Hz, $(500 \div f_g)$ Hz. The parameter Δf, ΔL, N, L_n, A_{+2} are determined at linear (frequency up to 500 Hz) and logarithmic (frequency above 500 Hz) scales of frequency changes.

6.2. *Transients*

6.2.1. Transient response. The measurement of the transient response is done under the same conditions as was the case with the transfer response. The following are adopted as parameters of the transient response:
– the parameters Δf, ΔL, N, L_n, A_{+2} determined under the same assumptions as was the case with the transfer response,
– the parameters of transient response D and D_r,
– the parameters $\Delta(\Delta f)$, $\Delta(\Delta L)$, ΔN, ΔL_n, ΔA_{+2}, which determine the diffrence of the absolute value of parameters Δf, ΔL, N, L_n, A_{+2}, transfer response and transient response for successive moments of time τ.

At the same time in the measurement system designed to determine transient responses, a possibility of the change of the gate opening time is envisaged in the microphone system which records the loudspeaker response to a given excitation by the Tone Burst signal, at a determined shape of the gate envelope.

Within the frequency range of Tone Burst below 500 Hz, the gate opening time is constant and equals 10 ms.

For frequencies higher than 500 Hz, time Δt decreases and $\dfrac{\Delta t}{T} = \text{const.} = 5.9$.

Hence,
for $f = 1$ kHz

$$\Delta t = 5.9 \text{ ms},$$

for $f = 5$ kHz

$$\Delta t = 1.2 \text{ ms}$$

for $f = 10$ kHz

$$\Delta t = 0.6 \text{ ms},$$

for $f = 16$ kHz

$$\Delta t = 0.4 \text{ ms}$$

6.2.2. Durations of the initial transient t_n and final transient t_z. Energy coefficients H_n, H_z

The measurements of these quantities are made under the same conditions as the measurement of the transient response in the measurement system with the gate width retunable in time.

7. Results

The investigations comprised three groups of loudspeakers: low, mid, and high tone ones, six loudspeakers in each case. The low tone loudspeakers are the same type of loudspeakers with changes in the design of the membrane mass and coil circuit. Individual mid- and high-tone loudspeakers differed with respect to their type.

7.1. *Results of subjective evaluations*

A sketch of the listening room in which subjective evaluations and part of objective evaluations were conducted is shown in Fig. 3.

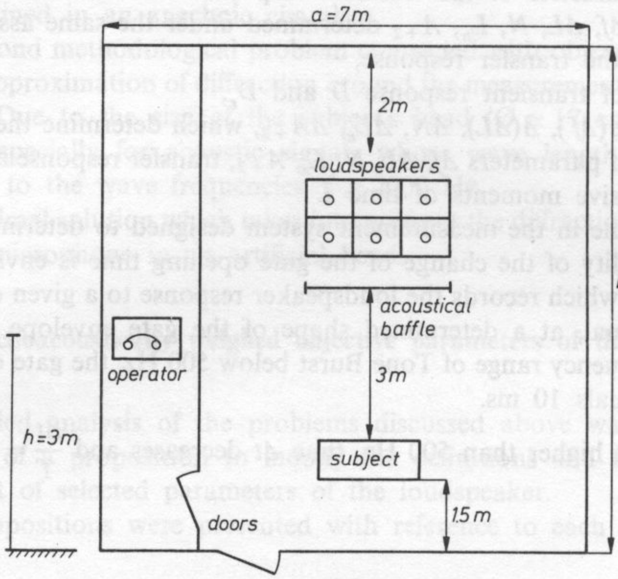


FIG. 3. A sketch of the listening room

The basic investigations were preceded by preliminary investigations which included a test of the subjects' selection and a selection of the listening material. They also helped define the influence of the source directivity on the results of the evaluations.

Following the preliminary investigations, in which the correlation of the ranks of each of the subjects was determined with respect to the subject for whom the value of the statistics tested was the highest, six subjects out of ten were selected to participate in the basic investigations. The results of this selection were verified using the MDS method. The compatibility of the evaluations of subjects was verified on the basis of Kendall and Babington-Smith's concordance index [3].

Once it was determined that the subjects evaluate the loudspeakers under investigation in a similar way, the differences in the arrangement of the loudspeakers with respect to different musical pieces were interpreted as a uniformity test of the sample material.

Table 1 includes a description of 6 musical pieces (duration ~ 10 s) together with stress values for particular dissimilarity matrices, respectively in three-, two- and one- dimensional perception space.

The criterion of 5% stress value was adopted as the measure of the adjustment

Table 1. Description of musical pieces

No	Type of music	Performer	Title	Stress [%] I	II	III
1	big beat	Jimmy Hendrix	Stepping stone	43.4	0.244	x
2	popular	Bert Kaempfer orchestra	Caravan	39.2	0.001	x
3	popular	Uriah Heep group	Sunrise	–	22.6	0
4	popular	Zygmunt Malik orchestra	Sweet Beat	34.6	0.126	x
5	big beat	Jimmy Hendrix	Midnight	43.2	0	x
6	big beat	Dżamble group	I've got to have a girl friend	–	22.6	0

of the perception space to the data in the dissimilarity matrix. It follows from the analysis of the data in Table 1 that for four musical pieces (1, 2, 4, 5) the perception space of sensations is two-dimensional, and for two musical pieces (3, 6) it is three-dimensional. Considering the percentage of the total variance connected with a given dimension, a hypothesis was advanced about the significance of two dimensions of the perception space which differentiate the loudspeakers under investigation.

In an attempt to explain any influence of the source directivity on the results of the subjective evaluations by the subjects, the results of the evaluations from two series of the investigations *A* and *B* were compared. The series differed only as regards the location of the loudspeakers with respect to each other (Fig. 4).

A

1	2	3
4	5	6

B

1	6	2
4	5	3

FIG. 4. Distribution of loudspeakers in series *A* and *B*

The investigations conducted using the MDS technique have shown that the impression of the directivity does not have any influence on the arrangement of the loudspeakers in series *A* and *B* along the 1st dimension of perception and can be almost neglected (the transposition of loudspeakers 2 and 3) during their arrangement along the 2nd dimension of perception.

Six subjects took part in the basic investigations. They evaluated three groups of loudspeakers (each time six loudspeakers) on the basis of the selected sample material. The method of stimuli comparison by triade was used. The same musical piece reproduced by one of the 3 loudspeakers switched on randomly by the experimenter was evaluated in the triad. The subject had at his disposal three buttons and made decision about the order in which he switched on one of the three loudspeakers. The subject's task was to point to the pair of loudspeakers most similar and most dissimilar with respect to each other. The results of the evaluation of the similarity between the sounds was processed by the method of multidimensional scaling of individual differences INDSCAL.

On the basis of the dissimilarity matrices defined for all subjects and for all pieces of music, a two-dimensional perception space and an appropriate arrangement of the loudspeakers along both dimensions were obtained for each group of loudspeakers.

7.2. *Results of objective evaluations*

A block diagram of the measurements system was shown in Fig. 5.

The use of the microcomputer helped fully automate the entire process of measurement and data processing. All loudspeakers from the three groups were

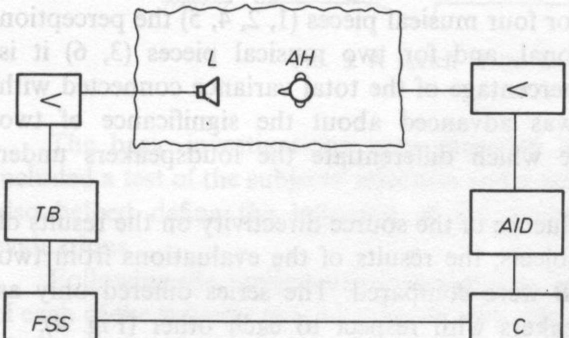


FIG. 5. Block diagram of the measurement system L – loudspeaker, AH – artificial head, A/D – analog-to-digital converter, C – computer, FSS – frequency sweep system, TB – tone burst generator

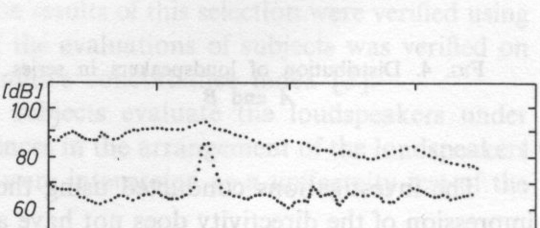


FIG. 6. Transfer response (top curve) and transient response (bottom curve) of a low tone loudspeaker

investigated. The same methodology as described in Section 5 was used and the same parameters as described in Section 6 were determined. Two cases were identified for transient responses: with constant and varying width of the time gate.

For example, Figure 6 and Table 2 show the results for one low tone loudspeaker with respect to the transfer response and transient response ($\tau = 1$ ms, varying width of the time gate).

Table 2. Values of low tone loudspeaker parameters determined from transfer response and transient response

	f_d f_g [Hz]	ΔL [dB]	A_{+2} [dB]	Δf [Hz]	N [son]	L_n [fon]	D [dB]
Transfer response	$(f_d \div 500)$	2.45	0.56	–	–	–	–
	$(500 \div f_g)$	3.28	1.46	–	–	–	–
	$(f_d \div f_g)$	3.21	1.38	3306	891.1	138.1	–
Transient response	$(f_d \div 500)$	5.71	3.01	–	–	–	9.5
	$(500 \div f_g)$	4.12	2.41	–	–	–	12.3
	$(f_d \div f_g)$	4.26	2.47	3240	508.2	129.9	12.1
Difference in the value of the parameters of both responses	$(f_d \div 500)$	3.26	2.45	–	–	–	–
	$(500 \div f_g)$	0.84	0.95	–	–	–	–
	$(f_d \div f_g)$	0.95	1.09	66	382.9	8.2	–

Table 3 shows the results of the measurements of the duration of transients t_n, t_z and energy coefficients H_n, H_z.

Table 3. Values of low tone loudspeaker parameters determined from transients

f_d, f_g [Hz]	t_n[T] [ms]	t_z [T]	H_n	H_z
$(f_d \div 500)$	2.28	2.33	0.91	0.30
$(500 \div f_g)$	4.85	4.53	0.93	0.110
$(f_d \div f_g)$	4.57	4.29	0.93	0.13

Numerical data of particular parameters of loudspeakers shown in Tables 2 and 3 helped arrange the loudspeakers in three successive groups, according to the growing numerical values of the parameters.

7.3. *Correlation of subjective and objective evaluations*

SPEARMAN'S [3] rank correlation index ϱ_s was used to investigate the correlation force between two arrangement scales. Adopting the significance level $\alpha = 0.05$, for the number of elements tested $n = 6$, we get the critical value of the rank correlation index $\varrho_k(\alpha, n) = 0.83$. Hence the value $\varrho_s > \varrho_k = 0.83$ decides on the existence of a dependence between the arrangements.

Matching the arrangements of loudspeakers obtained as a result of subjective and objective evaluations, each time in a given group of loudspeakers, the values of the rank coefficient ϱ_s were determined. In practice, this means a comparison of arrangements in the dimension of the perception space which are the result of the subjective evaluation with arrangements which result from the growing numerical values of the parameters of the loudspeakers.

Assuming the value $\varrho_s > 0.83$ as the necessary condition for the existence of a correlation between the evaluations, the following parameters of the loudspeakers were distinguished:

1. First group of parameters:
Parameter f, $\varrho_s = 1.00$, first dimension of perception,
Parameter D, $\varrho_s = 1.00$, second dimension of perception (parameter D – defined for the transient response determined at the constant width of the time gate).
2. Second group of parameters:
Parameter t_n (max), $\varrho_s = 0.94$, first dimension of perception
Parameter $t_z(f_d \div f_g)$, $\varrho_s = 0.83$, second dimension of perception.
3. Third group of parameters:
Parameter N, $\varrho_s = 1.00$, first dimension of perception,
Parameter L_n, $\varrho_s = 1.00$, first dimension of perception,
Parameter A_{+2}, $\varrho_s = 0.89$, second dimension of perception
Parameter ΔL, $\varrho_s = 0.89$, second dimension of perception.

A detailed analysis of the results helps formulate the following conclusions:

1. When evaluating different types of loudspeakers with assumed large differences in the transmission band range, a full correlation with the subjective evaluation is obtained by the band width of Δf (as defined in Section 3.1) – in the first dimension of the perception space and the transient response index D (as defined in Section 3.2) in the second dimension of the perception space.

2. When evaluating one type of loudspeakers (very small differences in the parameters of the design assumed), the best correlation with the subjective evaluation is obtained between the parameters of the loudspeaker connected with the transients, precisely with the duration of the initial transient t_n (as defined in Section 3.3; its maximal value in the entire transfer band) – in the first dimension of the perception space and the duration of the final transient t_z (as defined in Section 3.3: its mean value for the entire band of the space) – in the second dimension of this space.

3. The evaluation of different types of the same group loudspeakers leads to the statement that in this case the best correlation is found between the parameters of the loudspeaker connected with loudness N or the loudness level L_n (as defined in Section 6.1) – in the first dimension of the perception space and the area A_{+2} (as defined in Section 6.1) – in the second dimension of this space.

In the second dimension of the perception space the equivalent value of the correlation index is also provided by the non-uniformity of the transfer response ΔL (as defined in Section 6.1).

8. Conclusion

An attempt to correlate subjective evaluations of a sound emitted by loudspeakers with objective parameters of loudspeakers, on the basis of the technique of multidimensional scaling, subjective evaluations and the methods of the steady state and the impulse technique (objective evaluations) has shown that such a correlation exists at level $\varrho \geqslant 0.83$. Specific requirements pertaining to the methodology of investigations and the selection of appropriate objective parameters of loudspeakers, depending on the object of the investigations, have to be observed.

The weighing of the loudspeaker parameters on the basis of the resuts of psychoacoustic investigations has produced a positive result through the increase of the degree of the power of the correlated evaluations.

Presently, it seems extremely interesting to look for two independent attributes which describe the quality of the sound perceived, ascribed to two determines dimensions of the perception space.

The results of these investigations will help bind specific parameters of loudspeaker parameters with attributes of the perception space.

References

[1] R. Bücklein, *Hörbarkeit von Unregelmassigkeiten in Frequenzgängen bei akustischer Übertragung*, Frequenz **16**, 3, 103–108 (1962).

[2] R. Feldtkeller, E. Zwicker, *Das Ohr als Nachrichtenempfänger*, S. Hirzel Verlag, Stuttgart 1956.

[3] A. Gabrielson, *Perceived sound quality of sound reproducing systems*, JASA **65**, 4, 1019–1033 (1979).

[4] E. Hojan, *Detectability of pulse distortion in an acoustic loudspeaker field*, Archives of Acoustics **2**, 4, 281–296 1972.

[5] E. Hojan, *Analysis of the configuration of the acoustic field of loudspeaker systems excited with a sinusoidal signal*, Archives of Acoustics **5**, 1, 45–56 (1980).

[6] E. Hojan, M. Niewiarowicz, *Messystem für die Registrierung von Transientcharakteristiken der Lautsprecher*, Frequenz **39**, 11, 307–309 (1985).

[7] K. L. Kantor, A. P. Koster, *Psychoacoustically optimized loudspeaker*, 74th Convention, AES New York 1983.

[8] M. Karjalainen, *A new auditory model for the evaluation of sound quality of audio systems*, IEEE, **19**, 9.1–19.9.4, (1984).

[9] J. M. Kates, *A perceptual criterion for loudspeaker evaluation*, JAES **32**, 12, 938–944 (1984).

[10] M. Komamura, K. Tsuruta, M. Yoshiba, *Correlation between subjective and objective data for loudspeaker*, JASJ, 33, 3, 103–115 (1977).

[11] M. Niewiarowicz, P. Perz, *A method for determining selected parameters of a loudspeaker under the aspect of its subjective evaluation*, Frequenz **42**, 4, 113–116 (1988).

[12] J. Salmi, A. Weckstroem, *A new psychoacoustically more correct way of measuring loudspeaker frequency responses*, 73 rd Convention AES, Eindhoven 1983.

[13] H. Staffeldt, *Correlation between subjective and objective data for quality loudspeakers*, JAES, 22, 6, 402–415 (1974).

[14] Z. Sołtys, Z. Wąsowicz, *Methods of auditory evaluation of loudspeakers*, Scientific Papers of the Institute of Telecommunication and Acoustics of Wrocław Technical University, 35, 17, 1978 (in Polish).

[15] F. E. Toole, *Subjective measurements of loudspeakers — a comparison of stereo and mono listening*, 74th Convention AES, New York 1983.

Received June 5, 1989

ARCHIVES OF ACOUSTICS
16, 2, 283–293 (1991)

RELATION BETWEEN SELECTED ATTRIBUTES OF THE PERCEPTION SPACE AND THE EVALUATION OF SOUND FIDELITY

E. HOJAN, A. FURMANN

Institute of Acoustics, Adam Mickiewicz University
(60-796 Poznań, ul. Matejki 48/49)

The aim of the presented investigations was to determine the relations between the selected attributes of sound perception space and the subjective evaluation of sound "fidelity". Each sound examined was evaluated twice: for its "fidelity" with respect to the mental model and for the attributes of sound perception space ("clearness", "fullness", "spaciousness", "sharpness", "loudness", "lack of distortions"). We examined test signals (pink noise, white noise, speech signal and a musical piece) filtered by five kinds of filters and reproduced by a loudspeaker system. Hi-fi experts acted as subjects. Analysis of the results showed that "fullness", "sharpness" and "lack of distortions" determine the values of the subjective evaluation of sound "fidelity" when "fidelity" is evaluated with respect to a mental model.

1. Introduction

The quality of loudspeakers is evaluated by objective methods in which selected acoustic and electrical parameters are measured and by subjective methods in which the quality of reproduced sounds (of music and speech) is evaluated on the basis of a listening test. There are many methods of objective evaluation of loudspeakers. However, full correlation between the objective and subjective evaluations still lacks [1, 2]. This is why attempts are made to find such methods of objective evaluation of loudspeakers which would correlate with the subjective evaluation. In order to achieve this aim, we must first explain what physical parameters of a reproduced sound influence its evaluation in the listening test. To find these parameters must determine first how many independent criteria a subject uses to evaluate complex sound (such as speech or music). The authors have conducted investigations [3] on the basis of which they have identified six criteria of sound evaluation in a listening test: "fullness", "spaciousness", "sharpness", "clearness", "loudness" and "lack of distortions". These criteria have been called attributes of the sound perception space. It is known that auditory evaluation of loudspeakers is connected primarily with the evaluation of the "fidelity" of reproduction. Therefore, it was decided to determine which of the selected attributes of the sound perception space affect the auditory evaluation of its “fidelity”.

The “fidelity” of a sound transmitted through a loudspeaker is evaluated objectively by the measurement of the difference between the physical parameters of the input signal at the microphone and the output signal at the loudspeaker, and subjectively by the evaluation of the difference in the perception of a sound transmitted by the loudspeaker and the original sound. If the differences are below the perception threshold, the transmitted sound is considered to be of “absolute fidelity”. If few subjects perceive a difference between a sound radiated by the loudspeaker and the original sound, then the transmitted sound can be called a “high fidelity” sound [4].

The subjective evaluation of the “fidelity” with respect to the original sound is very rarely used in practice as such investigations are very laborious and costly. For these reasons, evaluations of “fidelly” are usually performed with respect to a certain model. The latter can be one of the following:

– an original sound remembered by the subject (mental model)

– a sound reproduced by a model loudspeaker

– a “simulated live” (a sound transmitted by the loudspeaker) the sound plays the role of the original sound [5].

In the present paper we discuss the results of investigations which allow to determine the relation between selected attributes of sound perception space and to evaluate its “fidelity” with respect to the mental model. The assumption is that this model of “fidelity” most closely approximates the practice of evaluation of the quality of reproduced sounds.

Knowing the physical parameters of sounds which affect the attributes of the sound perception space and knowing the relation between the evaluation of sound “fidelity” with the attributes of sound perception space, it is possible to determine which physical parameters of sound affect the evaluation of its “fidelity”. Explanation of these relations will help to find the correlation between the objective and subjective evaluation of loudspeakers.

2. Method

2.1. *Signals and subjects*

Four one-minute test signals were examined:

*P*1 – pink noise (–3 dB/octave), sound level 80 dB (A)

*P*2 – speech – speaking male voice of the master of ceremonics in a theater. Sound level 75 dB (A). Gramophone record: PRONIT PLP0035

*P*3 – wideband music – orchestra, excerpt from “Les Préludes” by F. Liszt, performed by Berlin’s Philharmonic Orchestra. Sound level about 75–85 dB (A). CD Deutsche Gramophone 413587-2 GH

*P*4 – white noise, sound level 80 dB (A).
All signals were filtered through each of the five filters whose frequency responses are shown in Fig. 1. The frequency responses of the filters have been selected to correspond to the frequency responses of various average class loudspeaker systems. After filtering, the number of signals to be investigated was 20, i.e., 4 test signals × 5 filters. They were reproduced through a stereo EXTRA FLAT loudspeaker system.

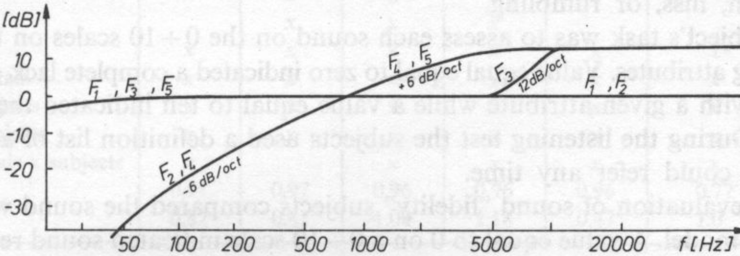


FIG. 1. Frequency responses of filters *F*1–*F*5

The sound levels given above refer to the level of the respective test signals as reproduced by the loudspeaker system with no filtering (*F*1) and measured by a sound level meter at the listener's position in the listening room.

The listening tests were presented in a listening room compatible with IEC standards [6].

Eight subjects (men aged 25–40 years) participated in the investigation. They were randomly chosen from among hi-fi experts.

2.2. *Procedure*

The evaluation of sounds was performed in two stages:

– at the first stage selected attributes of the sound perception space were evaluated

– at the second stage the "fidelity" of reproduced sound with respect to the mental model was evaluated.

The selection of sound attributes was based on the results of earlier investigations [3] in which it was found that the attributes may be defined as follows:

– "clearness" the sound is pure, clear; different instruments and voices can be easily distinguished, instruments and voices sound clear and pure without distortions, onsets, transients and other in the music details can be easily perceived

– "sharpness" the sound contains components whose mid- and high-frequency levels are too high

— "fullness" means that the sound contains the entire spectrum without any limitations, at least in the bass range. A sound which lacks the bass range is the opposite of a "full" sound

— "spaciousness" the reproductions is spacious, the sound is open, has width and depth, fills the room, gives the impression of the subject's presence in the space surrounded by sound

— "lack of distortions" indicates a pure sound, without distortions, one which is not harsh, hiss, or rumbling.

The subject's task was to assess each sound on the $0 \div 10$ scales on the basis of the following attributes. Value equal equal to zero indicated a complete lack of sensation connected with a given attribute while a value equal to ten indicated the maximum sensation. During the listening test the subjects used a definition list of attributes to which they could refer any time.

In the evaluation of sound "fidelity" subjects compared the sound with respect to a mental model. A value equal to 0 on a $0 \div 10$ scale indicated sound reproduction which differed most from the mental model while a value equal to 10 indicated an ideal reproduction of the mental model. Mental models of the white and pink noises were shaped in the subjects as a result of their long exposure to these types of sounds; the subjects were recruited from among designers of loudspeakers who are exposed to white and pink noises in their everyday work. At both stages of the investigations subjects evaluated each sound four times.

3. Results

The results of the subjective evaluation of sounds underwent a multidimensional analysis of variance (ANOVA) separately for each subject and for all subjects together [7]. A mixed model was used in which subjects were taken as a random factor, while test signals and filter characteristics were taken as a constant factor. The results of the evaluation of each attribute of sound were analyzed separately.

The analysis of variance makes possible estimates of reliability for each subject individually (the intra-individual reliability index r_w) and for all subjects together (the inter-individual reliability index r_b) [8]. The interpretation of reliability r_w and r_b should be considered with the random error MS_r. If $r_w > 0.70$ and $MS_r < 1.5$, the reliability is good, if $0.40 \leqslant r_w \leqslant 0.60$ and $MS_r < 1.50$, the reliability is satisfactory. However, if $r_w < 0.40$ and $MS_r \geqslant 1.50$, the reliability is not satisfactory.

The reliability index r_w ranged from 0.7 to 0.9 whereas MS_r ranged between 0.3 and 1.5, depending on the attribute and subject.

The obtained values of the reliability index r_w and random error MS_r show that the reliability for each subject is good.

Inter — individual variability r_b refers to the agreement between the ratings of different subjects. The results in Table 1 indicate that the subjects give similar weight to different perceptual dimensions.

Table 1. Results of the analysis of variance of group data, separately for particulary attributes of the perception space and for fidelity; x denotes statistically significant differences at the significance level $\alpha = 0.01$, r_b – reliability index for the agreement between subjects, MS_r – error variance

Source of variance	Clearness	Sharpness	Fullness	Spaciousness	Loudness	Lack of distortions	Fidelity
Filters	x	x	x	x	x		x
Signals		x	x	x			x
Subjects	x	x	x	x	x	x	x
Filters x signals	x	x	x	x	x	x	x
Filters x subjects	x	x	x	x	x	x	x
Signals x subjects	x	x	x	x	x	x	x
Filters x signals x subjects	x		x	x	x	x	x
r_b	0.91	0.97	0.96	0.96	0.96	0.77	0.97
MS_r	0.82	0.81	1.00	1.18	0.47	1.0	1.13

It follows from the analysis of individual and group data for a subjects that the evaluation of the attributes of the perception space of the sounds under investigation depends on:

a) the test signal – the sounds were differently evaluated depending on the test signal

b) subjects – different subjects tend to use somewhat different parts of the $0 \div 10$ scale

c) filter characteristics – the sounds were differentiated depending on the frequency responses of a filter through which the test signal was passed.

Table 1 shows results of the analysis of group data, separately for specific attributes of the perception space and "fidelity" (x denotes statistically significant differences between groups of sounds under comparison). Considering the fact that the analysis of variance showed statistically significant differences between the subjects, we can conclude that the subjects differed as to the absolute evaluation of sounds on the $0 \div 10$ scale. The results the subjective evaluation of sounds were therefore interpreted on the basis of the values of the median and not on those of the arithmetic mean, determined from the data obtained by all the subjects. Figure 2 shows values of the median of the subjective evaluation of sounds. The bottom and top quartils have been indicated.

It follows from Fig. 2 and Table 1 that:

– the evaluation of the "lack of distortions" of the sounds under investigations does not depend on the frequency response of a filter and the test signal but only on the interaction filter x test signal. This suggests that differentation of sounds is dependent on the kind of test signal filtered. This can be related to the fact that the influence of the filter frequency response on the spectrum of reproduced sound is dependent on the kind of test signal.

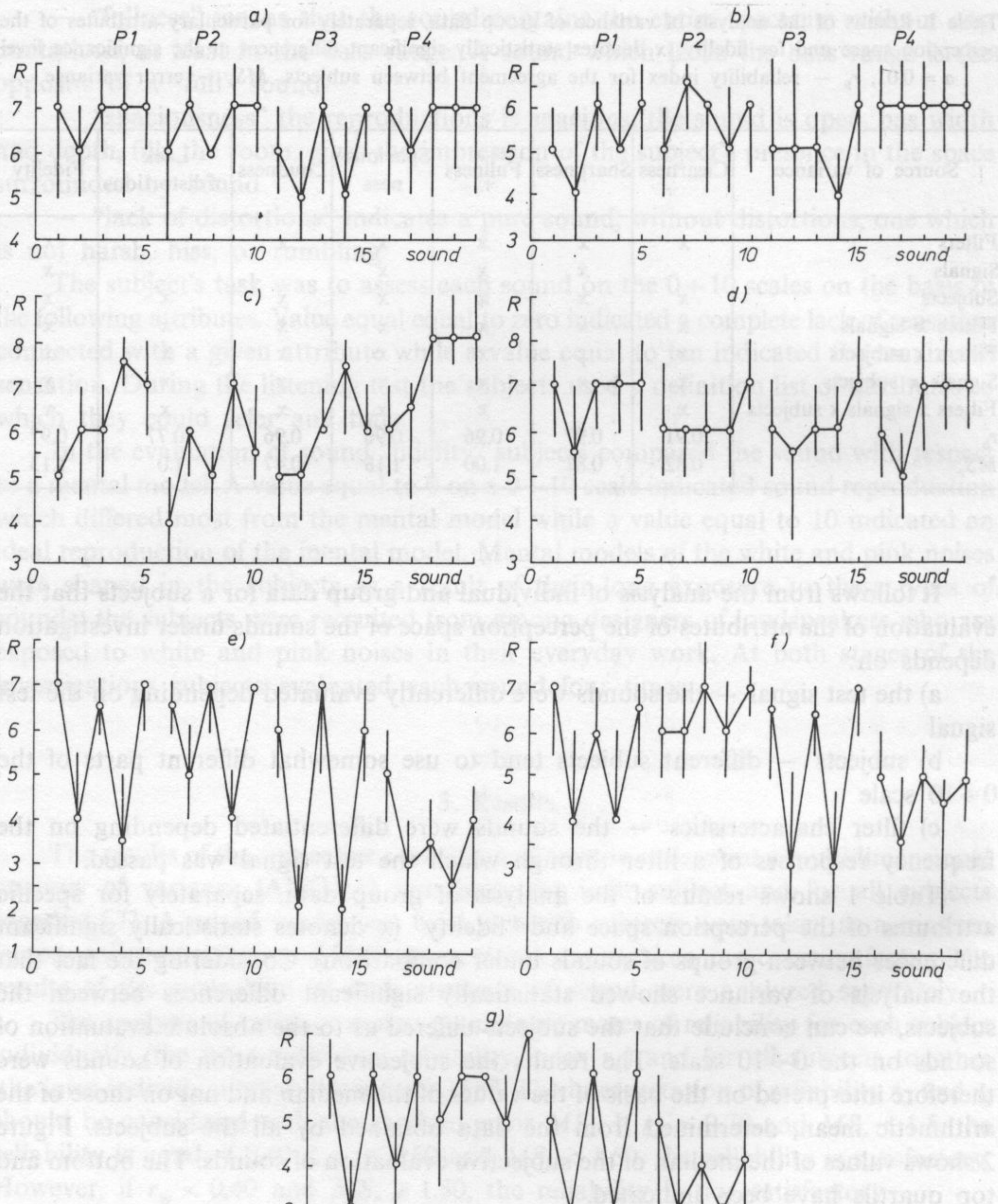


FIG. 2. Median values of the evaluation of attributes: a) "clearness", b) "lack of distortions", c) "sharpness", d) "loudness", e) "fullness", f) "spaciousness", g) "fidelity" of all sounds under investigation, where: R – denotes rating on the $0 \div 10$ scale $P1 \div P4$ — groups of sounds for particular test signals, correction with filters $F1 \div F5$. The bottom and top quartil have been marked

— the evaluation of "loudness" and "clearness" depends primarily on the frequency response of a filter through which the test signal was passed

— the evaluation of the remaining attributes of the perception space and "fidelity" is greatly affected by both the frequency response of a filter and the kind of test signals.

4. Analysis of the relation between the separate attributes of sound perception and the global evaluation of sound "fidelity"

4.1. *Methods of analysis*

The relation between the attributes of the perception space and sound "fidelity" has been examined by means of analysis of correlation and multiple regression. The analyses were made separately for three kinds of data:

1. The results of the evaluation of sounds for each test signal $P1 \div P4$ were considered separately. Since test signals were filtered through five different filtres $F1 \div F5$, five sounds corresponded to each test signal. Hence only five elements were included in the analysis of this kind of data. For this reason multiple regression could not be determined simultaneously for all attributes of the perception space (the number of elements was smaller than the number of variables examined). It was assumed that multiple regression must entail such attributes which have the greatest influence on the evaluation of "fidelity", i.e., whose variability determined the greatest number of variable "fidelity".

2. The results of the evaluation of all sounds under investigation were analysed, a total of 20 elements, i.e., 4 test signals x 5 filters. The results of the analysis could be affected in this case by the interaction filter x test signal (the influence of the filter frequency response on the evaluation of sound, dependent on the test signal).

3. The differences between median values of the sound evaluation and the mean value of median for the sound evaluation by the same test signal were considered.

Table 2 shows examples of the median values of the evaluation sound "fidelity" for the all sounds under investigation and arithmetic means of medians by the same

Table 2. Median values of the evaluation of sound "fidelity" for particular test signals P1 – P4, particular filters $F1 \div F5$ and mean values of median for particular test signals

	$F1$	$F2$	$F3$	$F4$	$F5$	Mean
$P1$	6	3	6	4	6	5
$P2$	5	6	6	5	7	5.8
$P3$	5	3	6	3	6	4.6
$P4$	5	3	4	3	4	3.8

test signal (given in the right-hand margin in the matrices in Table 2). The values shown in the rows reflect the results of the evaluation of sounds "fidelity", respectively, to the test signal $P1 \div P4$. The values shown in the columns reflect the results of the evaluation of sounds "fidelity" filtered through the filter with a specific frequency response, respectively, $F1 \div F5$. The data shown in Table 2 can be used to determine values equal to the differences between the median values of the evaluation of sounds under investigation and the mean value of these medians for the same test signal. The values are shown in Table 3.

Table 3. The differences between median values of the evaluation sound "fidelity" and the mean value of median for the sound evaluation, by the same test signal. The values were determined on the basis of the data shown in **Table 2**

	$F1$	$F2$	$F3$	$F4$	$F5$
$P1$	1	−2	1	−1	1
$P2$	−0.8	0.2	0.2	−0.8	1.2
$P3$	0.4	−1.6	2.4	−1.6	2.4
$P4$	1.2	−0.8	0.2	−0.8	0.2

These data, contrary to those of the second kind, are not affected by the interaction filter x test signal.

4.2. Results of the analysis

Table 4 shows the values of the linear correlation coefficient after Pearson [9], which determine the relationship between the values of the evaluation of "fidelity" and the values of the evaluation of specific attributes of the perception space of the sounds under investigation. The square value of this coefficient multiplied by 100 determines the percentage of variability of "fidelity" caused by the variability of the

Table 4. Coefficients of liner correlation after Pearson, determined between evaluation of „fidelity" and particular attributes of the percepion space for three kinds of data – 1, 2, 3

Kinds of data \ Attributes		Clearness	Sharpness	Fullness	Spaciousness	Loudness	Lack of distortions
1.	$P1$	0.79	−0.27	**0.93**	**0.94**	0.79	0.84
	$P2$	0.53	0.00	0.53	0.76	0.80	0.75
	$P3$	**0.88**	−0.08	**0.95**	**0.96**	0.75	0.39
	$P4$	0.37	−0.54	**0.93**	0.51	−0.06	—
2.		**0.45**	**−0.48**	**0.88**	**0.88**	0.34	0.21
3.		**0.68**	−0.18	**0.85**	**0.82**	**0.52**	**0.60**

values of specific attributes of the perception space of the sounds under investigation. The values of correlation are shown separately for each kind of the data analysed – the values determining a statistically significant correlation at the significance level of $\alpha = 0.05$ are shown in bold type.

The correlation between the values of "fidelity" and the values of "lack of distortions" has not been determined for the case of a white noise ($P4$). The sounds there always awarded the same evaluation with respect to the "lack of distortions".

The analysis of the results shown in Table 4 helps state that the values of the "fullness" of sounds correlate statistically significantly for almost all kinds of data (except for data for the speech signal) with values of the sound "fidelity". Among the attributes whose values most often correlate with the values of the sound "fidelity" are "spaciousness" and "clearness". "Sharpness" is the only attribute whose coefficient of correlation has a negative value. The value of the correlation coefficient in the case of input data of the second and third kind differ especially with respect to "loudness", "sharpness" and "lack of distortions". This proves that in the case of these attributes the influence of the interaction test signal x filter was significant.

Multiple regression considered separately for each kind of data showed that:

1. In the case of data of the first kind:

– in the case of the pink noise ($P1$) the values of attributes "spaciousness" and "lack of distortions" determined 97% of the variability of the value of the sound "fidelity"

– in the case of the speech signal ($P2$) no correlation has been found between any of the attributes under investigation

– in the case of a musical piece ($P3$), the values of the "spaciousness" of sounds reflected 90% of the variability of the values of "fidelity"

– in the case of the white noise ($P4$), the values of the "fullness" of sounds determined 82% of the variability of the "fidelity" of sounds.

2. In the case of data of the second kind, values of all the attributes of the perception space determined 74% of the variability of the values of the sound "fidelity", and the same values of the "fullness" of sounds determined 76% of the variability.

3. In the case of data of the third kind, values of all the attributes of the perception space determined 73% of the variability of the values of "fidelity", and the values of three of them: "fullness", "lack of distortions" and "sharpness" determined 78% of "fidelity".

5. Conclusions

1. The relation between the evaluation of "fidelity" and the attributes of the perception space depends on the test signal. In the case of speech no relation between the attributes of the sound perception space and the evaluation of the sounds "fidelity" with respect to the mental model was found. This can be due to the limited number of measured elements (the degree of freedom was 3) as well as to the lack of

differentiation of the evaluation of "fidelity" of sounds when speech was the test signal (see Fig. 2).

2. The relationship between the attributes of the perception space and the evaluation of the "fidelity" of sounds indicates which attributes of the perception space significantly determine its evaluation with respect to the mental model.

Correlation between values of the evaluation of "fidelity" and values of the evaluation of selected attributes of the sounds (Table 4) shows that such attributes as "spaciousness", "fullness" and "clearness" correlate with the evaluation of the sound "fidelity". It was also found that the selected attributes of the perception space are, for a specific test signal, mutually correlated. In the case of the pink noise and music piece there is a correlation between the values of "fullness", "spaciousness" and "clearness" and in the case of the white noise there is a correlation between "sharpness" and "loudness".

Elimination of the influence of test signals on the results of the regression analysis helped isolate attributes which, irrespective of the kind of test signal used, are a significant contribution to the evaluation of sound "fidelity" with respect to the mental model. These are: "fullness", "sharpness" and "lack of distortions".

3. When considering the influence of the attribute "lack of distortions" on the evaluation of "fidelity", one should also take into account different kinds of distortions, e.g., linear distortions and nonlinear distortions. In the case under discussion, all kinds of distortion were considered simultaneously.

4. It requires explanation whether the evaluation of sound "fidelity" made with respect to a model other than mental leads to a similar relationship with selected attributes of the perception space. Explanation of this problem should facilitate the search for a uniform relationship between the subjective evaluation of the "fidelity" of sounds reproduced by loudspeakers and their physical parameters.

Acknowledgements

The work has been sponsored by the project C.P.B.P. 02.03.07. The authors would like to thank Prof. H. Ryffert, project coordinator, for her valuable comments in the course of the investigations.

References

[1] F. E. Toole, *Subjective measurement of loudspeaker sound quality and listener performance*, J. Audio Eng. Soc., **33**, pp. 33–53, (1985 Jan).

[2] H. Staffeldt, *Correlation between subjective and objective data for quality loudspeakers*, J. Audio Eng. Soc., **22**, pp. 402–415, (1974 July/Aug).

[3] E. Hojan, A. Furmann, *Wahl der bei der Beurteilung akustischer Signale wesentlichen Attribute des Perzeptionsraumes*, Forschritte der Akustik – DAGA'87, pp. 589–592, (1987).

[4] G. Slot, *Audio Quality*, Philips Paperbacks, (1964).

[5] E. Villchur, *Simulated "Live vs Recorded". Test for Loudspeakers*, Audio, (November 1966).

[6] *International Electrotechnical Commission, Listening tests on loudspeakers*, Publ. 268–13: Sound systems equipment, Part 13, (1985).
[7] *Statgraphics, Statistical graphics system* by Statistical Graphics Corporation, Version 2.2, 1986.
[8] A. Gabrielsson, *Statistical treatment of data from listening tests on sound reproducing systems*, Reports from Technical Audiology, Karolinska Institute, Stockholm, No. 92, (1979).
[9] A. Luszniewicz, *General Statistics* (in Polish) PWE, Warszawa (1987).

Received September 28, 1989

ARCHIVES OF ACOUSTICS
16, 2, 295–303 (1991)

PREDICTION OF THE EARLY PART OF ECHOGRAMS INDISPENSABLE IN COMPUTER SIMULATION IN ROOMS

E. HOJAN

Institute of Acoustics, Adam Mickiewicz University
(60-796 Poznań, ul. Matejki 48/49)

CH. PÖSSELT

Lehrstul für allgemeine Elektrotechnik und Akustik Ruhr – Universität Bochum, Germany

The paper discusses the possiblility of an optimal selection of the early part of the room echogram during its calculation in computer simulation.

The existence of two such time segments in the echogram has been found.

In the first time segment it is necessary to follow precisely successive reflections. In the second time segment the reflections can be treated stochastically.

1. Introduction

Numerous works on speech acoustics, room acoustics, electroacoustics and psychoacoustics indicated the influence of transients on properties of transmitted sound signals [2, 3, 6, 7, 8, 9, 10, 14, 15, 16, 25, 26, 27]. Room acoustics employs a number of criteria for the evaluation of rooms. The criteria can be determined on the basis of their so-called room impulse response (echogram) [5]. The criteria differ with respect to the value of the early part of the echogram, if one considers the usefulness of successive reflections and the function of their weight [4, 11, 12, 13, 21].

In computer simulation of echograms of closed rooms it is very important to determine the size of the early part of the said echogram used to evaluate rooms, as it determines the calculation time [23]; there is no need to calculate the entire room echogram (theoretically during an indefinitely long time); it suffices to calculate its early part. Without going into details about the relations between particular criteria and whether or not they could be used in computer simulation [4, 24], we attempted to estimate the border temporal value of the early part of the echogram, in view of a change of an acoustic signal, perceived in subjective evaluation, connected with early parts of an echogram (of a different length) through the convolution function.

In order to accomplish the task specified above, speech and music signals recorded under the conditions of a free field, following their sampling, were fed into the computer. Having undergone the convolution operation there with the early parts of the room echogram (of a different length), obtained through computer simulation, following their digital-analog conversion, the signals were recorded at the computer output in the tape-recorder system.

The acoustic signals prepared as described above have been evaluated subjectively.

2. Recommended values of the time limit of the echogram's early part

Determination of the echogram of a real room requires that the following two factors be taken into consideration: disturbances which overlap the echogram (room's own noises, noises of the transmission track – the said problems are absent in computer simulation), and, on the other hand, the fact that it is practically possible to determine an echogram at a finite time T_i while its definition requires $T_i = \infty$.

Since an exponential function approximates the envelope of the decaying echogram, the value of time determined by the signal dynamics can be assumed to be the maximal time limit T_i. Given signal dynamics equal to 40 dB:

$$T_i = \frac{2}{3}T \qquad (1),\ [14]$$

where: T — room reverberation time.

In a cuboidal room with reflecting walls, the number of reflections ΔN per time unit Δt grows along with time square t:

$$\Delta N = \frac{4\pi c^2}{V}\cdot t^2\,\Delta t \qquad (2),\ [5]$$

where: c — sound propagation speed [m/s] V— room volume [m^3]. Designating the temporal density of reflections as

$$\bar{m} = \frac{\Delta N}{\Delta t} \quad [\mathrm{s}^{-1}], \qquad (3)$$

we can determine time t from Eq. (2):

$$t = \sqrt{\frac{V}{4\pi c^2}\cdot \bar{m}} \quad [\mathrm{s}] \qquad (4)$$

Substituting for $\bar{m}$ in the expression (4) the limit value resulting from the inability of the hearing organ to distinguish signals consisting of a number of impulses per second greater than $\bar{m}_{\max}$, we get

$$t_{\mathrm{st}} = \sqrt{\frac{V}{4\pi c^2}\cdot \bar{m}_{\max}} \quad [\mathrm{s}] \qquad (5)$$

CREMER [5] has interpreted the limit value t_{st} as the time of the beginning of static reverberation; for time values greater than t_{st} there is no point in precise studying of the amplitude-temporal structure of the room echogram. Substituting the value of 340 m/s for c in Eq. (5) we obtain

$$t_{st} = 4.5 \cdot 10^{-2} \sqrt{V \cdot \bar{m}_{\max}} \text{ [ms]} \tag{6}$$

The values of $\bar{m}_{\max}$ given in the literature, depending on the methodology of the investigations, are within the range of $\bar{m}_{\max} = 2000 \div 2800$ [s^{-1}] [5, 19, 13].

Assuming successive values of $\bar{m}_{\max} = 2000$ s^{-1}, and then $\bar{m}_{\max} = 2800$ s^{-1},

$$t_{st} = 2\sqrt{V} \text{ [ms]} \tag{7}$$

and

$$t_{st} = 2.4\sqrt{V} \text{ [ms]} \tag{8}$$

However the time limit which separates the useful echo from the useless one, in the sense of the disappearance of its perception, has been set by many authors at $40 \div 90$ ms. The value depends on the energy of all echoes coming from a given direction, the difference in the levels of the direct and reflected sounds, sound direction, the place of the sound reception in the room, the room reverberation time, the kind of sound signal (speech, music) [1, 5, 15, 18, 20, 22].

Hence we can state that there are two notions which define the early parts of the echogram from the standpoint of their psychoacoustic evaluation, T_i and t_{st}. The former defines echogram time T_i of the early part of the echogram, necessary for stabilized subjective evaluation of the acoustic sensation in a given room – an increase in the time segment T_i does not change this sensation. The latter defined duration t_{st} of the early part of the echogram, which is only a certain segment of the fragment limited by time T_i ($t_{st} < T_i$), in which a detailed amplitude-temporal structure of reflections must be taken into account. After this time t_{st} there is no need to take into account the detailed structure of the reflections; its stochastic distribution up to time T_i is sufficient.

In computer simulation of the room echogram detailed calculations are necessary within the time range of up to value t_{st}. The calculations relate to successive reflections approaching the signal registration point. Over t_{st} only a segment with a stochastic distribution of amplitude and duration $(T_i - t_{st})$ can be added to the echogram with time t_{st}. This distinction between time T_i and t_{st} helps considerably shorten the calculation time in computer simulation of the room echogram as well as the calculation time of the convolution function of the said echogram with a selected fragment of the acoustic signal.

Consequently, the procedure speeds up investigations with respect to the evaluation of the acoustic properties of rooms. Real signals of speech or music are evaluated in a computer simulated room whose physical parameters can be changed quickly and without any restrictions.

3. Computer simulation of room echogram

An attempt at the verification of the hypothesis about the possibility of distinguishing time T_i and t_{st} in an echogram, obtained by computer simulation of the distribution of the acoustic field in a room has been made. The method of virtual images has been applied.

An echogram was calculated for a room for which simultaneous binaural registration in the artificial head was simulated; the transfer response of the external ear was taken into account [17]. It was also possible to stimulate the change of the position of both the source and reception points of the sound.

The early parts of the room echogram obtained in the calculations, each of a different length, we convolved with a sampled test signal recorded under the conditions of a free field.

The test signal consisted of a fragment of speech (a sentence in German) or a fragment of music a (violin concerto) each lasting 10 s.

The evaluation of the acoustic signals, recorded at the computer output, following their digital-analog conversion, was conducted by the present authors through ear-phones. The main task was to determine whether the pairs of acoustic signals are different or not. The successive pairs of acoustic signals consisted of randomly combined signals, each of a different (growing) time T_i and then of a signal of a set time T_i combined with a signal of a different (growing) time t_{st}. Each pair was evaluated by each author ten times. The results of the measurements were analysed statistically.

4. Results

Three cuboidal rooms A, B and C were computer-simulated. Their dimensions, reverberation times and values of critical radius are given in Table 1.

During the first stage of investigations the initial fragment of the room echogram was calculated at the set localization of the sound source and reception point, for $T_i = (50, 100, 150, 200, 250, 300)$ ms; Fig. 1.

Figure 2 shows results of the calculations for $T_i = 100$ ms and Fig. 3 for $T_i = 300$ ms.

The difference in the subjective evaluation of musical signals of $T_i = 50$ and 200 ms is very clear and greater than that between 100 and 150 ms or 150 and 200 ms. The comparison of signals $T_i = 200$ and 250 ms and then of $T_i = 250$ and 300 ms indicates that the difference between them is imperceptible. The border value of the initial time of a fragment of the echogram T_i was estimated to be $T_i = 250$ ms.

In the case of speech signals, the general tendency in the evaluation of the differences is similar. One perceives not only differences in the global evaluation of signals but also with respect to specific attributes, e.g., spaciousness or timbre.

The investigations were repeated for room B of the same dimensions as room A but with a reduced reverberation time T.

Table 1. Values of reverberation time T and critical radius r in octave bands for computer simulated rooms, A, B, C

Room	Dimensions $b \times l \times h$ [m]		Octave mid frequencies [Hz] 125	250	500	1000	2000	4000	8000
A	15x20x5	Reverberation time [ms]	1033	1641	910	1669	1569	1309	1144
		Critical radius r [m]	2.2	1.7	2.3	1.7	1.8	1.9	2.1
B	15x20x5	Reverberation time [ms]	806	674	716	787	683	616	615
		Critical radius r [m]	2.5	2.7	2.6	2.5	2.7	2.8	2.8
C	13x20x4.5	Reverberation time [ms]	728	613	649	726	718	558	583
		Critical radius r [m]	2.3	2.5	2.4	2.3	2.5	2.6	2.7

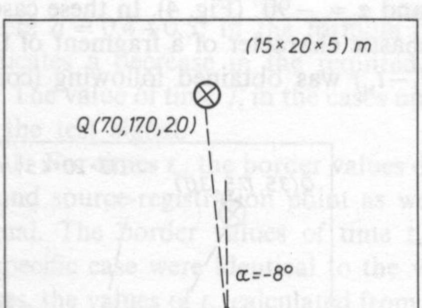


FIG. 1. Room A with source point Q and reception point H.

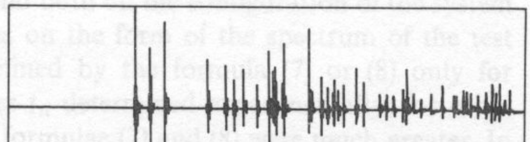


FIG. 2. Echogram of room A, at $T_i = 100$ ms.

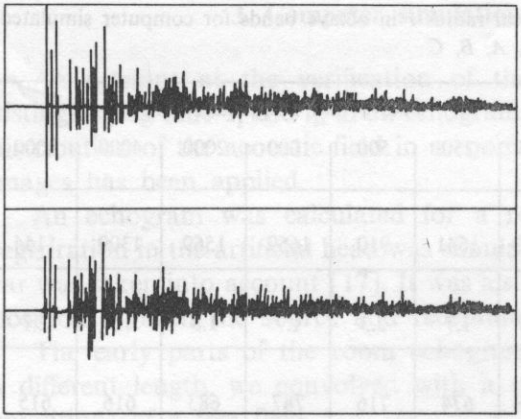


FIG. 3. Echogram of room A, at $T_i = 300$ ms.

Comparison of border values T_i with values of the rooms reverberation time T (Table 1) indicates that for the cuboidal room in question and for the given speech and music signals, the border values T_i are smaller than the values resulting from the formula (1): for example, in the octave band 500 Hz for room $A - T_i = 607$ ms, for room $B - T_i = 311$ ms.

In the second stage of the investigations the early parts of the echogram of room C were calculated (Table 1). The reciprocal configuration of the signal source and its reception point was such that the source with respect to the binaurally registered signal was localized at the angle of $\alpha = -15°$ and $\alpha = -90°$ (Fig. 4). In these cases, we adopted $T_i = 400$ ms = const, and the stochastic character of a fragment of the echogram of room C at the time segment $(T_i - t_{st})$ was obtained following (com-

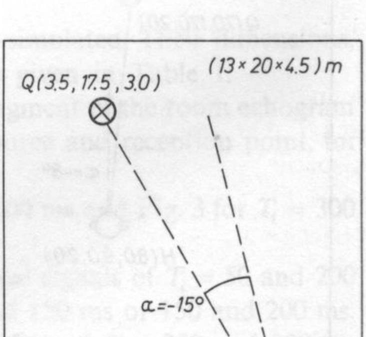


FIG. 4. Room C with source point Q and reception point H.

putational) alteration of the channels in the system of binaural registration after time $t_{st} = (5, 10, 15, 20, 25, 30, 50, 80, 100)$ ms. A subjective evaluation comparing acoustic signals obtained as a result of the convolution of test signals with fragments of echograms with different times t_{st}, after which they assumed stochastic character, showed that the border values of time t_{st} are within the range of $t_{st} = (80 \div 100)$ ms for speech signals and $t_{st} = (5 \div 10)$ ms for music signals — when $\alpha = -15°$. For $\alpha = -90°$, border values of $t_{st} = (5 \div 10)$ ms were obtained for both speech and music signals. The comparison of the border values of t_{st} with values of t_{st} resulting from the formula (7) — $t_{st} = 68$ ms and (8) — $t_{st} = 82$ ms indicates the lack of their uniform compatibility or incompatibility.

5. Final conclusions

The investigations help formulate the following conclusions:

1. The hypothesis about the possibility of distinguishing two early parts of the echogram with border times T_i and t_{st} was confirmed in the subjective evaluation of the acoustic signal prepered in computer simulation of the distribution of the acoustic field in a room.

2. The border numerical values of T_i indicates a possibility of modification of the formula (1) by the formula (9):

$$T_i = a \cdot T \tag{9}$$

where $a = 0.4 \div 0.5$; in the formula (1) the value of the coefficient $a = 0.67$. This indicates a decrease in the required signal dynamics to the order of 30 dB.

The value of time T_i in the cases under consideration does not depend on the kind of the test signal.

3. For times t_{st} the border values depend both on the configuration of the system sound source-registration point as well as on the form of the spectrum of the test signal. The border values of time t_{st} defined by the formula (7) or (8) only for a specific case were identical to the values t_{st} determined experimentally; for other cases, the values of t_{st} calculated from the formulae (7) and (8) were much greater. In each case the border values t_{st} are greater than the time after which the first reflection following a direct sound reaches the subject, what again confirms the assumptions about the significance of this reflection for the subjective evaluation of sound in closed rooms.

6. Conclusion

The problem of the estimation of the required duration of a fragment of the room echogram in computer simulation of the distribution of the acoustic wave in the same room is closely connected with the optimization of the conditions of such a simulation.

The results of the investigations indicate the possibility of distinguishing between two early time segments in the room echogram.

In the first segment, up to time point t_{st}, it is necessary to follow precisely successive reflection from room boundaries. At the time segment $T_i - t_{st}$, the reflections can be treated stochastically.

The formulae (1), (7), and (8) used to determine the value of time T_i and t_{st} generally give values much greater than the border values of these times obtained in the investigations under discussion.

Systematic investigations into the problem, which has only been outlined in the present paper, are necessary. They would help minimize the calculation time in computer simulation of acoustic processes in a closed room.

Acknowledgements

The authors would like to thank Prof. J. Blauert for his participation in discussions held during the investigations.

E. Hojan extends his gratitude to the von Humboldt Foundation for providing funds to cover the cost of his stay in Bochum.

References

[1] B. S. Atal, M. R. Schroeder, G. H. Sessler, *Subjective reverberation time and its relation to sound decay* ICA-5, G-32, Lüttich 1965.

[2] H. Backhaus, *Über die Bedeutung der Ausgleichvorgänge in der Akustik*, Z. für techn. Physik 13, 31–46 (1932).

[3] J. Blauert, *Spatial hearing*, MIT Press, Cambridge, Ma 1983.

[4] R. H. Bolt, P. E. Doak, *Tentative criterion for the short-time transient response of auditoriums*, JASA **22**, 4, 507–509 (1950)

[5] L. Cremer, H. A. Müller, *Die wissenschaftlichen Grundlagen der Raumakustik*. Band 1. S. Hirzel Verlag, Stuttgart 1978.

[6] L. F. Elfner, R. T. Tomsic, *Temporal and intensive factors in binaural lateralization of auditory transients*, JASA **43**, 4, 746–751 (1968).

[7] J. Florkowski, *Bedeutung der Transientsignale für Lokalisation der Schallquelle*. Dissertation, Universität–Poznań, 1980.

[8] E. Hojan, *Über die Hörsamkeit der Transientvorgänge mit verschieden Hüllkurven und Dauerlänge*. Dissertation Universität–Poznań, 1969.

[9] E. Hojan, J. Florkowski, M. Niewiarowicz, U. Kokowska, *Ausnutzung von Impulsmethoden für die Untersuchungen der Lautsprecher* (in Polish) Rozpr. Elektrotechniczne **23**, 1, 169–179 (1977).

[10] E. Hojan, M. Niewiarowicz, *Messsystem für die Regiestrierung des Einschwingsverhaltens von Lautsprechern*, Frequenz **39**, 11, 307–309 (1985).

[11] H. Kuttruff, *Über Autokorrelationsmessungen in der Raumakustik*, Acustica **16**, 3, 166–174, (1965–66).

[12] R. Kürer, *Zur Gewinnung von Einzahlkriterien bei Impulsmessungen in der Raumakustik*, Acustica **21**, 6, 370–372, (1969).

[13] H. Niese, *Die Messung der Nutzschall und Echogradverteilung zur Beurteilung der Hörsamkeit in Räumen*, Acustica **11**, 4, 201–213, (1961).

[14] P. LEHMANN, *Über die Ermittlung raumakustischer Kriterien und deren Zusammenhang mit subjektiven Beurteilungen der Hörsamkeit*, Dissertation, Technische Universität–Berlin, 1976.

[15] D. R. PERROTT, *Role of signal onset in sound localization*, JASA **45**, 2, 436–445 (1969).

[16] G. PLENGE, *Über die Hörsamkeit kleiner Änderungen der Impulsntwort eines Raumes*, Acustica **25**, 6, 315–325 (1971).

[17] C. PÖSSELT, J. SCHRÖTER, M. OPITZ, P. L. DIVENYI, J. BLAUERT, *Generation of binaural signals for research and home entertainment*, ICA-12, Bi-6, Toronto 1987.

[18] F. SANTON, *Numerical prediction of echograms and the intelligibility in rooms*, JASA **59**, 6, 1399–1045 (1976).

[19] L. SCHREIBER, *Was empfinden wir also gleichförmiges Rauschen*? Frequenz **14**, 12, 399–403 (1960).

[20] H. P. SERAPHIM, *Untersuchungen über die Unterschiedsschwelle Exponentiellen Abklingens von Rauschbandimpulsen*, Acustica **8**, 4, 280–284 (1958).

[21] H. P. SERAPHIM, *Über die Wahrnehmbarkeit mehrerer Rückwürfe von Sprachschall*, Acustica **11**, 2, 80–91 (1961).

[22] R. THIELE, *Richtungsverteilung und Zeitfolge der Schallrückwurfe in Raumen*, Acustica **3**, 4, 291–302 (1953).

[23] J. P. VIAN, *Different merits and their applications*, ICA–12, E4–10, Toronto 1987.

[24] M. VORLÄNDER, *Zur Genaudigkeit von Ray–Tracing–Berechnungen*, DAGA–87, Aachen 1987.

[25] W. M. HARTMANN, *Localization of sound in rooms, JASA* **74**, 5, 1380–1391 (1983).

[26] B. RAKERD, W. M. HARTMANN, *Localization of sound in rooms, II: The effects of a single reflecting surface*, JASA **78**, 2, 524–532 (1985).

[27] B. RAKERD, W. M. HARTMANN, *Localization of sound in rooms, III: Onset and duration* effects, JASA **80**, 6, 1695–1705 (1986).

Received February 14, 1989

ARCHIVES OF ACOUSTICS
16, 2, 305-323 (1991)

DETERMINATION OF PERCEPTUAL BOUNDARIES BETWEEN THE MALE FEMALE AND CHILD'S VOICES IN ISOLATED SYNTHETIC POLISH VOWELS[1]

J. IMIOŁCZYK

Department of Acoustic Phonetics, Institute of Fundamental Technological Research,
Polish Academy of Sciences
(61-704 Poznań, Noskowskiego 10)

The problem of perceptual boundaries between the male, female and child's voices was considered. The experimental material included 730 synthetic realisations of the six Polish oral vowels: /i/, /ɨ/, /e/, /a/, /o/ and /u/. „Target" male, female and child's utterances as well as a number of intermediate ones were obtained for each vowel by selecting appropriate combinations of FO and formant frequency values. Results of the listening test show that FO is the principal factor determining the perception of voice category and that, as such, it plays the key role in the perceptual normalisation of the speaker's vocal tract.

W pracy podjęto problem granic percepcyjnych między głosami męskim, kobiecym i dziecięcym. Jako materiał badawczy wykorzystano 730 syntetycznych wypowiedzi sześciu ustnych samogłosek polskich: /i/, /ɨ/, /e/, /a/, /o/ oraz /u/. Poprzez odpowiedni dobór wartości FO i częstotliwości formantowych zsyntetyzowano dla każdej z samogłosek „docelowe" realizacje męskie, kobiece i dziecięce oraz szereg realizacji pośrednich. Rezultaty badań odsłuchowych wskazują, że częstotliwość podstawowa jest zasadniczym czynnikiem decydującm o postrzeganej kategorii głosu i że odgrywa ona w związku z tym kluczową rolę w percepcyjnej normalizacji toru głosowego osoby mówiącej.

1. Introduction

Despite many years of continuous research and hundreds of experiments it has not yet been explained in what way variability of the speech signal, posing so great a problem e.g. in automatic speech recognition, is eliminated with remarkable ease and efficiency in the process of perception. In attempts to resolve this issue some authors (notably K. N. Stevens and S. Blumstein) have claimed that in the apparently variable signal there occur certain invariant features which provide guidelines to perceptual classification processes. Other investigators have considered this line of research unproductive and have suggested concentrating on attempts to find out how the listener deals with the variability which is there [11].

---

[1] This research was carried out within the CPBP 02.03 Problem.

One of the sources of variability facing the listener in speech perception are interspeaker differences. As noted by HOLMES [20], variations in the acoustic structure of a vowel spoken by a man, woman and child are so considerable that the phonetic equivalence of the three utterances **must** have a perceptual basis. He goes on to say that there probably exist **systematic relations** between acoustic features, which allow interpretation of markedly different acoustic signals as linguistically equivalent (p. 347). Thus, it seems valid to assume that (1) in order to efficiently eliminate the speaker-related variability from the speech signal the listener must be able to differentiate at least three major voice categories, viz. the male, female and child's voice and (2) the voice category information is contained in the signal in the form of appropriate relations between acoustic features.

The possibility of perceptual identification of a voice as male, female or child's implies the existence of more or less sharply defined boundaries between the three voice types. As can be expected, determining those boundaries and establishing which acoustic parameters affect them should contribute to a better understanding of the mechanism of perceptual normalisation of the speaker's vocal tract. These two aims have provided motivation for the present research.

2. Recording, analysis and resynthesis of natural vowels

At the initial stage, six Polish oral vowels, i.e. [i], [ɨ], [e], [a], [o] and [u], spoken by a male voice, were recorded under laboratory conditions. An RS 249–946 microphone and a cassette recorder Revox B710 were used for that purpose. Special care was taken to ensure that the level of the recording and the FO pattern (rising-falling) were the same for all the utterances.

The vowels were then low-pass filtered, sampled at 10 kHz and stored on disk of a MASSCOMP MC5400 computer. In their subsequent acoustic analysis, a specialised software packet, named AUDLAB, was used. Of the functions it included the following were applied:

- FO extraction
- calculation of momentary and average spectra
- preparation of 2- and 3-dimensional spectrograms
- measurement of segment duration.

On the basis of the data derived from the analysis, the vowels were (re)synthesized, using the Klatt software formant snythesiser[2]. Choosing the parallel branch of the synthesiser made it possible to control the amplitudes of individual formants. In order to make the vowels sound as natural as possible, the second of the two voicing sources („ss" = 2) was selected. The rate of D/A conversion was 10000 samples/sec.

Apart from the „standard" parameters of formant synthesis (eg. FO, bandwidths of spectrum envelope peaks etc.), two additional ones were used. One of them, "no"

---

[1] A later version of the programme described in [22] was used; see also [1].

corresponds to the number of samples in the open phase of the voicing source period and the other one, "tl", controls the spectral tilt. The principles of handling these two parameters will be described in Sections 3.2 and 4.2.

Vowels synthesized on the basis of the data from the analysis of natural vowels were analyzed again and the spectral characteristics of the two groups of vowels were compared. Necessary corrections in parameter values were introduced until the spectra of synthetic vowels closely matched those of their natural counterparts.

In order to minimize quality-unrelated differences between the particular synthetic vowels, their duration was made uniform (equal to 440 ms), and the same amplitude and FO pattern, extracted from the natural [e][3], were applied in all of them. The values of "no" and "tl" were also made identical.

Consequently, the differences between the synthetic vowels were limited to: (1) frequencies, (2) amplitudes and (3) bandwidths of formants as well as (4) the overall gain control which was used to make the vowels approximately equally loud.

In each vowel, the above-named parameters had constant values, typical of that vowel. Time-varying parameters such as amplitude of the voice source and FO were updated every 5 ms[4].

The FO contour is illustrated in Table 1. Changes between the discontinuity points ($t = 0$ ms, $t = 100$ ms etc.) are linear.

Formant frequencies of the six synthetic vowels are shown in Table 2.

Table 1. FO contour in synthetic male vowels

t [ms]	0	100	245	390	435
FO [Hz]	139.0	144.0	119.0	112.0	119.5

Table 2. Formant frequencies of synthetic male vowels

Formant [Hz]	Vowel i	ɨ	e	a	o	u
F1	250	365	600	800	620	340
F2	2170	1960	1720	1220	950	760
F3	3100	2530	2620	2760	2730	2250
F4	3770	3230	3050	3760	3920	3030
F5	4500	4100	4350	4850	4560	4140

---

3) Strictly speaking, both were aproximations of the patterns characterizing the natural [e].

4) Strictly speaking, those changes were delayed until the beginning of the next glottal pulse (cf. [22], p. 978).

3. Synthesis of female vowels

3.1. Formant frequencies

The vowel formant frequencies given in Table 2, assumed to be representative of a male voice, were subsequently scaled in such a way that for each original (male) vowel its phonetic equivalent was obtained with formant frequencies typical of a female voice. The *k* scaling factors given by FANT ([14], p. 87) were used for that purpose.

The length of the supralaryngeal vocal tract is known to be approx. 15...20 per cent shorter in the female than in the male (eg. [5], [14]). However, owing to a difference in proportion of the length of the oral to the pharyngeal cavity in both genders, the frequencies of the male and female formants cannot be related by means of a single factor. As shown by FANT ([14], [15]), in order to express the gender-related differences in the frequencies of the three lower vowel formants, three separate scaling factors (*k1*, *k2* and *k3*) are required which, in addition, depend on vowel type.

Table 3 presents the values of scaling factors adopted in the present work. They

Table 3. *k* scaling factors and formant frequencies of synthetic female vowels

Vowel	*k1* [%]	F1 [Hz]	*k2* [%]	F2 [Hz]	*k3* [%]	F3 [Hz]	*k4* [%]	F4 [Hz]	*k5* [%]	F5 [Hz]
i	8	270	22	2647	16	3596	17	4411	17	5265
ɨ	10	402	24	2430	20	3036	17	3779	17	4797
e	24	744	20	2064	19	3118	17	3569	17	5090
a	16	928	16	1415	16	3202	17	4399	17	5675
o	15	713	14	1083	15	3140	17	4586	17	5335
u	10	374	5	798	19	2678	17	3545	17	4844

are the result of a modification of FANT's data ([14], Table. 1), taking into account the articulatory-acoustic specificity of the Polish vowels. The values of *k4* and *k5*, not considered in [14], were in all cases equal to 17 per cent. Table 3 also shows the female formant frequencies obtained by appropriately increasing the male formants.

Owing to the differences in the *k* factors for the individual vowels, selecting the formant sets intermediate between the male and the female formant values required that each vowel should be treated separately. It was assumed that the differences in F1, F2 and F3 between the consecutive formant sets, corresponding to a shortening of the vocal tract, would not be greater than the difference limens given by FLANAGAN [16], equal to: ±20 Hz for F1, ±50 Hz for F2 and ±75 Hz for F3. Consequently, the following numbers of formant sets were obtained for each vowel: [i] – 11, [ɨ] – 11, [e] – 9, [a] – 8, [o] – 7 and [u] – 7. The sets are presented in Table 4.

Table 4. Male/female formant frequency sets

Vowel	Formant set	F1, ΔF1 [Hz]	F2, ΔF2 [Hz]	F3, ΔF3 [Hz]	F4, ΔF4 [Hz]	F5, ΔF5 [Hz]
	1	250	2170	3100	3770	4500
i	…	$\Delta = 2$	$\Delta = 47.7$	$\Delta = 49.6$	$\Delta = 64.1$	$\Delta = 76.5$
	11	270	2647	3596	4411	5265
	1	365	1960	2530	3230	4100
ɨ	…	$\Delta = 3.7$	$\Delta = 47.0$	$\Delta = 50.6$	$\Delta = 54.9$	$\Delta = 69.7$
	11	402	2430	3036	3779	4797
	1	600	1720	2620	3050	4350
e	…	$\Delta = 18.0$	$\Delta = 43.0$	$\Delta = 62.2$	$\Delta = 64.9$	$\Delta = 92.5$
	9	744	2064	3118	3569	5090
	1	800	1220	2760	3760	4850
a	…	$\Delta = 18.3$	$\Delta = 27.9$	$\Delta = 63.1$	$\Delta = 91.3$	$\Delta = 117.9$
	8	928	1415	3202	4399	5675
	1	620	950	2730	3920	4560
o	…	$\Delta = 15.5$	$\Delta = 22.2$	$\Delta = 68.3$	$\Delta = 111.0$	$\Delta = 129.2$
	7	713	1083	3140	4586	5335
	1	340	760	2250	3030	4140
u	…	$\Delta = 5.7$	$\Delta = 6.3$	$\Delta = 71.3$	$\Delta = 85.8$	$\Delta = 117.3$
	7	374	798	2678	3545	4844

In the synthesis of all the utterances representing the same vowel, formant amplitudes and bandwidths were constant and identical with those in the original ("male") set. Formant frequencies higher than 5 kHz were not considered.

3.2. *Fundamental frequency*

The fundamental frequency of a female voice is, on average, an octave higher than that a male voice (e.g. [13], p. 242) and amounts to approx. 220 Hz. In the present work, fundamental frequency equal to 1.7 that of the male was assumed, after [23], as typical of a female voice. The female FO pattern and 6 intermediate ones were obtained by multiplying FO values at each of the discontinuity points of the male pattern by an appropriate scaling factor (1.1, 1.2… 1.7). The average FO value was 128 Hz in the male pattern and 218 Hz in the female one (see Table 5).

As shown by a number of authors (eg. [19], [23], [28]), the proportion of duration of the open and the closed phase of the glottal period (the open qotient) is greater in the female voice and the shape of the glottal pulse is more symmetrical. As a consequence, the female voice is characterized by a steeper spectral tilt: whilst the rate of the spectral fall-off in a typical male voice averages −12 dB/octave ([12], [13], [17]), in a female voice, it can be as sharp as −15…−18 dB/octave ([23],

Table 5. Male/female FO contours and glottal period characteristics

scal. fact.	FO [Hz] Oms	100ms	245ms	390ms	435ms	mean FO [Hz]	no	% peri. dur.	tl
1.0	139.0	144.0	119.0	112.0	119.5	128	33	42%	0
1.1	152.9	158.4	130.9	123.2	131.5	141	33	47%	3
1.2	166.8	172.8	142.8	134.4	143.4	154	33	51%	6
1.3	180.7	187.2	154.7	145.6	155.4	166	33	55%	9
1.4	194.6	201.6	166.6	156.8	167.3	179	33	59%	12
1.5	208.5	216.0	178.5	168.0	179.3	192	33	63%	15
1.6	222.4	230.4	190.4	179.2	191.2	205	33	68%	18
1.7	236.6	244.8	202.3	190.4	203.2	218	33	72%	21

[28]). It has to be noted, however, that in the case of phonation at high pitch, the spectral tilt in a male voice can also approach −18 dB/octave ([28]).

In view of the above observations, the following two principles relating to the voice source characteristics were adopted in the present work:

(1) with an increase in mean FO, the spectral tilt became steeper; it was −12 dB/octave for "tl" = 0, and ca. −17 dB/octave for "tl" = 21.

(2) duration of the open phase of the period was in all the cases equal to 3.3 ms ("no" = 33), which means that the open quotient ranged from 42 to 72 per cent.

Each of the eight FO patterns (Table 5) was combined with each of the formant sets distinguished for the individual vowels (Tables 4...9). A total of 424 stimuli were synthesized at this stage.

4. Synthesis of child's vowels

4.1. Formant frequencies

The synthesis of child's vowels was based on FANT's claim (cf. [14]) that the differences in formant frequencies between female and child's vowels can be expressed by means of a single scaling factor independent of vowel type. All the female formant frequencies (i.e. the ones occuring in the last sets in Table 4) were consequently increased by the factor of $k = 18$ per cent (see Table 6).

Table 6. Child's formant frequencies

Vowel	F1 [Hz]	F2 [Hz]	F3 [Hz]	F4 [Hz]
i	319	3123	4243	5205
ɨ	474	2867	3582	4459
e	878	2436	3679	4211
a	1095	1670	3778	5191
o	841	1278	3705	5411
u	441	942	3160	4183

Apart from the "target" child's vowels, vowels with intermediate (between female and child's) formant frequencies were also synthesized. Similarly as before (cf. section 3.1), it was assumed that *Δk1* will be less than 20 Hz, *Δk2* < 50 Hz, *Δk3* < 75 Hz. As a result, the following numbers of formant sets were obtained for each of the vowels: [i] — 10, [ɨ] — 9, [e] — 8, [a] — 9, [o] — 8 and [u] — 7. The sets are presented in Table 7. Their numbering continues the numbering from Table 4.

Table 7. Female/child's formant frequency sets

Vowel	Formant set	F1, ΔF1 [Hz]	F2, ΔF2 [Hz]	F3, ΔF3 [Hz]	F4, ΔF4 [Hz]
i	12 … 21	275 Δ = 4.9 319	2695 Δ = 47.6 3123	3661 Δ = 64.7 4243	4490 Δ = 79.4 5205
ɨ	12 … 20	410 Δ = 8.0 474	2479 Δ = 48.5 2867	3097 Δ = 60.6 3582	3855 Δ = 75.5 4459
e	10 … 17	761 Δ = 16.7 878	2111 Δ = 46.4 2436	3188 Δ = 70.1 3679	3649 Δ = 80.3 4211
a	9 … 17	947 Δ = 18.5 1095	1443 Δ = 28.4 1670	3266 Δ = 64.0 3778	4487 Δ = 88.0 5191
o	8 … 15	729 Δ = 16.0 841	1107 Δ = 24.4 1278	3211 Δ = 70.6 3705	4689 Δ = 103.1 5411
u	8 … 14	384 Δ = 9.5 441	819 Δ = 20.5 942	2747 Δ = 68.8 3160	3636 Δ = 91.2 4183

As in the previous stage, formant amplitudes and bandwidths were not changed. Formant frequencies exceeding 5 kHz were disregarded.

4.2. Fundamental frequency

According to Fant ([13], p. 242), voice fundamental frequency in a child at the age of 10 averages 300 Hz, although the interspeaker variation may be considerable. Similar, though somewhat lower values are quoted by Hasek and Singh [18] for 5…10 year old boys and girls.

Considering the above observations, furhter five FO patterns were synthesized in the way described in section 3.2. The scaling factors used ranged from 1.8 to 2.2. For want of data concerning the glottal tone characteristics of a child's voice, the value of

Table 8. Female/child's FO contours and glottal period characteristics

scal. fact.	FO [Hz] Oms	100ms	245ms	390ms	435ms	mean FO [Hz]	no	% peri. dur.	tl
1.8	250.2	259.2	214.2	201.6	215.1	230	26	60%	21
1.9	264.1	273.6	226.1	212.8	227.1	243	26	63%	21
2.0	278.0	288.0	238.0	224.0	239.0	256	26	67%	21
2.1	291.9	302.4	249.9	235.2	251.0	269	26	70%	21
2.2	305.8	316.8	261.8	246.4	262.9	282	26	72%	21

"tl" was in all cases the same as in the "target" female pattern (cf. Table 5) and the open quotient varied from 80% to 72%. The relevant figures are given in Table 8.

Each of the five FO patterns and, additionally, the pattern with the scaling factor of 1.7 (see Table 5) were combined with each of the formant sets presented in Table 7. As a result, 306 stimuli were obtained.

5. Listening test

5.1. Test material. The manner of presentation

The total number of stimuli generated amounted to 730 (424+306). The set comprized: 148 [i]s, 142 [ɨ]s, 120 [e]s, 118 [a]s, 104 [o]s, and 98 [u]s. The material was randomized using a special procedure included in the software package used for the synthesis and recorded on a Revox B710. The inter-stimulus interval (ISI) on the test tape was 3 secs, which, according to [7], is the decay time of the auditory memory in a vowel discrimination task and the time after which context effects disappear in vowel identification.

The interval between groups of 10 stimuli was 4.5 secs.

20 subjects with no known hearing impairments participated in the listening experiment. Their task consisted in identifying each stimulus as one of the six vowels and classifying it as an utterance by a man, a woman or a child. This was done by putting in appropriate two-letter symbols (e.g. "im" – [i], male voice, "af" – [a], female voice, "uc" – [u], child's voice) in an answer sheet.

The material was presented to the subjects in two sessions a few days apart. At the beginning of each session the subjects were instructed as to their task and the set of 18 "target" male, female and child's vowels was played to them in random order.

5.2. Results

5.2.1. Vowel identification. Of the total of 14 600 responses 557 were incorrect (3.8%), [o] being by far the most frequently misidentified vowel. Error rates for the individual vowels are given in Table 9.

Table 9. Vowel identification errors

Vowel present.	No of stimuli	No of errors	% errors	misidentified* as
u	1960	2	0.1	−(2)
e	2400	4	0.2	−(4)
i	2960	18	0.6	ɨ (14), u(1), −(3)
a	2360	28	1.2	o(24), −(4)
ɨ	2840	190	6.7	i(145), e(38), −(7)
o	2080	315	15.1	a(312), −(3)

* "−" in this column denotes the lack of response in an answer sheet

Percentages of identification errors for [u], [e], [i] and [a] can be said to comply with the "norm". This is confirmed by the random distribution of incorrect responses obtained for these vowels. On the other hand, a considerable number of mistakes which occurred in the identification of [ɨ] and, especially, [o] seem to point to a "deficiency" in the acoustic structure of (some) stimuli representing these vowels. A probable explanation of this fact will be offered in section 6.1.

5.2.2. Recognition of voice category. When summing up the number of occurrences of voice qualifiers ("m" for male, "f" for female and "c" for child's) ascribed to the particular stimuli, all the responses containing a vowel identification error were ignored. With respect to some stimuli, especially of the [o] type, this considerably limited the number of "effective" voice category recognitions.

Figures 1...6 present the boundaries between the three voice categories deter-

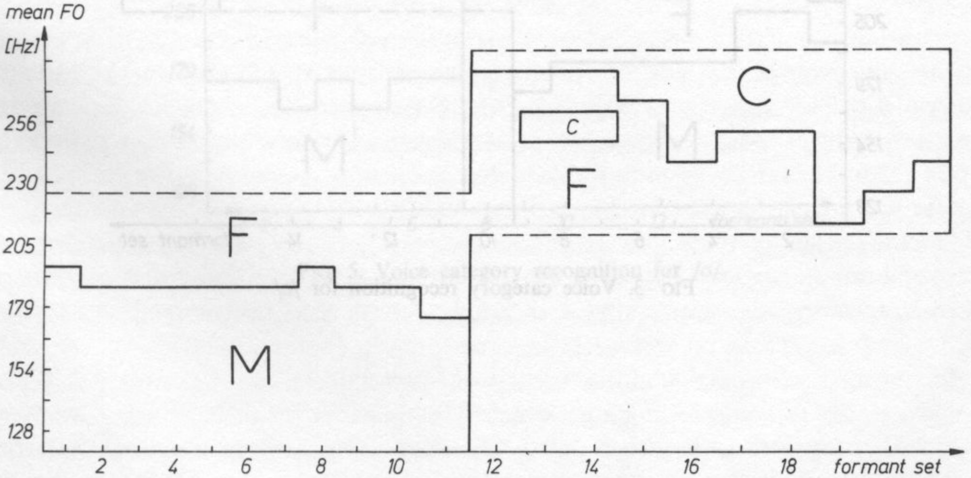


FIG. 1. Voice category recognition for /i/

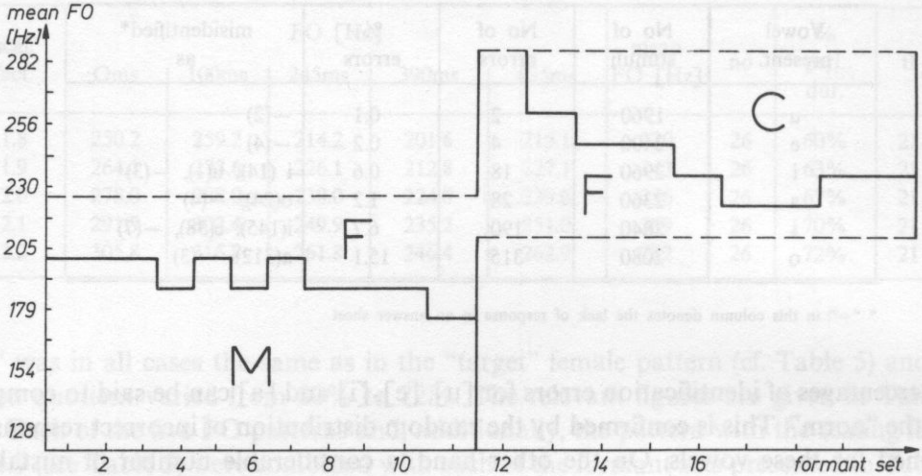


FIG. 2. Voice category recognition for /ɨ/

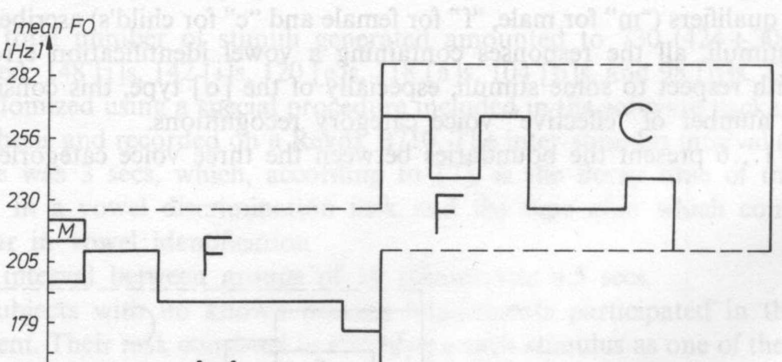


FIG 3. Voice category recognition for /e/

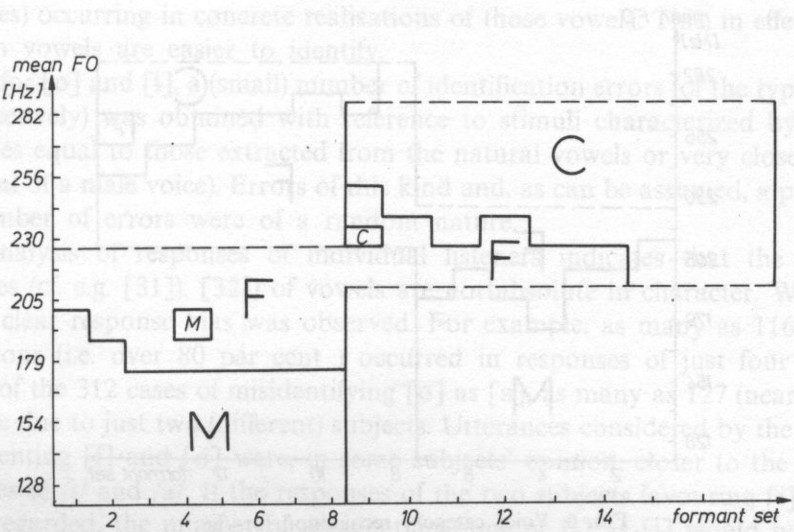


FIG. 4. Voice category recognition for /a/

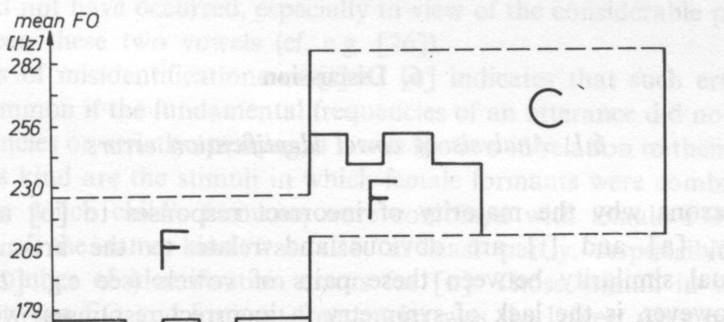


FIG. 5. Voice category recognition for /o/

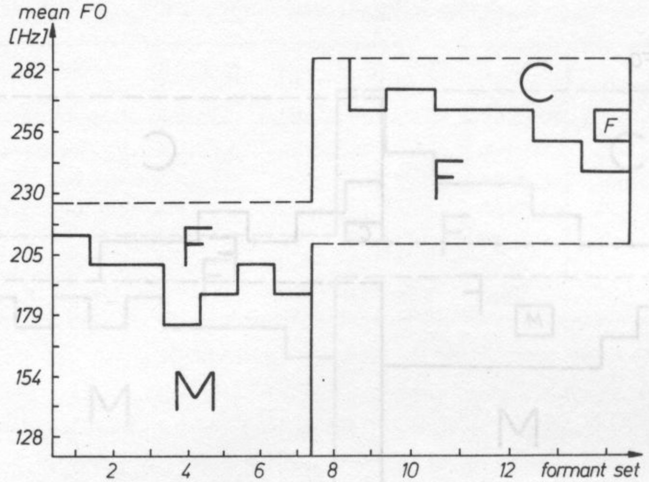


FIG. 6. Voice category recognition for /u/

mined for the individual vowels. The criterion used was the number of "m", "f" or "c" responses dominant for any given combination of vowel formants and an FO pattern.

6. Discussion

6.1. Analysis of vowel identification errors

The reasons why the majority of incorrect responses to [o] and [ɨ] were, respectively, [a] and [i] are obvious and relate to the articulatory-acoustic-perceptual similarity between these pairs of vowels (see e.g. [26]). What is striking, however, is the lack of symmetry in incorrect responses: whilst [o] was recognized as [a] as many as 312 times, the reverse substitution only occurred 24 times; similarly, [ɨ] was perceived as [i] 145 times, whereas [i] was taken for [ɨ] in only 14 cases. These facts seem to indicate that the scaling factors adopted for [o] and [ɨ] were too high, which resulted in excessive raising of F2 and F3 in the female and child's realisations of these vowels and, consequently, their increased acoustic and perceptual similarity to [a] and [i]. While such an explanation cannot be categorically dismissed (especially in the case of [o]), a few other facts should be noted which are of some relevance in this context.

According to Stevens' quantal theory of speech [33], [i], [a] and [u] have a special status in spoken language which is manifested by their forming discrete perceptual categories rather than being identified as points on a continuum. As a result, with such (quantal) vowels the perceptual classificatory mechanisms appear to be more tolerant to various acoustic deviations (from the appropriate phonetic prototypes) occurring in concrete realisations of those vowels. This, in effect, means that such vowels are easier to identify.

Both for [o] and [ɨ], a (small) number of identification errors (of the type [a] and [i], respectively) was obtained with reference to stimuli characterized by formant frequencies equal to those extracted from the natural vowels or very close to them (i.e. typical of a male voice). Errors of this kind and, as can be assumed, a part of the total number of errors were of a random nature.

An analysis of responses of individual listeners indicates that the phonetic prototypes (cf. e.g. [31]), [32]) of vowels are not absolute in character. With some subjects, clear response bias was observed. For example, as many as 116 [ɨ] → [i] substitutions (i.e. over 80 per cent) occurred in responses of just four persons; likewise, of the 312 cases of misidentifying [o] as [a], as many as 127 (nearly 42 per cent) were due to just two (different) subjects. Utterances considered by the majority as representing [ɨ] and [o] were, in some subjects' opinion, closer to the phonetic prototypes of /i/ and /a/. If the responses of the two subjects favouring [i] over [ɨ] were disregarded, the number of misidentifications or [ɨ] as [i] would not exceed 2 for any of the stimuli.

Apart from the 145 cases in which [ɨ] was perceived as [i], there also occurred 38 responses classifying this vowel as [e]. This would contradict the hypothesis that the scaling factors for [ɨ] were too high. If that indeed had been the case, reactions of the [e] type should not have occurred, especially in view of the considerable perceptual distance between these two vowels (cf. e.g. [26]).

An analysis of misidentifications of [ɨ] as [e] indicates that such errors were particularly common if the fundamental frequencies of an utterance did not "fit" the formant frequencies or, strictly speaking, if it was too low in relation to them. Typical example of this kind are the stimuli in which female formants were combined with male FO or in which child's formants were combined with female FO.

A "misfit" of the same kind was also, at least partly, responsible for the considerable number of identification errors for [o]. Those stimuli in which the discrepancy between FO and formant frequencies was the largest were most often identified as [a]. For example, stimuli combining female formants and male FO as well as child's formants and female FO were classified as [a] by as many as 16 subjects.

Undoubtedly, both FO itself and the distance F1–FO play a role in the perception of vowel height (e.g. [3], [9], [27], [35], [36], [37]). TRAUNMÜLLER [36] has shown, for example, that the perceived vowel openness changes with FO, even if formant frequencies remain constant. CARLSON et al. [3] conclude that for a vowel to retain its phonetic identity an increase in FO must be accompanied by an appropriate increase in formant frequencies.

It has to be noted here that the misfit between FO and formant frequencies in some stimuli is a consequence of the very design of the present experiment. The objective was not to generate "prototypical" (i.e. easily identifiable) vowel tokens but stimuli in which various formant sets would be combined with various FO patterns.

Considering the "non-prototypical" and, perhaps, even unrealistic nature of some stimuli (e.g. female formants with male FO), identification **must** have provided difficulties, especially with non-quantal vowels and those perceptually similar to others (e.g. [ɨ] and [o]). The difficulty of the task was additionally increased by the variety of "voices" (no two stimuli in the whole set were identical) occurring in random order (cf. [21]). This manner of presentation forced the listeners to treat each stimulus separately, i.e. without referring it to the ones heard previously.

From the acoustic point of view, realisations of the same vowel by a male and child's voice are markedly different ([20]). In spite of that, listeners can, of course, identify such two utterances as representing the same phoneme ([38]). This perceptual equivalence is arrived at by way of voice normalisation, the mechanism of which is not yet fully understood. An immediate proof of the reality of this process in speech perception is a longer reaction time in the identification of vowels produced by a number of different speakers in comparison with a single speaker vowel identification task ([34]). This means that voice normalisation (so frequent in the present experiment) makes vowel identification more difficult and can decrease its effectiveness.

Probably due to the random order of stimulus presentation, no vowel contrast effect occurred in the responses (cf. e.g. [4], [30], [31])[5].

An interesting aspect of the problem of discrepancy between FO and formant frequencies in an utterance is highlighted by [8], [24] and [38]. It appears that a skilful insertion of one word spoken by a man imitating a high FO of a child's voice into a recording of a phrase spoken by a child results in a considerable decrease in the recognition score of the vowel(s) contained in this word. This primarily confirms the importance of phonetic context in speech sound perception: the preceding context provides a reference frame which, under natural communication conditions, facilitates (optimizes) identification of the speech sounds that follow ([38]). In the case of a "mystification" such as was used in the papers quoted, the lack of agreement between the acoustic structure of the male utterance and the perceptual reference frame was bound to cause a high identification error rate (ca. 54 per cent in [8]). It has to be noted that while the imitation of child's FO by the male speaker was quite successful, the difference in formant frequencies between realisations of the same vowel by the two speakers was considerable. In [38], for example, male [ɛ] was most similar to child's [œ] and male [œ] showed greatest similarity to child's [y]. As can be expected, even if stimuli of this kind were presented in isolation, the discrepancy between FO and formant frequencies would result in erroneous recognition of the (intended) vowels. Indeed, this was the case with some [o] stimuli in the present experiment. The reason why some female [o]s were recognized as /a/ was that their formant frequencies only slightly differed from the formant frequencies of male [a], which, in turn, were almost identical with those of child's [o] (cf. Tables

[5] With 3 sec ISI, this effect should, in principle, be negligible (cf. [7]).

4 and 7). In those cases where FO was too low, this inevitably led to a change of phonetic category.

The perceptual identity of vowels is commonly claimed to be determined by their formant frequencies, especially F1 and F2. It is also emphasized that, unlike consonant perception, the perception of vowels is non-categorical, which means that more vowels can be discriminated than identified ([6], [25], [30]). Modifying formant frequencies within certain limits may, thus, lead to a change in perceived voice quality, but not necessarily to a change in phonetic category. Similar effects can be obtained by manipulating FO in an utterance with fixed formants ([36]). For a change in phonetic category to take place, the range of FO or formant shifts must exceed certain critical values (as can be expected, formant shifts play a more decisive role in this respect).

It follows from the above that vowel perception is determined not only by formant frequencies alone, but also by their relation to FO. This conclusion is contradicted by the findings of SUMMERFIELD and HAGGARD [34] who claim that while the significance of FO in perceptual identification of voices is beyond doubt, its role as a normalizing factor in vowel recognition is negligible. It has to be noted, however, that in the work quoted the difference in FO between voice I and the remaining ones amounted to just 20 Hz, which was probably too little for the normalizing "action" of this parameter to take place (all the four voices represented the same, i.e. male, voice category).

VAN BERGEM et al. [38] put forward a hypothesis according to which separate male, female and child's vowel templates exist in subjects' memory, formed on the basis of past language experience. Identification errors that occurred in their experiment are explained by the authors as resulting from the confusion of templates by the listeners.

No matter whether the above claim is valid or not, it seems certain that FO carries important information on voice type and, therefore, plays a crucial role in voice normalisation, thus making verbal communication more efficient. The results of perceptual voice categorisation, discussed below, support this supposition.

6.2. *Voice categorisation*

In order to ensure maximally objective conditions of the experiment and to avoid influencing the subjects' decision criteria in any way, no information was provided to the subjects as to the purpose of the experiment. In particular, they were not told that in the material to be presented they might come across high male voices or low female voices. Their decisions were thus fully independent in all cases.

Just as in vowel identification, response bias was observed in some listeners' answers, evidencing the relative nature of voice category "prototypes". Two subjects used the "woman" response as a reaction to stimuli judged by the majority to represent a male voice and one other consistently applied this label (i.e. "woman") in reference to stimuli predominantly classified as child's.

An analysis of the responses in Figures 1 ... 6 indicates that the preceding context, which proved not to be significant in the vowel identification task, had some effect on voice categorisation results. As can be seen, some stimuli clearly differ from their immediate surrounding with respect to the type of voice category qualifiers ascribed to them. Both contrast and attraction effects were observed. If, for example, the difference between two successive stimuli was considerable and the former was judged as "definitely male", the subjects more uniformly classified the latter as representing a female voice (contrast effect). If, on the other hand, the preceding stimulus was considered male and the difference between this stimulus and the next was not great, the tendency prevailed for this following stimulus to be classified as male as well (attraction effect).

6.2.1. *Male voice and female voice.* Voice pitch was undoubtedly the principal factor determining the listeners' classification of utterances as representing male or female voices. This is evidenced by the fact that (1) high formant frequencies in themselves did not guarantee that the "woman" response would predominate and (2) even utterances with low (i.e. potentially male) formants but high FO were classified as female. In the data obtained, the "woman" responses begin to appear at mean FO value of 179 Hz, and prevail at mean FO equal to 192 Hz. It is in this frequency range that the (perceptual) boundary between the male and female pitch probably lies.

It has to be noted that even in the case of stimuli which, owing to a misfit between FO and formant frequencies, were misidentified by a number of subjects, the voice qualifier in those erroneous responses was predominantly the same as in the majority of the correct ones. This fact confirms both the dominant role of fundamental frequency in determining voice category and the effect of FO on vowel identification.

6.2.2. *Female voice and child's voice.* The perceptual boundary between the female and child's voices seems somewhat more fuzzy than this between female and male voices. This may have resulted from the indefiniteness of the child's voice category or the indefiniteness of the very notion of the "child". Whilst the voice of a 5-year-old child is relatively easy to identify, differences in the acoustic structure of utterances produced by a woman and a 13-year-old boy may not be great. Obviously, it is difficult to establish what definitions of the "child" were adopted by the individual listeners for the purposes of the present experiment.

The indefiniteness of the child's voice caused, among others, greater contrats and attraction effects than those noted for stimuli around the male-female boundary (cf. Figures 1...6). Also, the choice between the "woman" and "child" responses seemed to be affected to a greater extent by formant frequencies of the stimulus, although at high mean FO values the „child" response predominated irrespective of formant frequencies. For the combined data, the perceptual boundary between the female and the child's voice lies in the range of mean FO values between 230 Hz and 243 Hz.

Slight deviation from this pattern can be observed for [u], in which case the boundary is somewhat shifted towards higher frequencies.

7.Conclusions

Contrary to what is still sometimes assumed, the perceptual phonetic identity of a vowel is not determined solely by its formant frequencies: they can only be interpreted on the basis of the information supplied by fundamental frequency. For a vowel to be perceptually distinct (and easily identifiable), its formant frequencies must combine with *appropriate* FO. If FO is too low in relation to the formants, the perceived vowel becomes more open; if, on the other hand, FO is too high, the vowel is perceived as more close.

Establishing voice category seems an indispensable condition of correct vowel identification. In other words, in order to understand **what** has been said it is first necessary to know **which voice category** the speaker represents. As the results obtained show, the largely necessary factor determining the perceptual classification of a given voice as male, female or child's is fundamental frequency. This goes to prove that FO guides the process of vocal tract normalization and thus makes spoken communication more efficient.

References

[1] J. ALLEN, M. S. HUNNICUTT, D. H. KLATT, *From Text to Speech: The MITalk System*, Cambridge University Press, Cambridge 1987.

[2] *AUDLAB User's Guide*, Centre for Speech Technology Research, University of Edinburgh, Edinburgh 1987.

[3] R. CARLSON, G. FANT, B. GRANSTRÖM, *Two-formant models, pitch, and vowel perception*, in: G. Fant and M. A. A. Tatham (eds.): Auditory Analysis and Perception of Speech, Academic Press, London 1975, pp. 55–82.

[4] K. CENTMAYER, *Interrelations of vowel perception and linguistic context*, in: G. Fant and M.A.A. Tatham (eds.): Auditiory Analysis and Perception of Speech, Academic Press, London 1975, pp. 143–152.

[5] T. CHIBA, M. KAJIYAMA, *The Vowel – Its Nature and Structure*, Tokyo 1941.

[6] L. CHISTOVICH, *Central auditory processing of peripheral vowel spectra*, J. Acoust. Soc. Amer., 77, 789–805 (1985).

[7] R. G. CROWDER, *Decay of Auditory Memory in Vowel Discrimination*, J. Exptl. Psychol.: Learning, Memory and Cognition, 8, No 2, 153–162 (1982).

[8] D. DECHOVITZ, *Information conveyed by vowels: a confirmation*, Haskins Laboratories Status Report, SR – 51/52, 213–219 (1977).

[9] M.-G. DI BENEDETTO, *On vowel height: acoustic and perceptual representation by the fundamental and the first formant frequency*, Proceedings of the 11th International Congress of Phonetic Sciences in Tallin, vol. 5, 198–201 (1987).

[10] P. D. EIMAS, J. L. MILLER, P. W. JUSCZYK, *On infant speech perception and the acquisition of language*, in: S. Harnad (ed.): Categorical Perception: The Groundwork of Cognition, Academic Press, Cambridge 1987, pp. 161–195.

[11] J. L. ELMAN, J. L. MC CLELLAND, *Exploiting Lawful Variability in the Speech Wave*, in: J. S. Perkell,

D. H. Klatt (eds.): Invariance and Variability in Speech Processes, Lawrence Erlbaum Assoc., Hillsdale 1986, pp. 360–380.

[12] G. FANT, *On the predictability of formant levels and spectrum envelopes from formant frequencies*, in: M. Halle, H. G. Lunt, M. McLean (eds.): For Roman Jakobson, Mouton, The Hague 1956, pp. 109–120.

[13] G. FANT, *Acoustic Theory of Speech Production*, Mouton, The Hague 1970 (2nd ed.).

[14] G. FANT, *A Note on Vocal Tract Size Factors and Nonuniform F-Pattern Scalings, in*: G. Fant: Speech Sounds and Features, The MIT Press, Cambridge, Massachusets, 1973, pp. 84–93.

[15] G. FANT, *Nonuniform Vowel Normalization*, Speech Transmission Laboratory QPSR, 2–3, 1–19 (1975).

[16] J. L. FLANAGAN, *Estimates of the Maximum Precision Necessary in Quantizing Certain "Dimensions" of Vowel Sounds*, J. Acoust. Soc. Amer., **29**, 533–534 (1957).

[17] J. L. FLANAGAN, *Speech Analysis, Synthesis and Perception*, Springer –Verlag, Berlin 1965.

[18] C. S. HASEK, S. SINGH, *Acoustic attributes of preadolescent voices*, J. Acoust. Soc. Amer., **68**, 1262–1265 (1980).

[19] E. B. HOLMBERG, R. E. HILLMAN, J. S. PERKELL, *Glottal airflow and pressure measurements for female and male speakers in soft, normal and loud voice*, paper presented at the 114th Meeting of Acoust. Soc. Amer., Florida 1987.

[20] J. N. HOLMES, *Normalisation in Vowel Perception*, in: J. S. Perkell, D. H. Klatt (eds.): Invariance and Variability in Speech Processes, Lawrence Erlbaum Assoc., Hillsdale 1986, pp. 346–357.

[21] D. KAHN, *On the identifiability of isolated vowels*, UCLA Working Papers in Phonetics, **41**, 26–31 (1978).

[22] D. H. KLATT, *Software for a Cascade/Parallel Formant Synthesizer*, J. Acoust. Soc. Amer., **67**, 971–995 (1980).

[23] D. H. KLATT, L. C. KLATT, *Voice quality variations within and across female and male talkers: implications for speech analysis, synthesis and perception*, paper presented at the 114th Meeting of Acoust. Soc. Amer., Florida 1987.

[24] P. LADEFOGED, P. BROADBENT, *Information Conveyed by Vowels*, J. Acoust. Soc. Amer., **29**, 98–104 (1957).

[25] A. M. LIBERMAN, F. S. COOPER, D. P. SHANKWEILER, M. STUDDERT-KENNEDY, *Perception of the Speech Code*, in: E. E. David, Jr., P. B. Denes (eds.): Human Communication: A Unified View, McGraw-Hill, New York 1972, pp. 13–50.

[26] P. ŁOBACZ, G. DEMENKO, *Dependence of the preception of segmental features of the Polish vowels on the structure of the long-term lexical-phonetic memory* (in Polish), Reports of the Institute of Fundamental Technological Research 40, Warsaw 1983.

[27] W. MAJEWSKI, J. ZALEWSKI, *The role of fundamental frequency in the process of perception of synthetic Polish speech signals* (in Polish), Reports of the Institute of Telecommunication and Acoustics, Technical University of Wrocław, **13**, 37–50, (1973).

[28] R. B. MONSEN, A. M. ENGEBRETSON, *Study of variations in the male and female glottal wave*, J. Acoust. Soc. Amer., **62**, 981–993 (1977).

[29] J. S. PERKELL, D. H. KLATT (eds.), *Invariance and Variability in Speech Processes*, Lawrence Erlbaum Assoc., Hillsdale 1986.

[30] B. H. REPP, A. F. HEALY, R. G. CROWDER, *Categories and Context in the Perception of Isolated Steady-State Vowels*, J. Exp. Psychol.: Human Perception and Performance, **5**, 1, 129–145 (1979).

[31] B. H. REPP, A. M. LIBERMAN, *Phonetic category boundaries are flexible*, in: S. Harnad (ed.): Categorical Perception: The Groundwork of Cognition, Academic Press, Cambridge 1987, pp. 89–112.

[32] A. G. SAMUEL, *Phonetic Prototypes*, Perception and Psychophysics, **31**, 4, 307–314 (1982).

[33] K. N. STEVENS, *The Quantal Nature of Speech: Evidence from Articulatory-Acoustic Data*, in: E. E. David, Jr., P. B. Denes (eds.): Human Communication: A Unified View, McGraw-Hill, New York 1972, pp. 51–66.

[34] A. Q. Summerfield, M. P. Haggard, *Vocal tract normalisation as demonstrated by reaction times*, in: G. Fant, M.A.A. Tatham (eds).: Auditory Analysis and Perception of Speech, Academic Press, London 1975, pp, 115–141.

[35] A. K. Syrdal, *Aspects of a model of the auditory representation of American English vowels*, Speech Communication, **4**, 121–135 (1985).

[36] H. Traumüller, *Perceptual dimension of openness in vowels*, J. Acoust. Soc. Amer., **69**, 1465–1475 (1981).

[37] H. Traunmüller, F. Lacerda, *Perceptual relativity in identification of two-formant vowels*, Speech Communication, **6**, 143–157 (1987).

[38] D. R. Van Bergem, L. C. W. Pols, F. J. Koopmans-Van-Beinum, *Perceptual Normalization of the Vowels of a Man and a Child in Various Contexts*, Speech Communication, **7**, 1–20 (1988).

Received April 30, 1990

ARCHIVES OF ACOUSTICS
16, 2, 325–342 (1991)

DESIGNING PROBLEMS IN COMPUTERIZED ACOUSTIC ANALYZERS

G. PAPANIKOLAOU

University of Thessaloniki, Greece

B. KOSTEK

Technical University of Gdańsk, Telecommunication Institute
(80-952 Gdańsk, ul. Majakowskiego 11/12)

An attempt to present considerable development in the domain of computerized acoustic devices is made. The synthesis refers mostly to computer-based acoustic analyzers leading in that domain because of their versatility. The mathematical basis as well as the main test functions are reviewed. Hardware and software assumptions of the most universal computerized equipment are shown. The example of an FFT analyzer realized on the basis of the LabVIEW software package for the Macintosh II computer is presented. Conclusions and index of bibliography are included.

1. Introduction

On the basis of the latest advances in computer technology, it is possible to combine all necessary test functions in a single multi-purpose instrument and at the same time to improve operation and measurement procedures. Such a system can be easy-adaptive to any kind of acoustical measurement.

The main problem which is to be solved prior to any design of the really

The main problem which is to be solved prior to any design of the really universal acoustic analyzer is the selection of its test functions. Equally important is the proper mathematical description of the analyzed physical phenomena. Mathematical formulas are to be represented by the algorithms used as the basis of procedures written in programming languages or machine codes. However, those procedures are not equally effective. This is of great importance when both the ease of programming and the speed of processing are concerned.

A computerized measuring device provides a powerful technique for collecting, analyzing and presenting data in an organized and systematic fashion. Nevertheless, a considerable amount of expertise in electronics and computers is necessary to configure a data acquisition system. That problem is particularly interesting because of the variety of configurations in existing measuring systems.

It is possible to discern a few types of computerized acoustic measuring devices. The first group comprises all multi-function analyzers dedicated to the domain of a single application: e.g., architectural acoustics, ultrasonics, studio technology, mechanical vibrations etc. The next group includes systems in which a computer is applied in the control of the measurement process and, additionally, makes possible communication with the user.

The introduction of personal computers to the fast processing of acoustic signals has affected in a significant way the topology of the typical measuring laboratory. A computer provided with data acquisition cards and supporting software becomes a powerful tool to perform even a very sophisticated analysis.

The mentioned solution can be considered to be the most advanced in the domain of digital measuring systems and, at the same time, economically recommended.

The main topics of the following discussion is the dependence of the system capabilities on its architecture and on the employed programming method.

2. Analysis of quantized acoustical signals

Acoustic measurements mostly require time and frequency analyses of signals. Moreover, the simultaneous analysis of two functions e.g. excitation and response signals, justifies the design of analyzers as dual-channel.

Analytical signals in the domain of time can be displayed in each channel in terms of: their real and imaginary part, their magnitude and phase versus time, their NYQUIST or NICHOLS plots.

The analytical signal corresponding to the real-valued time record of input can be represented by

$$x_z(n) = x(n) + j\tilde{x}(n) \tag{1}$$

where n — number of samples of signal record.

As the following time records $x(n)$ are to be averaged, the analytical signal can be expressed by

$$\bar{x}_z(n) = \bar{x}(n) + j\tilde{\bar{x}}(n) \tag{2}$$

$$\tilde{\bar{x}}_z(n) = \mathscr{H}[\bar{x}(n)] \tag{3}$$

$$\bar{x}_z(n) = \mathfrak{F}^{-1}[G(k)] \tag{4}$$

where k — frequency index, $\mathscr{H}$ — discrete Hilbert transform, $G(k)$ — instantaneous spectrum, $\mathfrak{F}^{-1}$ — inverse fast Fourier transform.

On the basis of Eq. (4) — the autospectrum function $\bar{G}_{11}(k)$ can be calculated as follows:

$$\bar{G}_{11}(k) = \bar{G}_1^*(k)\bar{G}_1(k) = |\bar{G}_1(k)|^2 \tag{5}$$

where $G_1^*(k)$ — complex conjugated spectrum.

The $|\bar{G}_1(k)|^2$ can be scaled and furtherly interpreted as: rms, power, power spectral density or energy spectral density of the signal $x(n)$. Those parameters provide an important representation of acoustical signals.

Calculation of the response of acoustical systems to different excitations is permitted on account of designing an analyzer as dual-channel.

Therefore the cross spectrum function of both channels should be calculated:

$$\bar{G}_{12}(k) = G_1^*(k)G_2(k) = \overline{G_2^*(k)G_1(k)}. \tag{6}$$

Using the latest equation various types of frequency responses could be calculated from

$$H_1(k) = \frac{\bar{G}_{12}(k)}{\bar{G}_{11}(k)}. \tag{7}$$

The importance of Eq. (7) results from its physical interpretation. The complex ration of the averaged cross spectrum $\bar{G}_{12}(k)$ and the averaged autospectrum $\bar{G}_{11}(k)$ reflect the gain and phase lag introduced by a linear system in response to the $x_1(n)$ excitation.

The frequency response can be expressed as magnitude, phase or plots in Nyquist or Nichols coordinates.

The convolution of the signals in both channels, represented in the frequency domain as a multiplication of spectra, can be used to determine the coherence between those signals.

As an indication of the quality of the measurement, one may use following formula:

$$C_{12}^2(k) = \frac{|\bar{G}_{12}(k)|^2}{\bar{G}_{11}(k)\bar{G}_{22}(k)}. \tag{8}$$

That factor expresses the level of linear dependence between the signals in both channels of the analyzer related by a linear frequency response of the measuring object to the excitation $x_1(n)$.

The next parameter, i.e., signal to noise ratio, determines one of the most important features of an audio channel. Moreover, like a coherence factor that parameter related by a linear frequency response function to the signal $x_1(n)$ may be used to indicate the quality of the performed measurement.

The signal to noise ration is defined by the expression

$$\frac{S}{N}(k) = \frac{C_{12}^2(k)}{1-C_{12}^2(k)} = \frac{\dfrac{|\bar{G}_{12}(k)|^2}{\bar{G}_{11}(k)\bar{G}_{22}(k)}}{1-\dfrac{|\bar{G}_{12}(k)|^2}{\bar{G}_{11}(k)\bar{G}_{22}(k)}} = \frac{|\bar{G}_{12}(k)|^2}{\bar{G}_{11}(k)\bar{G}_{22}(k)-|\bar{G}_{12}(k)|^2} \tag{9}$$

and may be used to describe the signal dynamics.

As it results from the modern auditory theory, the perception process cannot be properly considered without taking into account the binaural cross-correlation phenomena. Hence, providing the measuring system with the correlation analytical functions becomes substantial.

The basic function in the domain of correlation analysis is the autocorrelation function defined in the spectrum averaging mode as follows:

$$R_{11}(n) = \mathfrak{F}^{-1}[W(k)\bar{G}_{11}(k)] \tag{10}$$

where $W(k)$ is the weighting function of the autospectrum function.

In order to make possible the comparison of various signals, the normalization of autocorelation values is desired. That operation is defined by

$$\varrho_{11}(n) = \frac{R_{11}(n)}{R_{11}(0)} \tag{11}$$

Autocorrelation can be presented in terms of real and imaginary parts, magnitude and phase or pole coordinates.

Similarly the cross-correlation function is defined by

$$R_{12}(n) = \mathfrak{F}^{-1}[W(k)\bar{G}_{12}(k)]. \tag{12}$$

In that case the normalization leads to the definition of the cross-correlation coefficient as follows:

$$\varrho_{12}(n) = \frac{R_{12}(n)}{\sqrt{R_{11}(0)R_{22}(0)}} \tag{13}$$

where

$$R_{11}(0) = \frac{1}{N}\sum_{k=0}^{N-1} W(k)\bar{G}_{11}(k) \tag{14}$$

$$R_{22}(0) = \frac{1}{N}\sum_{k=0}^{N-1} W(k)\bar{G}_{22}(k) \tag{15}$$

According to the principles of the circuit and signal theory the impulse response can be used as the basic parameter of the transmission channel. Hence, its presence in the function menu of an analyzer cannot be neglected.

In the spectrum averaging mode, the impulse response can be defined by the equation

$$h(n) = \mathfrak{F}^{-1}[W(k)H_1(k)] \tag{16}$$

The impulse response of a linear system can be determined by exciting the system with white noise and cross-correlating the input and output. Although the most direct approach is to apply an impulsive excitation such as electronic spark gaps, pistol shots or exploding balloons, it is difficult, nevertheless, to provide that kind of signal with sufficient energy, repeatable in terms of its amplitude and directional characteristics.

In order to overcome these problems and to minimize the amount of computation required by the cross-correlation operation, the system can be excited by a binary maximum-length sequence, and the cross-correlation performed using the fast Hadamard transform [6], [7], [8], [23].

Several of the references explain how to generate maximum-length sequences based upon a primitive polynomial [9], [23].

The mathematical operations given above enable calculations of the majority of acoustical parameters, They permit, among others, to calculate sound power, sound absorption and insulation, reverberation time according to various standards [30], still spectrum manipulation such as: arithmetic operations, integration, weighting, differentiation, spectrum weighting etc.

Having calculated the discrete spectrum of an acoustic signal it is possible to display the measurement results in the form of octave or third-octave analysis by averaging spectral energy in the corresponding frequency bands.

In the domain of electroacoustic equipment technology, many additional test functions are required for quality test applications. These parameters which are presented in Fig. 1 can be calculated on the basis of the discrete FFT being the standard operation during signal processing in a digital analyzer.

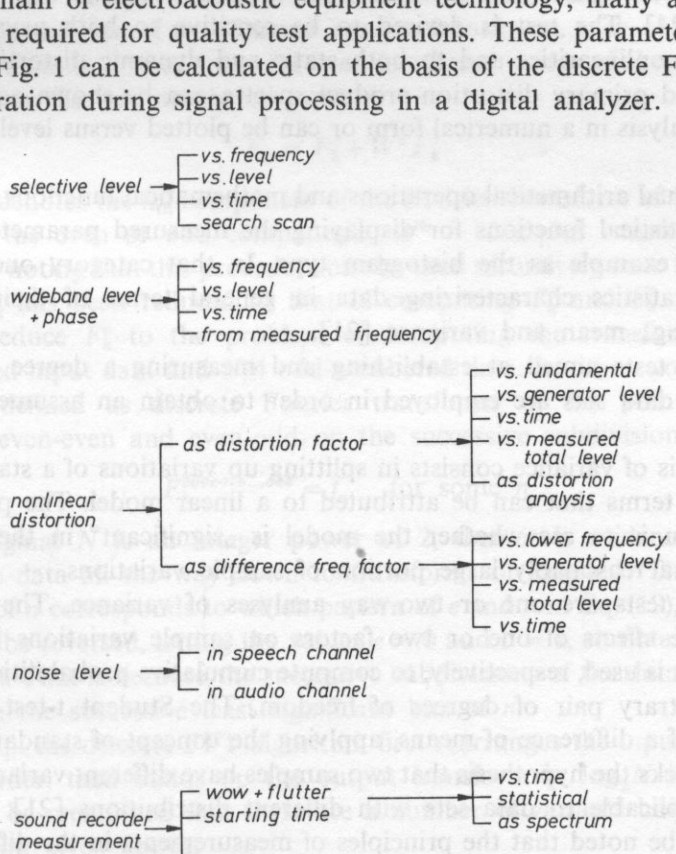


FIG. 1. Test functions for the electroacoustic equipment technology

As it is shown in Fig. 1, the amount of test functions is considerable, so the mathematical description will be restricted to the most important ones. The remaining parameters are defined in the bibliography [16], [22], [30].

The determination of the nonlinear distortion factor according to the definition given in Eq. (17) is a standard procedure when testing an audio channel

$$h_{0/0}=\frac{\sqrt{h_1^2+h_2^2+\ldots+h_n^2}}{n} \tag{17}$$

where $h_1 \ldots h_n$ are the subsequent harmonics of the signal extracted from the discrete spectrum.

In that kind of analysis the difference frequency factor is usually carried out. The IEC total difference-frequency distortion TDFD test is a recently proposed method for detecting and quantifying a broad range of nonlinear distortions in audio equipment [24]. The test is devised to be sensitive to both asymmetrical and symmetrical nonlinearities and to both static and dynamic distortions [24]. The test-signal and primary distortion-product spectra can be shown as results of the distortion analysis in a numerical form or can be plotted versus level, frequency, or time.

Besides usual arithmetical operations and mathematical functions, it is necessary to set up statistical functions for displaying the measured parameter in terms of statistics, for example as the histogram type. In that category one may include descriptive statistics characterizing data in general terms of their moments of distribution e.g., mean and variance [21].

Statistical tests aimed at establishing and measuring a degree of correlation between two data sets are employed in order to obtain an assumed accuracy of calculations.

An analysis of variance consists in splitting up variations of a statistical sample into a set of terms that can be attributed to a linear model. The purpose of this decomposition is to see whether the model is „significant" in the sense that it accounts for a reasonably large portion of sample variations.

ANOWA tests are one or two-way analyses of variance. They are used to determine the effects of one or two factors on sample variations [21].

The F-test is used, respectively, to compute cumulative probabilities and fractiles with an arbitrary pair of degrees of freedom. The Student t-test measures the significance of a difference of means, applying the concept of standard error while the F-test checks the hypothesis that two samples have different variances. There are also tests applicable to data sets with different distributions [21].

It should be noted that the principles of measurements in the different areas of acoustics are based on the same set of operational, mathematical and statistical functions, so it is possible to design an analyzer which is effective in any acoustic domain.

3. Algorithmic representation of the main analytical procedures

As it results from the quoted mathematical formulas, processing of the measured signals demands, first of all, the following operations in both frequency and time domains: FFT, FFT^{-1}, Hilbert and Hadamard transforms.

If the discrete FFT is performed using its definition, then the execution time requires n^2 multiplications, where n is the number of signal samples. In order to minimize the amount of computation, one can use the fast Fourier transform based upon the algorithms of Cooley and Tukey and, further, of Winograd [21]. The recently proposed methods eliminate the number of operations and as a result they speed up the processing hence addressing of data becomes more complicated. In fact, the discrete Fourier transform can be computed in $n\log_2 n$ operations.

The fast Fourier transform can be described in terms of a matrix multiplication. Its routine is based on the flow graph known also as butterfly graph [10, 20, 21].

Danielson and Lanczos showed that a discrete Fourier transform of length N can be rewritten as the sum of two discrete Fourier transforms, each of the length $N/2$ [21]. One of them is formed from the even-numbered points of the original N, the other one from the odd-numbered points:

$$F_k = F_k^e + W^k F_k^o \tag{18}$$

where F_k^e, F_k^o denotes the k_{th} component of the Fourier Transform of the length $N/2$ formed from the even or odd components, W^k — complex constant.

It is worth noting that this procedure can be used recursively once the problem of computing F_k has been reduced to that of cumputing F_k^e and F_k^o. Similarly it is possible to reduce F_k^e to the problem of computing the transform of its $N/4$ even-numbered input data and $N/4$ odd-numbered data. In other words, FK^{ee} and F_k^{oe} can be defined as discrete Fourier transforms of the points which are, respectively, even-even and even-odd on the successive subdivisions of the data

$$F_k^{eoeeoeo\ldots oee} = f_n \quad \text{for some } n \tag{19}$$

When original N is an integer power of 2, then it is evident that one can subdivide the data all the way down to transforms of length 1. Then, to identify which value of n corresponds to which pattern of e and o in Eq. (19), the pattern of e and o is to be reversed, letting the values $e = 0$ and $o = 1$, so obtaining in binary the value of n. That is because the successive subdivisions of the data into even and odd tests are the successive least significant bits of n.

To sum up: the discrete FFT algorithm first rearranges the input elements into bit-reverse order, then builds up the output transform in $\log_2 N$ iterations.

It should be mentioned that there are a number of variants built on the basic FFT algorithm given above.

As it has been said, the Hilbert transform is one of the basic operations required to process an analytical signal.

Unlike the Fourier transform which moves the independent variable of a signal

$x(t)$ from the time domain to the frequency domain, the Hilbert transform leaves the signal $x(t)$ in the same domain. The Hilbert transform $\tilde{x}(t)$ of a real-valued time signal $x(t)$ is another real-valued time signal, and the Hilbert transform of a complex-valued frequency function $X(f)$ is another complex-valued frequency function.

The Hilbert transform can be defined as a convolution integral' phase shift system or as the imaginary part of an analytic signal. In practice such an integral is to be replaced by finite summations [5].

Fast Madamard transform can be considered after FFT, as the most important transformation. When considering an acoustic measurement system excited by a pseudorandom noise, it is very efficient to use that transform in order to measure its impulse response [6], [8].

Like the discrete Fourier transform, the Hadamard transform can be described in terms of matrix multiplication [8], [9], [23]. The matrix that transforms the input vector is known as the Hadamard matrix H_n where n gives the number of rows or columns. The elements of the Hadamard matrix are all $\pm$ 1, and the matrix must satisfy the relation

$$H_n H_n^T = nI_n \tag{20}$$

where I_n − unit matrix

The algorithm applies only to the specific class of Hadamard matrices known as the Sylvester-type. The Hadamard matrix is defined recursively by [8]

$$H_1 = [1]$$

$$H_{2i} = \begin{bmatrix} \mathrm{H}_i & \mathrm{H}_i \\ \mathrm{H}_i & -\mathrm{H}_i \end{bmatrix}. \tag{21}$$

Only orders 2^k, where k is a nonnegative integer, exist.

The basic butterfly element for the fast Hadamard transform is shown in Fig. 2.

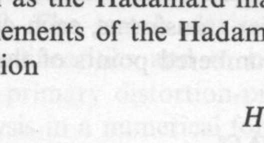


FIG. 2. Basic butterfly element.

It is evident that the flow graph is identical to the flow diagram for the FFT except that the twiddle factors are all unity, reflecting the fact that no multiplications are required. Then it is possible to adapt any Fourier transform routine to execute the fast Hadamard transform [8].

The fast Hadamard transform requires only about 2.5 $n\log_2 n$ operations which in this case are additions. Hence, there is no need for multiplications that represent a reduction in the execution time [8].

The speed of processing depends on the type of operation being performed (matrix multiplications or additions) and the type of computer beind used. The time required to execute FFT in one of the most powerful realization (see Paragraph 5), where the FFT function is defined by the following summations:

$$F(i) = \sum_{k=0}^{N-1} \text{array! } (k) * e^{-j2\pi ki/N}, \tag{22}$$

is about 7 s for the power 8, while VFFT (very fast Fourier transform) is approximately 100 times faster than FFT. The quoted parameters concern the McADIDS system on the Macintosh Plus computer [15].

4. System configuration problems

Contemporary analyzers develop, in qeneral according to the configuration of applied microprocessor systems.

The traditionally designed analyzer provides a detached measuring device which is assigned to collect and process measured values and to display results. The main insufficiency of such a device usually results from the lack of its versatility. Thus, in the case of more complex tests a certain number of similar devices has to be used simultaneously. As the tests results should be stored, printed or plotted, too many devices and too many interconnections between them are needed. This gives rise to a nonoptimal structure of the instrumentation system.

In the domain of acoustics, compactness and portability are of main importance. This issues from the fact that there is a need to transport those devices for outdoor tests in architectural acoustics and from the usual lack of space for storing them in sound recording studios. The principles of processing of measured signals considered in the previous section prove that the great majority of acoustic tests are based on a restricted number of processing routines. Thus, it is recommended to design acoustical analyzers as really universal ones.

Contemporary computer systems create the possibility of fulfilling those requirements but the optimal lay-out of an instrumentation system has not been defined so far. Thus that problem as being current and vital should be thoroughly examined.

Historically, data acquisition and control tasks were implemented either with large mainframe or with minicomputer system. Typically, there were powerful 16-bit machines that ran in time-sharing or multitasking modes. The development of the Personal Computer has enabled to take advantage of the flexibility and efficiency of computerized data acquisition and control. A universal complaint, however, is that a considerable amount of expertise in electronics and computers is necessary to configure a data acquisition system. Nevertheless, owing to a significant degree of standarization among PC and data acquisition and control systems, a large family of hardware and software tools has evolved. The general block diagrams of the computer data acquisition and control system are shown in Figs. 3 and 4.

An important part of any data acquisition system is its host computer. There are

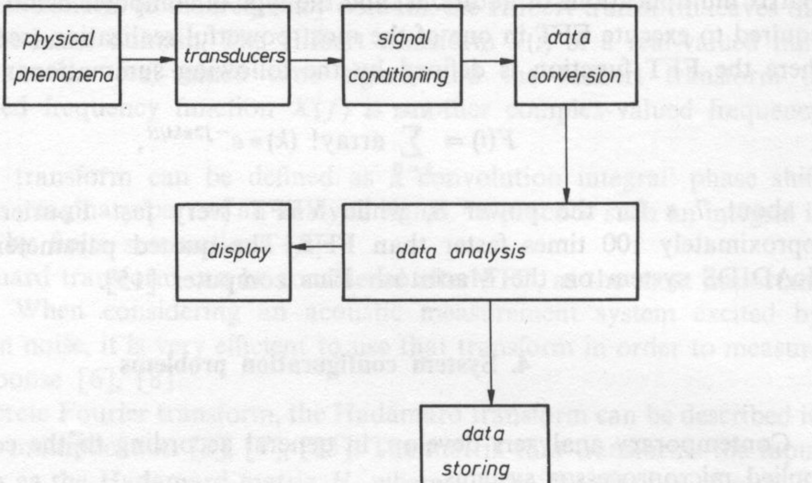


FIG. 3. Data acquisition and control flow diagram

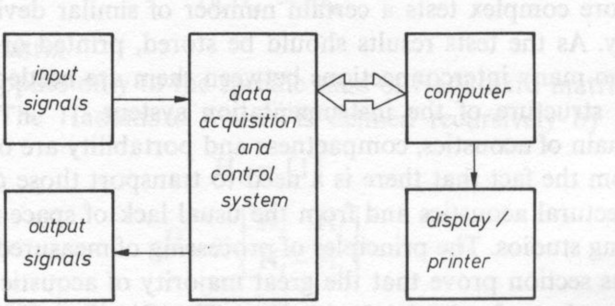


FIG. 4. Computerized data acquisition and control system

two possible ways to interface that system with a computer: connection via a standard interface such as RS-232 (RS-442) or IEEE-488 (see Fig. 5) or direct connection to PC bus (Fig. 6).

The first solution is recommended when the data acquisition system should have the possibility of being interfaced to any type of computer, could be placed remotely from it or when the configured system size is extra complex.

The GPIB (General Purpose Interface Bus) recognized as the IEEE-488 standard has been proven to be one of the most successful and widely accepted interfacing schemes [3]. First of all, the GPIB reduces the hardware requirements of the computer to one interface for each instrument so the main assumption concerning the preferable universal lay-out of the instrumentation system can be fulfilled.

Choosing an interfacing scheme remains one of the most troublesome tasks that engineers, researchers and system designers face, while the importance of the choice

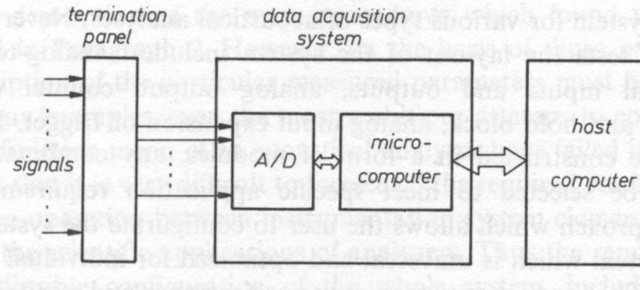


FIG. 5. Block diagram of an external bus data acquisition system

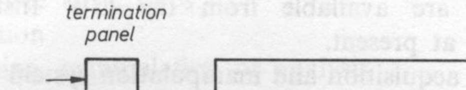


FIG. 6. Internal PC bus system

grows even greater as computers are applied increasingly to complex acoustical phenomena. An inflexible, slow or otherwise inappropriate interface can limit even the most powerful computer-based systems. Interfaces dating earlier than GPIB needed highly specialized hardware and were usually developed for a single purpose. Often those interfaces could not be adapted to even slightly different applications. In the end when the lay-out of universal acoustic instrumentation system is discussed only both previously standard interfaces should be taken into consideration. This follows from the prevalence of RS-232 and advantages of GPIB IEEE-488. According to that principle, several professional instrumentation systems have been designed [27], [29].

The main advantages of a direct connection of the data acquisition system to the PC bus are a data speed of high transmission and instrumentation system compactness. The last mentioned feature is particularly obvious, when the system is composed of computer cards which can be mounted directly into the PC mainframe. The PC expanded in that way provides a powerful platform for collecting, analyzing and presenting data in an organized and systematic fashion. The expansion card can use its own signal processor responsible for the acquisition and fast processing of data [25].

However, the key to the full universality of the discussed system in the domain of acoustics is its modularity. It results from the idea of re-configuring slightly the instrumentation system for various types of acoustical analyses. Nevertheless, in the great majority of tests the lay-out of the system includes: analog-to-digital data acquisition, digital inputs and outputs, analog output, counter (timer) pulse generator, sample and hold block, analog input expansion or trigger. Providing the quoted blocks are constructed in a form of modules, any combination of those instruments can be selected to meet specific application requirements. This is a considerable approach which allows the user to configurate the system by himself and leads to a system which is universal and optimized for individual requirements at the same time. That opportunity maximizes performance and minimizes cost since only those modules are to be installed, which are required for a satisfying performance of the entire range of the involved measurement tasks. The modular instrumentation systems are available from the GW Instruments and the Burr-Brown Corporation at present.

The MacADIOS data acquisition and manipulation system [18] elaborated by the GW Instruments (Massachusetts Institute of Technology) turns the 32-bit Macintosh II computer into a powerful analyzer with capabilities surpassing even the most specialized acoustical measurement systems. The MacADIOS II carrier card plugged into one of the Macintosh expansion slots provides analog input and output channels, digital input and output channels and counter/timer channels. There are also available additional modules, of which up to three can be installed on the carrier to perform such tasks as high-speed sampling and multi-channel input and output.

The IBM compatible personal computer can also co-operate with the similar kind of data acquisition and control system: Burr-Brown's PCI-20000 [27]. The basic configuration of that system is also modular.

As it can be seen from the above review of possibilities given by modern computer technology systems which consist of individual instruments having a front panel with its own combination of indicators, knobs and switches are going to be in defensive. Therefore producers of similar traditional instruments are compelled to revise their approach to the designing of acoustic analyzers.

As to the example of 2032, 2034, 2123 and 2133 Brüel and Kjaer analyzers [11], [17], [26], it becomes obvious that the difference between a computer and an analyzer decreases dramatically. The last mentioned model includes a screen monitor, keyboard and diskette drive. Nevertheless, the versatility of a properly extended standard PC seems to be much better than that of such an analyzer. As a result further development of acoustical analyzers based on PC may be expected.

It is obvious that the capabilities of any computer analyzer are strictly related to their software performance. These problems will be discussed below.

5. Software organization

The most important problem which is to be solved when designing a multi-purpose acoustic analyzer is the proper selection of its test functions. The mathematical description of the main procedures which found the test functions was presented in Paragraph 2. However, on the basis of those procedures a more detailed description of the particular measured parameters must be defined. As it is proven by many examples, even the most ambitious attemps to construct the really universal test function menu of an acoustical analyzer have failed in practice. This is due to the fact that it is very difficult to foresee all the required configurations, ranges and ways of co-operation between instrumentation system elements. This concerns, in particular, the scientific applications of analyzers. Thus the remaining solution is to choose a flexible configuration of the whole system, including the software organization.

The best strategy for high performance is to construct independent procedures for the following three basic steps:

– data acquisition

– data processing, manipulation or analysis

– presenting the data in tabular, graphical or report format.

Each of the acquisition procedures should support a corresponding instrument module which is available in the system. Thus user-callable routines should allow users to specify parameters such as gain, range etc., and to define input/output channels. Those routines are expected to be executable as high-level commands in high-level language such as Basic INPUT statement, and then receive a data string from the acquisition unit.

To enable the processing, manipulation and analysis of data, the instrumentation system based on a personal computer should be designed in such a way that the users have the opportunity to write specialized programs for data acquisition, storage display, logging and control in high-level languages. That possibility demands the use of a special supporting software which links high-level languages to data acquisition subroutines.

Three classes of that software can be discerned: program development tools, function subroutine librairies and complete application packages.

Program development tools and function libraries are used to facilitate writing the unique software. They usually include drivers that provide the interface to the hardware acquisition system. Those software interfaces are intended to establish communication with a data acquisition system using Basic, C or Turbo-Pascal commands. The specially designed data analysis languages such as ASYST [3] or MacADIOS [19] are based on function libraries which include basic procedures used in data acquisition and processing, e.g., data reading, signal filtering, FFT etc.

Complete application software packages are designed to get the system working immediately, usually with no programming required. However, those routines can also offer users facilities to modify the software to meet their own needs. Nevertheless, they are less flexible than the other classes of software.

The most advanced solutions, making possible to create optional instrumentation systems by the user, not necessarily introduced to the programming of a computer, are offered by the software developed by National Instruments Inc. [15],

[27]. Each of these products are intended to run on one of the two main leading PC's: IBM and Macintosh.

The LABTECH is the vast program package which enables the reduction of data acquisition to menu-driven choices. Programming options give users automation and customization of advanced analysis. Digital input data can be recorded and displayed in real-time. Since LABTECH is menu-driven and extremely easy to learn and use, it requires very little computer skills on the part of the operator.

LabVIEW, which is designed for Macintosh computers, replaces conventional programming techniques with intuitive diagramming techniques. The central concept underlaying LabVIEW is the virtual instrument. The virtual instrument is similar in modularity and functions to a physical instrument; however, it is built out of software. This instrument is hierarchical, i.e., one virtual instrument can include a variety of other instruments or components. LabVIEW supports operating hierarchical instruments to perform complex tests and measurements in a manner similar to physical instruments connected with each other. Diagramming techniques give users opportunities to create a block diagram of an instrument and its front panel for interactive control using a computer mouse. Drawing an instrument diagram is directly associated with programming data acquisition, processing and displaying. That is performed entirely in a graphic programming language invented specifically for this task. That language uses data flow concepts, traditional program – control structures and a set of instructions consisting of graphical elements.

The signal processing functions in LabVIEW provide users with the most frequently used signal processing operations, e.g. FFT, fast convolution, power-spectrum analysis and many others. The FFT analyzer shown in Figs. 7 and 8 can serve as an example of a virtual instrument created using the LabVIEW system. The lay-out of that instrument was programmed and experimentally verified in the Music and Acoustic Laboratory of the University of Thessaloniki. Figure 7 shows the designed front panel of the virtual instrument. The program of that instrument written in the graphic language is presented in Fig. 8. The flow of data specified in the block diagram is created by selecting and arranging graphical objects and wiring them together. The FASTIO procedure performs very high speed input and output operations. It is specified by the parameters listed below: *trigger*, *points*, *loops*, *mult*, *device*, *type*, *offset*, *slot*, *error*. The *trigger* parameter determines the external triggering mode. Setting to zero indicates no external triggering is to be used, so the input or output operation begins immediately after calling FASTIO. The *points* parameter specifies the number of samples that are digitized. The number of data elements involved in processing is defined by the power of 2, where the range of power is assigned from 7 to 17. For input operations the *loops* parameter should be set to one. The *mult* parameter is used to determine the cycle time which corresponds to the duration between consecutive samples, and is calculated as *mult* × 0.2 μs. It is determined by the computer being used and should not exceed the minimum and maximum time specifications which, in case of Macintosh II, are consequently 1.2 μs and 26 μs. The *slot*, *device*, *type*, *offset* parameters specify the installation of the

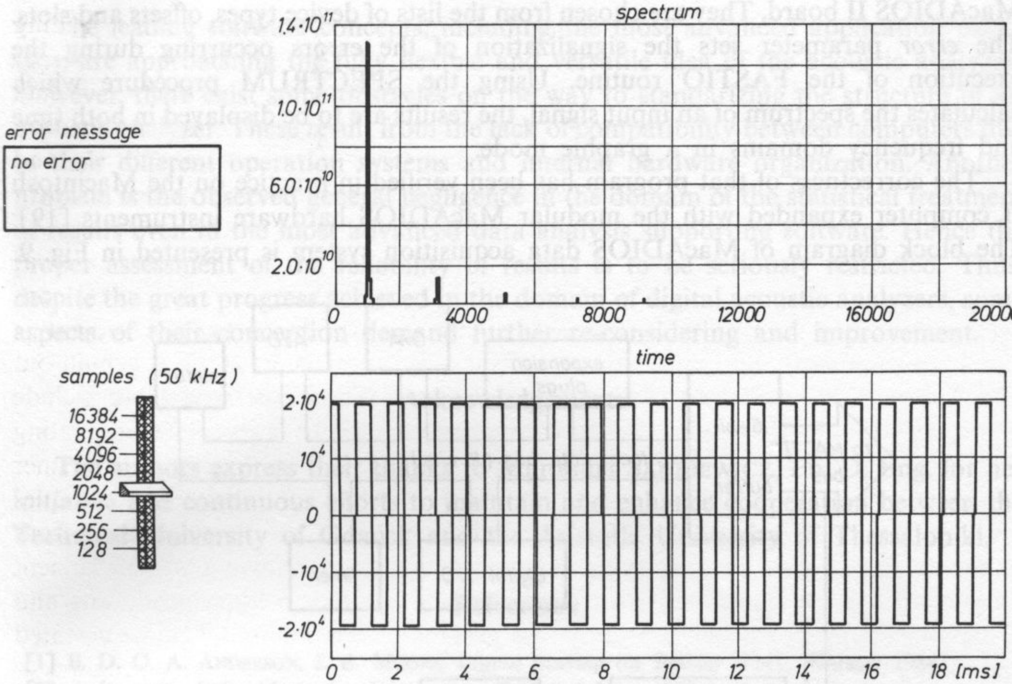


FIG. 7. Front panel of the FFT analyzer designed using the LabVIEW

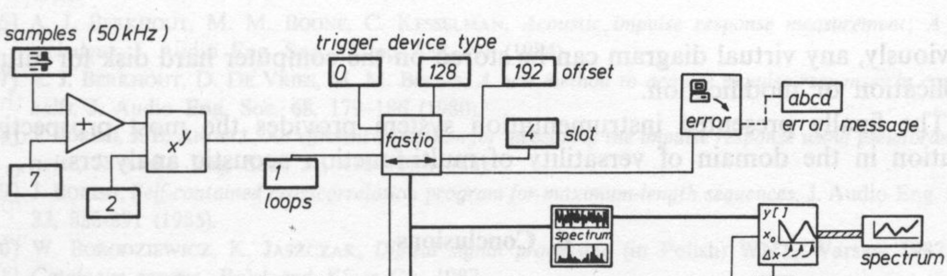


FIG. 8. Lay-out of the FFT analyzer

MacADIOS II board. They are chosen from the lists of device types, offsets and slots. The *error* parameter sets the signalization of the errors occurring during the execution of the FASTIO routine. Using the SPECTRUM procedure which calculates the spectrum of an input signal, the results are to be displayed in both time and frequency domains in a graphic mode.

The correctness of that program has been verified in practice on the Macintosh II computer expanded with the modular MacADIOS hardware instruments [19]. The block diagram of MacADIOS data acquisition system is presented in Fig. 9.

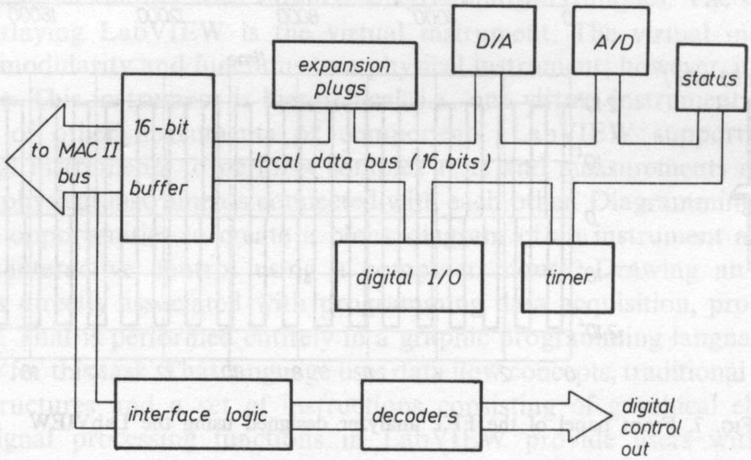


FIG. 9. Block diagram of the MacADIOS system

Obviously, any virtual diagram can be stored on the computer hard disk for future application or modification.

The finally presented instrumentation system provides the most prospective solution in the domain of versatility of multi-function acoustic analyzers.

6. Conclusions

As it follows from the presented analysis of concepts of contemporary digital acoustic analyzers, these are based on differentiated hardware and software solutions. The existence of such a large variety of ideas proves that a final consensus on the acoustic measurement system structure has not been reached so far. Nevertheless, as it has been shown above, the great majority of measured parameters can be defined using limited collection of basic functions.

On the other hand, the topology of the computerized analyzer is based on the restricted number of reconfigurable units responsible for the acquisition and processing of data. The two circumstances mentioned above lead to the modular conception of the analyzer in both hardware and software domains.

The leading software concepts, including the most advanced application packages, are approaching the fully flexible and versatile idea of the acoustic analyzer. However, there exist some obstacles on the way to standarizing the structure of an acoustic analyzer. These result from the lack of compatibility between computers due to their different operation systems and internal hardware organization. Another problem is the observed general negligence in the domain of the statistical treatment of results even in the most advanced data analysis supporting software. Hence the proper assessment of the reliability of results is to be seriously restricted. Thus, despite the great progress achieved in the domain of digital acoustic analyzers, some aspects of their conception demand further re-considering and improvement.

Acknowledgements

The authors express their thanks to Marianna Sankiewicz, Ph. D. Eng. for her initiative and continuous efforts to maintain and enhance cooperation between the Technical University of Gdańsk and the Aristotle University of Thessaloniki.

References

[1] B. D. O. A. Anderson, J. B. Moore, *Digital filtering* (in Polish) WNT, Warsaw 1984.
[2] *Audioscope, Audio Measuring Instruments. Technical Note 1989.*
[3] *Asyst, A scientific system*, Macmillian Software Company 1986.
[4] *Avec le progiciel Aria, greffez une oreille electronique directement sur votre micro-ordinateur, Adverting notes*, 01 dB 1979.
[5] J. S. Bendat, *The Hilbert transform and applications to correlation measurements*, Brüel and Kjaer 1986.
[6] A. J. Berkhout, M. M. Boone, C. Kesselman, *Acoustic impulse response measurement: A new technique*, J. Audio Eng. Soc. **32**, pp. 888–891 (1984).
[7] A. J. Berkhout, D. De Vries, M. M. Boone, *A new method to acquire impulse responses in concret halls*, J. Audio Eng. Soc. **68**, 179–186 (1980).
[8] J. Borish, J. B. Angell, *An efficient algorithm for measuring the impulse response using pseudorandom noise*, J. Audio Eng. Soc., **31**, 478–487 (1983).
[9] J. Borish, *Self-contained crosscorrelation program for maximum-length sequences*, J. Audio Eng. Soc., **33**, 888-891 (1985).
[10] W. Borodziewicz, K. Jaszczak, *Digital signal processing* (in Polish) WNT, Warsaw 1987.
[11] *Catalogue resume, Brüel and Kjaer* Co. 1987.
[12] *Committed to excellence Tektronix Co. Catalogue* 1989.
[13] *Data equisition on the Macintosh. Technical Note* GW. Instruments, Corp. 1988.
[14] *Electronic measurement, Design, Computation*, Hewlett Packard Catalogue 1987.
[15] *LabVIEW Software and Demonstration disks, user's guide*, GW. Instruments, Inc., 1987.
[16] E. Leinonen, M. Otala, J. Curl, *A method for measuring transient intermodulation distortion* TIM, J. Audio Eng. Soc., **25**, 170–177 (1977).
[17] *Les mesures acoustiques de demain sous vos yeux*, Technical Note Brüel and Kjaer, M65/88.
[18] *MacADIOS II Hardware Manual*, GW. Instruments, Inc., 1987.
[19] *MacADIOS II Software Manual*, GW. Instruments, Inc., 1987.
[20] A. W. Oppenheim, R. W. Schafer, *Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, 1975.
[21] W. E. Press, B. P. Flannery, S. A. Teukolsky, V. T. Vetterling, *Numerical recipes. The art of scientific computing*, Cambridge University Press 1986.

[22] R. B. RANDALL, *Application of Brüel and Kjaer equipment to frequency analysis*, Brüel and Kjaer 1977.

[23] D. D. RIFE, J. VANDERKOOY, *Transfer-function measurement with maximum-length sequences*, J. Audio Eng. Soc., 37, 419–444 (1989).

[24] R. H. SMALL, *Total difference-frequency distortion: practical measurements*, J. Audio Eng. Soc., 34, 427–436 (1986).

[25] *SYSid TM, the complete acoustic test instrument*, Information note, ARIEL 1988.

[26] *The New Leaders in 2-channel FFT analysis... types* 2032 *and* 2034, *Advertising notes* Brüel and Kjaer 1988.

[27] *The handbook of personal computer instrumentation for data acquisition. Test measurement and control*, Burr-Brown Corporation 1986.

[28] *The test for audio measurements: The* NFA-1 Audio Analyzer, Information note Wandel and Goltermann 1988.

[29] B. ZIMIŃSKI, *New generation of phonic measurement instruments* (in Polish) Przegląd techniki, Radio i Telewizja 1–2 1987.

[30] Z. ŻYSZKOWSKI, *Foundations of electroacoustics*, WNT: Warszawa 1984.

Received December 13, 1989

ARCHIVES OF ACOUSTICS
16, 2, 343–353 (1991)

APPLICATION OF ACOUSTIC METHODS TO ASSESSMENT OF CONCRETE HUMIDITY INFLUENCE ON THE PROCESS OF CONCRETE DESTRUCTION

J. PYSZNIAK, J. HOŁA

Technical University Wrocław
(50-370 Wrocław, Plac Grunwaldzki 11)

Owing to application of acoustic research methods, i.e. the ultrasound and the acoustic emission methods, it was pointed out in the paper that concrete humidity is of an essential influence on the process of concrete destruction. It was also proved that the totting sum of the acoustic emission recorded during the process of destruction of concrete is clearly determined by bulk humidity of the material under investigation.

The examination results submitted here refer to plain concrete subjected to an influence of innediate loads during tests of compressing it quasi-axially and stretching according to Brazilian method.

Poprzez zastosowanie akustycznych metod badawczych to znaczy metody ultradźwiękowej i emisji akustycznej, wykazano w pracy, że wilgotność betonu wpływa w istotny sposób na przebieg jego niszczenia. Wykazano także, że rejestrowana podczas procesu niszczenia betonu suma zliczeń emisji akustycznej jst wyraźnie uzależniona od wilgotności masowej badanego materiału.

Prezentowane wyniki badań dotyczą betonu zwykłego o różnej wilgotności masowej, poddanego działaniu obciążeń doraźnych w próbie ściskania quasi-osiowego oraz rozciągania metodą brazylijską.

1. Introduction

Presence of water in a capillary-porous material, as is concrete, shapes its building and functional properties both at the stage of its forming and during its exploitation. Water is introduced to concrete at the technological stage and partly remains in it while its curing, hardening and utilizing. The other source is humidity concrete acquires and which results from sorption processes or capillary pull-up from the surrounding. Depending on how concrete constructions are exploited, the capillary flow of water may take place within a relatively broad range of concrete humidity, i.e. from sorption humidity as far as till a full saturation state [6].

Since constructions made of concrete, depending on the assignment work in various humidity conditions, it is essential to determine the influence of the material's humidity on the course of its destruction caused by a load. The fact that humidity plays an important role in the destruction of loaded concrete was emphasized in, among others, the papers of [10, 11, 12].

However, exhaustive empirical data in this scope, have not yet been obtained.

It is also important to determine the influence of humidity concrete contains on acoustically-measured parameters, especially those obtained owing to the emission method, which serve as the basis for determining the scale or the course of load-caused destruction of this material. The problem should be recognized as crucial as, among others, since 1981 the acoustic emission method has found wider and wider application to studies on concrete in our country [3, 4, 5].

In order to explain the above issues, the investigations were carried out for various bulk humidity concretes subjected to Brazilian method quasi-axial compression. A wide range of the concrete's bulk humidity was achieved by its appropriate storage which rendered it possible for the material to remain in a dry state, a state of maximal sorption humidity and a state of full saturation. The ultrasound and acoustic emission methods were made use of as the investigation techniques.

2. Description of the investigations

The investigations referred to normal concrete of the compression resistance comprised within the B20-class, widely applied to concrete constructions.

The investigations were accomplished after a 90-day curing on samples $100 \times 100 \times 100$ mm and $150 \times 150 \times 150$ mm in size, made of a concrete mix, the formula of which was, in terms of 1 m^3, as follows:

— portland cement "35", coming from The Groszowice Cement Plant	— 320 kg,
— natural gravel "Proszowice"	— 1188 kg,
— river sand "Wrocław"	— 716 kg,
— municipal water	— 177 l.

Depending on the way of storage of the samples, various bulk humidity concrete was obtained at the moment of examination, i.e.:

— concrete of bulk humidity $W_m = 0{,}00\%$	— the samples had been stored for 86 days in a climatic chamber at a temperature of $+18°C$ ($\pm 1°C$) and relative air humidity of ca 95%, and then dried at a temperature of 105°C to reach constant weight (dry state).
— concrete of bulk humidity $W_m = 2{,}28\%$	— the samples had been stored for 28 days in a climatic chamber at a temperature of $+18°C$ ($\pm 1°C$) and relative air humidity of ca 95%, and then, till the 90th day, in the laboratory at a temperature of $+18°C$ ($\pm 3°C$) and relative air humidity of ca 65% (sorption humidity state).
— concrete of bulk humidity $W_m = 2{,}56\%$	— the samples had been stored for 90 days in a climatic chamber at a temperature of $+18°C$ ($\pm 1°C$) and relative air humidity of ca 95% (maximal sorption humidity state).

– concrete of bulk humidity $W_m = 5{,}84\%$ – the samples had been stored for 86 days in a climatic chamber at a temperature of $+18°C$ ($\pm 1°C$) and relative air humidity of ca 95%, and then water-saturated to reach a full saturation state.

Depending on bulk humidity of the concrete, the samples were divided into 4 groups and named series. The following series denotations were accepted:

– the "As" – series, $W_m = 0{,}00\%$;
– the "Ap" – series, $W_m = 2{,}28\%$;
– the "Ak" – series, $W_m = 2{,}56\%$;
– the "Aw" – series, $W_m = 5{,}84\%$.

The accomplished investigations employed the ultrasound and the acoustic emission methods, the samples being compressed quasi-axially and subjected to tension according to Brazilian method. The axial compression was accomplished having withdrawn the friction at the contact spots of the samples' surfaces and the tester's pressure plates. To achieve this, the surfaces has been grinded and then lubricated with cup grease.

The testing stand comprised the following elements:

– Instron 1126 – type tester,
– American, made by Acoustic Emission firm, acoustic emission measurement set,
– acoustic emission detector, AC175-model, of resonance frequency of 175 kHz;
– two-channel X Y_1 Y_2 – recorder manufactured by a Japanese firm Riken Denshi;
– Polish-made ultrasound testing set, Unipan 541-type;
– ultrasound heads of frequency of 500 kHz.

While testing the $100 \times 100 \times 100$ mm concrete samples, the totting sum of the acoustic emission, effective voltage of the acoustic emission, RMS, as well as the transition time of longitudinal ultrasound waves along the direction perpendicular to that of the force action were recorded as a function of stress increment. The

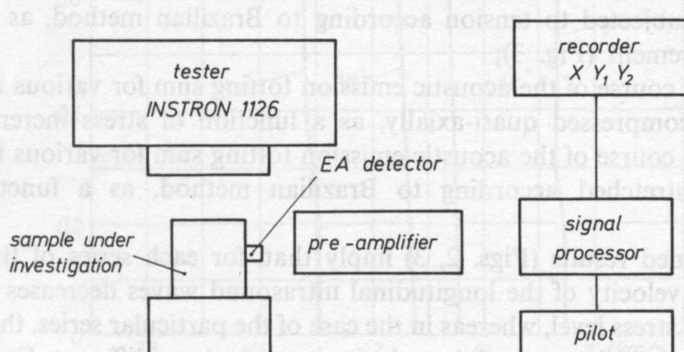


FIG. 1. A diagram of the research stand to measure the acoustic emission.

investigation were carried on at the air temperature of 20°C (±2°C) and relative air humidity of 55% (±5%).

3. Investigation results and their analysis

On the grounds of the classical resistance examinations of the concrete in question, having tested the samples 150×150×150 mm in size, mean compression and tension resistances after a 90-day curing were determined. The obtained results have been displayed in Table 1.

Table 1. Mean compresion and stretching resistances of concrete, determined on 150×150×150 mm samples after a 90-day curing

Concrete series denotation	Concrete humidity	Bulk humidity of concrete W_m [%]	Mean compression resistance $\bar{R}$ [MPa]	Mean stretching resistance $\bar{R}_r$ [MPa]	Softening coefficient while compressing [−]
1	2	3	4	5	6
"As"	dry state	0,00	31,50	1,97	1,00
"Ap"	damp state	2,28	27,00	2,30	0,86
"Ak"	damp state	2,56	27,30	2,40	0,87
"Aw"	water-saturated state	5,84	23,49	2,64	0,75

According to the investigation results obtained with the help of the ultrasound and the acoustic emission methods, the following relationships were determined:
- a velocity change of the longitudinal ultrasound waves for various bulk humidity concrete compressed quasi-axially, as a function of stress increment (Fig. 2);
- a velocity change of the longitudinal ultrasound waves for various bulk humidity concrete subjected to tension according to Brazilian method, as a function of stress increment (Fig. 3);
- variability course of the acoustic emission totting sum for various bulk humidity concrete compressed quasi-axially, as a function of stress increment (Fig. 4);
- variability course of the acoustic emission totting sum for various bulk humidity concrete stretched according to Brazilian method, as a function of stress increment.

The obtained results (Figs. 2, 3) imply that, for each series of the concrete in question, the velocity of the longitudinal ultrasound waves decreases together with an increase of stress level, whereas in the case of the particular series, the stress values at which a distinct decrease of the velocity is merked are different. For instance, the tension test for the "Aw" − series displays the phenomenon at the level of 0.9 $\frac{\sigma}{R}$, for

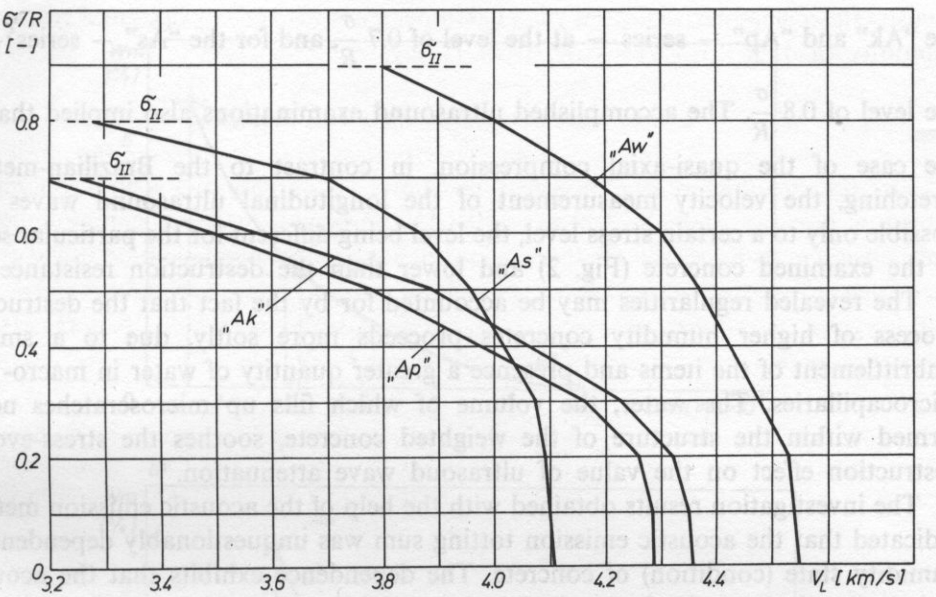


FIG. 2. Velocity change of longitudinal ultrasound waves, for the "As", "Ap", "Ak" and "Aw"-series concrete compressed quasi-axially, as a function of stress increment.

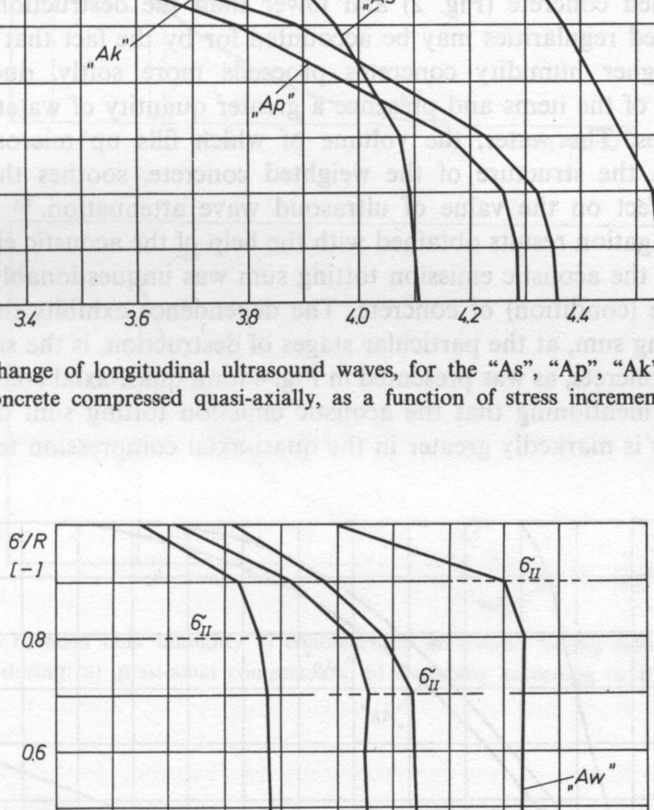


FIG. 3. Velocity change of longitudinal ultrasound waves, for the "As", "Ap", "Ak" and "Aw", series concrete stretched according to Brazilian method, as a function of stress increment.

the "Ak" and "Ap" – series – at the level of $0.7\ \frac{\sigma}{R}$, and for the "As" – series – at the level of $0.8\ \frac{\sigma}{R}$. The accomplished ultrasound examinations also implied that, in the case of the quasi-axial compression, in contrast to the Brazilian-method stretching, the velocity measurement of the longitudinal ultrasound waves was possible only to a certain stress level, the level being different for the particular series of the examined concrete (Fig. 2) and lower than the destruction resistance.

The revealed regularities may be accounted for by the fact that the destruction process of higher humidity concretes proceeds more softly, due to a smaller embrittlement of the items and presence a greater quantity of water in macro- and microcapillaries. The water, the volume of which fills up microscratches newly formed within the structure of the weighted concrete, soothes the stress-evoked destruction effect on the value of ultrasoud wave attenuation.

The investigation results obtained with the help of the acoustic emission method indicated that the acoustic emission totting sum was unquestionably dependent on humidity state (condition) of concrete. The dependence exhibits that the acoustic emission totting sum, at the particular stages of destruction, is the smaller the more humid is the concrete, as was presented in Fig. 4 for a quasi-axial compression test. It is also worth mentioning that the acoustic emission totting sum of the examined concrete series is markedly greater in the quasi-axial compression test as compared

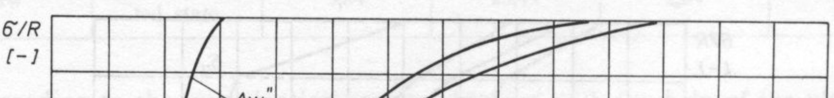


FIG. 4. Variability course of the acoustic emission totting sum, for the "As", "Ap", "Ak" and "Aw"-series concrete compressed quasi-axially, as a function of stress increment.

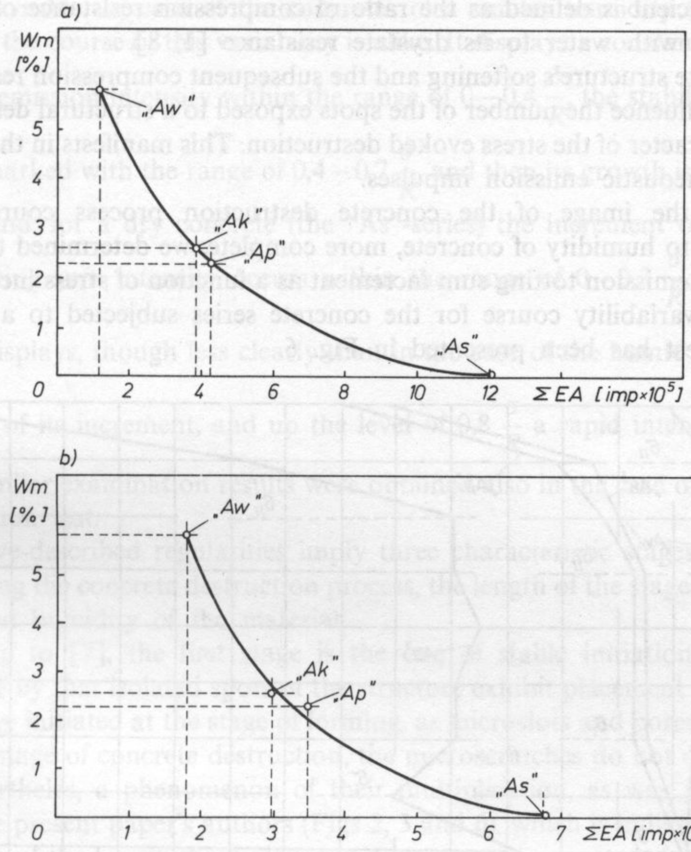


FIG. 5. Relationship between bulk humidity of concrete and an overall totting sum of the acoustic emission occurring during: a) quasi-axial compression, b) stretching according to Brazilian method.

with the tension test according to Brazilian method. This fact is exemplified in Fig. 5, which displays the relationship between bulk humidity of the examined concrete and the overall totting sum of the acoustic emission occuring during quasi-axial compression (Fig. 5a) and stretching according to Brazilian method (Fig. 5b).

Concrete humidity appears to have an influence on the value of the acoustic emission totting sum, the emission occuring during concrete destruction. The influence is associated with a weakening of connections between the structural lattice crystals. The weakening results from the fact that the connections, as the material's humidity increases, are partially dissolved causing a gradual decrease of the concrete's compression resistance [1]. In the case of the examined concrete series, there was also observed a decrease of this resistance, characterized by the walues of the softening coefficient, drawn up in Table 1. It is worth mentioning that the softening coefficient is defined as the ratio of compression resistance of a concrete fully saturated with water to its drystate resistance [1, 8].

The concrete structure's softening and the subsequent compression resistance loss immediately influence the number of the spots exposed to a structural destruction, as well as the character of the stress evoked destruction. This manifests in the number of the recorded acoustic emission impulses.

To make the image of the concrete destruction process course changes, corresponding to humidity of concrete, more complete, we determined the intensity of the acoustic emission totting sum increment as a function of stress increment. For instance, this variability course for the concrete series subjected to a quasi-axial compression test has been presented in Fig. 6.

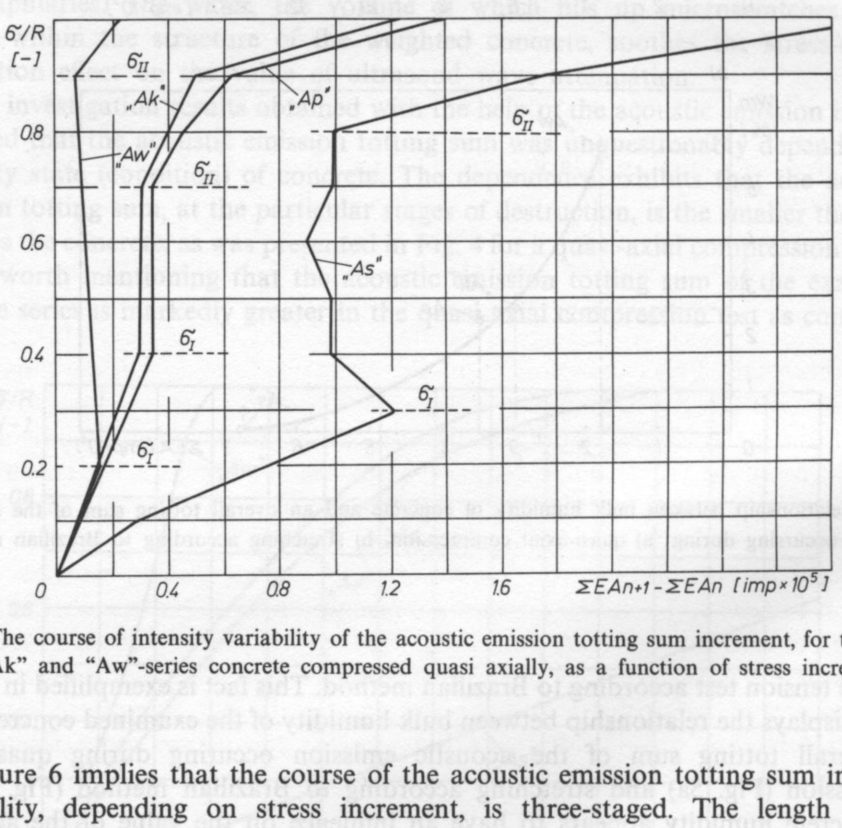


FIG. 6. The course of intensity variability of the acoustic emission totting sum increment, for the "As", "Ap", "Ak" and "Aw"-series concrete compressed quasi axially, as a function of stress increment.

Figure 6 implies that the course of the acoustic emission totting sum intensity variability, depending on stress increment, is three-staged. The length of the particular stages, however, varies with respect to an examined series and depends on the concrete's bulk humidity.

Thus, for water-saturated concrete (the "Aw"-series), the range of $0-0{,}2\ \frac{\sigma}{R}$ displays a constant increment of the acoustic emission totting sum, the range of $0{,}2-0{,}9\ \frac{\sigma}{R}$ exhibits stabilizing of its increment, and above $0{,}9\ \frac{\sigma}{R}$ its repeated increment is visible. In case of the concretes of a similar humidity (the "Ak" and "Ap"-series), the course of this variability is alike. It displays a constant increment of the acoustic emission intensity within the range of $0-0{,}4\ \frac{\sigma}{R}$, the stabilization of this intensity is marked with the range of $0{,}4-0{,}7\ \frac{\sigma}{R}$, and then its growth is observed. On the other hand, for a dry concrete (the "As"-series) the increment of the acoustic emission totting sum intensity occurs within the range of $0-0{,}3\ \frac{\sigma}{R}$; the range of $0{,}3-0{,}8\ \frac{\sigma}{R}$ displays, though less clearly than in the case of the humid concretes, the stabilization of its increment, and up the level of $0{,}8\ \frac{\sigma}{R}$ a rapid intensity growth is observed. Similar examination results were obtained also in the case of the Brazilian method tension test.

The above-described regularities imply three characteristic stages to be distinguished during the concrete destruction process, the length of the stages undoubtedly depending on humidity of the material.

According to [7], the first stage is the one of stable initiation of scratches, characterized by that isolated spots of the structure exhibit placement of microscratches that were initiated at the stage of forming, as microslots and pores. It is peculiar that, at this stage of concrete destruction, the microscratches do not develop. There occurs, nevertheles, a phenomenon of their multiplication, as was proved by the studies of the present paper's authors (Figs 2, 3 and 6), which is indicated by a slight velocity drop of the longitudinal ultrasound waves and a constant increment of the acoustic emission totting sum intensity. An increase of the load makes the concrete destruction process enter the stage of stable propagation of the scratches, during which occur propagation of the microscratches formed at the first stage as well as formation of new, stable microscratches – resulting from either the destruction of adhesion between the aggregate grains and the slurry or the one of the slurry itself. [7]. This state is indicated by a considerable attenuation of the ultrasound waves and a stabilization of the acoustic emission totting sum intensity increment. Provided there is a further increase of the load, the concrete destruction process enters the stage of catastrophic destruction. The stage exhibits formation of distinct, wide scratches spreading out unstably to reach a total destruction of the material. The phenomenon, as was pointed out by the investigations is particularly visible in the quasi-axial compression test and marked by a total loss of measurability of the longitudinal ultrasound wave velocities, and also by a rapid increment of the acoustic emission totting sum intensity.

The boundaries between the particular stages of concrete destruction are determined by the critical stresses, σ_{I} and σ_{II}, which according to, among others, [2, 11] are identified with fatique and long-term resistances of concrete, respectively.

Values of these stresses, which characterize the qualitative changes occuring in the structure of the examined material as the load increases, were determined for the particular series of the concrete under investigation with the help of the ultrasound and the acoustic emission methods, according to the criteria provided in [3]. Then the values were plotted in Figs 2, 3 and 6. Additionally, Fig. 7 submits a matching to

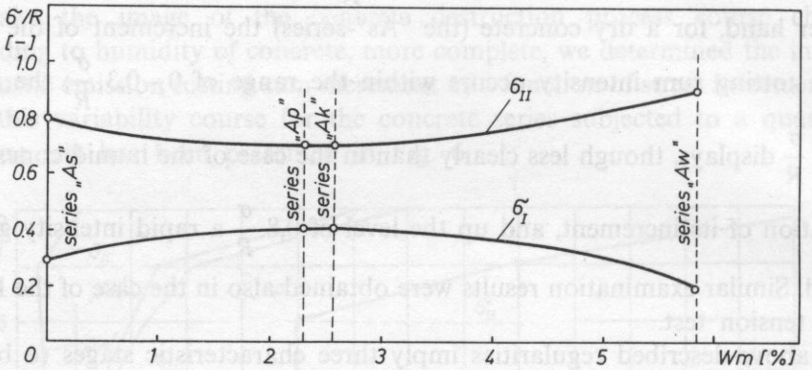


FIG. 7. Variability course of the critical stresses' values, σ_{I} and σ_{II} in relation to bulk humidity of concrete.

show the variability course of the critical stresses, σ_{I} and σ_{II}, in relation to humidity of a concrete under investigation. An analysis of Fig. 7 implies that concrete humidity is of a crucial influence on these stresses' values. Hence, as humidity of a concrete increases, the values of the critical stresses, σ_{I}, initially grow and then, down the bulk humidity value of ca 2,60%, decrease. For the critical stresses, σ_{II}, the relationship is reverse.

It need be noted that the accomplished studies pointed out a great usefulness of the acoustic emission method when determinig the critical stresses' values, σ_{I} and σ_{II} of a concrete, the values distinguishing between the particular stages of concrete destruction. Hitherto, the stresses have been determined on the grounds of knowing the variability course of strain characteristics obtained due to measurements of longitudinal and cross-sectional strains of concrete [2].

5. Summary

The application of acoustic research methods, i.e. the ultrasound and the acoustic emission ones, has proved concrete humidity to have a substantial influence on the process of destruction of concrete. The results obtained owing to these methods, and especially those of the acoustic emission method, clearly prove that three characteristic stages are distinguishable during the concrete destruction process, and that the duration of the stages unquestionably depends on humidity of concrete. It has also been pointed out that the values of the critical stresses, σ_{I} and σ_{II}, constituting boundaries between the particular concrete destruction stages are related to humidity of the material in question.

The valus of the critical stresses, σ_{I}, have been stated to grow initially as concrete humidity increases, from 0,3 to 0,4 $\frac{\sigma}{R}$ and then, starting from bulk humidity of ca 2,60%, to decrease to 0,2 $\frac{\sigma}{R}$. For the critical stresses, σ_{II}, the relationship is reverse, yet initially the values decrease from 0,8 to 0,7 $\frac{\sigma}{R}$ and then grow to 0,9 $\frac{\sigma}{R}$. The occurrence of differences in the course of the process of concrete destruction, the differences being connected with concrete humidity, is also evidenced by differences in totting sum quantities of the acoustic emission recorded during the process of destruction of concrete.

It has also been revealed that the acoustic emission totting sum, the emission being recorded during the destruction process, depends substantially on bulk humidity of the material under investigation. The sum is the smaller the greater bulk humidity of the concrete. The fact need be taken into account when interpreting the results of studies on concrete obtained with the help of the acoustic emission method.

References

[1] Г. Д. Цискрели, *О растяженности армированных бетонов*, Бетон и железобетон, 3, 124–127, (1963).

[2] K. Flaga, K. Furtak, *Wpływ rodzaju kruszywa na poziomy naprężeń krytycznych w betonie ściskanym*, Arch. Inż. Ląd., **26**, 4, 653–666. (1982).

[3] J. Hoła, A. Moczko, *Analiza procesu niszczenia wybranych struktur betonowych przy wykorzystaniu metody ultradźwiękowej i emisji akustycznej*, Praca doktorska, PWr, Wrocław, (1984).

[4] J. Hoła, A. Moczko, J. Pyszniak, *Wpływ wilgotności i naprężeń w betonie ściskanym na jego emisję akustyczną*, XXX Konferencja Naukowa KILiW PAN, Krynica, 3, 175–180. (1984).

[5] J. Hoła, A. Moczko, J. Pyszniak, *Możliwości wykorzystania metody emisji akustycznej w badaniach betonu.* Przegląd budowlany 11, 477–479. (1988).

[6] M. Leśniewska, A. Pogorzelski, *Wpływ porowatej struktury materiałów budowlanych na ruch kapilarny wody*, Arch. Inż. Ląd. **27**, 1, 123–140. (1981).

[7] K. Newman, J. B. Newman, *Failure theories and design criteria for plain concrete.* Civil Engineering Materials Conference: Structure Solid Mechanice and Engineering Design. Procceedings of the Sauthampton, ref. 83, (1969).

[8] J. Pyszniak, *Wpływ wilgotności betonu na ocenę jego wytrzymałości i prędkości fal ultradźwiękowych.* Praca doktorska, PWr, Wrocław. (1963).

[9] П. А. Ребиндер, *Физико-химическая механика как основа закрепления грунтов в дорожном строительстве и производстве строительных материалов на основе грунтов*, Труды совещания по теорителеским основам техники грунтов, МГУ, Москва, (1961).

[10] H. Rüsch, *Physikalische Fragen der Betonprüfung*, Zement – Kalk – Gips, **12**, 1, 1–9, (1959).

[11] S. P. Shah, S. Chandra, *Critical stress volume change and microcracking of concrete.* ACJ Journal, **65**, 9, 770–781, (1968).

[12] P. H. Wittmann, *The structure of hardened cemente paste. A basic for a better understanding of the materials properties.* Proceedings of a Conference held at University of Sheffield, Cement and Concrete Association, London, 96–117. (1976).

Received August 17, 1989

ARCHIVES OF ACOUSTICS
16, 2, 355–367 (1991)

APPLICATION OF CORRELATION METHODS FOR AN INVESTIGATION OF THE ACOUSTIC FIELD IN A ROOM

K. ŚRODECKI

University of Gdańsk, Laboratory of Applied Acoustics and Spectroscopy

A. ŚLIWIŃSKI

University of Gdańsk, Institute of Experimental Physics
(80-952 Gdańsk, ul. Wita Stwosza 57)

The aim of the present work is presenting the possibilities of applying the autocorrelation function of a room defined by Kuttruff for the estimation of some acoustic properties of rooms with various geometries. Within the work measurements tending to specify and perfect Kuttruff's method of autocorrelation function measurement of a room were carried out.

The authors focused, first of all, on the determination of the effect of frequency bandwidth of the signal used for room excitation on the autocorrelation function measured there. The work also presents the measurement results of temporal diffusion Δ – the parameter introduced by Kutruff for a quantitative estimation of autocorrelograms registered in the room. The measurements were varried out in three different rooms.

The investigation prove that the autocorrelation function measurement is particularly useful in the estimation of the number and of the time distribution of strong reflections or sequences of reflections, subjectively perceived as echo, flutter echo or sound coloration. It allows also to determine the degree of sound dispersion in a room and thus it may help to estimate the quality of the reverberation decay of sound energy in the room.

Celem niniejszej pracy było wykazanie możliwości zastosowania funkcji autokorelacji pomieszczenia, zdefiniowanej przez Kuttruffa, do oceny niektórych własności akustycznych wnętrz o różnej geometrii. W ramach pracy przeprowadzono badania mające na celu uściślenie i udoskonalenie metody Kuttruffa pomiaru funkcji autokorelacji pomieszczenia.

Skoncentrowano się przede wszystkim na określeniu wpływu szerokości pasma częstotliwości sygnału używanego do pobudzenia pomieszczenia na mierzoną w tym pomieszczeniu funkcję autokorelacji. W pracy przedstawiono również wyniki pomiarów dyfuzyjności czasowej Δ – parametru wprowadzonego przez Kuttruffa dla ilościowej oceny rejestrowanych w pomieszczeniu autokorelogramów. Pomiary te wykonano w trzech różnych pomieszczeniach.

Na podstawie przeprowadzonych badań stwierdzono, że pomiar funkcji autokorelacji szczególnie przydatny jest do oceny liczby i rozkładu czasowego silnych odbić czy też ciągów odbić, odczuwanych subiektywnie jako echo, echo drżące lub zmiana barwy dźwięku. Pozwala również na określenie stopnia rozproszenia dźwięku w pomieszczeniu a zatem służyć może do oceny jakości pogłosowego zaniku energii dźwiękowej w pomieszczeniu.

1. Introduction

An important parameter determinig the acoustic quality of a room is the degree of sound diffusion in a room as well as the amplitude and time structure of subsequent reflections, particularly early ones, reaching the listener [7, 5]. Both characteristics can be found by measuring the autocorrelation function of acoustic signals in a room. This function, as one of four statistic functions, describes the main properties of random signals such as sound of music or speech spreading around the room.

In 1966 KUTTRUFF [9] suggested the application of the defined by him autocorrelation function of a room to investigations of the time structure of the room response to pulse excitation. The introduced by him parameter — "temporal diffusion" Δ was to determine the degree of sound diffusion in a room basing on the measured autocorrelation function of a room. Further works by BILSEN [8] using the assumptions of Kuttruff's theory and the results of investigations made by SCHRODER, ATAL and KUTTRUFF [4] allowed to find the hearing threshold of the disadvantageous phenomena subjectively perceived as sound coloration, echo or flutter echo.

In recent years, due to the fast development of modern measurement techniques, based on digital analysis of acoustic signals, it has become possible to intensify the investigations concerning the application of autocorrelation function in room acoustics. The progress is reflected particularly in ANDO's works [1, 2, 3] who, basing on subjective investigations, introduced correlation criteria of time structure preference of early reflections reaching the listener and optimal reverberation time.

The investigation undertaken by the authors have in view a verification and specification of Kuttruff's method of autocorrelation function measurement of a room with a particular regard to the effect of frequency bandwidth of the signal exciting the room on the measured autocorrelation function.

The results of these investigations are presented in the first part of the present work. The second part is an attempt of showing some possibilities of the application of autocorrelation function to estimate some acoustic properties of selected rooms.

2. Main assumptions

2.1. *Autocorrelation function of a room*

Let us assume that the investigated room can be treated as a linear, stationary dynamic system. While exciting it with a short duration pulse, theoretically determined e.g. by Dirac pulse $\delta(t)$, we obtain the response of the room as:

$$h_0(t) = \sum_k a_k \delta(t - t_k) \tag{2.1}$$

where t_k — time moments in which pulse reflections from the surfaces limiting the room reach the measurement point. They are of stochastic character and the more randomly located they are on the time axis, the more advantageous it is for a good acoustics of the room. (We assume at the moment that the reflected pulses do not undergo frequency deformation and are merely amplitude weakened – factor a_k is the measure of this weakening).

The autocorrelation function of the room has been defined by KUTTRUFF [9] and described by the formula:

$$\phi(t) = \int_{-\infty}^{+\infty} h_0(\tau)h_0(\tau+t)d\tau = \int_{-\infty}^{+\infty} h_0(-\tau)h_0(t-\tau)d\tau \tag{2.2}$$

Substituting Eq. (2.1) into Eq. (2.2) we can determine the autocorrelation function by the following dependence:

$$\phi(t) = \sigma(t)\sum_k a_k^2 + \sum_{k\neq l}\sum a_k a_l \delta(t+t_k-t_l) = \phi_1(t)+\phi_2(t) \tag{2.3}$$

The above equation presents the sum of subsequent multiple reflections of the signal $\delta(t)$. Component $\phi_1(t)$ determines the main maximum of autocorrelation function for $t = 0$ whose height amounts to $\sum a_k^2$ and $\phi_2(t)$ expresses the remaining part of the autocorrelogram, the so called "tail" of the autocorrelation function with mean height $\overline{a_k \cdot a_l}$; it is absolutely valid for completely stochastic reflections.

In the case of reflections of the character determined in the autocorrelogram (multiple reflections in intervals $t_l - t_k$) apart from the main maximum, lateral maxima will appear for $t = t_l - t_k$. It is quite obvious because of the properties of autocorrelation function due to which we can separate the periodic component from the random signal.

Thus the autocorrelation of the pulse response of a room for a given point of the field perfectly accounts for the time structure of this response. For the quantitative estimation of the time structure of the room response to pulse excitation Kuttruff introduced the parameter "temporal diffusion" and its measure – the ratio of the main maximum $\phi_1(t)$ for $t = 0$ to the subsequent magnitude maximum, belonging to $\phi_2(t)$ [9]:

$$\Delta = \frac{\phi_1(0)}{\mathrm{Max}(\phi_2)} \tag{2.4}$$

Equation (2.2) is equivalent to the following dependence [6], [10]:

$$\phi(t) = \mathscr{F}^{-1}\{S(f)\} \tag{2.5}$$

where $S(f)$ is the power spectrum of a time signal $h_0(t)$ and $\mathscr{F}^{-1}$ – the inverse Fourier transform. The power spectrum $S(f)$ is found here as the squared modulus of Fourier transform of the signal $h_0(t)$:

$$S(f) = |\mathscr{F}\{h_0(t)\}|^2 \tag{2.6}$$

The method of finding the autocorrelation function basing on Eq. (2.2) is a conventional one, while the method using Eq. (2.5) is an indirect method consisting in, first of all, the calculation of the power spectrum by means of the forward Fourier transformation and then calculation of the inverse Fourier transform of the power spectrum.

3. Experiment

3.1. The investigated rooms

In the investigations concerning the verification of Kuttruff's method of autocorrelation function measurement of a room and in those concerning the possibility of the application of autocorrelation function for the estimation of some acoustic properties of rooms, three rooms were used: the room of the Mathematics-Physics-Chemistry Department Council of Gdańsk University (room A), Auditorium 3 of Physics Institute (room B), and Auditorium I of the Electronics Department of Gdańsk Technical University (room C).

Room A of volume 204 m^3 has perpendicular walls, its length – 10.3 m, width – 5.65 m, height – 3.5. It is a typical conference room whose scheme and the location of measurement points is presented in Fig. 1 The recommended value of reverberation time T for such room amounts to 0.75 s ($\pm$ 10%) and should be constant in the frequency function [7].

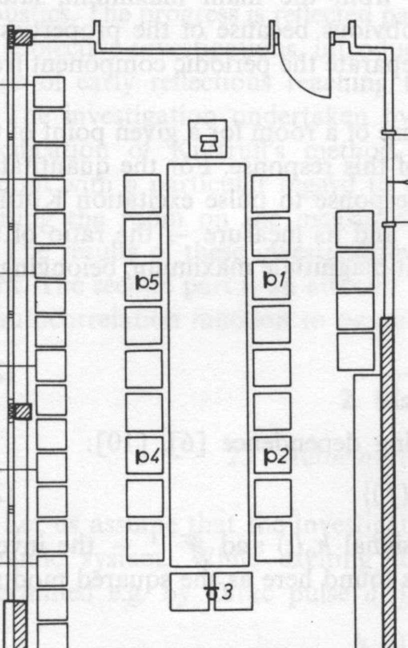


FIG. 1. Horizontal section of room A

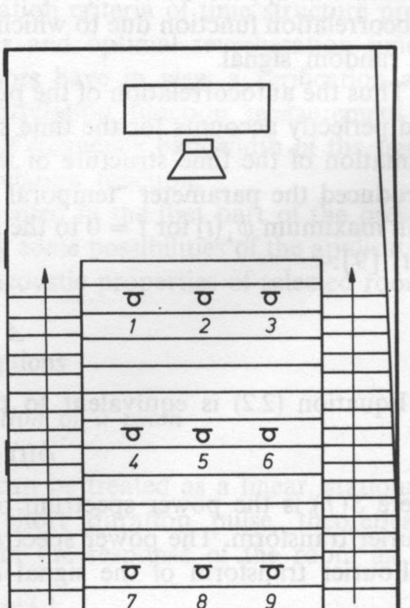


FIG. 2. Horizontal section of room B

Room B of volume 876 m^3 is amphitheatric. Its length amounts to 17.65 m, width in the widest place – 10 m, in the narrowest one – 9,4 m. The height on zero level (at the entrance) amounts to 2.85 m. The floor slopes down reaching the level of – 3.35 m on the opposite side of the room. The scheme and measurement points location is presented in Fig. 2. The recommended value of reverberation time for such room amounts to 0.8 s ($\pm 10\%$).

Room C of volume 1195 m^3 is also amphitheatric. Its length – 13.5 m, width in the widest place – 17.5 m, in the narrowest 12 m. The height on zero level (at the entrance) amounts to 3.5 m. The floor slopes down reaching the level of – 3 m on the opposite side of the room.

The scheme with the measurement points is presented in Fig. 3 The recommended value of reverberation time for such room is 0.85 s ($\pm 10\%$).

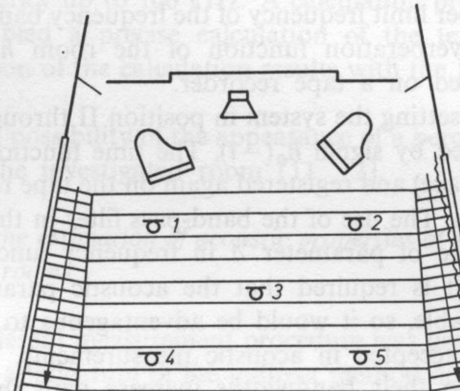


FIG. 3. Horizontal section of room C

3.2. Measurement procedure applied to verification of Kuttruff's theory of autocorrelation function estimation

The measurement of autocorrelation function in the investigated room by Kuttruff's method can be carried out in a simple way. The autocorrelation function is registered on a tape recorder just at a given measurement point as a response of the room to the signal being the reciprocal of the pulse response $h_0(-t)$ registered earlier at this measurement point. So, Kuttruff's metod is a conventional one (Eq. (2.2)) in which the investigated room plays the role of a correlator.

The schematic diagram of the instrumentation for the registration of the autocorrelation function by Kuttruff's method is presented in Fig. 4.

The first stage of the measurement is carried out in the position of connections I. The rectangular pulse from the generator is transferred to a band-pass filter and a loudspeaker.

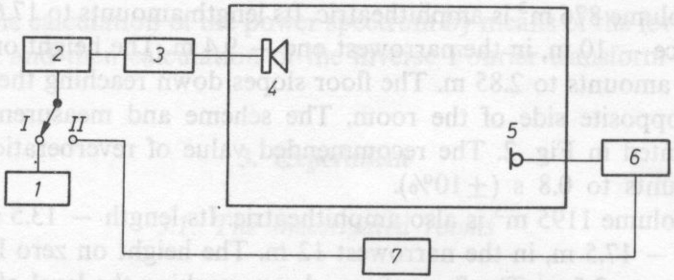


FIG. 4. Schematic diagram of instrumentation used in autocorrelation function measurement (Kuttruff's method): 1 — pulse generator, 2 — band-pass filter, 3 — power amplifier, 4 — isotropic sound source, 5 — condenser microphone, 6 — measuring amplifier, 7 — tape recorder.

The pulse duration should be shorter than a half of the period corresponding to the upper limit frequency of the frequency band of the filter used in the measurement. The reverberation function of the room h_0 (t) received by the microphone is registered on a tape recorder.

On setting the system in position II through the filter and loudspeaker the room is excited by signal $h_0(-t)$. The time function received by the microphone (in the same place) and registered again on the tape recorder is the required autocorrelation function. The use of the band-pass filter in the measurement system enables to find the value of parameter Δ in frequency function.

As it is required that the acoustic parameters of the room should be easily measurable, so it would be adventageous to use filters of width 1/1 or 1/3 octave widely accepted in acoustic measurements. Those filters are constant percentage filters so their bandwidths increase with the increasing measurement frequency.

Kuttruff tried to find the effect of the frequency-band-widening on the measured autocorrelation function. His theoretical formula [9]:

$$S = \sqrt{\frac{B'}{B}} \tag{3.1}$$

where B and B' are widths of the successive frequency bands of the filter used in the measurement, determines this effect in the case of an almost ideal sound diffusion. Values Δ obtained for higher frequencies should then be corrected — i.e. divided by coefficient S to compare them with values Δ obtained for lower frequencies.

In order to work out a univocal method of temporal diffusion measurement so that it can be widely used for objective estimations of subjectively perceived phenomena in rooms, it was necessary to carry out experiments which could specify the dependence of coefficient Δ on the bandwidth of the signal exciting the room. These investigations were made in room B.

The measurements were made with band filters of centre frequencies $f_0 = 500$, 1000, 2000 and 4000 Hz. For each measurement frequency the bandwidth B of the used filter was changed. The values B used in the investigations are presented in Table 1. The analysis of the room autocorelation functions was made by means of an

Table 1

Center frequency of band-pass filter f_0 [Hz]	Frequency bandwidth of filter used in measurements B [Hz]				
500	65	115	200	350	620
1000	130	230	400	700	1230
2000	260	460	800	1400	2500
4000	530	920	1600	2800	4900

MSM microcomputer (compatible with IBM XI) coupled with a 12-bit analogue--to-digital converter of sampling frequency up to 100 kHz. A calculation programme has been worked out which enabled a precise calculation of the temporal diffusion value Δ as well as a comparison of the calculation results with the Bilsen's weighting function [8].

On that basis, we could determine a possibility of the appearance of a perceptible flutter echo or sound coloration in the investigated room [11, 12].

3.3. Measurement procedure applied to the estimation of acoustic properties of selected rooms

In these investigations a slightly different measurement procedure was used than that described in ch. 3.2 basing on the application of the indirect method described by Eq. 2.5 [13]. The apparatus described in ch. 3.2 (Fig. 4) was used in the investigations but the measurement was limited only to the first stage which corresponds with the connection system marked by I in Fig. 4. Thus in the investigated room a registration of the room response to pulse excitation $h_0(t)$ was made and then a computer estimation of autocorrelation function by FFT method was carried out.

The worked out calculation programme for the estimation of autocorrelation function also enabled a precise analysis of the obtained autocorrelograms. The above investigations were carried out in rooms A, B and C. The investigated rooms were band-excited at each measurement point while the centre frequencies of the used 1/3 and 1/1 octave filters amounted to: 500, 1000, 2000 and 4000 Hz.

4. Results of investigations

4.1. Measurement of the dependence of temporal diffusion value Δ on the frequency bandwidth of the exciting signal

The dependence of the parameter value Δ on the bandwidth of the signal used for the excitation of the room for selected measurement points and frequencies is

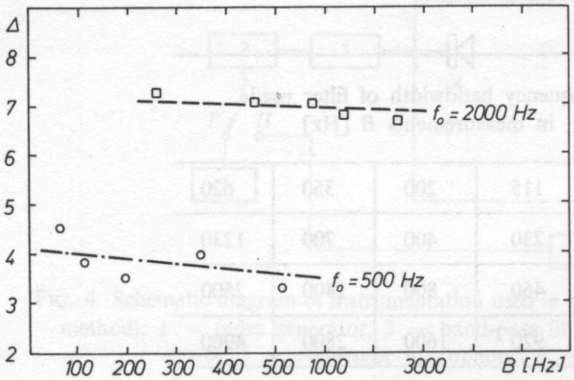


FIG. 5. Dependence of temporal diffusion Δ on the frequency bandwidth of signal used for room excitation (room B, measuring point 3). o: $f_0 = 500$ Hz, □: $f_0 = 2000$ Hz

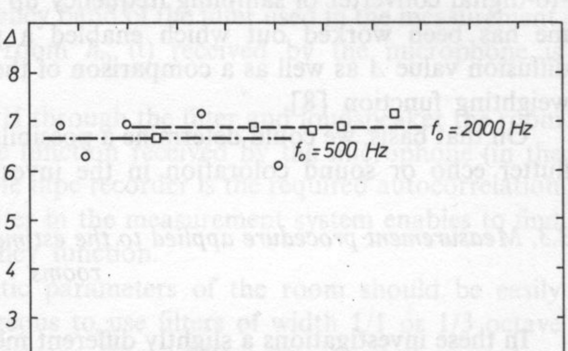


FIG. 6. Dependence of temporal diffusion Δ on the frequency bandwidth of signal used for room excitation (room B, measuring point 6). o: $f_0 = 500$ Hz, □: $f_0 = 2000$ Hz

presented in Fig. 5 and 6. This choice has been conditioned by the assumptions of Kuttruff's theory [9] according to which formula (3.1) is valid for a large temporal diffusion of the room. Kuttruff suggested a correction of the measured values of the temporal diffusion coefficient for $\Delta > 3.5$.

From the dependencies presented in Fig. 5 and 6 we can see a clearly insignificant effect of the bandwidth of the signal exciting the room on the value of temporal diffusion for the same measurement frequency. In the case of the measurements points and frequencies for which Δ was smaller than 3, the dependences of temporal diffusion values on the frequency bandwidth of the exciting signal were similar. Apparently it agrees with Kuttruff's theory according to which formula (3.1) in not valid for small temporal diffusion of the room.

4.2. Application of autocorrelation function for the estimation of acoustic properties of selected rooms

The results of investigations concerning the above subject are presented in Fig. 7. Each figure contains mean values of temporal diffusion Δ in frequency function for

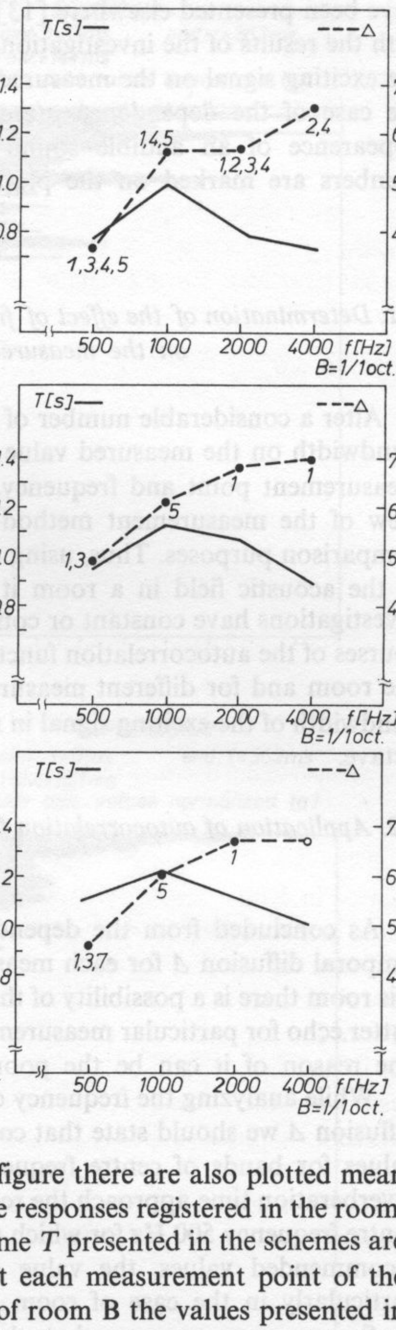


FIG. 7. Reverberation time T and temporal diffusion Δ in octave bands: a) room A, b) room C, c) room B.

each investigated room (rooms A, B, C). On each figure there are also plotted mean values of reverberation time T found from the pulse responses registered in the room. The values of the coefficient Δ and reverberation time T presented in the schemes are arithmetic means of values Δ and T measured at each measurement point of the investigated rooms (Fig. 1, 2, 3). Only in the case of room B the values presented in Fig. 7 are arithmetic means of the values measured only at measurement points 1, 3, 5, 7 and 9.

The dependence presented in Fig. 7 are plotted for octave bands. The respective values of temporal diffusion Δ and reverberation time T found for 1/3 octave bands have been presented elsewhere [13]. The differences are slight which is in agreement with the results of the investigations concerning the effect of frequency bandwidth of the exciting signal on the measured autocorrelation function presented in ch. 4.1. In the case of the dependence presented in Fig. 7 black circles mark the possible appearence of an audible sound coloration at the measurement points whose numbers are marked on the plot.

5. Conclusions

5.1. Determination of the effect of frequency bandwidth of the signal exciting the room on the measured value of temporal diffusion

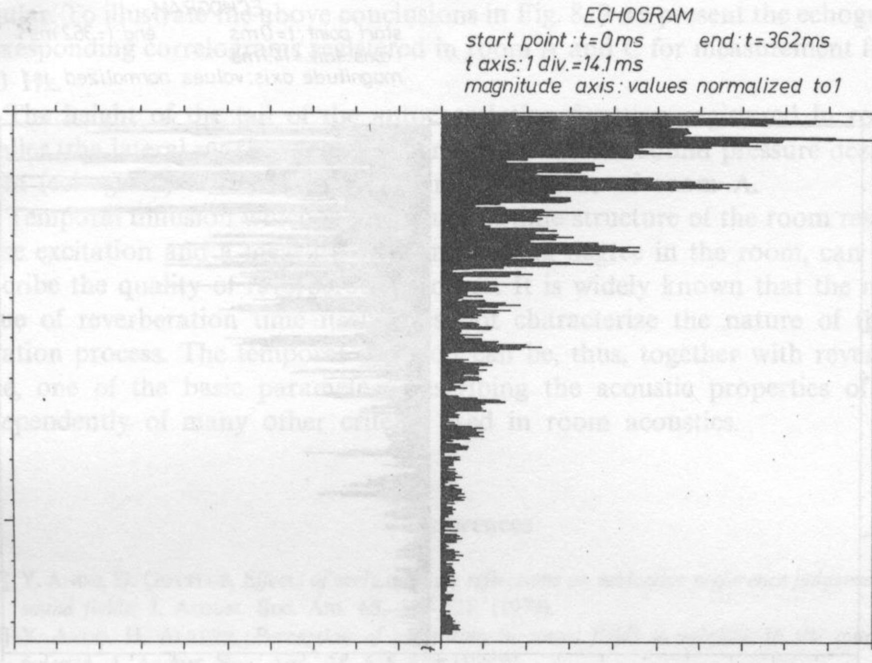
After a considerable number of measurements carried out, no effect of frequency bandwidth on the measured value of temporal diffusion was observed for the same measurement point and frequency. It is an essential conclusion from the point of view of the measurement method and the procedure of finding coefficient Δ for comparison purposes. Thus, using the autocorrelation function for the investigation of the acoustic field in a room it is not essential whether the filters used in the investigations have constant or constant percentage bandwidth. When analysing the courses of the autocorrelation function for different bandwidths of the signal exciting the room and for different measurement frequencies it was found that the optimal bandwidth of the exciting signal in measurement of this type is the width equal to 1/1 octave.

5.2. Application of autocorrelation function for the estimation of acoustic properties of rooms

As concluded from the dependences presented in Fig. 7, the largest values of temporal diffusion Δ for each measurement frequency were obtained in room C. In this room there is a possibility of the appearance of a perceptible sound coloration or flutter echo for particular measurement points. A reverse situation occurs in room. A. The reason of it can be the poor geometry of room A.

While analyzing the frequency characteristics of reverberation time and temporal diffusion Δ we should state that coefficient Δ in the investigated rooms reaches high values for bands of centre frequency 2000 and 4000 Hz for which the values of reverberation time approach the recommended values. Still in the case of the band of centre frequency 500 Hz for which the values of reverberation time also approach the recommended values, the value of temporal diffusion decreases. It is observed particularly in the case of room A.

So, we may suppose that the value of temporal diffusion is not so much connected with the value of reverberation time but rather with its "quality" — the bigger the temporal diffusion coefficient, the higher the quality of reverberation course, that is, the course of sound energy decay in the room is smoother, more



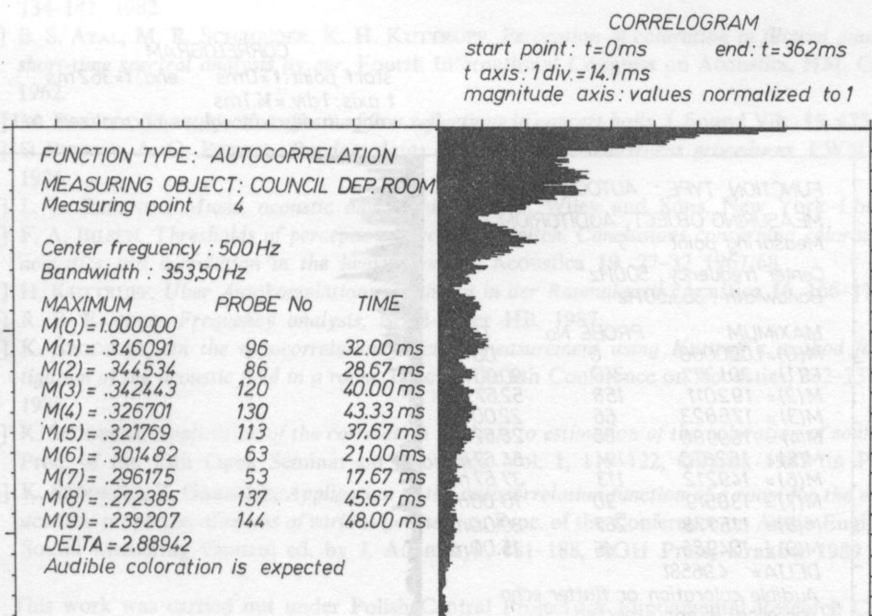


FIG. 8. Echogram and correlogram obtained from the response of room A.

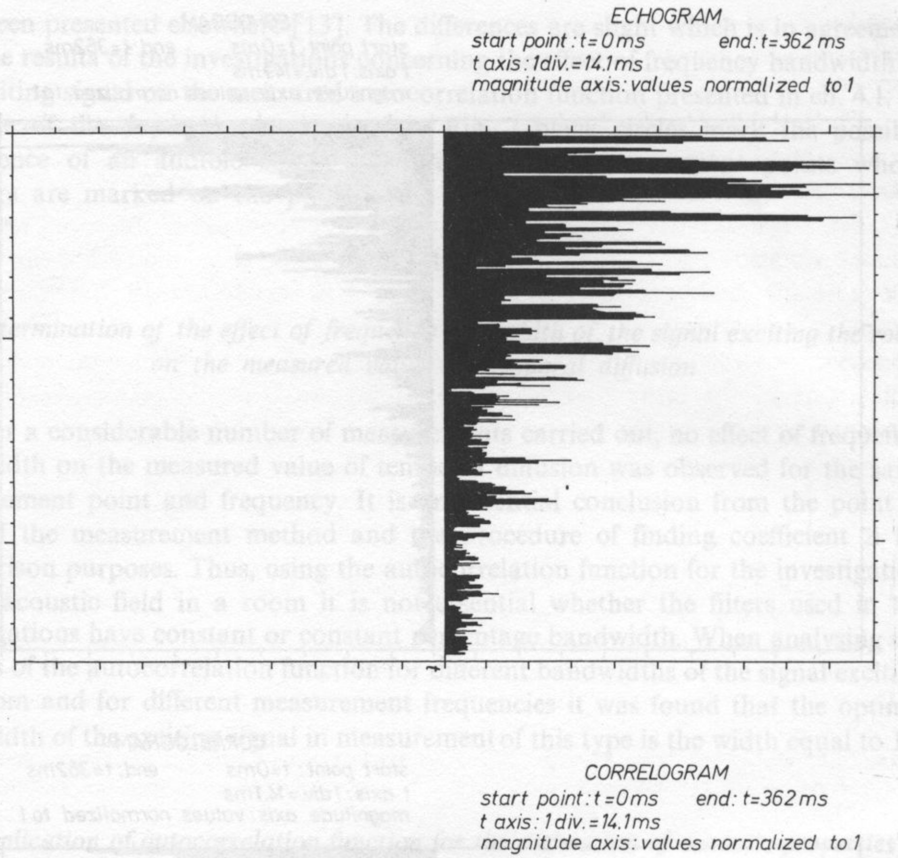


FIG. 9. Echogram and correlogram obtained from the response of room C.

regular. To illustrate the above conclusions in Fig. 8, 9 we present the echograms and corresponding correlograms registered in room A and C for measurement frequency 500 Hz.

The height of the tail of the autocorrelation function registered in room C is smaller (the lateral maxima are less significant) and the sound pressure decay in the room (echogram) is more regular than in the case of room A.

Temporal diffusion which is a measure of time structure of the room response to pulse excitation and a measure of sound mixing degree in the room, can precisely describe the quality of reverberation course. It is widely known that the measured value of reverberation time itself does not characterize the nature of the reverberation process. The temporal diffusion can be, thus, together with reverberation time, one of the basic parameters describing the acoustic properties of a room independently of many other criteria used in room acoustics.

References

[1] Y. Ando, D. Gottlob, *Effects of early multiple reflections on subjective preference judgements of music sound fields*, J. Acoust. Soc. Am. **65**, 524–527, (1979).

[2] Y. Ando, H. Alrutz, *Perception of coloration in sound fields in relation to the autocorrelation function*, J. Acoust. Soc. Am., **71**, 616–618 (1982).

[3] Y. Ando, M. Okura, K. Yuasa, *On the preferred reverberation time in auditoriums*, Acustica **50**, 134–141, 1982.

[4] B. S. Atal, M. R. Schroeder, K. H. Kuttruff, *Perception of coloration in filtered gaussian noise; short-time spectral analysis by ear*, Fourth International Congress on Acoustics, H**31**. Copenhagen 1962.

[5] M. Barron, *The subjective effects of first reflections in concert halls*, J. Sound Vib., **15**, 475–494, 1971.

[6] S. Bendat, A. G. Piersol, *Random data: analysis and measurement procedures*, PWN, Warszawa 1976.

[7] L. L. Beranek, *Music, acoustic and architecture*, J. Wiley and Sons, New York–London 1962.

[8] F. A. Bilsen, *Thresholds of perception of repetition pitch. Conclusions concerning coloration in room acoustics and correlation in the hearing organ*, Acoustica **19**, 27–32 1967/68.

[9] H. Kuttruff, *Uber Autokorrelationsmessungen in der Raumakustik*, Acustica **16**, 166–174, 1965/66.

[10] R. B. Randal, *Frequency analysis*, Bruel-Kjaer HB, 1987.

[11] K. Środecki, *On the autocorrelation function measurement using Kuttruff's method for an investigation of the acoustic field in a room*, Proc. of the 9th Conference on Acoustics, 232–237, Budapest 1988.

[12] K. Środecki, *Application of the correlation analysis to estimation of the coloration of sound in rooms*, Proc. of the 13th Open Seminar on Acoustics, Vol. **1**, 119–122, Gdańsk 1983 (in Polish).

[13] K. Środecki, Z. Garnusz, *Application of the autocorrelation function of a room for the estimation of acoustic properties of rooms of various geometries*, Proc. of the Conference on Audio Engineering and Sound Mastering Control ed. by J. Adamczyk, 181–188, AGH Press, Kraków 1989 (in Polish).

This work was carried out under Polish Central Project for Fundamental Research CPBP 02.03.

Received September 12, 1989

ARCHIVES OF ACOUSTICS
16, 2, 369–381 (1991)

DIRECT APPLICATION OF LINGUISTIC VARIABLE TO THE SPEECH SEGMENT DESCRIPTION IN ISOLATED WORD RECOGNIZER

W. W. WIĘŹLAK

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

A simple model of speech recognition based on identification of broad articulatory classes is presented. Direct application of the notion of linguistic variable to description of acoustic parameters evolutions and recognition of broad articulatory classes is discussed. The recognition algorithm is based on a fuzzy automatic network. An example of application of the described method to isolated word recognition is given.

1. Introduction

One of the most essential and important problems in the automatic speech recognition task is that of finding the relations between acoustic signal continuum and the string of phonetic classes which is related to the signal. More precisely, the relations between acoustic cues for the description of certain physical characteristics of the speech signal and phonetic features employed for characterization of speech units, mainly phonemes. Years of research in phonetics and speech physiology have shown that the solution of the "acoustic cues vs. phonetic features" problem is neither simple nor unambiguous and the efforts to established automatic phonetic transcription of the acoustic speech signal have been succesful only for strictly limited conditions. One of the main reasons of such results was that we have to find the complex relations between psychological, i.e. rational rather than physical, units as phonemes, which are defined as the smallest distinctive elements of a word for a particular language, and physical, i.e., measurable parameters like speech spectrum, LPC coefficients etc. Although very sophisticated and precise methods of speech signal measurements and description have been developed, one has to tackle the problem of the descriptive character of phonetic concepts. This vagueness of description results mainly from the approximate definitions of phonetic classes in terms of acoustic cues as pointed above, but another source of vagueness are the coarticulation effects which make the consecutive speech segments influence one another, changing their acoustic characteristics without modifying their phonetic meanings. In addition, individual and statistical differences across speakers and tokens make the whole task very complicated.

For automatic recognition of isolated or connected words, various strategies of solving this problem have been proposed, from the deterministic pattern recognition methods used by Bezdel [1], Sambur and Rabiner [2] or Weinstein and al. [3] to statistical pattern recognition, e.g., [4] and Bayesian classifier [5]. The fuzzy theoretical approach has also attracted the attention of researchers in the last 10 years, offering philosophical concepts and mathematical apparatus which was thought to formalize and overcome the inherent vagueness of the relations between acoustic cues and phonetic features [6], [7].

The study presents how to use the concept of a linguistic variable in describing and recognizing certain phonetic classes in the Isolated Word Recognizer (IWR). Apart from the application of the fuzzy theoretical framework to the system, i.e., direct case of the notion of linguistic variable and fuzzy automata, it is also shown how a careful analysis of a very poor set of parameters, only three used in the present study, supported by an important amount of subjective and objective knowledge incorporated in the recognition algorithm could be used for successful identification of certain broad phonetic classes. General speech description strategy is described elsewhere [7], [8]; however, it consists mainly in detecting and describing certain acoustic events related to the occurrence of broad articulatory phonetic classes.

2. Speech material and parameter extraction

The vocabulary used in this study consists of 60 words (10 digits + 50 command words) provided for a voice controlled minicomputer. Each 60 word was spoken at normal speed by 6 subjects (5 men, 1 woman), two of them with slurred pronunciation. This set made up the learning group. The algorithm was tested on another set of 6 speakers (5 men, 1 woman), 5 of them were new to the system.

The block diagram of IWR is shown in Fig. 1. The incoming speech signal is analysed by the Parametric Speech Analyzer, built at the Speech Acoustic Laboratory IFTR. At the output of the analyzer, a sequence of parameters is sent to the PC microcomputer every 10 ms. Seven parameters were measured:

1. AO — log overall amplitude envelope,
2. LP — log amplitude in low frequency band, 80–800 Hz,
3. HP — log amplitude in high frequency band, 4.5–8 kHz,
4. $F1$, $F2$ — first and second formant frequencies,
5. ZCR — zero-crossing rate,
6. FO — fundamental frequency.

However, only the first three of them AO, LP, HP were used in the study (Fig. 1).

3. The set of speech segments categories

The set of labels used for word description in IWR consisted of 9 names of broad acoustic articulatory classes which resulted from the general classification of

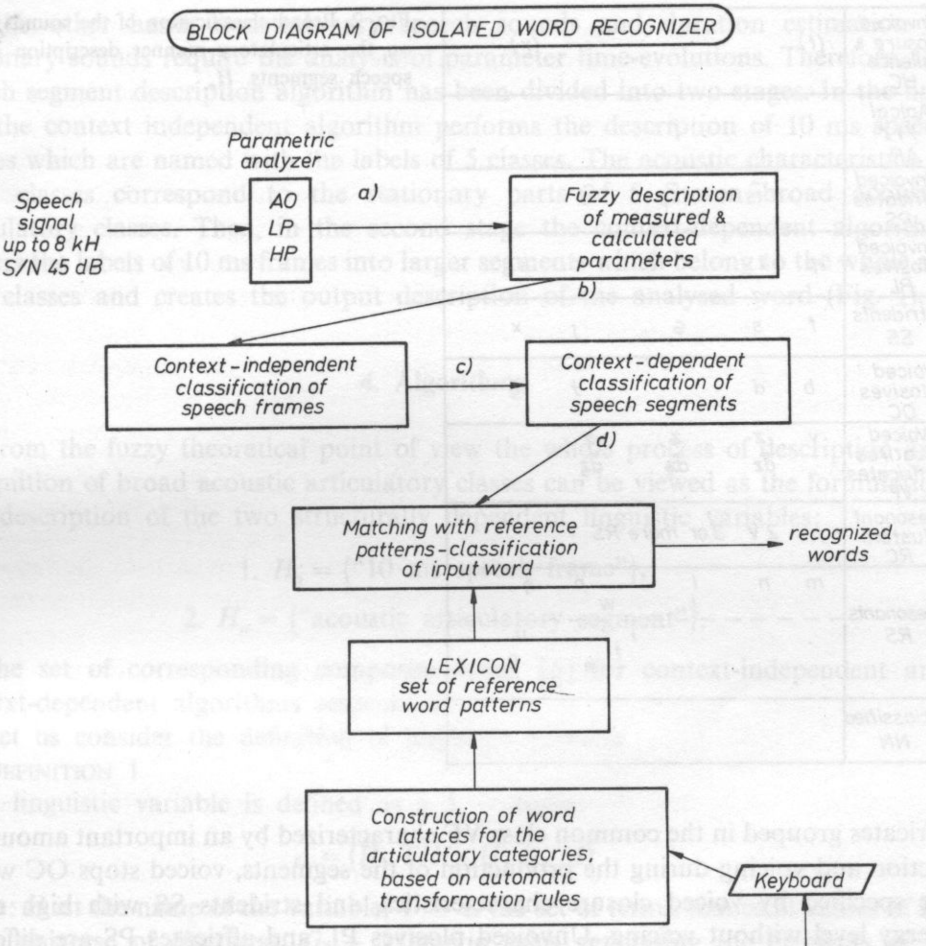


FIG. 1. Block diagram of the isolated word recognizer a) data matrix, b) description of data values with linguistic terms as "low", "medium", "high",... c) string of composite terms describing linguistic variable $H_s = \{10 \text{ ms speech frame}\}$, d) string of composite terms articulatory categories — manner describing linguistic variable $H_a = \{\text{articulatory segment}\}$.

speech sounds based on the manner of articulation, as it is shown in Fig. 2. Resonants are defined as sounds produced with such a suitable shaping of the vocal tract that the airflow through the mouth and/or nostrils is free. This class contains vowels, nasals, glides and laterals. They are divided into two classes RS and RC on the basis of their duration. All other sounds are called obstruents, made by obstructing the flow of air. The are characterized by an extreme narrowing or constriction at some point of the vocal tract resulting in blocking of the air stream, or in its turbulent character. The definition of obstruents implies various acoustic phenomena to be analysed, within the obstruent part of the speech signal. Therefore they have been divided into 7 classes of segments. There are voiced fricatives and

Unvoiced closure & silence UC	(f) (x)
Apical /r/ AR	r
Unvoiced affricates PS	t͡s t͡ɕ t͡ʃ
Unvoiced plosives PL	p t c k
Stridents SS	f s ɕ ʃ x
Voiced plosives OC	b d ɟ g
Voiced fricatires & affricates VF	v z ʑ ʒ d͡z d͡ʑ d͡ʒ
Resonant clusters RC	2 V 3 or more RS
Resonants RS	m n l ɲ ŋ j w - - - - - - - - - - - - i u ɨ o e a
Unclassified NN	

FIG. 2. Broad classification of the sounds based on the articulatory manner description of the speech segments. H_a

affricates grouped in the common class VF characterized by an important amount of friction and voicing during the production of the segments, voiced stops OC which are specified by voiced closure characteristics, and stridents SS with high noise energy level without voicing. Unvoiced plosives PL and affricates PS are differentiated through the amount of noise and its duration after the unvoiced closure UC. Apical /r/ (AR), the most common realization of the /r/ sound in Polish, presents itself in a different way in the acoustic signal but its salient feature is a sequence of short resonant-obstruent units. In intervocalic position it shows a distinct short dip in the overall amplitude of the signal and before and after resonants it results in saw-tooth-like amplitude variations on the raising or falling edge. So-called unvoiced closure UC is mainly caused by blocking of the airstream but sometimes also by weak aspiration or delay of voicing onset in the sequence of segments with strongly different places of articulation.

Among obstruents one can find two types of sounds: quasi-stationary ones, these which can be produced without restricted timing, e.g., fricatives underlaying the classes SS and VF, and nonstationary speech sounds which without restricted timing loose their phonemic character or cannot be produced, e.g., plosives, /r/ sound. Certainly, stationary segments like SS, UC, partially VF as well as resonants may be identified only through the analysis of instantaneous values of acoustic parameters.

On the other hand, nonstationary speech sounds and duration estimation of stationary sounds require the analysis of parameter time-evolutions. Therefore, the speech segment description algorithm has been divided into two stages. In the first one the context independent algorithm performs the description of 10 ms speech frames which are named with the labels of 5 classes. The acoustic characteristics of these classes correspond to the stationary parts of 5 chosen broad acoustic articulatory classes. Then, in the second stage the context-dependent algorithm gathers the labels of 10 ms frames into larger segments which belong to the whole set of 9 classes and creates the output description of the analysed word (Fig. 1).

4. Algorithms

From the fuzzy theoretical point of view the whole process of description and recognition of broad acoustic articulatory classes can be viewed as the formulation and description of the two structurally dependent linguistic variables:

1. $H_s = \{$"10 ms speech frame"$\}$,
2. $H_a = \{$"acoustic articulatory segment"$\}$,

by the set of corresponding composite terms [6] for context-independent and context-dependent algorithms respectively.

Let us consider the definition of linguistic variable.

DEFINITION 1

A linguistic variable is defined as a 5 – tuple:

$$\Lambda = (H, T(H), U, G, M)$$

where: H ist the name of the variable, T – is the set of terms, names of values of H, U – universe of discourse, G – syntactic rules generating the elements of T, M – semantic rules generating the meaning m for each element $t \in T$.

From the formal definition of linguistic variable given above, in real applications the important question arises how to represent the syntactic and semantic rules which describe the set of terms T. When structural relationship between the terms and variable is not sophisticated, the simplest solution to the problem is to apply fuzzy naming relations describing the elements t in the most straightforward form of syntactic and semantic rules. Moreover, in such a case there exist learning procedures assuring optimal efficiency of the relations in the sense of minimum description error of. [6]. In other cases generative grammars are frequently applied, hence the main drawback of this approach is the hypothesis-and-test structure resulting in the top-down "active" algorithms. Another possibility is to use fuzzy automata (FA) which could make the "passive" network realizations more suitable for bottom-up analysis of acoustic cues. Therefore, taking into account the bottom-up organization of the recognition procedures in IWR, fuzzy naming relations and FA network were chosen for the description of the linguistic variables

H_s and H_a. The main advantage of such a representation of syntactic and semantic rules lies in the simplicity of the recognition algorithm.

4.1. *Context-independent algorithm*

The following assumptions are made:

1. The name of the variable

$$H_s = \{10 \text{ ms speech frame}\} \tag{2}$$

2. The set of composite terms consists of the names of 5 broad categories corresponding to the stationary parts of 5 chosen broad acoustic articulatory classes and the complementing term '**n**':

$$T_s = \{\text{'}\mathbf{r}\text{'}, \text{'}\mathbf{o}\text{'}, \text{'}\mathbf{u}\text{'}, \text{'}\mathbf{s}\text{'}, \text{'}\mathbf{v}\text{'}, \text{'}\mathbf{n}\text{'}\} \tag{3}$$

3. The universal set U consists of measured and computed parameters

$$U_s = \{U_1 \dots U_7\} \tag{4}$$

where: $U_1 = A\phi_i$; $U_2 = LP_i$; $U_3 = HP_i$;
$U_4 = A\phi_{\max} - A\phi_i = DA\phi$,
$U_5 = LP_{\max} - LP_i = DLP$,
$U_6 = HP_i - A\phi_i = DHA$,
$U_7 = HP_i - LP_i = DHL$, $i = 1 \dots \mathrm{N}$

N – number of frames, max – denotes maximum values of the corresponding parameter for each word.

The relations defining 5 classes of composite terms and the complementing term are expressed as follows:

DEFINITION 2

The term '**r**' describing the acoustic characteristics of the class RS

$$\mathbf{r} = l(DAO_n) \circ [l(DLP_n) \vee l(DHL_n)]. \tag{5}$$

DEFINITION 3

The term '**s**' describing the acoustic characteristics of the class *SS*

$$\mathbf{s} = h(DHL_n) \wedge h(DHA_n). \tag{6}$$

DEFINITION 4

The term '**u**' describing the acoustic characteristics of the class *UC*

$$\mathbf{u} = l(AO_n) \wedge l(LP_n) \wedge l(HP_n). \tag{7}$$

DEFINITION 5

The term '**o**' describing the acoustic characteristics of the quasi-stationary part of the class OC

$$\mathbf{o} = \bar{\mathbf{s}} \wedge \bar{\mathbf{u}} \wedge [m(DAO_n) \odot m(DLP_n)] \tag{8}$$

where $\bar{\mathbf{s}}$ and $\bar{\mathbf{u}}$ have the meaning no **s** and no **u** classes.

DEFINITION 6

The term '**v**' describing the acoustic characteristics of the quasi-stationary part of the class VF

$$\mathbf{v} = m(DLP_n) \wedge h(HP_n) \wedge 0. \tag{9}$$

DEFINITION 7

The term '**n**' complementing in the fuzzy sense the description of the variable H_s

$$\mathbf{n} = \overline{\mathbf{r} \vee \mathbf{s} \vee \mathbf{u} \vee \mathbf{o} \vee \mathbf{v}} \tag{10}$$

where l( ), m( ), h( ) – "low", "medium", "high" functions defined over the measured and calculated parameters, $\odot$ – bounded product, [9], $\wedge$ – intersection.

Equations (4)–(10) are used to compute the numerical values of the possibility that the actually analysed speech frame belongs to the classes terms from the set T_s, representing different manners of articulation. It was assumed, that each frame receives the label of the term with the highest possibility

$$h_i \in H_s \quad h_i \text{ receives the label } \mathrm{k} \Leftrightarrow \mathrm{poss}_k > \mathrm{poss}_l \tag{11}$$

where: $k, l \in T_s$,

The membership functions which are used in Eqs. (4)–(10) were estimated heuristically on the set of 120 tokens and then verified on the whole material from the learning group (Fig. 3).

4.2. Context-dependent algorithm

For the context-dependent algorithm, the following assumptions are made:

1. The name of the variable is,

$$H_a = \{\text{"acoustic articulatory segment"}\},$$

2. The set of composite terms which are the labels of 9 broad classes from Fig. 2,

$$T_a = \{'RS', 'RC', 'SS', 'VF', 'OC', 'AR', 'UC', 'PL', 'PS', 'NN'\}$$

3. The universal set U_a

$$U_a = \{'\mathbf{r}', '\mathbf{o}', '\mathbf{u}', '\mathbf{s}', '\mathbf{v}', '\mathbf{n}', U_1\}$$

4. Syntactic rules have the form of the FA network where the input alphabet is U_a and output alphabet is T_a.

5. Semantic rules are the output description functions which assign to each

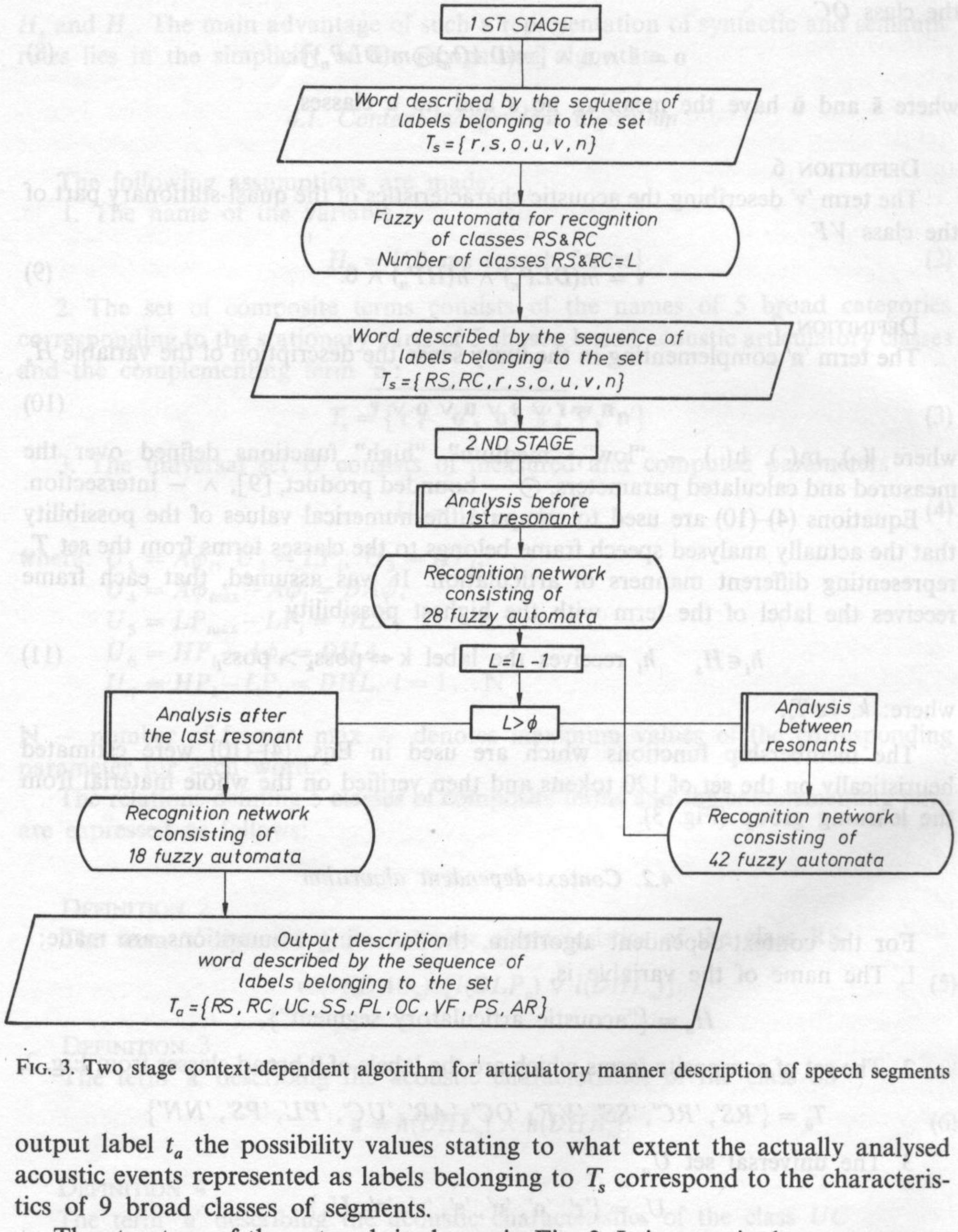


FIG. 3. Two stage context-dependent algorithm for articulatory manner description of speech segments

output label t_a the possibility values stating to what extent the actually analysed acoustic events represented as labels belonging to T_s correspond to the characteristics of 9 broad classes of segments.

The structure of the network is shown in Fig. 3. It is organised as a two-stage network where the whole set of input labels, the result of the previous algorithm, is analysed twice. At the first stage the fuzzy automaton recognizes the classes *RS* and *RC*. Both classes of segments are treated as the strong, "anchor" points of the analysis (they were detected at 1% error) and around them a more careful analysis of the previously labeled segments is performed to identify classes more complex to detect. At the second stage the algorithm has three separate parts of word analysis viz., before the first resonant in the word, between two consecutive resonants and after the last resonant. The resulting description consists in formulating sequence of labels belonging to the set T_a, with corresponding possibilities of the existence of a class in a given part of the word. At each position of the output string, fuzzy description one or more labels (up to three in the actual network). An example of such a representation of the word /drukarka/ (printer) uttered by a male voice is shown in Fig. 4.

EXAMPLE OF ARTICULATORY (MANNER) DESCRIPTION OF THE WORD

DRUKARKA

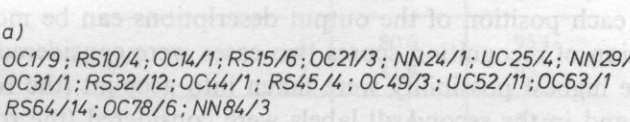
a)

OC1/9 ; RS10/4 ; OC14/1 ; RS15/6 ; OC21/3 ; NN24/1 ; UC25/4 ; NN29/2
OC31/1 ; RS32/12 ; OC44/1 ; RS45/4 ; OC49/3 ; UC52/11 ; OC63/1
RS64/14 ; OC78/6 ; NN84/3

b)

$$\begin{pmatrix} OC_{0.9} & AR_{0.88} & RS_1 & UC_1 & RS_1 & X_{0.75} & UC_1 & RS_1 \\ X_{0.1} & X_{0.12} & & & & AR_{0.25} & & \end{pmatrix}$$

c)



d)

	Poss (O_i) $i=1 \ldots 8$
O_1 = OC AR RS UC RS UC RS	0.75
O_2 = OC AR RS UC RS AR UC RS	0.25
O_3 = OC RS UC RS UC RS	0.12
O_4 = OC RS UC RS AR UC RS	0.12
O_5 = AR RS UC RS UC RS	0.1
O_6 = AR RS UC RS AR UC RS	0.1
O_7 = RS UC RS UC RS	0.1
O_8 = RS UC RS AR UC RS	0.1

FIG. 4. Example of articulatory manner description of the word /drukarka/ (printer) a) after context-independent analysis, b) after context-dependent analysis, c) graph representation of possible output descriptions, d) chart of output sequences with corresponding degrees of possibility

There are several important advantages of such an approach. The analysis of acoustic events which correspond to the chosen classes of segments is carried independently of the word which is currently analysed, but it depends on their phonetic context. Therefore, the rules of description and, in consequence, the recognition automata are vocabulary-independent. Another important feature of the algorithm is that the same automaton can be used in different contexts so it reduces the amount of memory required to save the state transition tables of automata. The total number of automata in the network is 52 with 500 states and 1100 paths between them. Finally, the structure of the network is designed in such a way that one can add a new automata in the process of learning when a certain sequence of segments cannot be recognized by the existing network. The actual network was elaborated during the learning process based on a 60 — word vocabulary spoken by 6 persons, containing about 1500 segments.

5. Experimental results

The algorithm was tested on 3120 segments from the 60 word vocabulary spoken by 12 persons (learning + test groups). To verify the efficiency of the algorithm, the output descriptions were compared with the reference descriptions of words made from phonetic transcriptions in accordance with the classification scheme shown in Fig. 2. Because on each position of the output descriptions can be more than one label in the evaluation of recognition scores two cases were considered. In the first only labels with the highest possibility were taken into account (the rows denoted with (1) in Tab. 1) and in the second all labels were considered (the rows (1)). The results are presented in terms of two coefficients taking into account three types of errors committed by the algorithm viz: omissions, insertions and substitutions. It seems very inconvenient to give the number of percentage of each error type for each class of segments used in the system. It is interesting for the system designer just to be apply a good basis for comparison of the quality of such a process. For this purpose two coefficients were proposed, which could play a role of effective parameters describing the process of phonetic segmentation and labeling in speech recognition applications.

The coefficients are defined as follows:

1. The correctness of recognition (p_r) of a chosen class of segments gives the relative number of segments from the reference description which was correctly put in the output description by a given algorithm.

2. The correctness of description (p_0) of a chosen class of segments gives the relative number of segments which were correctly put in the output description by a given algorithm.

The interpretation of either coefficient is the following. The correctness of recognition (p_r) says how difficult the class is for recognition by a given algorithm while the correctness of the description (p_0) shows to what degree one can be sure that the labels representing a given class in the output description belongs in reality to this class of segments.

The results of recognition obtained for 3120 segments (60 words and 12 speakers) are shown in Fig. 5. The average recognition scores were 97% for correctness of recognition and 96% for correctness of description for all the analyzed material. Errors were uniformly spread between resonants and obstruents. For resonants 1 to 5% of segments were badly described, for obstruents this value reached 20 to 25%. Omissions which were the main source of errors were caused by the segments of

Table 1. Results of recognition of broad articulatory classes

		P_r%	P_0%
RS	1	98.3	97.9
	2	99.1	99.3
RC	1	88.8	94.1
	2	99.3	100.0
AR	1	41.7	93.5
	2	56.2	97.5
OC	1	79.4	78.6
	2	88.2	90.0
VF	1	80.6	95.1
	2	81.9	95.2
SS	1	92.6	92.6
	2	95.3	95.2
PS	1	81.3	82.6
	2	89.2	94.3
PL	1	69.1	92.2
UC	2	70.8	95.8

where: P_r – *correctness of recognition* – percentage of input segments correctly recognized,
p_o – *correctness of description* – percentage of segments correctly described in output description of word,
1 – first candidate description,
2 – all descriptions.

classes AR and PL (83% of all omissions). The results are considered to be quite good, given the very simple parametric representation applied in the speech signal.

6. Concluding remarks

Very simple methodological assumptions have led us to the theoretically simple and computationally efficient method of identification of broad classes of segments in IWR. Direct application of the notion of linguistic variable and the process of its description in the bottom-up strategy of word recognition give good results also on the lexical level.

The recognition of the 30–word set was done on 360 utterances spoken by 12

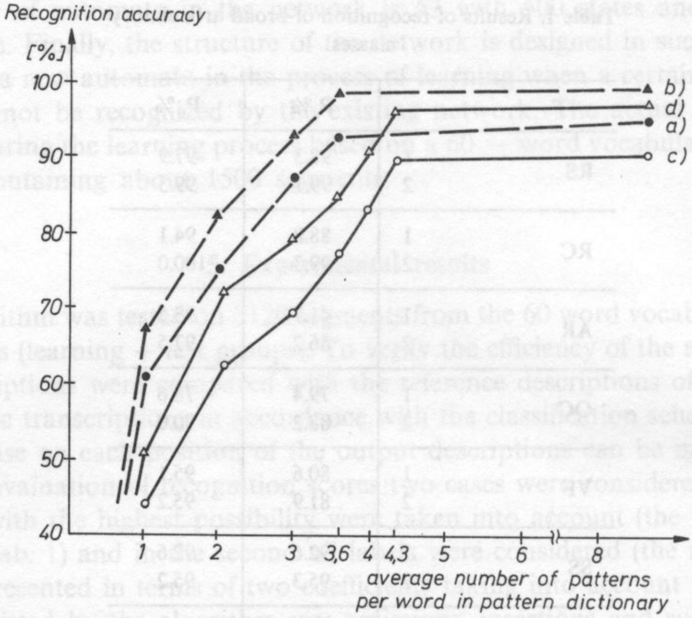


FIG. 5. Results of recognition for 30 words' set a recognition accuracy (R.A.) for the learning group-deterministic approach, b) R.A. for the learning group-fuzzy approach, c) R.A. for test group-deterministic approach, d) R.A. for test group-fuzzy approach

speakers of which 6 formed the training group and 6 speakers were new to the system. When only the highest possibilities were taken into account in the output description case (1), individual recognition scores were between 83–100%; however, when all the descriptions were considered, case (2), they reached 90–100%. Average recognition accuracies were from 91–95% respectively. The 5% increase of the word recognition score is consistent with the results given by De Mori [6]; however, the increase of the recognition rates of certain segment classes is even higher up to 10% (in individual cases to 25%).

To maintain the straightforward bottom-up system architecture, we avoided, in more complex cases, the very useful concept of knowledge source. The subjective and objective knowledge is contained in the FA network which is designed during the learning process. The explicit application of knowledge source will be needed when other levels of speech signal description are incorporated in the system.

An important feature of the method is that the segmentation problem was overcome in such a way that the algorithm estimates the possibility of the presence of a class in a word without explicit description of segment boundaries.

References

[1] W. BEZDEL, *Some problems in man-machine communication using voice*, Int. Journ. Man–Machine Stud. 2, 257–269 (1970).

[2] M. R. Sambur, L. R. Rabiner, *A speaker-independent digit recognition system*, Bell Syst. Techn. Journn. **54**, 1, (1975).

[3] C. J. Weinstein, S. S. Mc Candless, L. F. Mondshein, Y. W. Zue, *A system for acoustic-phonetic analysis of continuous speech*, IEEE ASSP **23**, 1, 54–67 (1975).

[4] S. Makino, K. Kido, *Recognition of phonemes using time-spectrum pattern*, Speech Communication **5**, 255–287 (1986).

[5] T. G. von Keller, *An on-line recognition system for spoken digits*, JASA 49, 1288–1296 (1971).

[6] R. De Mori, *Computer models of speech using fuzzy algorithms*, Plenum Press, New York, London 1983.

[7] R. Gubrynowicz, *Application de la théorie des sous-ensembles flous à l'analyse et à la reconnaissance automatique de la parole*, Note Technique CNET-LANNION NT/LAA/TSS/157 1983.

[8] W. W. Wieźlak, *Application of approximate articulatory description of speech signal to the recognition of limited set of isolated words* (in Polish) Dr. Eng. Thesis, Institute of Fundamental Technological Research, Polish Academy of Sciences, Warsaw 1987.

Received August 10, 1989