

ISSN 0137-5075



P O L I S H   A C A D E M Y   O F   S C I E N C E S  
I N S T I T U T E   O F   F U N D A M E N T A L   T E C H N O L O G I C A L   R E S E A R C H  
C O M M I T T E E   O N   A C O U S T I C S   •   P O L I S H   A C O U S T I C A L   S O C I E T Y

# ARCHIVES of ACOUSTICS

QUARTERLY

Vol. 38, No. 1, 2013

WARSZAWA



# ARCHIVES of ACOUSTICS

QUARTERLY, Vol. 38, No. 1, 2013

## In Memoriam

- P. Ranachowski, F. Rejmund, J. Etienne *A look at the life and activity of a prominent Polish acoustician of the 20th century Professor Ignacy Malecki (1912–2004) on the occasion of the centenary anniversary of his birth* 3

## Research Papers

- S. Merchel, M. E. Altinsoy, *Music-induced vibrations in a concert hall and a church* ..... 13
- R. Barański, G. Wszolek, *Educational implementation of a sound level meter in the LabVIEW environment*.... 19
- F. Zotter, M. Frank, *Efficient phantom source widening* ..... 27
- P. Song, L. Zhao, Y. Bao, *Spectral mapping using kernel principal components regression for voice conversion* . 39
- M. Meissner, *Evaluation of decay times from noisy room responses with pure-tone excitation* ..... 47
- S.-N. Yao, L. J. Chen, *HRTF adjustments with audio quality assessments* ..... 55
- M. Karpiński, *Acoustic features of filled pauses in Polish task-oriented dialogues*..... 63
- E. B. Skrodzka, B. B. J. Linde, A. Krupa, *Modal parameters of two violins with different varnish layers and subjective evaluation of their sound quality* ..... 75
- P. Kiełczyński, *Power amplification and selectivity in the cochlear amplifier*..... 83
- L. Dąbek, T. Hornowski, A. Józefczak, A. Skumiel, *Ultrasonic properties of magnetic nanoparticles with an additional biocompatible dextrane layer* ..... 93
- E. Kozaczka, G. Grelowska, S. Kozaczka, W. Szymczak, *Detection of objects buried in the sea bottom with the use of parametric echosounder*..... 99
- D. Zhao, Z. Huang, S. Su, T. Li, *Matched-field source localization with a mobile short horizontal linear array in offshore shallow water*..... 105

## Technical Notes

- N. Garg, A. Kumar, S. Maji, *Practical concerns associated with single-number ratings in measuring sound transmission loss properties of partition panels*..... 115
- T. Gudra, D. Banasiak, K. Herman, K. Opieliński, *Ultrasonic method for monitoring environmental risks associated with precipitation*..... 125

## Chronicle

- Winter School organized by the Upper Silesian Division of the Polish Acoustical Society ..... 137

## Calendar of Events

- Conferences ..... 153

## Editorial Board

**Editor-in-Chief:** NOWICKI Andrzej (Institute of Fundamental Technological Research PAN, Warszawa)

**Deputy Editor-in-Chief:** MEISSNER Mirosław (Institute of Fundamental Technological Research PAN, Warszawa)

### Associate Editors

BATKO Wojciech (AGH University of Science and Technology, Kraków)

DANICKI Eugeniusz (Institute of Fundamental Technological Research PAN, Warszawa)

GRELEWSKA Grażyna (Polish Naval Academy, Gdynia)

GUBRYNOWICZ Ryszard (Polish-Japanese Institute of Information Technology, Warszawa)

GUDRA Tadeusz (Wrocław University of Technology, Wrocław)

KUJAWSKA Tamara (Institute of Fundamental Technological Research PAN, Warszawa)

KULKA Zbigniew (Warsaw University of Technology, Warszawa)

MAKAREWICZ Rufin (Adam Mickiewicz University, Poznań)

MIŚKIEWICZ Andrzej (The Fryderyk Chopin University of Music, Warszawa)

PREIS Anna (Adam Mickiewicz University, Poznań)

RDZANEK Wojciech P. (University of Rzeszów, Rzeszów)

SNAKOWSKA Anna (AGH University of Science and Technology, Kraków)

**Secretary:** ŻYCHOWICZ–POKULNIEWICZ Joanna

## Advisory Editorial Board

**Chairman:** DOBRUCKI Andrzej (Wrocław University of Technology, Wrocław)

BJØRNØ Leif (Technical University of Denmark, Lyngby)

BLAUERT Jens (Ruhr University, Bochum)

BRADLEY David (The Pennsylvania State University, Pennsylvania)

CROCKER Malcomml J. (Auburn University)

DEMENKO Grażyna (Adam Mickiewicz University, Poznań)

ELLIOTT Stephen J. (University of Southampton, Southampton)

ENGEL Zbigniew (Central Institute for Labour Protection, Warszawa, retired prof. AGH)

HANSEN Collin (University of Adelaide, Adelaide)

HESS Wolfgang (University of Bonn, Bonn)

KOZACZKA Eugeniusz (Naval University of Gdynia, Gdynia)

LEIGHTON Tim G. (University of Southampton, Southampton)

LEWIN Peter A. (Drexel University, Philadelphia)

PUSTELNY Tadeusz (Silesian University of Technology, Gliwice)

RAKOWSKI Andrzej (The Fryderyk Chopin University of Music, Warszawa)

SUNDBERG Johan (Royal Institute of Technology, Stockholm)

ŚLIWIŃSKI Antoni (University of Gdańsk, Gdańsk)

TITTMANN Bernhard R. (The Pennsylvania State University, Pennsylvania)

TORTOLI Piero (University of Florence, Florence)

VORLÄNDER Michael (Institute of Technical Acoustics, RWTH Aachen)

WESOŁOWSKI Zbigniew (Polish Academy of Science, Warszawa)

ZWISŁOCKI Jozef J. (Syracuse University, Syracuse)

### Editorial Board Office:

Pawińskiego 5B, 02-106 Warszawa (Poland)  
fax (48-22) 826 98 15, phone (48-22) 826 12 81 ext. 206  
e-mail: akustyka@ippt.pan.pl <http://acoustics.ippt.pan.pl>

PRINTED IN POLAND

Indexed and abstracted in Science Citation Index Expanded (also known as SciSearch®)  
and Journal Citation Reports/Science Edition

Abstracted in Acoustics Abstracts, in Applied Mechanics Reviews  
and indexed in Journal of the Acoustical Society of America

Recognised by the International Institute of Acoustics and Vibration, I IAV

Edition co-sponsored by Ministry of Science and Higher Education

## In Memoriam

# A Look at the Life and Activity of a Prominent Polish Acoustician of the 20th Century Professor Ignacy MALECKI (1912–2004) on the Occasion of the Centenary Anniversary of His Birth

Przemysław RANACHOWSKI, Feliks REJMUND, Jerzy ETIENNE

*Institute of Fundamental Technological Research, Polish Academy of Sciences*  
Pawińskiego 5B, 02–106 Warszawa, Poland; e-mail: pranach@ippt.pan.pl

(received October 24, 2012; accepted February 5, 2013)

Ignacy Malecki was born on November 18, 1912, at the family estate of Pokiewnia near Vilnius. After Poland regained its independence in 1918, the town was at the border of the Second Republic of Poland and Central Lithuania. The church and village were located in the Polish area, while the family estate on the Lithuanian side. It took some time before those parts could be reunited.

Father Jan was a wealthy landowner and a bank clerk (President of the Tax Chamber) in Vilnius. He was also Treasury Minister of Central Lithuania. He came from an old noble family whose genealogy papers dated from 1627. Mother Emilia came from the artistic family of Witkiewicz. Stanisław Witkiewicz – art theorist, painter, and architect, promoter of the Zakopane highland style – was her uncle.

In the years 1924–1930 Ignacy Malecki studied at the prestigious Joachim Lelewel State Gymnasium in Vilnius, where he lived with his family. After completing his education he took the final exam in mathematics and natural sciences with very good results. Then he passed difficult entrance exams to Warsaw University of Technology. Despite the large competition – seven candidates for a place – he became a student of the university, however, not at the desired Architecture but at the Division of Weak Electrical Currents. After two years of studies, in October 1932, he received the first certificate (Fig. 1). The first diploma exam consisted of 13 subject exams, completions of numerous exercises and laboratory tasks, as well as two factory practices – mechanical and electrical. These practices took place at railway workshops and the power plant in Vilnius. In the summer of 1934, he had foreign practices in Paris and Marseille. After having passed



Fig. 1. Certificate of the first diploma exam of Malecki by Electric Department of Warsaw University of Technology, 1932.

with very good results the final exam he graduated from the Telecommunication Division of the Electrical Department of Warsaw University of Technology. He received the degree of electrical engineer (Fig. 2).



Fig. 2. Certificate of degree of Electrical Engineer obtained at the Faculty of Warsaw University of Technology, 1935.

His diploma thesis was prepared under the supervision of prof. Janusz Groszkowski, known worldwide as a radio communication and electronic specialist. The work concerned the radiation produced by rotating electromagnetic fields. After finishing the polytechnic studies in 1935, Malecki took a half year long training in AVA Radio Company. This firm designed and produced radio equipment for the army, as well as the encoding and decryption apparatus.

The prominent personality of prof. Groszkowski had a great influence on the choice of further scientific way of the excellently promising engineer. He persuaded Malecki to take part in the competition for the financial support of the National Culture Fund. A grant for further studies in Germany, which he then received, was an important moment in the life of Ignacy Malecki. It was the moment when he became devoted to acoustics for the rest of his professional life. At Heinrich Hertz Institute in Berlin, a young engineer gained the knowledge and experience in the field of ultrasound. His supervisor, prof. Edwin Mayer educated the entire generation of German acousticians. Malecki conducted studies of ultrasound vibrations of frequency of 200 kHz, high at that time. These vibrations were generated using a tourmaline crystal. Unfortunately, the stay in Berlin (January–August, 1936) coincided with the violent change of social attitudes

of the German society. Due to a growing nationalistic ideology, the name of Heinrich Hertz Institute was changed into Institut für Schwingungsforschung (Vibration Research Institute). The young Polish engineer was made to leave the Third Reich for fear of his own safety. Fortunately enough, he was offered to complete his foreign training at the laboratory of Philips firm in Eindhoven.

After returning to Warsaw, on the 1st of October 1936, he was appointed for his first responsible position. The Executive Director of the Polish Radio Stock Company employed him at the Study Office of the Polish Radio in Warsaw. The task of the engineer was to create an acoustic laboratory. From the 1st of June 1938, he became the head of the Technical Department of the company. The staff of the Department was engaged in measurements and improvement of conditions of the acoustic chambers in the building of a studio at 25, Zielna Street. In addition, measurements of the acoustic isolation, tests of microphones were conducted, as well as theoretical and experimental work in the field of room acoustics. New radio studios were organized in connection with launching, in March 1937, of another radio channel – Warsaw II.

Under the direction of engineer Malecki a full technical and acoustic design of recording chambers in the premises of the Polish Radio was prepared. The com-

plex of buildings, according to prof. Bohdan Pniewski project, was to be built at the corner of Batorego and Puławska Street. The works began in March 1939. However, before the war only foundation trenches were made.

While working for the Polish Radio, the engineer carried out research in the field of architectural acoustics at the Research Faculty of Construction of the Architecture Department at Warsaw University of Technology. The dean of the faculty was the famous constructor prof. Stefan Bryła. The study, started already in 1936 under the guidance of prof. Bryła, would constitute for many years the essential scientific and practical interests of Ignacy Malecki.

Apart from the scientific capacity, the young engineer revealed personal culture and a diplomatic talent. In the framework of the work for the Polish Radio he was appointed the representative of Poland in the *Union Internationale de Radiodiffusion* (International Union of the Radiophony). His efforts in this organization were concentrated on the allocation of the most favourable frequency bands for the rapidly expanding Polish radio broadcasting. As he recalled later, most disputes were with the representatives of the Third Reich.

Several years of commitment to the work on development of the Polish Radio had an unusual final when the war started. The engineer Malecki headed the technical work of the studio at 25, Zielna Street. After the Germans seized the long-wave radio station in Łazy near Raszyn on September 7, 1939, the regional station Warsaw II took over its responsibilities. It had at its disposal a 10 kW transmitter at Mokotów Fort. It was connected by an underground cable with the main building in Zielna Street. The transmitter was assisted by another, 1.5 kW power, one, located at Warsaw University of Technology. From September 8, using just these two transmitters, President of Warsaw Stefan Starzyński regularly communicated with the citizens of Poland. At noon on September 23, the Warsaw power plant in Powiśle district was bombed down. Since that time, the broadcast station went silent.

During the years of German occupation Malecki led an underground teaching and scientific work at the Department of Architecture, Warsaw University of Technology, under the supervision of prof. Bryła. This activity was carried out behind the facade of the research facility, and later a vocational school. An underground educational work was involved with a significant risk. Prof. Bryła, for participating in the conspiracy, was executed on December 3, 1943. In the first period – from 1940 – Eng. Malecki worked at the Research Institution of Technical Physics. Under the pretext of an official technical expertise, the works for the purpose of the conspiracy were carried out, including teaching activities. With the opening by the Germans in 1942 of the Public Higher Technical School in the buildings

of University of Technology he got employed in it. In the years 1942–1943 he was delivering underground lectures on room acoustics for the 3rd year students, Faculty of Architecture. Illegal activities, including teaching, were facilitated by Professor Gütinger who was at that time the German rector of the school. He pretended not to see the underground activities of the Polish employees.

In January 1942, Malecki defended his PhD thesis entitled *Physics of acoustic porous materials*. Its subject was the issue of sound absorption and especially the mechanisms of penetration of acoustic waves through materials of capillary structure. The defense was held up against the team of rector Kazimierz Drewnowski, with the participation of prof. Stefan Bryła as a scientific advisor and prof. Roman Trechciński (outstanding electrotechnician, teletechnician, and inventor in the field of telephony). The thesis originated from the research carried out at the Institute of Heinrich Hertz and was completed before the war began in 1939.

In May 1943, Malecki, PhD, defended his habilitation work (DSc thesis) entitled *Distribution of acoustic field in the closed space*. Professors Stefan Bryła, Roman Trechciński, and Mieczysław Wolfke (world-wide famous physicist, precursor of television and holography) were the reviewers of this work. In the work, the author applied an original statistical method to calculate acoustic non-stationary states in the closed acoustic volume. After the successful defense Malecki, PhD, obtained the right to deliver lectures at universities (*venia legendi*) as a docent, equivalent to associated professor. The received title was equivalent to the present-day post-doctoral degree (habilitation). The dissertation was extended and published 6 years later at Gdańsk University of Technology, under a modified title: *The mechanism of propagation of sound waves in the halls* (150 pages). Official recognition of the habilitation by the Senate of Warsaw University of Technology took place on 30 January 1946.

During the Warsaw Uprising Ignacy Malecki found the living extremely difficult. He needed food to support his young wife and a new born son Wojtek. Unfortunately his wife was shot down together with other residents of the apartment building by the soldiers of RONA – Russian National Liberation Army, brigade of Bronisław Kamiński. An unusual twist of fate saved the baby, lying under the body of his mother. Than Malecki and the baby were taken to the temporary camp in Pruszków, from where he escaped and hid from the Germans in the property of the writer Jarosław Iwaszkiewicz and his wife Anna in Stawisko, nearby Podkowa Leśna.

In the spring of 1945, he declared to be in disposition of the Provisional Government of the Republic of Poland. The new authority appointed him, as its representative, in the Lower Silesia. Soon he was di-

rected to Gdańsk, destroyed during the war, with the aim of rebuilding the electricity system of the region. He succeeded in this task very well and remained at that position till 1948 as the executive director of the Energy Division of the Pomeranian Union. In October 1945, he actively participated in the establishment of the branch of the Association of Polish Electrical Engineers (SEP). He was elected Chairman of the Board of the Baltic Sea Coast SEP Division for the 1945–1946 period. Apart from working for the electric power industry, he participated in reactivation of Gdańsk University of Technology. Since 1945 he gave lectures on electrical engineering to students of general faculties. A year later he was appointed Professor of Gdańsk University of Technology and organized the Department of Applied Electrical Engineering and Acoustics, where he was head up till year 1951.

At that time he established a well equipped acoustic laboratory with tools for measurements of absolute acoustic pressure using a Rayleigh disc and an echo-free chamber. A little later, the Professor organized, in cooperation with the Navy, research on German torpedoes steered with phase sensible acoustic systems and on acoustic and magnetic mines. The results of these studies were developed in the first PhD promoted by the Professor. The thesis of Zbigniew Zubelewicz was entitled *Some acoustic and magnetic devices in the sea telemechanics* (1949). The author used until 1948 the pseudonym of Jan Góra taken during the conspiracy times.

In 1950, Professor ended his activity in the field of electric power industry and gradually moved with his family back to Warsaw. The need for reconstruction of many damaged public buildings made Professor engage in the process of their reconstruction with regard to their acoustic properties. He promoted the modeling technique of closed spaces, the theory of a heterogeneous distribution of the acoustic field in halls, and optimization of the reverberation effect. He introduced the parameter of “distinctness” as an evaluation criterion of the acoustic quality of halls. Based on previous research and publications, he developed a method for determining the optimum profile of meeting halls using the method of “spatial sources” and the theory of application of perforated sound absorbing elements. The above issues were presented in a monograph, entitled *Architectural Acoustics*, published in 1949 (ed. Building Research Institute). A number of concert halls and theatres projects under the direction of Professor were designed. The principles established by Malecki were followed in the design of the National Theatre, the Parliament Hall, and government buildings in Aleje Ujazdowskie (at that time the building of the Council of State). Further works concerned the methods of increasing the acoustic insulation of prefabricated building constructions. Rules were developed to take into account the acoustic conditions in urban design. More-

over, the measurement systems of recording the acoustic shock wave were improved too. These issues were widely covered in the book *Fighting Noise in Industrial Plants* (PWT 1954).

From 1948, Ignacy Malecki was a permanent member of the Technical Council of the Polish Radio. The demand for professionals in the field of electroacoustics grew rapidly, especially from the Polish Radio, film studios and the Ministry of Posts and Telegraphs (now Ministry of Communication). It led to the establishment by the Minister of Higher Education, on the 1st of October 1949, the Chair and Institute of Electroacoustics at the Faculty of Communication, Warsaw University of Technology. Film and broadcasting technologies were the first directions of students’ education. The books and papers by Malecki became the first handbooks used in the teaching process (*Radio and Film Acoustics*, 1950, PWT and *The Technique of Recording and Playing Back Sounds*, 1953, PWT).

In the initial period of its activity the Faculty used the equipment of the Polish Radio and Polish Film and the Main Institute of Technical Physics, which was established in 1949. Prof. Malecki was its employee from February 1950. The Institute occupied the rooms in the building of Physics of Warsaw University of Technology. This made convenient to conduct a joint research. The Faculty of Electroacoustics was managed by professor from February 1950 till 1969. Witold Straszewicz and Stefan Basiński were his students and then coworkers. After the reorganization, the Faculty was incorporated into the Institute of Radioelectronics, which was headed by prof. Stanisław Ryzko. In 1950, Malecki was appointed professor at Warsaw University of Technology in the field of electroacoustics, and on April 3, 1951, became an ordinary professor. He was bound with Warsaw University of Technology, as a researcher and lecturer, up to 1983 (1973–1983 part-time professor). One of his closest colleagues then was Jerzy Narkiewicz-Jodko, PhD.

In the years 1951–1952 Professor served as the dean of the Faculty of Communication. During this time he was also prorector for science at Warsaw University of Technology. In the period 1950–1951 he organized an acoustic laboratory at the Main Institute of Technical Physics. Professor had some colleagues and collaborators at the laboratory. Among them were students and next PhD students like Leszek Filipczyński, Jerzy Wehr, and Waclaw Koltoński.

Professor Malecki took part in the 1st Congress of Polish Science which was held in the period of 29 June – 2 July 1951. During the Congress the decision to create the Polish Academy of Sciences (Polska Akademia Nauk – PAN) was made. Since that time PAN took over previous academy corporations, existing till then for science, as well as art and literature associations. In the range of the technical sciences, the idea was to organize an institute of the Polish Academy of Sciences



Fig. 3. Prof. Malecki among the collaborators on the plane during a business trip to Moscow (meeting of acousticians); engineers J. Ranachowski and Z. Pawłowski are standing, Eng. J. Wehr and Prof. I. Malecki are sitting, January 1962.

on the basis of already existing facilities. The Institute was formed gradually, as a result of a closer cooperation, and a formal merge of scientific units which differed among themselves in the number of research workers and the scope of studies.

At the turn of 1951/52 the Main Institute of Technical Physics, after less than three years of operation, was divided between Warsaw University of Technology and Jagiellonian University, Cracow. The acoustic laboratory was transformed into the Department of Examination of Vibration of the Polish Academy of Sciences. This division, headed by prof. Malecki, led a close cooperation with the Division of Continuous Media Mechanics (head prof. Waław Olszak) and with the Division of Electronics (head prof. Janusz Groszkowski). In the field of fundamental research, the subject of study was complementary to a large extent. The Scientific Secretariat of PAN promoted the cooperation, as the basis for the founding of a new institute. Professor Witold Nowacki, Secretary of Department of Technical Sciences PAN, played the essential role in the creation and development of the Institute. He gathered a strong group of researchers which constituted the core of the Division of Continuous Media Mechanics. They worked on problems of isotropic and un-isotropic elastostatics, thermoelasticity, elastodynamics, non-linear elasticity, elastic composites, and elasto-optics. In June 1952, the Scientific Secretariat of PAN appointed a joint Scientific Council for the three mentioned above units. In addition, the Division of Metals in Cracow was established at

the Department of Technical Sciences PAN, with prof. Aleksander Krupkowski as its head. By resolution of the Scientific Secretariat of PAN, in December 1952, four of these units along with a smaller Division of Theoretical Electrotechnics (head prof. Paweł Szulkin) and a Laboratory of Astronautics (head prof. Kazimierz Zarankiewicz), were combined to form the Institute of Fundamental Technological Research PAN (Instytut Podstawowych Problemów Techniki – IPPT PAN). Professor Ignacy Malecki was the first director of the Institute, until 1961. For the first 5 years the institute was located in the Staszic Palace in Warsaw. Resolution of the Government Presidium in September 1953 officially confirmed the establishment of the institute. New coworkers of professor Malecki in the institute were Janusz Kacprowski, MSc, Wincenty Pawjewski, MSc, Jerzy Ranachowski, MSc, Stefan Czarniecki, MSc, and from 1962 Zdzisław Pawłowski, MSc. The Professor was again appointed director of IPPT PAN from 1973 to 1982, when he retired.

In addition, professor Malecki held a series of other positions at the institute. In the period 1961–1962, he was head of the Division of Vibrations Research. As a result of a reorganisation in 1969 the Division of Physical Acoustics was established. Prof. Malecki became its head until 1972. After the retirement, Professor was still employed part-time at the mentioned above division until 1995. In the following years his longtime collaborator prof. Feliks Rejmund was the head of the division. Professor took an active part in scientific studies and research projects of the Commit-



Fig. 4. The celebration of the 25th anniversary of the Institute of Fundamental Technological Research. Director of the institute, prof. Malecki has a speech. Next from the left: professors W. Gutkowski, J. Ranachowski, S. Kajfasz, 1977.

tee of Scientific Research. The last project in which he participated was the work *Corundum Ceramics of New Generation*. It was realized in cooperation with the Institute of Glass and Ceramics, in the years 2000–2002. The Professor was then 90 years old (!).

Prof. Malecki used to take an active part in the organization of cyclic conferences on the ceramic and composite materials, under the auspices of the Euro-

pean Materials Research Society (E-MRS) (Fig. 5). Prof. Jerzy Ranachowski, his friend and collaborator was their initiator. These conferences are still organised nowadays. For the rest of his life, Professor was a very friendly supervisor of postdoctoral and doctoral students, to whom he had always given great support and help. Shortly before his death, he delivered the review of the postdoctoral thesis by Jan Žera, PhD,



Fig. 5. The Conference *New Trends in Technologies and Materials Research* in Białowieża. From the left, professors: A. Rakowski, A. Śliwiński, I. Malecki, J. Ranachowski, A. Opilski, June 1998.

at the Council of the Faculty of Electronics and Information Technology of Warsaw University of Technology. The defense took place in 2004, after the death of the Professor.

Ignacy Malecki was the educator of two generations of the Polish acousticians. During the work at Warsaw University of Technology he supervised about 70 master theses in acoustics and electroacoustics. His most famous PhD students were: Leszek Filipczyński (1955) – an ordinary member of the PAN, director of the IPPT (1969–1974); Janusz Kacprowski (1957) – professor at the IPPT; Stefan Czarnecki (1959) – professor at the IPPT; Wacław Kołtoński (1959) – professor at the IPPT; Zenon Jagodziński (1960) – professor at Gdańsk University of Technology; Jerzy Wehr (1961) – professor at the IPPT; Andrzej Rakowski (1963) – an ordinary member of PAN, professor and rector of The Fryderyk Chopin University of Music in Warsaw (1981–1987); Witold Straszewicz (1965) – associate professor at Warsaw University of Technology, the recognized specialist in the field of the room acoustics, the designer of many concert halls, opera houses, and multifunctional objects; Ryszard Płowiec (1970) – professor at the IPPT, manager of the Division of Physical Acoustics (1996–1999). Among the 25 doctors promoted by Professor, 11 were employees of the IPPT. The first was Leszek Filipczyński, the last one was Przemysław Ranachowski (2001). Prof. Andrzej Nowicki, the longtime director of the IPPT, was also a student of prof. Malecki. He specialized in acoustics under Professor's care.

Apart from holding a number of positions at Warsaw University of Technology and at the IPPT, prof. Malecki carried out a wide activity at the Polish Academy of Sciences. In 1954, Professor became a Corresponding Member, in 1958 he was appointed an Ordinary Member of PAN. From 1962 to 1968 he was the Scientific Deputy Secretary of PAN, and since 1980, a member of the Presidium. In addition, he was an Ordinary Member of the Warsaw Scientific Society (Warszawskie Towarzystwo Naukowe – WTN, 1984) and a foreign member of New York Academy of Sciences.

Professor was also active in international organizations related to the evolution and development of the science policy. From 1963 to 1967 he was the Vice President of the International Council of Scientific Union (ICSU). In the years 1969–1973 he was the Director of the Department of Scientific Policy of the United Nations Educational, Scientific and Cultural Organization (UNESCO) in Paris. Apart from this, he was a member of the Executive Council of the International Council for Science Policy Studies (ICSPS) in the period 1976–1998. The diplomatic passport facilitated Professor's frequent travels abroad.

Prof. Malecki was one of the founders of the Polish scientology. As a founder member of the Scientific So-

ciety of Praxeology, he contributed to the creation of the Committee of Scientology PAN, of which he was the Chairman in 1963–1968 and 1973–1989. Then from 1990 he was its Honorary Chairman. The professor analysed the impact of fundamental research on the socio-economic development and the perspective planning methodology in science. He developed methods for evaluation of results of research projects. He also participated in the global development of scientology, meant as a federation of scientific disciplines. His important achievement was the study of the evolution of the scientology conception. In addition, since 1987 he presided the Editorial Council of the “Scientology Issues”. He was lucky enough to put into practice the solutions of the problems that were the subject of interest to him. In the frames of the Polish Academy of Sciences he took part in the creation of a system of perspective research planning. At the Council for the Technique he participated in the application of research results for the technical progress. As it was mentioned above, he also acted at the international organizations associated with carrying out scientific policies.

Professor belonged to the founders of the Polish Acoustical Society (Polskie Towarzystwo Akustyczne – PTA). Countrywide meetings of acousticians (Open Seminars on Acoustics – OSA) were held since 1954, initiated by prof. Marek Kwiek from Poznań University. Cooperation and information exchange between national centres was at that time quite weak. It resulted from various sorts of administrative obstacles and limited the possibility of publication of research results. The thematic range of these seminars was fairly narrow and largely due to the personal interest of the organizers. However, with time, the thematic range and number of participants steadily grew. The process was nevertheless gradual and only after seven years the breakthrough came. It happened during the VIII-th Open Seminar on Acoustics in Szczecin, on August 22, 1961. The formal creation of the PTA was done by a group of 38 founder-members, connected with scientific or professional acoustics. Within this group, apart from Professor, were among others prof. Marek Kwiek, assistant prof. Leszek Filipczyński, Antoni Śliwiński, PhD, and Alexander Opilski, MSc. Despite the tragic death of prof. Kwiek (on December 19, 1962), the General Founding Assembly of PTA was held on March 4, 1963, in Poznań, where the first General Assembly of Delegates was prepared. They represented larger national centers dealing with acoustic research. The Congress was held on July 18, 1963, also in Poznań, which is the registered office of the organization up till now. The first four divisions of PTA: Poznań, Warsaw, the Upper Silesian and Wrocław were established. Polish Acoustical Society was affiliated at IV-th Department of Technical Sciences of the Polish Academy of Sciences. Later other divisions in Gdańsk (1966), Rzeszów (1973) and Cracow (1980) were cre-

ated. Prof. Malecki was awarded by the General Assembly of Delegates PTA, the highest dignity of Honorary Member of the organization in 1979. Regardless of the active engagement in the Society, from 1963 to 1969, he served as the Chairman of the Committee on Acoustics PAN. From 1989 he became its Honorary Chairman. In the years 1982–1987 he was Chief Editor of the quarterly journal *Archives of Acoustics*. In 2002 Professor participated at 49th OSA conference, organized by Warsaw Division of the Polish Acoustical Society in Stare Jabłonki (Fig. 6). Then, as the Honorary Chairman, he actively worked at the Scientific Committee of the seminar. The last conference in which Professor Ignacy Malecki participated, a year before passing away, was 50th Open Seminar on Acoustics. It was held in Szczyrk, in September 2003. During 59th OSA conference, in September 2012, on the occasion of the 100th anniversary of his birth, an opening session was devoted to the Professor.

The activity of Professor was not restricted only to the national acoustic organizations. In the years 1966–1972 he chaired the worldwide organization, the International Commission for Acoustics (ICA). He was also a member of the European Acoustics Association (EAA) and Acoustical Society of America (ASA). He received the title of a Honorary Member of the Federation of Acoustical Societies of Europe (FASE, 1982), of which he was the Vice President in the period 1978–1982, as well as Acoustic Societies: Latin American (1968), Spanish (1972), and Indian (1980).

Beyond the scientific, organizational, professional, and teaching activities prof. Malecki managed to

find time and strength for social activities, both at home and abroad. In the years 1963–1968 he was a member of the Constant Committee of the Pugwash Conference – activity of scientists and Nobel laureates for disarmament and peace. He was a member of the General Council of the Main Technical Organization (Naczelna Organizacja Techniczna – NOT, 1949–1963). He served as the Chairman of the Scientific Council of the Central Institute for Labour Protection (Centralny Instytut Ochrony Pracy – CIOP, 1953–1969). He was the Chairman of the Technical Sciences Section for State Awards (1978–1982), then the Chairman of the Scientific and Technical Council of the Radio and Television (1982–1988). In the years 1991–2002 he served as a member of the Council of the International Biography Centre (IBC – Cambridge). In addition, in 1954–1958, he was a Councilor of the Warsaw-Centre district. He was also a board member of the circle of the Combatants and Former Political Prisoners Association at PAN (1989–2001).

In the personal life Professor had some very severe experiences. His only son Wojtek died tragically at the age of 22 on 15 August 1966, during climbing together with a friend (Marysia Konopacka) on Hrubý Wierch in the Slovak High Tatra Mountains. The cause of the accident could be a sudden weather breakdown. Wojtek was passionate about mountaineering and spent a lot of free time hiking. He was of extraordinary intelligence and inherited his father's ability in the field of science. Two months earlier (June 10) he passed his MSc exam at the Faculty of Physics of Warsaw University. Thirteen years later, on 23 December 1979, the



Fig. 6. Prof. Ignacy Malecki surrounded by the members of Warsaw Division of PTA during 49th OSA seminar in Stare Jabłonki, September 2002.

second wife of the Professor – Maria Tomczycka, died after a long and severe illness. The professor was very attached to her. She looked after Wojtek, whom she officially adopted, as her own son.

Difficult experiences and numerous responsibilities did not deprive Professor of cheerfulness. He was known for the ability of a very accurate assessment of the situation. Till the end of his life he retained the clarity of thinking and an unusual sense of humor. In a joking way, he willingly recalled anecdotes from his rich biography. Seeing a nice, but neglected lady Doctor, he jokingly referred to his fellows – “What this science can do of a woman”. Several times Professor traveled by plane to the United States. Being really busy, he usually did not have time to collect and arrange things carefully into his suitcase. One time, at the airport in New York, Professor’s luggage disappeared, as the only one of a large group of passengers. The airport service apologized and promised to deliver the suitcase to the hotel as soon as it would be found. Indeed, after a few hours the lost luggage was brought to him. How surprised Professor must have been seeing all his things neatly folded and arranged in the suitcase. Nothing was missing. The employees of the US Secret Service apparently felt obliged to make order after searching his luggage. At an already advanced age (90 years old) he met in Świętokrzyska Street a younger, retired scientist, who, however, barely moved with the aid of crutches. He recognized Professor with difficulty, stopped and asked “Are you maybe professor Malecki?”. “It’s really me.” answered friendly Professor. “And are you still alive?” at last exclaimed the scientist with an undisguised astonishment. Professor did not feel offended and humorously retold the situation to his friends.

Scientific activities of prof. Malecki covered a number of fields of acoustics, theory of vibration, and scientology. His publications concerned above all: room acoustics, architectural acoustics, protection against noise, electroacoustics, ultrasonic technique, quantum acoustics, and acoustic emission. His scientific output embraces 11 books and monographs and more than 230 articles (in part as co-authoring). Among the most important books in the chronological order are the following: *Architectural Acoustics* (ed. Building Research Institute, 1949), *Radio and Film Acoustics* (PWT, 1950), *Technique of Recording and Playing Back Sounds* (under the editorship, PWT, 1953), *Fighting Noise in Industrial Plants* (together with W. Kołtoński and W. Straszewicz, PWT, 1954), *Problems of Coordination of Scientific Research* (PWN, 1960), *Theoretical Fundamentals of Quantum Acoustics* (PWN, Library of Applied Mechanics, 1972), *Contemporary Acoustics and its Quantum Presentation* (PAN – Ossolineum, 1975), *Acoustic Emission – Sources, Methods, Applications* (under the editorship, ed. IPPT PAN, 1994). The last of the above monographs was the first in

the country and one of the few in the world, containing an attempt of a comprehensive presentation of the acoustic emission (AE), treated on the one hand as a physical phenomenon and on the other hand, as the research method for a wide range of applications. The most important publication of the Professor is the large monograph *Theory of Waves and Acoustic Systems* (PWN, 1964, 675 pages). It includes a theoretical approach to all problems of technical acoustics. Apart from the classical theory of acoustic fields and mechanical systems, the book contains original considerations of the Author. They concerned the ultrasonic technique, molecular acoustics, room acoustics, and hydro-location. This work was translated into English and published under the title *Physical Foundations of Technical Acoustics* (Pergamon Press – PWN, 1969, 743 pages) – (Fig. 7).



Fig. 7. View of the title page of the most important monograph by prof. I. Malecki, 1969.

A scientific activity has brought Professor a full recognition, both in the country and abroad. He received the honorable title of doctor *honoris causa* at Budapest University of Technology (8 September 1965), Academy of Mining and Metallurgy in Cracow (23 September 1982) and Gdańsk University of Technology (7 June 2002). The most important awards in the country included the individual Polish State Award of III degree for the experimental work on room acoustics (1952) and a II degree one for the group work in the field of ultrasonic waves propagation (1966, together with L. Filipczyński, Z. Pawłowski, and J. Wehr). In addition, Professor received the Golden Cross of Merit (1954), Crosses of the Order of the Restora-

tion of Poland: Officer's (1957), Commander's (1964), Commander's with Star (1973), as well as Golden Badges SEP (1959), NOT (1966), Warsaw University of Technology (1978), and the medals of various national institutions. The most important foreign awards are *French Order of Academic Palms* by the Prime Minister of France (1985), *Transenster* medal – for the scientific and technical achievements (University of Liège, Belgium, 1969), diplomas of International Societies of Non-destructive Testing – NDT (Tokyo, 1960 and Montreal, 1967), Honorary Diploma of the International Institute of Acoustics and Vibration – the IIAV (USA, Auburn University, 2003).

Professor Ignacy Malecki passed away suddenly on 12 June 2004 in Warsaw. He was buried in the Old Powązki Cemetery in Warsaw, place number 274, row 6 in the tomb of his second family Tomczyccy, together with his wife Maria Tomczycka, her relatives and the son Wojtek. The monument, made according to a special project, was chiseled out from granite after the tragic death of Wojtek in the Tatra Mountains. In 2005, Professor was commemorated with a special inscription on the board for the scientists of merit, placed at the second gate (of Saint Honorata) of Old Powązki Cemetery.

The authors would like to thank *Jerzy Narkiewicz-Jodko, PhD*, for collecting and arrangement of the materials for the biography and scientific activities of prof. Ignacy Malecki. At the same time, the authors point out that the source materials contain numerous inaccuracies and discrepancies. They concern primarily the biographical data, dates and titles of publications. Therefore, it was necessary to verify some of the data. Wherever a version was impossible to verify, the authors assumed the most probable one.

## References

1. *Who is Who in Poland*, Biographical Information (1984), [in Polish: *Kto jest kim w Polsce*, Informator biograficzny], Ed. I, p. 581, Wyd. Interpress, Warszawa.
2. RANACHOWSKI J. [Ed.] (1989), *Ignacy Malecki; The List of Publications of prof. I. Malecki, PhD; PhDs promoted under supervision prof. I. Malecki*, [in Polish: *Ignacy Malecki; Spis publikacji Prof. dra I. Maleckiego; Doktorzy promowani przez Prof. dra I. Maleckiego*], *Problemy i Metody Współczesnej Akustyki*, pp. 7–17, PWN, Warszawa – Poznań.
3. *Contribution of Prof. I. Malecki to the Polish School of Acoustics* (1993), *Archives of Acoustics*, **18**, 4, 487–505.
4. *The Golden Book of The Polish Science* (1999), [in Polish: *Złota Księga Nauki Polskiej*], p. 205, Wyd. HE-LION, Gliwice.
5. *Ignacy Malecki*, Professors and Docents of the Department of Electronics and Informational Techniques of Warsaw University of Technology, [in Polish: *Ignacy Malecki, Profesorowie i Docenci Wydziału Elektroniki i Technik Informatycznych Politechniki Warszawskiej 1951–2000*], pp. 128–129, Oficyna Wyd. PW, 2000.
6. *Who is Who in Poland*, Biographical Information (2001), [in Polish: *Kto jest kim w Polsce*, Informator biograficzny], Ed. IV, p. 563, Wyd. PAI.
7. *Contemporary Polish Scientists*, Biographical Dictionary, [in Polish: *Współcześni Uczni Polscy*, Słownik Biograficzny], Wyd. Ośrodek Przetwarzania Informacji, Vol. III, pp. 61–62, Warszawa 2001.
8. *A Tribute from prof. Michał Białko – promoter of prof. Ignacy Malecki's honoris causa doctorate*, [in Polish: *Laudacja Prof. Michała Białki – promotora doktoratu honoris causa Prof. Ignacego Maleckiego*, Pismo Politechniki Gdańskiej, No. 6, pp. 4–6, 2002.
9. *The Department Book, on the 50th jubilee*, [in Polish: *Księga Wydziału, wydana z okazji Jubileuszu 50-lecia, 1952–2002*, Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, Gdańsk 2002.
10. *50 years of the Institute of Fundamental Technological Research of the Polish Academy of Sciences 1952–2002*, [in Polish: *50 lat Instytutu Podstawowych Problemów Techniki Polskiej Akademii Nauk 1952–2002*, Wyd. IPPT PAN, Warszawa 2002.
11. *Professor Ignacy Malecki – In memoriam* (2004), *Archives of Acoustics*, **29**, 3, 393–394.
12. AUGUSTYŃSKA D., ENGEL Z. (2004), *Prof. Ignacy Malecki, PhD – Recollections*, [in Polish: *Prof. zw. dr hab. Ignacy Malecki – Wspomnienie*], *Bezpieczeństwo Pracy – nauka i praktyka*, 7/8, 10–11.
13. *In Memory of Professor Ignacy Malecki (1912–2004)* (2004), *Acta Acustica united with Acustica*, **90**, 5, 994–995.
14. *MALECKI Ignacy (1912–2004)*, Biographical Dictionary of Polish Technicians, [in Polish: *MALECKI Ignacy (1912–2004)*, Słownik Biograficzny Techników Polskich], Wyd. Federacja Stowarzyszeń Naukowo-Technicznych, Vol. 18, pp. 92–95, Warszawa 2007.
15. *Ignacy Malecki (1912–2004)*, a website, [in Polish: *Ignacy Malecki (1912–2004)*, strona internetowa Instytutu Radioelektroniki Politechniki Warszawskiej, Nauczyciele Akademicy Wydziału Elektroniki i Technik Informatycznych Politechniki Warszawskiej 1951–2011, biogramy: [http://www.elka.pw.edu.pl/pol/content/download/9539/63847/file/Malecki\\_Ignacy.pdf](http://www.elka.pw.edu.pl/pol/content/download/9539/63847/file/Malecki_Ignacy.pdf).
16. *Professor Ignacy Malecki*, Distinguished Polish Electricians, Directory of the Polish Society of Theoretical and Applied Electrotechnics, [in Polish: *Professor Ignacy Malecki, Zasłużeni Elektrycy Polscy*, Informator Polskiego Towarzystwa Elektrotechniki Teoretycznej i Stosowanej], No. 17, Poznań, December 2009.
17. ENGEL Z. (2012), *Recollections about Professor Ignacy Malecki at the centennial anniversary of his birth*, [in Polish: *Wspomnienie o Profesorze Ignacym Maleckim w setną rocznicę urodzin*], Proc. 59th Open Seminar on Acoustics, pp. 11–14, Poznań – Boszkowo.

## Research Papers

# Music-Induced Vibrations in a Concert Hall and a Church

Sebastian MERCHEL, Mehmet Ercan ALTINSOY

*Chair of Communication Acoustics, TU Dresden*  
01062 Dresden, Germany; e-mail: sebastian.merchel@tu-dresden.de

(received June 22, 2012; accepted September 25, 2012)

Sound and vibrations are often perceived via the auditory and tactile senses simultaneously, e.g., in a car or train. During a rock concert, the body vibrates with the rhythm of the music. Even in a concert hall or a church, sound can excite vibrations in the ground or seats. These vibrations might not be perceived separately because they integrate with the other sensory modalities into one multi-modal perception.

This paper discusses the relation between sound and vibration for frequencies up to 1 kHz in an opera house and a church. Therefore, the transfer function between sound pressure and acceleration was measured at different exemplary listening positions. A dodecahedron loudspeaker on the stage was used as a sound source. Accelerometers on the ground, seat and arm rest measured the resulting vibrations. It was found that vibrations were excited over a broad frequency range via airborne sound. The transfer function was measured using various sound pressure levels. Thereby, no dependence on level was found. The acceleration level at the seat corresponds approximately to the sound pressure level and is independent of the receiver position. Stronger differences were measured for vibrations on the ground.

**Keywords:** auditory-tactile music perception, room acoustics, transfer function measurements.

## 1. Introduction

This paper addresses sound and vibration during real-life situations. A typical example is the vibration perceived during an organ performance in a church. The following questions come to mind: Can vibrations be perceived even in a conventional classical concert hall during a concert performance? Do the acceleration levels exceed the perception threshold? Which dynamic range can be expected? Is there a linear relationship between the sound pressure level and the acceleration level for whole-body vibrations? How does the vibration intensity vary depending on receiver location? To answer these questions, comprehensive sound and vibration measurements were undertaken in two exemplary locations: a classical concert hall and a church.

## 2. Literature

The interest in multimodal reproduction has increased in the audio community over the recent years (RUMSEY, 2010). Especially, the field of auditory-tactile perception has gained importance in the context of virtual environments (ALTINSOY, 2012) or audio hardware (MERCHEL *et al.*, 2012). However, only

a few studies have been published that involve vibration measurements in a musical context. DAUB (2003) measured sound and vibrations in two different venues using musical instruments as sound generators. It was therefore difficult to separate the contribution of the sound source from the transfer characteristics of the room.

A study by SIMON *et al.* (2009) reported audio-induced vibrations in a car generated by a music sequence, which was played back via the automotive audio system. They found that high vibration levels between 50 Hz and 75 Hz were excited in the seat and floor. However, it is not clear whether this characteristic results from the spectral content of the source material. They also reported a relatively linear relationship between sound and vibration in this frequency range (a 4.5 dB increase in the bass level resulted in an approximately 4.5 dB higher acceleration level).

Abercrombie and BRAASCH (2010) measured structural impulse responses on different stage-floors, which were excited using a sledgehammer. They found that the acceleration level decreases and the propagation time increases with increasing distance from the source (maximum 4 m measured). Both were strongly dependent on the stage construction. Unfortunately, it is not

possible to separate air- from structure-borne vibrations in their measurements and to predict resulting vibrations in the auditorium.

For this reason, vibro-acoustical measurements in the auditorium of the Semperoper Dresden and the Lutherkirche Radebeul have been conducted.

### 3. Concert hall

This section discusses the relation between sound and vibration for frequencies up to 1 kHz in the Semperoper Dresden (Fig. 1). Therefore, the transfer function between sound pressure and acceleration was measured at different listening positions.



Fig. 1. Auditorium of the Semperoper Dresden with a lifted orchestra pit.

#### 3.1. Setup

A dodecahedron loudspeaker (Outline, Globe Source with Subwoofer) was used as the sound source. It was placed on the lifted orchestra pit 4 m from the edge of the stage and 1.5 m to the side of the middle axis. Six receiver positions were selected, three in the parquet (R1–R3), one in the loge (R4) and two in the balconies (R5 and R6). This is illustrated in Fig. 2. To measure room impulse responses, a measurement microphone at ear height (B&K, 4188), a spherical microphone array and a Kemar dummy head with blocked ear canals were used. The last-mentioned recordings can be used to reproduce the opera house virtually in the lab, e.g., using wave field synthesis or binaural reproduction, to conduct perceptual experiments. In this study, only the data from the measurement microphone are used as a reference.

The measurement could only be made at night in the empty concert hall. However, compared with a situation when two-thirds of the seats are filled, the reverberation time below 2 kHz was only prolonged by 0.5 s, which resulted in an approximately 3 dB increase in the sound pressure level averaged over the receiver

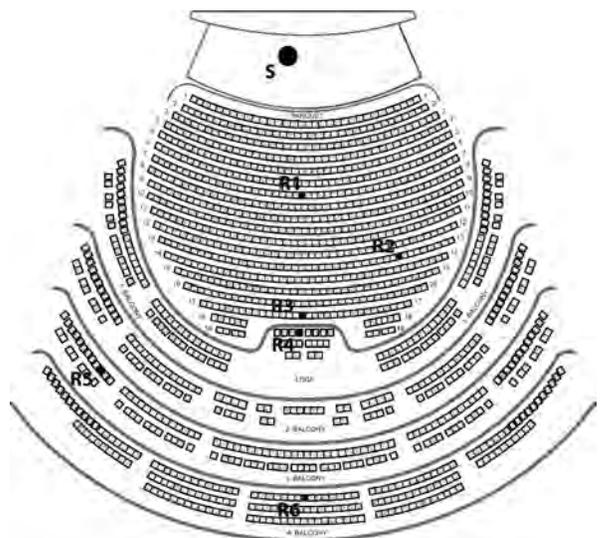


Fig. 2. Seating plan of the Semperoper Dresden with receiver position R1 to R6 and position S of the sound source.

positions R1 to R6 (KRAAK, 1984). Typically, the stage is narrowed for orchestral concerts using moveable wall elements to build a concert dome. Because this dome was not available, the measurements were taken with the safety curtain closed. The reverberation time in both situations is very similar (KRAAK, 1984).

White noise was used as a measurement signal and reproduced via the loudspeaker. However, the sound source shows some coloration. Magnitude spectra of the resulting sound pressures at all receiver positions at ear height are shown in Fig. 3. A homogeneous energy distribution over all receiver positions can be seen. However, a strong increase toward lower frequencies is observable, which results mainly from the characteristics of the sound source. This characteristic does not distort further results because the transfer function between the sound pressure at ear height and the acceleration at different surfaces will be calculated. The influ-

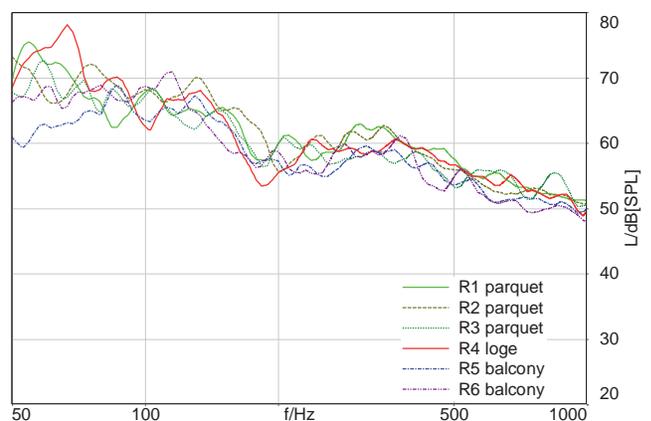


Fig. 3. Magnitude spectra (FFT 65536, 1/24th octave intensity averaging) of the sound pressure at ear height for all receiver positions.

ence of the overall sound pressure level on this transfer function will be discussed later.

To measure the vibrations at different surfaces, accelerometers (Kistler, 8636C10) were mounted to small metal plates with a 10 cm diameter and placed on the ground, on the seat and on the arm rest (see Fig. 4). The measurement position was then loaded with a person (80 kg). It should be noted that all other seats were unoccupied. The influence of a larger audience on the measured vibrations cannot be assessed easily. However, the presence of a second person in an adjacent seat did not change the results of a test measurement specifically.

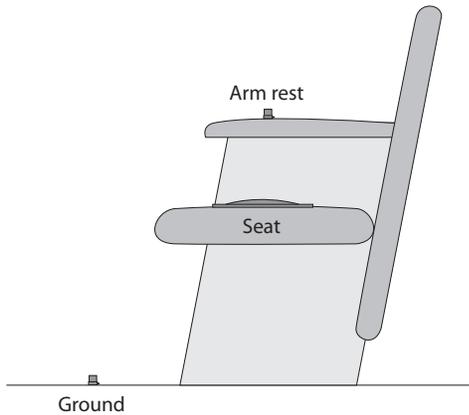


Fig. 4. Measurement setup with accelerometers on the ground, the seat and the arm rest.

### 3.2. Results

Figure 5 shows the transfer function between the acceleration on the ground and the sound pressure at ear height for the same measurement position. This corresponds to the difference between the acceleration level and the sound pressure level  $L_{acc} - L_{SPL}$ . A horizontal line at 0 dB represents equal levels for sound

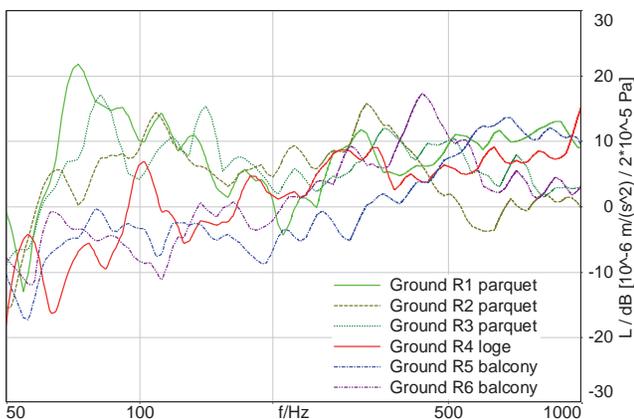


Fig. 5. Transfer functions (FFT 65536, 1/24th octave intensity averaging) between acceleration on the ground and the sound pressure at ear height at all receiver positions.

at ear height and vibrations on the ground. The overall sound pressure level at each measurement position was approximately 90 dB. It can be seen that higher acceleration levels were measured in many cases, especially in the parquet (R1–R3) for frequencies below 200 Hz. Interestingly, this location-dependent difference disappears at the seat surface (Fig. 6). The frequency response is relatively homogeneous over a broad frequency range with a slight decrease toward lower frequencies. Only a few positions showed isolated resonances (e.g., 100 Hz at receiver position R4).

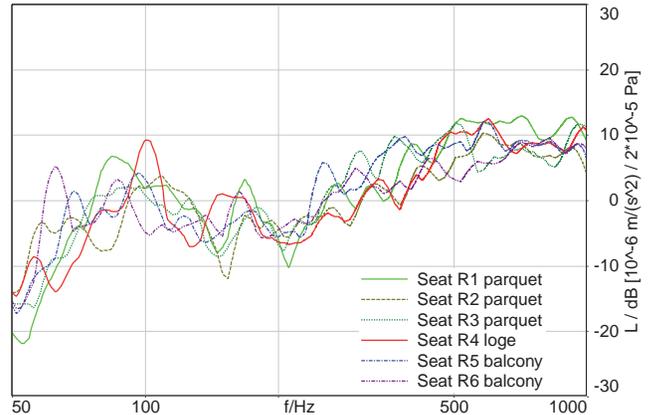


Fig. 6. Transfer functions (FFT 65536, 1/24th octave intensity averaging) between acceleration on the seat and the sound pressure at ear height at all receiver positions.

No distinct dependence of the overall level on the distance between the receiver and source was observed, e.g., there was not much difference between the accelerations at positions R1 and R3. This indicates that the vibrations are not transmitted via the ground from the loudspeaker to the seat. It is hypothesized that the airborne sound excites the surface near the listener. To test this hypothesis, the loudspeaker was decoupled from the stage floor using a sheet of foam (55 cm  $\times$  45 cm  $\times$  16 cm). The vertical resonance frequency of this system was measured to be approximately 8 Hz, resulting in effective vibration isolation in the interesting frequency range above 50 Hz. The exemplary transfer functions at position R4 (seat) in Fig. 7 show no considerable difference with and without the isolated loudspeaker. This finding supports the hypothesis that the vibrations are excited via air-borne vibrations in the auditorium. Air-borne transmission could also explain the lower levels on the ground for positions R4–R6 due to the smaller floor areas in the balconies.

If the sound pressure level rises, the excited vibrations should grow accordingly. However, it is not clear whether there is a linear relationship between both levels. Therefore, a few measurements were taken at different sound intensities. Figure 8 shows two exemplary transfer functions at position R4 (seat) with a 30 dB difference in the sound pressure level. Both curves are

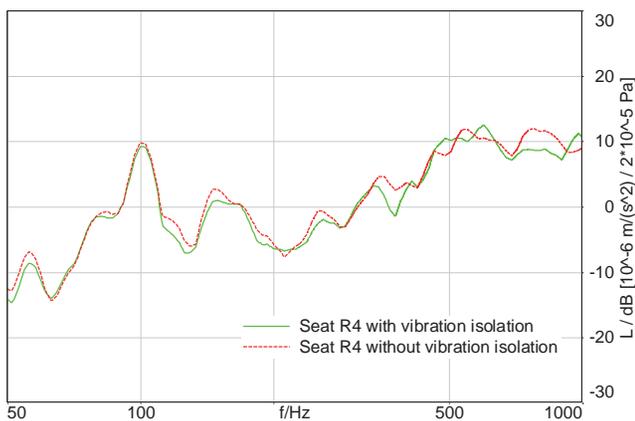


Fig. 7. Transfer functions (FFT 65536, 1/24th octave intensity averaging) between acceleration at the seat and the sound pressure at ear height at the receiver position R4 with and without vibration isolation of the sound source.

almost identical. This proves a linear relationship between the two physical variables, which is a prerequisite for meaningful transfer functions.

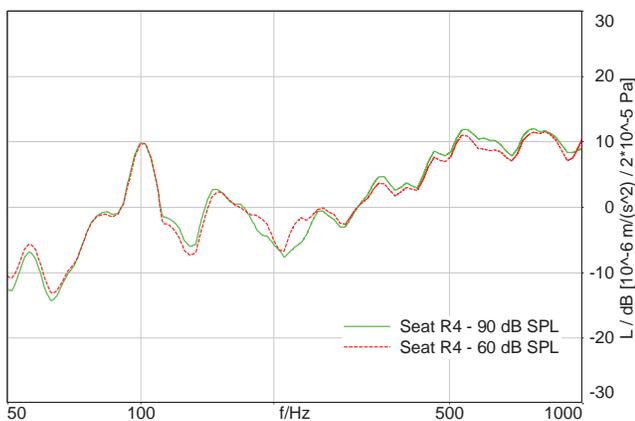


Fig. 8. Transfer functions (FFT 65536, 1/24th octave intensity averaging) between acceleration at the seat and the sound pressure at ear height at the receiver position R4 for different sound pressure levels.

### 3.3. Discussion

Typical sound pressure levels in a concert hall for fully orchestrated passages are between 80 dB and 90 dB (forte), depending on the instrumentation and the room (MEYER, 2009). The fortissimo can reach average sound pressure levels approximately 10 dB higher (WEINZIERL, 2008). For example, there have been measurements in the Semperoper with  $L_{AI}$  from 96 dB to 98 dB for themes from Wagner's Lohengrin (Kraak, 1984). The peak level at low frequencies can reach even higher values. Taking into account the perception threshold for sinusoidal seat vibrations, which is approximately  $L_{acc} \approx 90$  dB for frequencies below 150 Hz (MERCHEL *et al.*, 2011), the above

measurements indicate perceivable vibrations for the forte and fortissimo parts of orchestral music. However, these vibrations might not be perceived separately because of the integration of the tactile sense with the other sensory modalities into one multimodal concert event. During subsequent concert visits to the Semperoper, the author paid special attention to music-induced vibrations and could clearly identify them during a classical concert and a jazz concert. Interestingly, other concert visitors, who had been unaware of music-induced vibrations in the concert hall before, did confirm these findings.

The measurements suggest that differences between positions and the influence of local resonances should be clearly perceivable because they are considerably larger than the perceivable difference in the acceleration level, which is approximately 1.5 dB (FORTA, 2009). In addition, the dynamic range for vibration perception is quite small, which results in strongly perceived vibration intensity differences even for small changes in acceleration level. It is expected that these differences increase further between different venues. Therefore, a second measurement series was taken in a church.

## 4. Church

This section discusses the relation between sound and vibration in the Lutherkirche in Radebeul, a typical church build in 1892 with massive bearing walls and a stone floor. An outline of the church can be seen in Fig. 9. An organ loft is located in the back of the

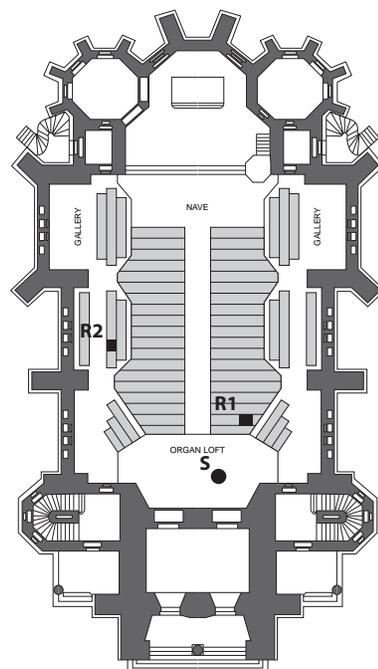


Fig. 9. Floor plan of the Lutherkirche Radebeul with positions of the receivers and the source.

church. Wooden pews can be found in the nave and on the wooden galleries. Again, transfer functions between sound pressure and acceleration were measured at different exemplary listening positions.

The same method and equipment was used as in the concert hall measurement described above. The dodecahedron source S1 was placed in the organ loft to simulate organ stimulation. Various measurement positions have been selected, but only two exemplary receiver positions (R1 in the nave and R2 in the gallery) will be discussed here.

#### 4.1. Setup

Again, different microphone setups were used to record various room impulse responses. Additional vibration impulse responses were measured on the ground, foot rest, seat and back rest of the wooden pews using accelerometers (Kistler, 8636C10). This is illustrated in Fig. 10. The measurement position was then loaded with the same person (80 kg) as before.

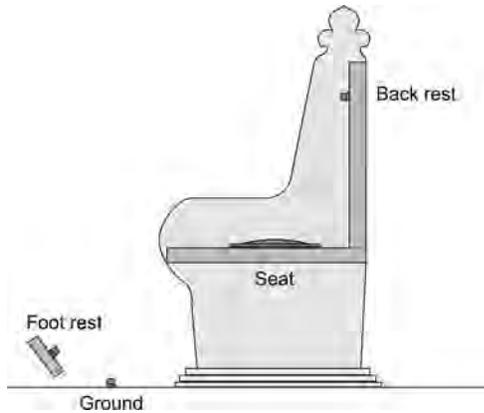


Fig. 10. Measurement setup with accelerometers on the ground, foot rest, seat and back rest.

#### 4.2. Results

Figure 11 plots the transfer function between acceleration on the ground and the sound pressure at ear height for the same position. The acceleration on the ground, which is excited by the sound, differs significantly between positions, due to the massive stone floor in the nave and the wooden construction of the gallery. Again, a broad vibration spectrum is excited with a slight roll-off toward lower frequencies.

This frequency spectrum differs completely for measurements at the foot rest, a long wooden board which is mounted only at its ends. Figure 12 shows strong resonances for foot rest vibrations at both receiver positions. This resonance pattern is also dependent on the position of the accelerometer along the board, which is not plotted here. The acceleration level varies considerably with frequency; however, the overall level is similar in both conditions.

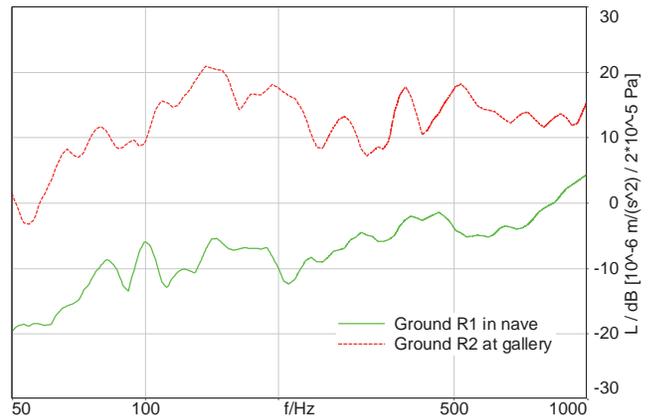


Fig. 11. Transfer functions (FFT 65536, 1/24th octave intensity averaging) between acceleration on the ground and the sound pressure at ear height in the nave and in the gallery.

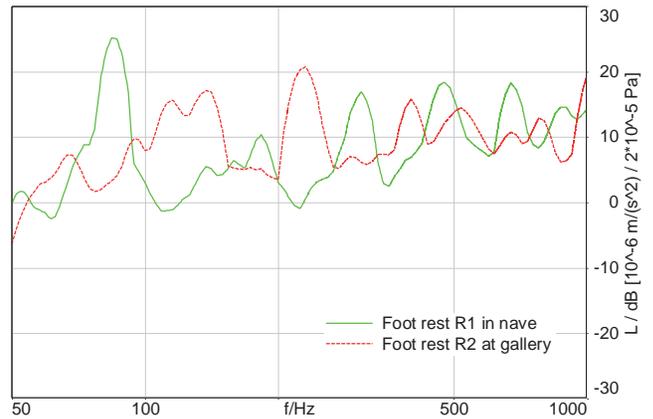


Fig. 12. Transfer functions (FFT 65536, 1/24th octave intensity averaging) between acceleration at the feet and sound pressure at ear height in the nave and in the gallery.

Comparable levels at both positions were also measured at the seat (see Fig. 13) and back rest. Compared with the concert hall, the overall acceleration level at

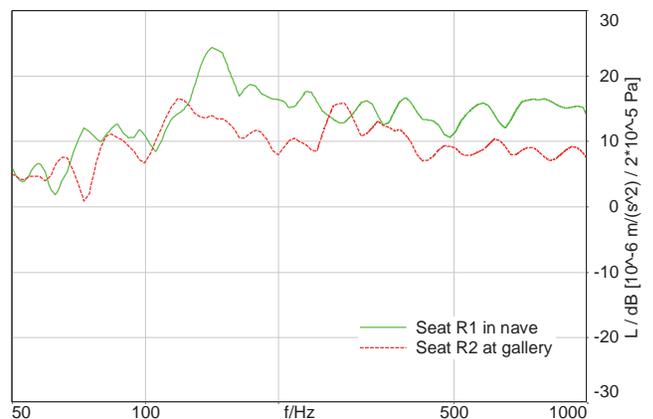


Fig. 13. Transfer functions (FFT 65536, 1/24th octave intensity averaging) between acceleration at the seat and the sound pressure at ear height in the nave and in the gallery.

the seat is significantly higher. This might be due to the missing seat upholstery and the large continuous surfaces, which can be excited by the sound more intensively.

Similar to the concert hall, no change of the transfer function was found when the subwoofer was decoupled from the ground, supporting the hypothesis of the dominance of airborne vibrations in the auditorium. There was also no dependence on the overall sound pressure level, confirming the linear relationship between sound pressure and acceleration discussed before.

#### 4.3. Discussion

Organ-generated sound pressure levels in a church can be quite high. A sample sequence with a significant low-frequency content (Max Reger, *Introduktion d-Moll*) performed by organ player Gottfried Treppe in the Lutherkirche Radebeul reached sound pressure levels of 90 dB(A) at both receiver positions. Clear-to-strong vibrations (100 dB) were excited most of the time in a broad frequency range.

### 5. Conclusions

It was shown that sound can excite perceivable surface vibrations. It can also be seen that our experience with audio-induced vibration can vary heavily. Even within one venue, the vibration intensities and frequency spectra are strongly dependent on the listener position. However, the measured sound-induced vibrations are only exemplary in nature.

The measurements reveal nothing about the perceived quality of such music-induced vibrations. No ideal sound-to-vibration transfer function can be deduced. To identify which vibrations are favorable, comprehensive listening tests are necessary, such as those described by MERCHEL and ALTINSOY (2008) and MERCHEL and ALTINSOY (2009). If an ideal sound-to-vibration transfer function exists, it might be possible to improve the concert experience by modifying vibrations through architectural changes or artificial generation. The latter case is especially interesting for audio reproduction systems but could improve the music experience even in a classical concert. It is expected that vibrations have a strong influence on the listener's presence or envelopment – parameters which are of vital importance for the quality of concert halls (CERDÁ *et al.*, 2012).

### Acknowledgments

The authors wish to thank the responsible people at the Semperoper Dresden and the Lutherkirche Rade-

beul for permission to conduct measurements in the respective venues.

Part of this work was presented at the 59th Open Seminar on Acoustics (Poland, Boszkowo, September 10–14, 2012).

### References

1. ABERCROMBIE C., BRAASCH J. (2010), *Auralization of audio-tactile stimuli from acoustic and structural measurements*, J. Audio Eng. Society, **58**, 10, 818–827.
2. ALTINSOY M. E. (2012), *The quality of auditory-tactile virtual environments*, J. Audio Eng. Society, **60**, 1/2, 38–46.
3. CERDÁ S., GIMÉNEZ A., CIBRIÁN R. M. (2012), *An objective scheme for ranking halls and obtaining criteria for improvements and design*, J. Audio Eng. Society, **60**, 6, 419–430.
4. DAUB M. (2003), *Cross-modal correlation – relationship between musically produced whole-body vibrations and auditory perception* [in German], Diploma Thesis, Institute of Communication Acoustics, Ruhr-University, Bochum.
5. FORTA N. G. (2009), *Vibration intensity difference thresholds*, Ph.D. Thesis, Institute of Sound and Vibration Research, University of Southampton.
6. KRAAK W. (1984), *Report on the acoustical testing of the Semperoper Dresden* [in German], TU Dresden.
7. MERCHEL S., ALTINSOY M. E. (2008), *5.1 or 5.2 Surround – Tactile enhancement of surround sound* [in German], DAGA, Dresden.
8. MERCHEL S., ALTINSOY M. E. (2009), *Vibratory and acoustical factors in multimodal reproduction of concert DVDs*, HAID, LNCS 5763, Springer, Berlin.
9. MERCHEL S., ALTINSOY M. E., STAMM M. (2011), *Equal intensity contours for whole-body vibrations compared with vibrations cross-modally matched to iso-phones*, HAID, LNCS 6851, Springer, Berlin.
10. MERCHEL S., ALTINSOY M. E., STAMM M. (2012), *Touch the sound: audio-driven tactile feedback for audio mixing applications*, J. Audio Eng. Society, **60**, 1/2, 47–53.
11. MEYER J. (2009), *Acoustics and the performance of music*, originally published in German by Edition Bochinsky, Springer, Berlin.
12. RUMSEY F. (2010), *Audio in multimodal applications*, J. Audio Eng. Society, **58**, 3, 191–195.
13. SIMON G., OLIVE S., WELTI T. (2009), *The effect of whole-body vibrations on preferred bass equalization of automotive audio systems*, 127th Conv. of AES, pp. 1–18, New York.
14. WEINZIERL S. [Ed.] (2008), *Handbook of audio technology* [in German], Springer, Berlin.

## Educational Implementation of a Sound Level Meter in the LabVIEW Environment

Robert BARAŃSKI, Grażyna WSZOŁEK

*AGH University of Science and Technology*

al. Mickiewicza 30, 30-059 Kraków, Poland; e-mail: {robertb, grazyna.wszolek}@agh.edu.pl

*(received November 4, 2011; accepted October 10, 2012)*

As a consequence of recent implementations of EU Directives related to noise protection more and more students of various AGH-UST programs are introduced to the basics of acoustic measurements. Students at various levels of theoretical background in the field of acoustic measurements are offered practical training in measurements using digital sound analyzers. The situation would be optimal if each student could have a device at his/her own disposal. Unfortunately, such a situation is not possible at the moment because of various reasons.

With the above problem in mind, a dedicated software package has been developed, implemented in the LabVIEW environment, which allows detailed studies of problems related to the acoustic signal measurement using sound level meters, as well as tasks in spectral analysis (1/1 and 1/3 band filters) and narrow-band (FFT) analysis. With such organization during the introductory laboratory classes each student is offered a direct individual contact with a virtual device that is properly pre-programmed for realization of a well-constructed learning process. It definitely facilitates understanding of the essence of acoustic signal measurements and provides a good basis for further laboratory work carried out as a team-activity.

**Keywords:** LabVIEW, acoustics, sound level meters, microphones, free-field investigation, frequency responses.

### 1. Introduction

Because of the recent implementation of EU directives, which put a special emphasis on the noise protection problems, there is a growing need for education and training of students in that field. At AGH-UST, in addition to the specialized classes for students of the Vibroacoustics program, there are many classes for students with a much weaker background in acoustics (e.g. from Mining or Geology Departments). Even after listening to the lectures on Acoustic Metrology, the practical measurements of the acoustic signal and their interpretation often seem to present serious problems for these students.

Laboratory activities related to acoustic measurements have to start from a general introduction covering the metrological features and functioning of sound level meters. Only after such introduction the student can start the routine acoustic measurements (e.g. in their work environment or external locations), or some more specialized tasks (e.g. measurements of the acous-

tic power, reverberation times etc.). It would be optimal if every student could have a separate measuring instrument at his/her disposal. Individual execution of the measurement exercises by each student adds much to the effectiveness of the learning process. Unfortunately, because of the rapid growth in the number of laboratory groups and the number of students in each group and because of considerable cost of professional devices for sound measurement such a situation is not expected to happen in the near future.

At present, in the Polish technical practice the predominant share of professional equipment for acoustic measurements consists of devices made by two Polish companies: Svantek and Sonopan, as well as two well-known global brands: Danish Brüel & Kjær and the Norwegian Norsonic. Most frequently those devices are digital sound analyzers with built-in sound level meters whose individual cost is very high – ca. several thousands of Polish zlotys. The detailed price depends on the included software and the number of additional functions of the device.

A professional sound level meter must conform to the Polish standard PN-EN 61672 (2005) and is subject to legal metrological verification – obligatory type approval according to the regulation of the Polish Ministry of Economy (2007). At present, the cheapest professional devices conforming to the above-mentioned requirements are produced by the Sonopan company. It manufactures only sound level meters with a function set conformant to the function set required by the standard PN-EN 61672-1 (2005) and the government regulation by the Polish Ministry of Economy (2007). Still the cost of such a device exceeds many times the prices of non-professional devices made in China, like AZ8921 or Monacor SM-2. The Chinese devices include the manufacturer's declaration of conformity to the standard PN-EN 61672-1 (2005) which is printed on their cases. There are also companies that issue calibration certificates for such devices. The combination of price, declaration of conformity, and/or calibration certificate creates a situation when these devices are frequently purchased, used, and even applied for the teaching practice. Only those well aware of the regulations related to the legal metrology and professional measurement are able to evaluate the actual value of such devices. In practice these measuring devices do not conform to many standard and legal requirements and cannot be calibrated by standard methods (PN-EN 61672-3, 2007). Thus, one cannot ensure the reliability of the results obtained from such measurements.

The problems with the lack of availability of a sufficient number of professional sound level meters and that prevents an effective realization of laboratory activities have lead to the concept of implementation of an educational sound level meter in the LabVIEW environment. We decided to use LabVIEW because its functionality and exhibility has been proved in many different research studies (TROJANOWSKI, WICIAK, 2010; BARAŃSKI, BATKO, 2011). An attempt has been made to construct and implement a laboratory measurement setup which could offer the students a possibility to gain a thorough and more detailed knowledge of the problems related to the acoustic signal measurement using sound level meters, as well as frequency band analysis (1/1 and 1/3 octave filters) and narrow-band Fourier analysis (FFT). The software package written for the above-mentioned measurement setup has been labeled as LVSLM.EDU.v1 (shortened from: LabVIEW\Sound.Level.Meters\EDUcation\version1). The actual implementation is based on the solution presented on 15th International Conference Noise Control 2010 (BARAŃSKI R., *Low Cost Sound Level Meter Based On LabVIEW*).

The software presented on 15th International Conference Noise Control 2010 has been subject (in combination with the soundcard of Dell E1405 laptop computer) to the tests that are routinely applied

to professional sound level meters during their periodical check-ups, according to the standard PN-EN 61672-3 (2007) and respective measurement procedures used in the Laboratory of Vibroacoustics (WSZOLEK, KŁACZYŃSKI, 2008). The results of those professional tests for such an unprofessional measuring device have been found to be quite satisfactory and enhanced the authors' motivation to a further activity in the proposed direction.

The present paper describes the effects of adaptation of the above-mentioned software to the requirements of teaching practice and detailed investigations of microphones used for the computerized measurements. Due to implementation of the LVSLM.EDU.v1 software each student (each of 15 people in a laboratory group) will be able to have a separate device for his/her own use. Visualization of selected problems will definitely facilitate the process of assimilation of the necessary knowledge related to acoustic measurements.

## 2. The software concept – basics of acoustic measurements

After assimilation of the basic ideas related to the acoustic phenomena the students proceed to studying devices used for acoustic measurements and respective measurement methods. The basic concepts used in the introduction are: sound (and the related subjective impressions), acoustic pressure (and the decibel scale), sound spectrum, acoustic field and types of acoustic signals. The basic knowledge regarding measuring devices has to include the sound level meter, measurement microphone, and band filters. If this basic knowledge is purely theoretical the student, after laying his/her hands on the device (at present often an advanced digital sound analyzer), will carry out measurements without actually knowing what he/she is doing at the moment. There are some digits flashing on the display of the device that in most cases are rather hard to understand and properly interpret. Therefore, it is very important that the freshly acquired theoretical knowledge is instantly implemented in practice. Namely, the student should instantly “feel, hear, and see” the effects of the device's settings on the measurement results.

The practical teaching experience of the authors of the present work indicates that the best effects in teaching sound level measurements are achieved by a simultaneous presentation to the students of both analog and digital meters. In such a situation students more promptly acquire the proper way of configuring the device in the context of a specific type of measured signal.

The learning process referring to the scope of the basic ideas used in the acoustic measurements has to start from an introduction to the sound level meter and the following actions:

- switching the device on (various solutions in various device types by various manufacturers sometimes make the process not-so-simple),
- setting the proper measurement range,
- setting the proper frequency weighting (A, C, LIN/Z),
- setting the proper time weighting (Fast, Slow, Impulse),
- configuration of the measured (and/or read-in) parameters of the measured quantity (momentary effective value SPL, maximum sound level, minimal sound level, peak-value sound level, equivalent sound level, sound exposure level),
- setting up the measurement time,
- device calibration using a reference sound source (acoustic calibrator) to the sound level declared by the manufacturer of a given device, in its reference range (if the device offers more than one measurement range).

Understanding of the dependence between the measurement results of individual parameters shown by the device and the generated and simultaneously heard acoustic signal is more easily gained when the analog indicator scale is observed. Thus, students learn faster how to set up properly the time characteristics or the measurement range.

Also, a visual presentation of the spectral analysis of simple signals (noise) without any correction curve (LIN/Z) followed by the same presentation with specific frequency weightings (A, C) leads to a fast comprehension of the effects of frequency characteristics on the measurement results and the connection between the correction effects and the human subjective impressions.

The decibel scale and understanding of the dependence between an equivalent value of the sound level and a momentary measured value are also frequently difficult to grasp for students. This is much easier to understand if the knowledge is presented in a visual mode.

Only when the student comprehends the above-mentioned basics of measurements, when his/her knowledge has been verified by practical measurements of various signal types under the supervision of a teacher, he/she is ready for the individual (in practice in a student's group) laboratory practice dedicated to specific types of measurements (e.g. traffic noise, industrial noise, acoustic power).

Implementation of the educational sound level meter in the LabVIEW environment (appended by the signal spectral analysis and FFT functions) allows the student to study in detail the properties of sound level meters and perform simple, evaluating measurements on his/her own using a dedicated computer setup.

### 3. Educational setup – VSLM\_EDU\_V1 software package

The main objective of the elaborated software for simulating functioning of a sound level meter, spectral analysis, and narrow-band frequency analysis is to prepare the user for working with real, professional measuring devices. Therefore, the program is equipped with the most frequently used measurement functions encountered in standard digital sound analyzers.

The student equipped with a PC computer with the LVSLM\_EDU\_V1 software, soundcard, microphone, and a set of headphones for monitoring the generated signals, can easily and without excessive costs study the basics of acoustic measurements and signal analysis. The software allows the analysis of exemplary acoustic signals selectable from a pull-down menu (e.g. real recordings of the traffic noise, industrial room environment noise) and the analysis of signals received by a microphone in the actual laboratory environment (e.g. signals generated by the extra equipment used by the supervising teacher).

Such a solution allows direct observation of the reaction and reading changes on the virtual measuring device caused by an arbitrary acoustic signal registered by the measuring system. Practical experience suggests that such a visual presentation provides the fastest progress in understanding the work of a professional sound level meter.

The authors of the paper, being aware of the errors occurring in the readings of the virtual meter, have carried out a series of tests concerning the accuracy analysis of the results obtained from a virtual instrument based on a low-cost soundcard and a microphone routinely used in standard measurement solutions. The results of the electrical tests executed by measurement procedures (WSZOLEK, KLACZYŃSKI, 2008) conformant to the standard (PN-EN 61672-3, 2007) have been presented on 15th International Conference Noise Control 2010. The acoustic studies of low-cost, generally available microphones dedicated to the present measurement setup are described in Sec. 5 of the present paper.

### 4. Description of the LVSLM\_EDU\_V1 application

The LVSLM\_EDU\_v1 program has been developed in the LabVIEW 2010 (32-Bit) environment (LabVIEW 2010, Help, National Instruments Corp., 2010). Additionally, functions implemented in the Sound and Vibration Toolkit ver. 2010 (NI Sound and Vibration Measurement Suite, 2009) have been used. The main application window of LVSLM\_EDU\_v1 is shown in Fig. 1. Further-on the complete computer measurement setup together with the necessary software will be labeled as a single term “measuring device”.

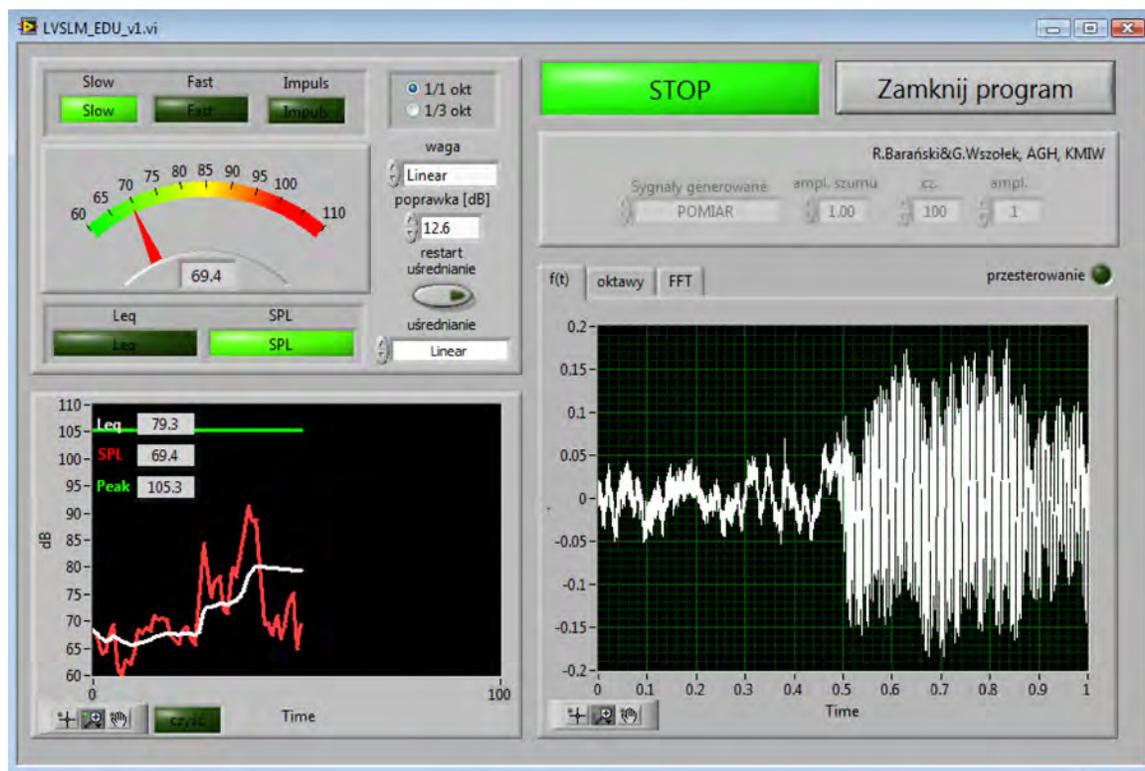


Fig. 1. Main application window of LVSLM\_EDU\_v1.

The main application window has been split into four sections: SIGNAL, SETTINGS, ACOUSTIC PARAMETERS, SPECTRA (see Fig. 1). Each panel (excluding the SIGNAL) reflects the basic measurement functions implemented in the sound level meter and its enhancements covering the spectral and narrow-band signal analyses.

Additionally, the main window includes two buttons executing respectively switching the meter on and off (START/STOP) and closing the main application window (Exit program).

#### 4.1. The “Signal” section

In order to ensure the ability to simulate the functioning of a real device the virtual signal meter has been prepared to operate with several signal types, both simulated and real (Fig. 2).

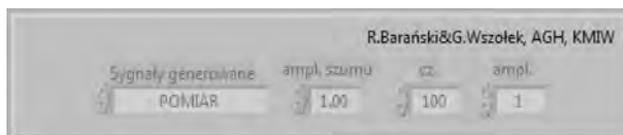


Fig. 2. SIGNAL section.

The list of signal types that can be analyzed by the virtual meter looks as follows:

- sinusoidal signal,
- sinusoidal + pseudorandom signal (Gaussian noise distribution),

- pink noise,
- traffic noise (registered in the Mickiewiczza alley in Kraków),
- industrial noise at a construction site (work of an excavator and pneumatic hammer).

For every generated signal type there is a possibility of changing the settings (parameters like amplitude and/or frequency).

The last option in the SIGNAL section is the MEASUREMENT. It allows registration of the sounds present in the local neighborhood, using the computer's soundcard and a connected external microphone.

#### 4.2. The “Settings” section

The panel of the SETTINGS section is shown in Fig. 3. That section is used for a proper configuration of the measurement session i.e.:

- Setting the proper time weighting (Fast, Slow, Impulse),
- Selection of the measured and read-in parameters of the measured quantity (momentary effective value SPL, peak-value sound level, equivalent sound level),
- Activation of the proper frequency weighting (A, C, LIN/Z),
- Selection of the averaging type (Linear, Exponential),

- Calibration of the device using a reference sound source (acoustic calibrator),
- Selection of the frequency bandwidth for spectral analysis.

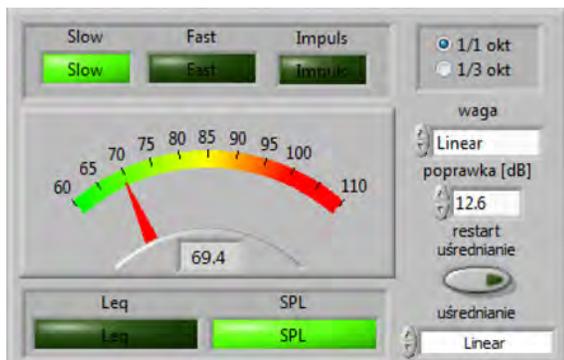


Fig. 3. Window of the SETTINGS section configuration.

In the central part of the SETTINGS section the analog and digital indicators are located, which is a direct reference to the indicators on the sound level meters – both in the analog (pointer on a scale) and digital (a window displaying the results in dB) versions. Such a visual presentation of changes of the measured quantity (by both pointer movement and changing digits) allows the students to observe the effects of setting the respective time characteristics (Fast, Slow, Impulse) on the measurement results. It also allows a direct observation of establishing the *equivalent level* ( $L_{eq}$ ) or *momentary value* of *Sound Pressure Level* (SPL) for the analyzed signal.

An essential element is the averaging used in the process of calculating the  $L_{eq}$  value. The available averaging modes *Linear* or *Exponential* are restarted automatically after each initiation of the measurement or manually by using the *restart averaging* button.

Since the device offers the ability to perform measurements with an arbitrary type of a soundcard plugged into the computer, a dedicated element has been implemented, responsible for the calibration of the measuring system. For that purpose a special calibration correction, specified in dB, is taken into account in the signal processing line. During the system calibration, using a professional acoustic calibrator (e.g. B&K type 2231), the student specifies the proper correction value, so that the LVSLM\_EDU\_v1 software can adjust the result to the proper calibration value (e.g. 94 dB). In order to ensure a proper calibration of the signal processing line, a special matching adapter has been prepared, matching the atypical microphone size to the housing in the professional calibrators. The authors of the paper are fully aware of the importance of matching the proper adapter to the calibrator. It has been proved that the adapter directly affects calibration results. However, for educational purposes, the accuracy of calibration described

above is quite sufficient. The calibration procedure itself is intended to develop in students the habits of professional measuring personnel – the awareness of performing the device calibration before and after measurements.

#### 4.3. The “Acoustic Parameters” section

That section allows observation of the history of changes for the basic parameters used in the current acoustic measurement, like  $L_{eq}$ , SPL, or Peak (Fig. 4). Such a presentation offers the user the possibility to understand the differences in calculation of these three parameters. Additionally, by comparison it is easy to notice how the parameters influence each other and what the possible consequences of this are.

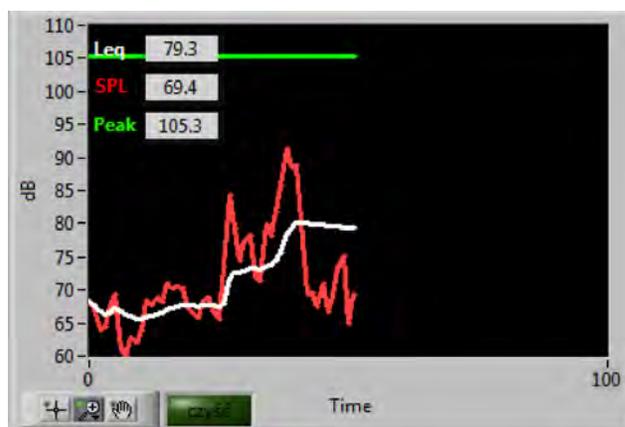


Fig. 4. Window of the ACOUSTIC PARAMETERS section.

The best example is provided by observation of the strong effects of a very high momentary SPL value on the values of  $L_{eq}$  and their changes in consecutive averaging results.

#### 4.4. The “Spectra” section

That section allows observation of the history (Fig. 5). In order to introduce the student to identification of individual characteristic frequencies in the analyzed signals, a separate FFT tab allows observation of *Fast Fourier Transform* (FFT) analysis results. It is particularly useful in the analysis of signals with strong monoharmonic components like those during the calibration of the measurement system. Another tab called Octave allows presentation of the frequency spectrum of the analyzed signal using the octave filters (1/1 octave) or tertiary filters (1/3 octave).

The program allows presentation of the frequency spectra of the analyzed signal, calculated with various frequency weightings A, C, or LIN. The user is offered the choice to switch on/off an arbitrary frequency correction characteristic.

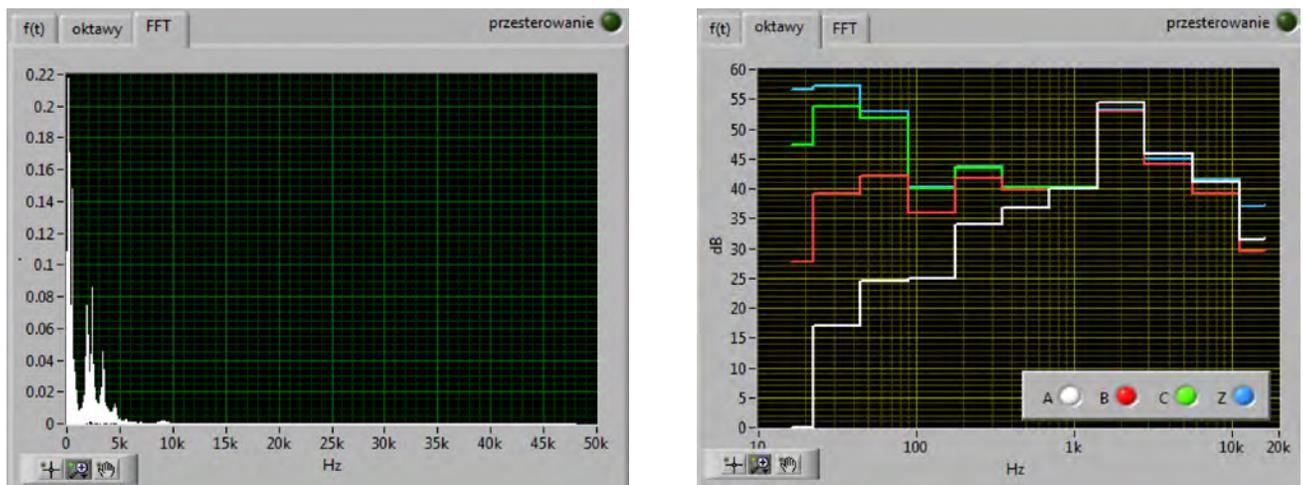


Fig. 5. Windows of the SPECTRA section for FFT and Octave tabs.

## 5. Examination of the microphones in free field

As mentioned before the tests of the implemented sound level meter SLM\_LV.1.2c, which provided the basis for the development of the LVSLM.EDU.v1 software, have been presented on 15th International Conference Noise Control 2010. They have been carried out according to Sec. 3 of the standard PN-EN 61672 dedicated to the periodical check-ups. They did not include determination of the frequency response of the implemented device or examination of the microphones.

The present paper describes the examination of two types of widely available low-cost microphones that can be used in the described educational measurement setup. Because they are not measurement microphones, conformant to the standard – they cannot be examined using typical indirect methods (e.g. by electrostatic actuator method) according to the above-mentioned standard (PN-EN 61672-3, 2007) or dedicated standards for calibration of measurement microphones (PN-EN 61672-3, 2007). The microphones had to be examined in free field conditions (in anechoic chamber), by a direct determination of the frequency response of a given microphone or the whole meter. Examinations have been carried out for two types of microphones: MM10 and Digitus. Figure 6 presents the overall view of the whole measurement setup in the anechoic chamber.

The microphones have been tested according to professional rules applied for determination of device characteristics during acoustic measurements in free field conditions (WSZOLEK, ENGEL, 2004; ENGEL, WSZOLEK, 2004; WSZOLEK, BARWICZ, 2007; WSZOLEK, 2008). During the tests, fully professional equipment has been used including a measurement system used for calibration of acoustic devices, based on the *PomAk* (WSZOLEK *et al.*, 2006) computer program. An LW2 class microphone – Brüel & Kjær type 4191 – was used as the reference (standard) one.



Fig. 6. Overall view of the measurement setup in the anechoic chamber.

In order to determine the “averaged” frequency characteristics for the examined microphones, 11 pieces of each microphone type have been examined. During the microphone tests the integrated soundcard of a Lenovo laptop computer has been used as a preamplifier. The installed ASIO drivers were responsible for transferring the signal to the headphone outputs where further analysis took place. The microphone characteristics have been determined for  $0^\circ$  wave incidence angle.

Figure 7 presents the average frequency response for both types of microphones. The figure also includes the tolerance limits for frequency characteristics of class 2 sound level meters, conformant to the requirements of part 1 of the standard PN-EN 61672 (2005).

The microphone characteristics can be subject to small changes due to the detailed features of its connec-

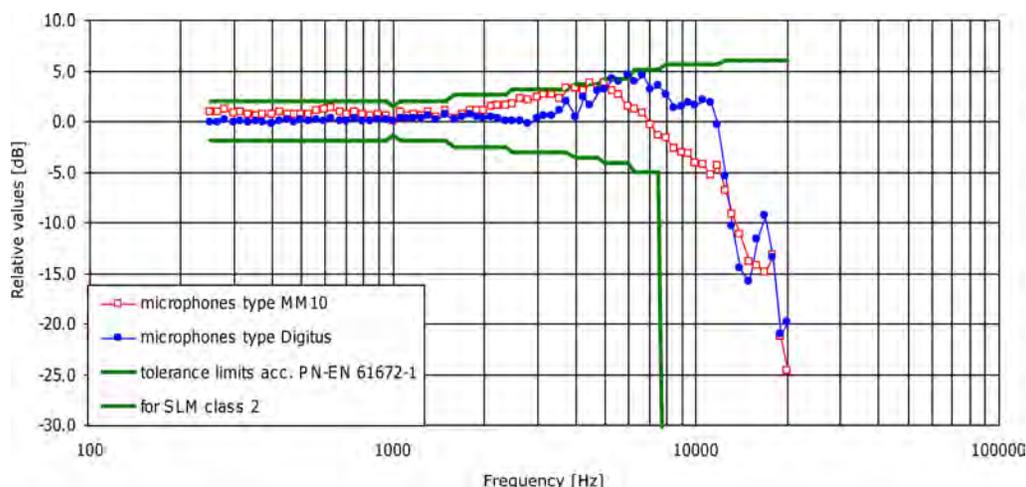


Fig. 7. Average frequency responses (typical characteristics) for MM10 and Digitus microphones, determined in free field conditions, for wave incidence angle  $0^\circ$ .

tion to the measurement system (the soundcard). However, the results of the study seem to be satisfactory and there are no objective arguments precluding such low-end microphones from being used in education-dedicated laboratory setups.

## 6. Conclusions

The present work is a continuation of the study presented on 15th International Conference Noise Control 2010. It describes the effects of adaptation of a specialized software developed in the LabVIEW environment for educational applications. The presented computer-based measurement setup is intended to offer the students a thorough and more detailed understanding of the problems related to acoustic signal measurements using sound level meters and frequency band analysis (1/1 and 1/3 octave filters), as well as the narrow-band Fourier analysis (FFT).

The concept of elaboration of such a virtual measurement setup has been born during the teaching activities in the Chair of Mechanics and Vibroacoustics, namely, during the laboratory classes with students with strongly varying levels of theoretical knowledge in the field of metrology and acoustics. Most frequently, the knowledge concerning the acoustic metrology is not sufficient for its correct implementation in practice. The flaws in proper understanding of the basic concepts and errors in a practical configuration of a sound level meter (setting the time constants and meter calibration) result in errors during execution of measurement and further interpretation of the results.

During the laboratory activities the actual number of students taking an active part in the measurement is limited because of the small number of available measuring devices, while the number of people in the laboratory groups is growing. In consequence, after completing the laboratory class, the majority of students

still do not possess the necessary skills in operating the sound level meter at a competence level necessary for carrying out sound measurements (both routine and more specialized ones) and proper interpretation of the results.

Application of the LVSLM\_EDU\_v1 virtual sound level meter with the implemented additional measurement functions will allow a wider number of students to have an individual contact with the measuring device starting from the first introducing laboratory class. The educational scope of the described application covers the necessary knowledge in the field of acoustic metrology. Understanding and internalization of the concepts, practical knowledge of the device settings and its calibration will definitely have positive effects during the following classes and performing of specific sound measurement (e.g. in external environment, on a specific working post, determination of the acoustic power of devices etc.). The laboratory classes can be organized using standard PC computers available in the computer rooms used for teaching information technology.

The proposed software is dedicated purely to educational purposes. The important part is to inform the users that the software cannot replace any professional measuring devices in the same way as professional measuring devices cannot be replaced by low-cost, non-professional devices for sound measurements that do not conform to the requirements of the respective legal acts and standard regulations.

## Acknowledgment

The research has been realized under AGH-UST Statutory Research contract nr 11.11.130.885 in year 2011. The work is an extended version of the paper delivered at the 58th Open Seminar on Acoustics, 13–16 September 2011, Gdańsk-Jurata.

## References

1. BARAŃSKI R., BATKO W. (2011), *The Spectral Response of Hand-Arm System*, Acta Physica Polonica A, **119**, 913–915.
2. ENGEL Z., WSZOŁEK G. (2004), *Investigations of Sound Level Meter in Free Field Conditions* [in Polish: *Badania mierników poziomu dźwięku w polu swobodnym*], Proceedings of Congress of Metrology, KM 2004, Vol. 2, pp. 683–686, Wrocław.
3. LabVIEW 2010, Help, National Instruments Corp., 2010.
4. *Manufacturer's data from NI Sound and Vibration Toolkit*, NI Sound and Vibration Measurement Suite, 2009, available at [http://www.ni.com/pdf/products/us/cat\\_svmsuite.pdf](http://www.ni.com/pdf/products/us/cat_svmsuite.pdf), dated 2011-06-11.
5. PN-EN 61672 *Electroacoustics – Sound level meters*; Part 1: *Specifications* (2005), Part 2: *Pattern evaluation tests* (2005), Part 3: *Periodic tests* (2000) [in Polish: *Elektroakustyka. Mierniki poziomu dźwięku*; Część 1: *Wymagania* (2005), Część 2: *Badania typu* (2005), Część 3: *Badania okresowe* (2007)].
6. PN-EN 61094 *Measurement microphones – Part 4: Specifications for working standard microphones* (2000), *Part 6: Electrostatic actuators for determination of frequency response* (2005) [in Polish: *Mikrofony pomiarowe. Część 4: Wymagania dla roboczych mikrofonów wzorcowych*, Część 6: *Pobudniki elektrostatyczne do wyznaczania charakterystyki częstotliwościowej*].
7. *Regulation by the Ministry of Economy dated May, 28th 2007 r. concerning the requirements imposed on the sound level meters and a detailed specification of examinations and check-ups carried out during legal metrological inspection of these measurement devices* (Dz.U. nr 105, poz. 717) [in Polish: *Rozporządzenie Ministra Gospodarki z dnia 28 maja 2007 r. w sprawie wymagań, którym powinny odpowiadać mierniki poziomu dźwięku, oraz szczegółowego zakresu badań i sprawdzeń wykonywanych podczas prawnej kontroli metrologicznej tych przyrządów pomiarowych* (Dz.U. nr 105, poz. 717)].
8. WSZOŁEK G., ENGEL Z. (2004), *Investigations of Uncertainty of Acoustical Measuring Instruments Applied to Noise Control*, Archives of Acoustics, **29**, 2, 283–295.
9. WSZOŁEK G., BARWICZ W., DUMINOV S. (2006), *Automatic Measuring System for Acoustical Devices Calibration in the Free-Field*, Archives of Acoustics, **31**, 4, 397–402.
10. WSZOŁEK G., BARWICZ W. (2007), *Comparative Calibrations of Measuring Microphones Performed in a Free Field*, Archives of Acoustics, **32**, 4 (Supplement), 305–312.
11. WSZOŁEK G., KŁACZYŃSKI K. (2008), *Calibration of Sound Level Meters according to the PN-EN 61672-3 Standard, Procedure S3/B, Accredited with the PCA Calibration Laboratory*, third edition dated 15.12.2008, Cracow [in Polish: *Wzorcowanie mierników poziomu dźwięku zgodnie z PN-EN 61672-3, procedura S3/B akredytowanego w PCA laboratorium wzorcującego*, wyd. 3 z 15.12.2008, Kraków].
12. WSZOŁEK G. (2008), *Selected Problems of Uncertainty Assessment in the Calibration of Acoustic Measuring Instruments* [in Polish: *Wybrane zagadnienia szacowania niepewności przy wzorcowaniu przyrządów do pomiarów akustycznych*], Proceedings of 50th Open Seminar on Acoustics OSA 2008, pp. 343–348, Wrocław-Szklarska Poręba/Piechowice.
13. TROJANOWSKI R., WICIAK J. (2010), *Preliminary Results of Laboratory Tested System for Active Control of Plates via LabVIEW and Piezoelectric Elements*, Acta Physica Polonica A, **118**, 168–173

# Efficient Phantom Source Widening

Franz ZOTTER, Matthias FRANK

*Institute of Electronic Music and Acoustics, University of Music and Performing Arts Graz*  
Inffeldgasse 10/3, 8010 Graz, Austria; e-mail: {zotter, frank}@iem.at

(received July 9, 2012; accepted December 6, 2012)

We present a highly efficient filter structure to create power-complementary filter pairs for phantom source widening. It either introduces frequency-dependent phase or amplitude differences in a pair of loudspeaker signals. We evaluate how the perceptual effect is influenced by off-center listening positions in a standard  $\pm 30^\circ$  loudspeaker setup. The evaluation of the phantom source widening effect is based on measurements of the inter-aural cross-correlation coefficient (IACC), which is justified by its pronounced correlation to the perceived phantom source width in prior listening test results.

**Keywords:** source widening, phantom source, decorrelation, IACC.

## 1. Introduction

A pair of loudspeakers symmetrically arranged with regard to the listener and driven by the same signal evokes a perceived auditory event in the middle between the loudspeakers. The location of this auditory event, the so-called phantom source (WENDT, 1963), is generally known to be influenced by time-delay and/or level differences of the loudspeaker signals (WENDT, 1963; PLEWA, KLECZKOWSKI, 2011). In general, it appears to have a width (KIN, PLASKOTA, 2011) that can be different to the width of a real sound source (ZOTTER *et al.*, 2011). The literature about (pseudo-)stereophony (GERZON, 1993; SCHRÖDER, 1958; ORBAN, 1970b), stereo decorrelation (KENDALL, 1995; POTARD, BURNETT, 2004; BOUÉRI, KYRIAKAKIS, 2004), and parametric spatial audio (LAITINEN *et al.*, 2012; POTARD, 2006; SZCZERBA *et al.*, 2011) describes that nonuniform amplitude differences or phase differences over frequency are suitable for controlling the phantom source width. Phantom source widening provides a way of shaping the spatial salience of a sound signal with only little effect on its location, reverberation, or coloration.

A recently presented sinusoidal-phase all-pass method (ZOTTER *et al.*, 2011) could be tested in varying parameters. It seemed to be quite efficient and effective according to listening tests that were done at the central listening spot. Moreover, a pronounced negative correlation between perceived source width and the inter-aural coherence (measured by the inter-aural

cross-correlation coefficient, IACC) could be experimentally proven. Nevertheless, the question remains how good the method works for outside the central listening spot and how much audible spectral corruption is produced.

Among the known methods, not all are phase-based. The known amplitude-based approaches are not similar to the phase-based method (ZOTTER *et al.*, 2011) in various ways (response structure, computational efficiency, power-complementarity). If a more similar amplitude-based phantom source widening algorithm was found, it would be useful to pose the question, whether amplitude-based or phase-based widening would be favorable.

This article therefore reviews suitable deterministic phantom source widening strategies in Sec. 2 with regard to their differences, considering computational effort, power-complementarity, and phasiness documented by GERZON (1993). A further simplified version of the phase-based method from (ZOTTER *et al.*, 2011) is presented in Sec. 3, making its implementation easier and its highly efficient nature clearly obvious. As alternative, Sec. 4 presents a novel amplitude-based method that uses the same filter structure as the phase-based one and re-uses its coefficients. Section 5 is a review of the experimental results from (ZOTTER *et al.*, 2011) and the IACC as a predictor, and it discusses the dependency of the IACC on the inter-channel cross-correlation coefficient (ICCC). The prediction is applied to laterally shifted listening positions in Sec. 6 with both the phase-based and amplitude-

based method, using the IACC (ISO, 2009) measured with a dummy head. Hereby, the question can be dealt with, how good both phantom source wideners work in comparison to each other and outside the central listening position. Additionally, Sec. 6 observes in how far the algorithms introduce perceivable spectral corruption by measuring composite binaural third-octave level deviations. The deviations from the levels of the unaltered phantom source indicate possible coloration effects.

## 2. Review of phantom source widening filter pairs

*Raised cosine filter responses* are surprisingly simple to achieve. A straightforward implementation is done as a pair of discrete-time systems. In the  $z$ -transform domain, a delay by a positive time-shift  $\delta(t - T)$  is written as  $z^{-N}$  (OPPENHEIM *et al.*, 1999), with  $N = T f_s$ , the adjustable time-delay  $T$ , and the sample rate  $f_s$ ,

$$H_{1,2}(z) = \frac{1}{2} \pm \frac{\hat{\phi}}{2} [z^N + z^{-N}] \quad (1)$$

and eventually yields the linear time-invariant (LTI) frequency responses of a raised cosine with a pair of signs

$$H_{1,2}(\omega) = \frac{1}{2} \pm \hat{\phi} \cos(\omega T). \quad (2)$$

To avoid out of phase signals, the parameter  $\hat{\phi}$  should be used within the interval  $-1/2 \leq \hat{\phi} \leq 1/2$ . Although this filter pair would produce widened phantom sources, it is not surprising that it was not discussed a lot. Its result is perceived as severely corrupted and colored when  $|\hat{\phi}|$  is close to one half. This is largely because the frequency responses are only amplitude complementary,  $H_1(\omega) + H_2(\omega) = 1$ , but not power-complementary.

The so-called *Lauridsen network* (SCHRÖDER, 1958; ORBAN, 1970a,b) is another simple pair of comb filters aiming at creating a widened phantom source. Combining positive and negative time-shifts with different signs, the pair of Lauridsen filters

$$H_{1,2}(z) = \frac{1}{2} [z^N \pm z^{-N}] \quad (3)$$

has Fourier transforms that are obtained after inserting  $z = e^{i\omega/f_s}$ ,  $N = T f_s$ , and applying Euler's identities

$$H_1(\omega) = \cos(\omega T), \quad H_2(\omega) = i \sin(\omega T). \quad (4)$$

These filters are LTI and power-complementary,

$$|H_1(\omega)|^2 + |H_2(\omega)|^2 = 1, \quad (5)$$

due to  $\cos^2 + \sin^2 = 1$ . One could expect that their effect is less annoying. However, this is not the case: their effect was perceived as even more annoying. GERZON (1992) reported it to be phasey and assumed that this is due to the phase shift of  $90^\circ$  between the filter responses.

*Frequency-dependent amplitude variation* that is time-invariant has been considered by GERZON (1992, 1993) in order to reduce the effects of phasiness. Among various other phase-based, even time-variant methods he proposed, this is done by a frequency dependent and power-complementary amplitude variation between the channels using the parameter  $-\pi/4 \leq \tilde{\varphi}(\omega) \leq \pi/4$

$$\begin{aligned} H_1(\omega) &= \cos(\pi/4 + \tilde{\varphi}(\omega)), \\ H_2(\omega) &= \sin(\pi/4 + \tilde{\varphi}(\omega)). \end{aligned} \quad (6)$$

Gerzon noted that this part of his idea would be more complicated to implement than all-passes. The impulse responses need to be made causal and would tend to be long. In his opinion, the curve  $\tilde{\varphi}(\omega)$  should describe a to-and-fro contour on a Bark frequency scale.

*Frequency-varying Vector-Base Amplitude Panning* (PULKKI, 1997; LAITINEN *et al.*, 2012) on a pair of loudspeakers located at  $\pm\alpha$  can be expressed by Eq. (6) using energy-normalized panning (*tangent law*, PULKKI (1997)). The angle  $\tilde{\varphi}(\omega)$  is dependent on the panning angle that lies between the loudspeakers  $-\alpha \leq \varphi(\omega) \leq \alpha$

$$\tilde{\varphi}(\omega) = \arctan \left\{ \frac{\tan[\varphi(\omega)]}{\tan \alpha} \right\}. \quad (7)$$

LAITINEN *et al.* (2012) proposed to employ a triangular panning curve over frequency, which lies between  $-\hat{\phi} \leq \varphi(\omega) \leq \hat{\phi}$  and has a random starting parameter  $\varphi_0$  for low frequencies,

$$\varphi(\omega) = \hat{\phi} \left[ 1 - \frac{2}{\pi} \arccos[\cos(\omega T + \varphi_0)] \right]. \quad (8)$$

Normally implementations of the resulting frequency responses are done by employing block convolution, which allows full freedom of adjustment, considering the width parameter  $\hat{\phi}$ , the loudspeaker angles  $\pm\alpha$ , and the parameter for the zero frequency  $\varphi_0$ . If only two loudspeakers are used, the curve  $\varphi(\omega)$  is limited to lie between  $-\alpha \leq \varphi(\omega) \leq \alpha$ . However, if more than one loudspeaker pair is available, the curve  $\varphi(\omega)$  may run throughout a single loudspeaker pair. Nevertheless, the implementation of such elaborated real-valued, all-positive linear-phase filters requires a certain level of computational complexity.

*All-pass approaches with frequency-dependent phase or time-delays* were proposed by various authors (ORBAN, 1970a,b; GERZON, 1992, 1993; KENDALL, 1995; POTARD, BURNETT, 2004; BOUÉRI,

KYRIAKAKIS, 2004; POTARD, 2006) who applied IIR all-passes or random-phase Fourier-based FIR filters. A recent work (ZOTTER *et al.*, 2011) investigated a sinusoidal-phase all-pass with a sparse and efficient FIR implementation in the time-domain. Hereby, a static cosine contour of the inter-channel time-delay with regard to frequency was achieved, which reliably and efficiently produced widened phantom sources. Due to the convincing results of the listening test, this method is a reference here, and it will be simplified to a reasonable approximation in the following section.

### 3. Phase-based: efficient sinusoidal phase all-pass pair

The widening method in ZOTTER *et al.* (2011) inserts a cosine-modulated inter-channel time-delay with the periodic modulation interval of  $\Delta f = 1/T$  over frequency and a peak magnitude of  $2\hat{\tau}$  between the pair of channels. Despite the algorithm applied the time-delay modulation depth  $\hat{\tau}$  as a parameter, a closer inspection of the listening test results in Sec. 5 reveals that rather the phase modulation depth  $\hat{\phi} = \hat{\tau}/T$  is a relevant parameter. In particular, the width indicated by the listeners and the IACC was related to the product  $\hat{\tau} \Delta f = \hat{\tau}/T = \hat{\phi}$ , see also Table 1. Therefore  $\hat{\phi}$  is favored as parameter here, also because it appears to control the ICCC independently of the time-delay  $T$ .

Table 1. Parameters for both phase-based and amplitude-based widening, resulting identical ICCC and IACC<sub>E3</sub> values for the central listening position.

| $\hat{\phi}$ | $\Delta\phi_{\max} = 2\hat{\phi}$ | $\Delta L_{\max}$ | ICCC | $P$ : IACC <sub>E3</sub> | $A$ : IACC <sub>E3</sub> |
|--------------|-----------------------------------|-------------------|------|--------------------------|--------------------------|
| 0.00         | 0.0°                              | 0.0 dB            | 1.0  | 0.88                     | 0.88                     |
| 0.31         | 35.6°                             | 5.8 dB            | 0.9  | 0.78                     | 0.78                     |
| 0.45         | 52.0°                             | 9.3 dB            | 0.8  | 0.68                     | 0.67                     |
| 0.57         | 66.3°                             | 13.6 dB           | 0.7  | 0.58                     | 0.57                     |
| 0.66         | 77.2°                             | 19.0 dB           | 0.6  | 0.50                     | 0.48                     |

The FIR responses from ZOTTER *et al.* (2011) use cylindrical Bessel functions and are re-written in the  $z$ -transform with  $J_{-l}(\hat{\phi}) = (-1)^l J_l(\hat{\phi})$  and  $N = T f_s$ ,

$$H_{1,2}(z) = \frac{1}{\sqrt{2}} \sum_{l=0}^{\infty} J_l(\hat{\phi}) \left[ z^{\pm lN} + (-1)^l z^{\mp lN} (1 - \delta_{l,0}) \right]. \quad (9)$$

The Kronecker delta  $\delta_{l,0}$  was inserted to avoid a factor of two for  $l = 0$ ; it is zero for  $l \neq 0$  and 1 otherwise. The Fourier transform of the equation was demonstrated in ZOTTER *et al.* (2011) and simply is

$$H_{1,2}(\omega) = \frac{1}{\sqrt{2}} e^{\pm i\hat{\phi} \sin(\omega T)}. \quad (10)$$

The impulse responses are sparse and real-valued, and with small arguments  $\hat{\phi} \leq \pi/2$ , the infinite sum can be easily truncated. According to the appended series expansion Eq. (23), the filter pair is simplified with reasonable accuracy by the equation

$$\sqrt{2} H_{1,2}(z) = g_0 \pm g_1 [z^N - z^{-N}] + g_2 [z^{2N} + z^{-2N}] \quad (11)$$

with the weights

$$g_0 = 1 - \frac{\hat{\phi}^2}{4}, \quad g_1 = \frac{\hat{\phi}}{2} - \frac{\hat{\phi}^3}{16}, \quad g_2 = \frac{\hat{\phi}^2}{8}. \quad (12)$$

Figure 1 shows the signal flow graph for causal phase-based phantom source widening applied on a mono signal, and its frequency responses are depicted in Fig. 2. The power  $|H_1|^2 + |H_2|^2$  of the simplified filter pair deviates from unity by less than 0.1 dB, which

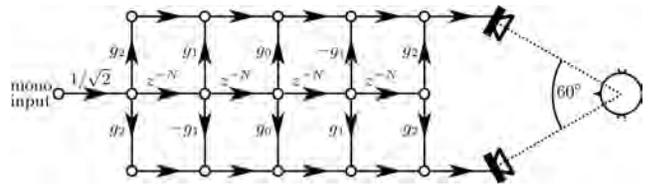


Fig. 1. Signal flow graph of a simplified and causal phase-based phantom source widener with extremely low computational demands. Suitable weights  $g_l$  are specified in Eq. (12) and  $N = T f_s$ .

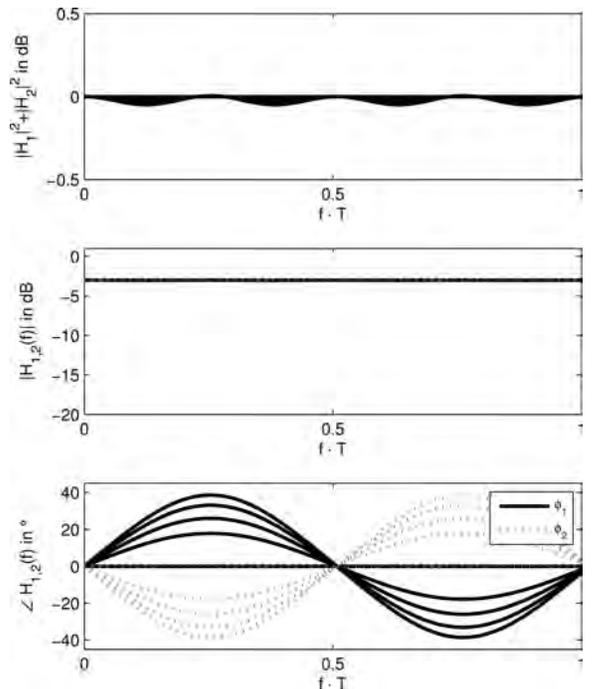


Fig. 2. The nearly constant power  $|H_1|^2 + |H_2|^2$  and magnitudes  $|H_{1,2}|$  in dB are shown, and the sinusoidally varying phase responses  $\angle H_{1,2}$  of the parametric phase-based widening filters of Eqs. (11) and (12) for all  $\hat{\phi}$  settings in Table 1. The frequency period  $1/T$  is adjusted by  $T = N/f_s$ .

is considered to be sufficiently small (KARJALAINEN *et al.*, 1999). The frequency period of the sinusoidal phase is adjustable by the time-delay as  $1/T$ .

#### 4. Amplitude-based: efficient cosine of raised cosine pair

A simple way to obtain a power-complementary zero-phase filter for widening would be to apply the square root to the raised cosine filter Eq. (1), see appended Eq. (24). A similar power-complementary response pair that is easier to approximate is obtained by insertion of  $\tilde{\varphi} = \hat{\phi} \cos(\omega T)$  into Eq. (6), which yields

$$H_{1,2}(\omega) = \cos\left[\pi/4 \pm \hat{\phi} \cos(\omega T)\right]. \quad (13)$$

This response pair is related to the phase-based one of Eq. (10) as formulated in the appended Eqs. (25) and (26), and it is therefore surprisingly simple. As Eq. (9), it also involves the Bessel functions as coefficients

$$\sqrt{2} H_{1,2}(z) = \sum_{l=0}^{\infty} J_l(\hat{\phi}) \sqrt{2} \cos\left(\frac{\pi}{4} \pm \frac{\pi}{2} l\right) \cdot [z^{lN} + z^{-lN}(1 - \delta_{l,0})], \quad (14)$$

but now two sign alteration patterns

$\langle +, -, -, +, +, -, -, \dots \rangle$  and  $\langle +, +, -, -, +, +, \dots \rangle$

expressed by  $\sqrt{2} \cos\left(\frac{\pi}{4} \pm \frac{\pi}{2} l\right)$  equally determine the signs of the symmetric delays  $z^{\pm lN}$ . Because this is the only difference, the infinite sum is truncated as in the section above, and the coefficients are approximated with Eq. (12),

$$\sqrt{2} H_{1,2}(z) = g_0 \mp g_1 [z^N + z^{-N}] - g_2 [z^{2N} + z^{-2N}]. \quad (15)$$

Note that also the filter structure is exactly the same as for the phase-based method, and it even creates the same decorrelation, see appended Eq. (28).

Figure 3 shows a causal signal flow graph for the new amplitude-based method of phantom source widening applied on a mono signal. The frequency responses are depicted in Fig. 4 and exhibit the deviation from a power-complementary response by less than 0.1 dB, as before. There is a curve for each  $\hat{\phi}$  value of Table 1.

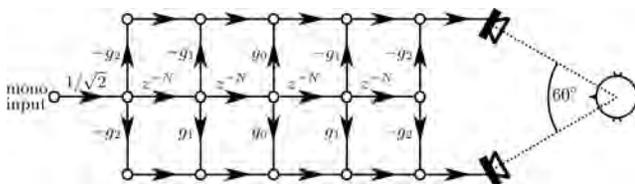


Fig. 3. Signal flow graph of an efficient and causal amplitude-based phantom source widener. Suitable weights  $g_l$  are specified in Eq. (12) and  $N = T f_s$ .

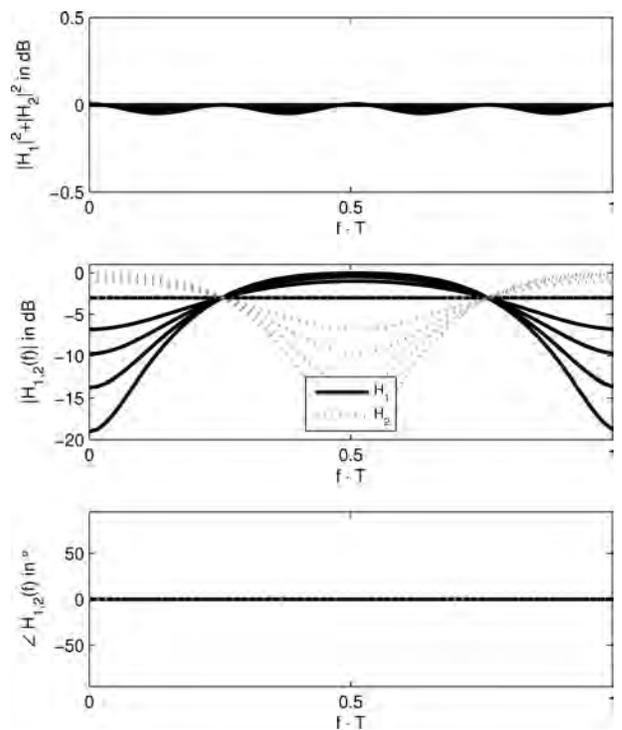


Fig. 4. The nearly constant power  $|H_1|^2 + |H_2|^2$ , the cosine of raised cosine magnitude responses  $|H_{1,2}|$  in dB, and the constant phase responses  $\angle H_{1,2}$  of the parametric amplitude-based widening filters of Fig. 3 for all parameters  $\hat{\phi}$  in Table 1. The frequency period  $1/T$  is adjusted by  $T = N/f_s$ .

#### 5. Evaluation at the central listening position

This section discusses the effect of the presented efficient widening algorithms in terms of known experimental data. These data are well predicted by the IACC, which is affected by the ICC of the two playback channels. Both technical measures are introduced and applied to compare both widening effects.

##### 5.1. Experimental results from a previous listening test

The phase-based approach was tested with a listening experiment in a preceding article (ZOTTER *et al.*, 2011). The analysis presented here uses experimental data thereof that was obtained for pink noise processed with the phase-based widening structure using the parameters<sup>1</sup> given in Table 2.

<sup>1</sup>In the print version of ZOTTER *et al.* (2011) the parameters  $\hat{\phi}$  and  $T$  were expressed equivalently as  $\hat{\tau} = \hat{\phi} T$  and  $\Delta f = 1/T$ . However, we listed erroneous values of  $\hat{\tau}$  in contrast to those actually employed in the experiment. Retrieval of  $T$  and  $\hat{\phi}$  from the original stimuli pairs  $X_{1,2}$ , i.e. the phase of  $\frac{X_1(\omega) X_2^*(\omega)}{\|X_1(\omega)\| \|X_2(\omega)\|}$ , clearly revealed that exactly half the stated value was employed. The erratum was forwarded to the DAFx-11 editors and the on-line publication includes a dated erratum remark.

Table 2. Listening test conditions of the phase-based approach that was evaluated in a former article (ZOTTER *et al.*, 2011). For the present article only the conditions C2, C4, C6, C7, and C8 are interesting. The  $(\hat{\tau}, \Delta f)$  parameters from the previous nomenclature were re-expressed using  $\hat{\phi} = \hat{\tau}/\Delta f$  (mind errata in (ZOTTER *et al.*, 2011)) and  $T = 1/\Delta f$ .

|    | $\hat{\phi}$ | T      | ICCC( $\hat{\phi}$ ) | IACC <sub>E3</sub> |
|----|--------------|--------|----------------------|--------------------|
| C2 | 0.0          | –      | 1.00                 | 0.824              |
| C4 | 0.3          | 5.0 ms | 0.91                 | 0.690              |
| C6 | 0.6          | 2.5 ms | 0.67                 | 0.554              |
| C7 | 0.6          | 5.0 ms | 0.67                 | 0.521              |
| C8 | 0.9          | 1.7 ms | 0.34                 | 0.468              |

**Experimental setup.** The experiment was set up in a 11 m×11 m×5 m room with a mean reverberation time of  $RT_{60} = 470$  ms, using 2 Genelec 8020 loudspeakers at  $\pm 30^\circ$ . The listening distance was 2 m, which is within the effective critical distance of the room. Originally, 8 conditions were compared in a full pairwise comparison that was performed twice within each stimulus set: one generated from 5 s pink noise, the other from 22 s male speech. Participants were familiarized with the noise stimuli that were leveled to 65 dB(A), took about 15 min to finish their 56 comparisons tasks (different random order for each participant) with noise, took an intermission, and then the entire procedure was repeated from the start but with speech stimuli. The subjects responded by either selecting the sound in each pair that was perceived as wider or stating that both were equal, using three marked keys of a keyboard on their lap. Seamless switching within the comparison pair while listening was possible at any time. All 11 participants were members of a trained listening panel and had participated in an experiment about source width before (FRANK *et al.*, 2011). Below, only results for the noise stimuli using the phase-based FIR algorithm with settings of  $T = \{1.7, 2.5, 5\}$  ms were used, as those for speech are similar anyway. The condition C5 with  $T = 0.8$  ms was excluded as it caused other effects as well, C1 (mono) and C3 (IIR widening) are excluded as they do not match the topic of the present article.

**Results.** The individual repetitions were averaged, and the Thurstone scales (THURSTONE, 1994) were recalculated based on the full pairwise comparison response matrices for the selected conditions of Table 2. The scales were calculated for each participant’s individual responses to estimate inter-subjective confidence intervals, and one overall scale was built upon a response matrix pooled from all participants as a consistency check. Figure 5 shows the significant effect of  $\hat{\phi}$  on the perceived source width scale. Conditions C6 and C7 only differ in the amount of delay T and yield significantly equal results.

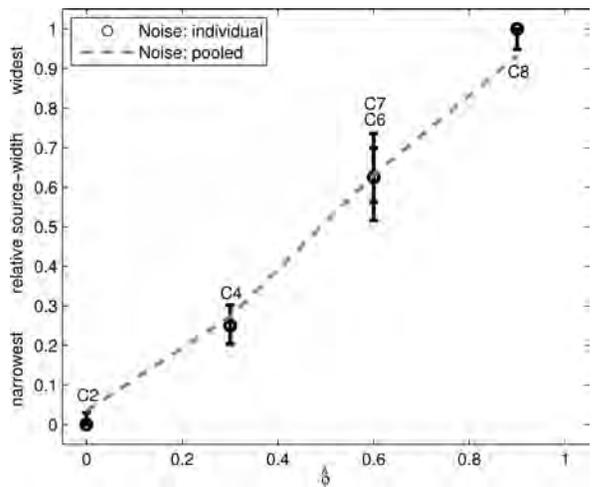


Fig. 5. Thurstone scales for noise conditions C2, C4, C6, C7, and C8 (ZOTTER *et al.*, 2011): median and inter-subjective 95% confidence interval using participants’ individual scales, compared to the over-all scale using answers pooled from all participants. The scales are plotted as a function of  $\hat{\phi}$ .

## 5.2. Technical prediction

**Inter-Aural Cross-Correlation Coefficient (IACC).** Originally conceived to describe the apparent source width in room acoustics, the IACC also suitably describes the width change of the phantom source after the application of widening. It is defined as the maximum of the inter-aural cross-correlation function (IACF), cf. ISO (2009),

$$\text{IACF}(\tau) = \frac{\int_{t_1}^{t_2} x_{\text{left}}(t)x_{\text{right}}(t+\tau) dt}{\sqrt{\left[ \int_{t_1}^{t_2} x_{\text{left}}^2(t) dt \right] \left[ \int_{t_1}^{t_2} x_{\text{right}}^2(t) dt \right]}}, \quad (16)$$

$$\text{IACC} = \max_{|\tau| \leq 1 \text{ ms}} |\text{IACF}(\tau)|, \quad (17)$$

which is calculated from the binaural signal pair  $x_{\text{left}}(t), x_{\text{right}}(t)$ . In our evaluation, the signal pair was replaced by the binaural impulse response pair due to an impulse at the system input, and the early part ( $t_1 = 0$  ms and  $t_2 = 80$  ms) is of particular interest, averaged over the octave bands 500 Hz, 1 kHz, and 2 kHz as defined in HIDAKA *et al.* (1995), denoted as IACC<sub>E3</sub>. The IACC<sub>E3</sub> exhibits a 97...98% statistical correlation to the perceived source width across all conditions of the experiment (speech, noise, other filtering structure). Therefore, we assume that the phantom source width can be estimated by measuring the IACC<sub>E3</sub>.

In order to document the ear signals of the experiment, they were recorded using a B&K 4128C Head-and-Torso simulator at the listening position, and binaural loudspeaker impulse responses were measured.

These,  $h_{1,\text{left}}(t)$ ,  $h_{2,\text{left}}(t)$ ,  $h_{1,\text{right}}(t)$ ,  $h_{2,\text{right}}(t)$ , can be used to calculate ear signals due to any pair of loudspeaker signals  $s_{1,2}(t)$  by convolution

$$x_{\text{left}}(t) = h_{1,\text{left}}(t) \star s_1(t) + h_{2,\text{left}}(t) \star s_2(t),$$

and

$$x_{\text{right}}(t) = h_{1,\text{right}}(t) \star s_1(t) + h_{2,\text{right}}(t) \star s_2(t).$$

### Inter-Channel Cross-Correlation Coefficient (ICCC).

The phantom source widening algorithms above directly control the correlation between the loudspeaker signals  $s_{1,2}(t)$ . The ICCC is the maximum of the inter-channel cross-correlation function (ICCF)

$$\text{ICCF}(\tau) = \frac{\int_{-\infty}^{\infty} s_1(t)s_2(t+\tau) dt}{\sqrt{\left[ \int_{-\infty}^{\infty} s_1^2(t) dt \right] \left[ \int_{-\infty}^{\infty} s_2^2(t) dt \right]}}, \quad (18)$$

$$\text{ICCC} = \max_{|\tau| \leq P} |\text{ICCF}(\tau)|. \quad (19)$$

To obtain a generalized estimation, the signal pair  $s_{1,2}(t)$  is replaced by the impulse response pair  $h_{1,2}(t)$  of phantom source widening. In the tested cases  $0 \leq \hat{\phi} < \pi/2$ , this yields for the period  $P > 0$  according to the appended Eq. (28)

$$\text{ICCC} = J_0(2\hat{\phi}), \quad (20)$$

which approximately equals  $\cos(\hat{\phi}\sqrt{2})$ . Table 2 shows the ICCC of Eq. (20) for the test conditions: The bigger the modulation depth parameter  $\hat{\phi}$  of the phantom source widener the smaller the ICCC.

**Relation between IACC and ICCC at the central listening position.** Judging from the values in Table 2, not only the IACC but also the ICCC seems to be an excellent predictor for the perceived width. Nevertheless, the ICCC only depends on the algorithmic parameter  $\hat{\phi}$ . Therefore, it is unable to estimate changes under more general circumstances, under which room and setup will have a natural impact on the apparent source width. However, in the acoustic situation of the experiment, the ICCC indirectly maps to the  $\text{IACC}_{\text{E3}}$  in a monotonic curve, see solid curve in Fig. 6. Linear regression (dashed curve) yields

$$\text{IACC}_{\text{E3}} \approx 0.77 \cdot \text{ICCC} + 0.03. \quad (21)$$

The relationship given in Fig. 6 is still valid if, instead of the phantom-source widening algorithms, variably correlated white noise signals are convolved with the binaural loudspeaker responses (dash-dot curve). This was also verified for another room (Sec. 6, Fig. 7a) with dominant direct sound (dotted curve). Noise signals of

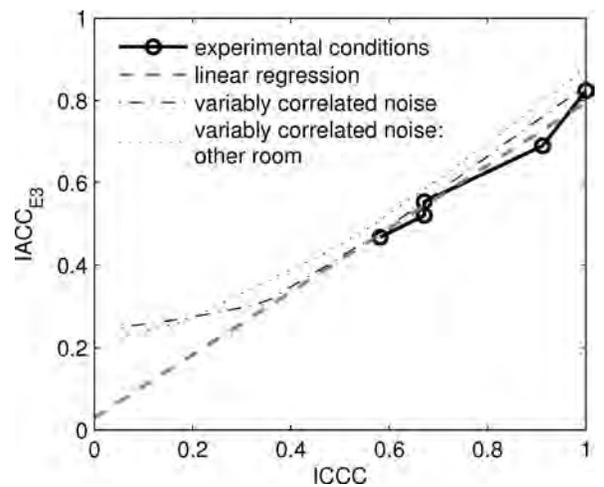


Fig. 6. The inter-aural cross-correlation coefficient (IACC) monotonically depends on the inter-channel cross-correlation coefficient (ICCC) in the experiment (ZOTTER *et al.*, 2011). The black curve is the relation for phase-based phantom source widening, the dash-dot curve for a variably correlated, white stereo noise. The dashed line is a linear regression, and the dotted line expresses the same tendency for stereo noise in a different room.

adjustable ICCC and equal variance are created from uncorrelated noise signals  $a(t)$ ,  $b(t)$  of equal variance, see appendix, cf. BLAUERT, LINDEMANN (1986), by

$$\begin{aligned} s_1(t) &= a(t), \\ s_2(t) &= \text{ICCC} a(t) + \sqrt{1 - \text{ICCC}^2} b(t). \end{aligned} \quad (22)$$

**Phantom source width controllability.** At the central listening position, the  $\text{IACC}_{\text{E3}}$  directly increases with the ICCC. Their linear relationship has been shown before. According to measurements done with the setup described in the next section (Fig. 7a), this is true for both amplitude- and phase-based widening. The corresponding  $\text{IACC}_{\text{E3}}$  values were measured to differ by 0.02 at most (rightmost columns in Table 1). Therefore amplitude and phase modification is expected to achieve similar widening in listening experiments; an expectation that goes along with our listening experience. For controlling the phantom source width, we ideally wish to achieve at least a just noticeable difference (JND) in the  $\text{IACC}_{\text{E3}}$  when increasing the ICCC by 0.1. As the JND for the  $\text{IACC}_{\text{E3}}$  is roughly 0.05 (BLAU, 2002), this means that we desire a slope of at least  $\frac{\Delta \text{IACC}_{\text{E3}}}{\Delta \text{ICCC}} \geq 0.5$ .

## 6. Evaluation at off-center listening positions

For the purpose of further evaluation, a series of binaural measurements have been taken in a different room and at various off-center positions. In particular, the measurement data shall clarify whether the

amplitude-based widening method is more robust to lateral shifts.

**Measurement setup.** A pair of Genelec 8020 loudspeakers has been set up in a standard  $\pm 30^\circ$  stereo arrangement with 1.8 m distance to the central listening position in a  $3.7\text{ m} \times 3.7\text{ m} \times 3.2\text{ m}$  room with a reverberation time of  $RT_{30} = 50\text{ ms}$ . The binaural impulse responses of the loudspeakers have been taken with the B&K dummy head. The responses of the filters in Figs. 1 and 3 convolved with binaural responses allow to predict the binaural signals for several settings of the algorithms. The series of binaural impulse responses included shifts of the dummy head to lateral positions by a shift  $d$  ranging from  $-90\text{ cm}$  to  $90\text{ cm}$  in  $10\text{ cm}$  steps, e.g., for  $d = 0\text{ cm}$  in Fig. 7a and  $d = 90\text{ cm}$  in Fig. 7b.

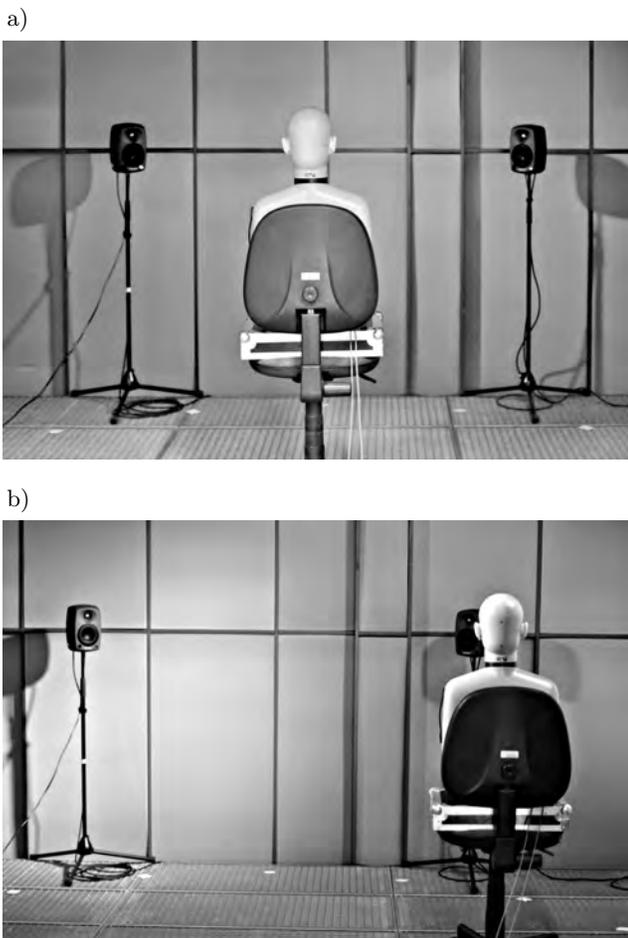


Fig. 7. The experimental setup for evaluating the effect of phantom source widening on IACC employs a dummy head in an acoustically damped room. The head is also employed at side-wise shifted locations  $d = \{-90, -80, \dots, +80, +90\}$  cm to gather information about off-center listening locations: a) central position, b) 90 cm shifted position.

**Results for IACC at different shifts and ICCCs.** Figure 8 shows which  $IACC_{E3}$  values are accessible at different shifts, playing back correlated and correlated

noise, i.e.  $ICCC = \{0, 1\}$ . This implies that any attempt to widen the phantom source could be ineffective at large shifts  $d \geq 40\text{ cm}$ .

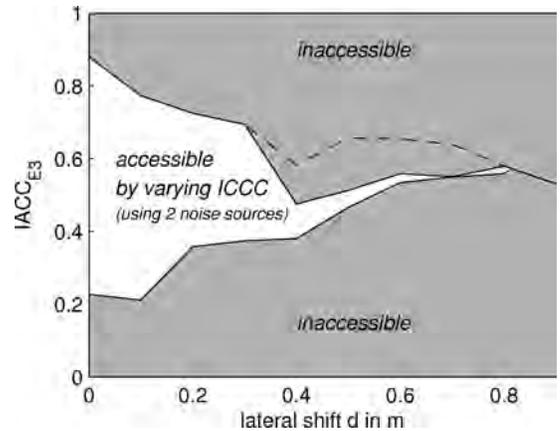


Fig. 8.  $IACC_{E3}$  values that can be created at different lateral shifts  $0\text{ cm} \leq d \leq 90\text{ cm}$  by controlling the ICCC between 0 and 1. (The dashed curve is for a lag within  $|\tau| \leq 2\text{ ms}$  or more in Eq. (17).)

Figure 9 shows the particular dependency of the  $IACC_{E3}$  on the ICCC for all measured positions and both algorithms using a time-delay of  $T = 5\text{ ms}$ . Table 1 summarizes the settings of both phantom source wideners to produce an ICCC of values between 0.6

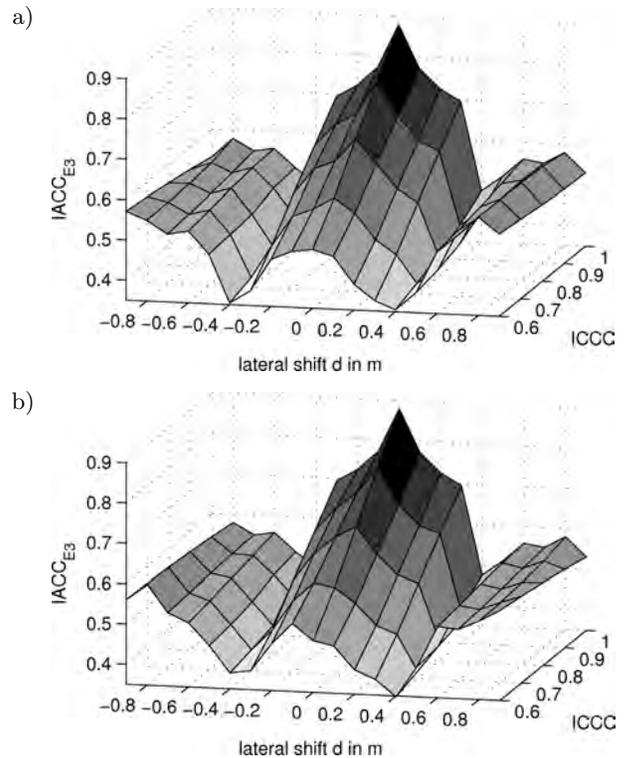


Fig. 9.  $IACC_{E3}$  measured for phantom source widening using lateral listening position shifts  $d = \{-90, -80, \dots, +80, +90\}$  cm. The ICCC is variable from 0.6 to 1 in 0.1 steps and should control the  $IACC_{E3}$ : a) phase-based,  $T = 5\text{ ms}$ ; b) amplitude-based,  $T = 5\text{ ms}$ .

and 1, in steps of 0.1. The ICCC value is independent of the particular time-delay value  $T$ . For the central listening position  $d = 0$  cm, the diagram shows the linear controllability already observed above. The desired slope of  $\frac{\Delta \text{IACC}_{E3}}{\Delta \text{ICCC}} \geq 0.5$  for controllability is achieved for all  $|d| < 40$  cm. Thus the phantom source widening is expected to work within this range of listening positions. This is true for all investigated settings and also various time-delays  $T = \{1.6, 2.5, 5\}$  ms of both filter structures. However, for  $|d| \geq 40$  cm the  $\text{IACC}_{E3}$  is obviously not controlled by the ICCC; for both algorithms. As the  $\text{IACC}_{E3}$  is related to the perceived source width, the perceived width at the farthest lateral listening positions seems to be indifferent to ICCC, as in Fig. 8. Nevertheless, different ICCCs seem to produce other perceivable differences according to informal listening experience.

**Coloration: Maximum third-octave level difference.** In order to observe whether the algorithms introduce perceivable spectral coloration when decreasing the ICCC, i.e. decorrelation, binaural third-octave levels are also measured. Levels are determined after summation of the third-octave powers across the both ear signals (ONO *et al.*, 2001), which was done for each band from 200 Hz to 12.5 kHz, here. The differences of the resulting third-octave levels to those of the unaltered phantom source should stay small. According to KARJALAINEN *et al.* (1999), a third-octave level deviation of 1 dB is audible in loudspeaker equalization.

The maximum of the third-octave level deviations is plotted in Figs. 10a to 10f for varying ICCC

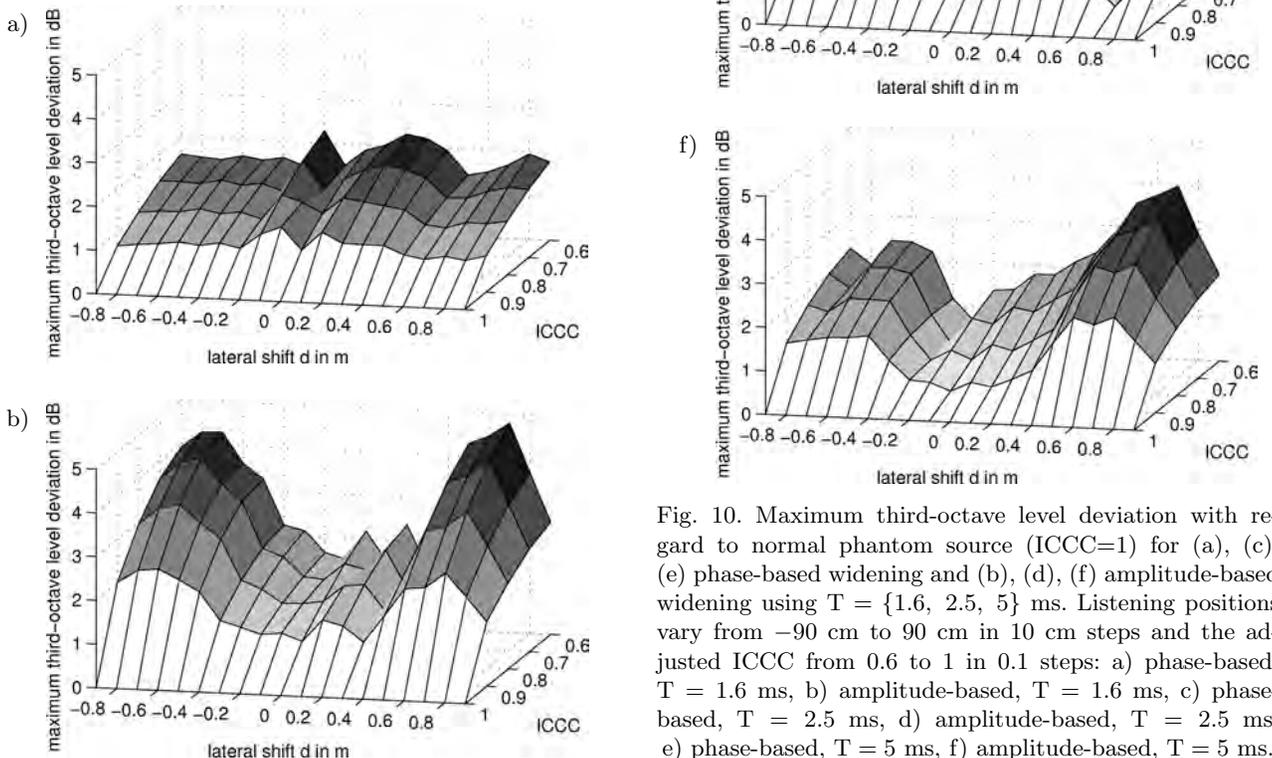


Fig. 10. Maximum third-octave level deviation with regard to normal phantom source (ICCC=1) for (a), (c), (e) phase-based widening and (b), (d), (f) amplitude-based widening using  $T = \{1.6, 2.5, 5\}$  ms. Listening positions vary from  $-90$  cm to  $90$  cm in  $10$  cm steps and the adjusted ICCC from  $0.6$  to  $1$  in  $0.1$  steps: a) phase-based,  $T = 1.6$  ms, b) amplitude-based,  $T = 1.6$  ms, c) phase-based,  $T = 2.5$  ms, d) amplitude-based,  $T = 2.5$  ms, e) phase-based,  $T = 5$  ms, f) amplitude-based,  $T = 5$  ms.

and lateral shifts  $d$ , for both algorithms and  $T = \{1.6, 2.5, 5\}$  ms. Clearly, the maximum deviation as a measure of sound coloration increases with the amount of decorrelation produced between the channels. Coloration also depends on the listening position. For small time-delays, the coloration of the phase-based approach is quite low, but there might be some phasiness perceived. For larger time-delay settings, both approaches become similar.

## 7. Conclusion

We presented a highly efficient filter structure for phantom source widening based on a frequency-dependent modification of either phase or amplitude. Both ways of controlling the structure are power-complementary at a high accuracy, yield the same inter-channel cross-correlation coefficient (ICCC), and use the same number of operations (4 delays of equal size, 8 additions, 10 multiplications depending on 3 variable weights). The algorithmic parameter  $\hat{\phi}$  allows to adjust the correlation of a mono signal used in stereo to a desired degree, with the relation  $\text{ICCC} \approx \cos(\hat{\phi}\sqrt{2})$ . Thinkable target applications are, e.g., stereo effect processors in audio mixing, the research of spatial hearing, and auditory interfaces in which the salience of sounds shall be controlled by spatial sharpness or width.

We took dummy head measurement at the central listening position to show the linear proportionality of the inter-channel coherence to the inter-aural coherence measured by the IACC, whose relation to the perceived phantom source width is known from former experimentation. Moreover, we could show that the controllability of the phantom source width has to do with the listening position. It strongly influences the slope in the dependency of the inter-aural on the inter-channel coherence.

By measurements at laterally shifted positions, we showed the ability of both amplitude and phase decorrelation to control the phantom source width for displaced listening positions within  $|d| \leq 30$  cm. This is largely independent of the size of the algorithmic time-delay within the interval of 1.6...5 ms. Neither of the algorithms controls the phantom source width outside the range of lateral shifts, according to the inter-aural coherence.

Both amplitude and phase decorrelation seem to cause sound coloration that increases with the parameter  $\hat{\phi}$ . The corresponding third-octave level differences to an unaltered phantom source are maximally 5 dB in all bands between 200 Hz and 12.5 kHz. This dependency is observed for all measured listening positions albeit the coloration of both algorithms behaves differently in its detail. For small time-delays the phase-based method seems to be better.

## Appendix

**Sinusoidal-phase all-pass.** The series expansion of the cylindrical Bessel function inserted into Eq. (9) yields

$$\begin{aligned} H_1(z) &= \frac{1}{\sqrt{2}} \sum_{l=0}^{\infty} J_l(\hat{\phi}) [z^{lN} + (-1)^l z^{-lN} (1 - \delta_{l,0})] \\ &= \frac{1}{\sqrt{2}} \sum_{l=0}^{\infty} \sum_{n=|l|}^{\infty} \hat{\phi}^n \frac{(-1)^{\frac{n}{2}-\frac{l}{2}} \delta_{n \bmod 2, l \bmod 2}}{\left(\frac{n}{2} - \frac{l}{2}\right)! \left(\frac{n}{2} + \frac{l}{2}\right)! 2^n} \\ &\quad \cdot [z^{lN} + (-1)^l z^{-lN} (1 - \delta_{l,0})] \\ &= \frac{1}{\sqrt{2}} \sum_{l=0}^{\infty} g_l [z^{lN} + (-1)^l z^{-lN} (1 - \delta_{l,0})]. \end{aligned} \quad (23)$$

In the range  $0 \leq \hat{\phi} \leq \pi/4$ , a reasonable approximation with  $-0.1 \leq |H_{1,2}(\omega)| \leq 0.02$  dB is obtained for  $|l| \leq 2$  and  $n \leq 3$  with the weights in Eq. (12).

**Root raised cosine.** In order to achieve the square root response of the raised cosine filter pair in Eq. (2), the binomial formula

$$(x + y)^n = \sum_{k=0}^n \frac{n!}{k!(n-k)!} x^{n-k} y^k$$

is used to represent

$$\sqrt{1+x} = \sum_{n=0}^{\infty} \frac{(-1)^n (2n-1)!!}{(1-2n)n!2^n} x^n.$$

Consequently, the application of the square-root on the  $z$ -domain expression in Eq. (1) can be re-expressed after applying the binomial formula and re-arranging the sums

$$\begin{aligned} H_1(z) &= \sqrt{1 + \hat{\phi}(z^N + z^{-N})} / \sqrt{2} \\ &= \frac{1}{\sqrt{2}} \sum_{n=0}^{\infty} \frac{(-1)^n (2n-1)!!}{(1-2n)n!2^n} [\hat{\phi}(z^N + z^{-N})]^n \\ &= \frac{1}{\sqrt{2}} \sum_{n=0}^{\infty} \hat{\phi}^n \frac{(-1)^n (2n-1)!!}{(1-2n)n!2^n} \\ &\quad \cdot \sum_{k=0}^n \frac{n!}{k!(n-k)!} z^{(2k-n)N} \\ &= \frac{1}{\sqrt{2}} \sum_{n=0}^{\infty} \sum_{l=-n}^n \hat{\phi}^n \frac{(-1)^n (2n-1)!!}{(1-2n)2^n \left(\frac{n}{2} + \frac{l}{2}\right)! \left(\frac{n}{2} - \frac{l}{2}\right)!} \\ &\quad \cdot z^{lN} \delta_{n \bmod 2, l \bmod 2} \\ &= \frac{1}{\sqrt{2}} \sum_{l=-\infty}^{\infty} \left[ \sum_{n=|l|}^{\infty} \hat{\phi}^n \right. \\ &\quad \cdot \frac{(-1)^n (2n-1)!!}{(1-2n)2^n \left(\frac{n}{2} + \frac{l}{2}\right)! \left(\frac{n}{2} - \frac{l}{2}\right)!} \\ &\quad \left. \cdot \delta_{n \bmod 2, l \bmod 2} \right] z^{lN}. \end{aligned} \quad (24)$$

The Kronecker delta  $\delta_{n \bmod 2, l \bmod 2}$  was introduced in line three to ensure that the factorials stay integer as in the lines before. This is only possible if  $n$  and  $l$  are both exclusively either even or odd. A simple approximation is obtained by truncating the sum over  $l$  to  $|l| \leq 2$  and the sum over  $n$  to  $n \leq 2$  and the weights  $\left[1 - \frac{\hat{\phi}^2}{4}, \frac{\hat{\phi}}{2} - \frac{3\hat{\phi}^3}{16}, -\frac{\hat{\phi}^2}{8}\right]$ .

**Cosine of raised cosine pair.** The response

$$\cos\left[\pi/4 \pm \hat{\phi} \cos(\omega T)\right]$$

of Eq. (13) can be re-written as

$$H_{1,2}(\omega) = \left[ e^{i\pi/4} e^{i \sin(\omega T \pm \pi/2)} + e^{-i\pi/4} e^{-i \sin(\omega T \mp \pi/2)} \right] / 2, \quad (25)$$

and thus it is related to Eq. (10) by

$$H_{1,2}(\omega) = \left\{ e^{i\pi/4} \tilde{H}_1[\omega \pm \pi/(2T)] + e^{-i\pi/4} \tilde{H}_2[\omega \mp \pi/(2T)] \right\} / \sqrt{2}.$$

Applying the corresponding frequency shift of  $\pm\pi/(2N)$ , the  $\pi/4$  phase shift, and Euler's formula  $\cos \alpha = (e^{i\alpha} + e^{-i\alpha})/2$  on  $H(z) = \sum_{l=-\infty}^{\infty} J_l(\hat{\phi}) z^{-lN}$ , cf. Eq. (9), yields

$$\begin{aligned} H_{1,2}(z) &= \frac{1}{2} \sum_{l=-\infty}^{\infty} J_l(\hat{\phi}) z^{-lN} \left[ e^{i\pi/4} (\pm i)^l + e^{-i\pi/4} (\mp i)^l \right] \\ &= \sum_{l=-\infty}^{\infty} J_l(\hat{\phi}) z^{-lN} \cos\left(\frac{\pi}{4} \pm \frac{\pi}{2} l\right). \end{aligned} \quad (26)$$

Using  $J_{-l}(\hat{\phi}) = (-1)^l J_l(\hat{\phi})$  and  $(-1)^l \cos\left(\frac{\pi}{4} \mp \frac{\pi}{2} l\right) = \cos\left(\frac{\pi}{4} \pm \frac{\pi}{2} l\right)$ , we find that the involved cosine expresses a factor  $1/\sqrt{2}$  and the alternating sign sequence  $\langle 1, \mp 1, -1, \pm 1 \rangle$  for  $\pm l = 0, 1, 2, 3$ . The response pair is therefore symmetrical with regard to time, which is consistent with Eq. (13) being zero-phase. Simplified by multiplication with  $\sqrt{2}$ , we get

$$\begin{aligned} \sqrt{2} H_{1,2}(z) &= J_0(\hat{\phi}) \mp J_1(\hat{\phi}) [z^{-N} + z^N] \\ &\quad - J_2(\hat{\phi}) [z^{-2N} + z^{2N}] \\ &\quad \pm J_3(\hat{\phi}) [z^{-3N} + z^{3N}] \\ &\quad + J_4(\hat{\phi}) [z^{-4N} + z^{4N}] \mp \dots \end{aligned} \quad (27)$$

Obviously, the same approximated Bessel functions can be used as weights as suggested two paragraphs above, but now with different signs. Due to the same mathematical origin, the hereby obtained filter pair yields the same accuracy of power-complementarity as suggested two paragraphs above.

**The ICCC of both structures** is obtained from the normalized frequency domain cross-correlation. The inter-channel cross-correlation function (ICCF) in the time domain becomes at a variable lag  $\tau = nT$ ,  $n \in \mathbb{Z}$ ,

$$\begin{aligned} \text{ICCF}[nT] &= \frac{\int_0^{2\pi} H_1(\omega) H_2^*(\omega) e^{i n \omega T} d(\omega T)}{\sqrt{\left[ \int_0^{2\pi} |H_1(\omega)|^2 d(\omega T) \right] \left[ \int_0^{2\pi} |H_2(\omega)|^2 d(\omega T) \right]}} \\ &= \frac{1}{2\pi} \int_0^{2\pi} e^{i [2\hat{\phi} \cos(\omega T) + n\omega T]} d(\omega T) \\ &= \frac{1}{2\pi} \int_0^{2\pi} \sin\left[n\omega T + \frac{\pi}{2} + 2\hat{\phi} \cos(\omega T)\right] d(\omega T) \\ &= J_n(2\hat{\phi}). \end{aligned} \quad (28)$$

Thereof, the ICCC is the maximal value within some time period  $P$ , i.e.  $\text{ICCC} = \max_{|nT| \leq P} J_n(2\hat{\phi})$ , which yields just  $J_0(2\hat{\phi})$  if  $2\hat{\phi} < \pi/2$  or  $P$  is short.

**Variably correlated noise pair.** Assume two zero-mean random signals  $a$  and  $b$ ,  $E\{a\} = E\{b\} = 0$ , that are uncorrelated  $E\{ab\} = 0$  and whose variance equals  $\sigma^2 = E\{a^2\} = E\{b^2\}$ ;  $E\{\cdot\}$  denotes the expected value. The cross-correlation of  $s_1 = a$  with  $s_2 = \gamma a + \sqrt{1-\gamma^2} b$  from Eq. (22) is easily shown to equal  $\gamma$

$$\begin{aligned} \text{ICCC} &= \frac{E\{s_1 s_2\}}{\sqrt{E\{s_1^2\} E\{s_2^2\}}} \\ &= \frac{\gamma \sigma^2 + \sqrt{1-\gamma^2} E\{ab\}}{\sqrt{\sigma^2 [\sigma^2 \cancel{\gamma^2} + \sigma^2 (1-\cancel{\gamma^2}) + 2\gamma \sqrt{1-\cancel{\gamma^2}} E\{ab\]}} \\ &= \gamma. \end{aligned} \quad (29)$$

## Acknowledgments

We gratefully thank Alois Sontacchi and Georgios Marentakis for their contributions in our former paper on which this article is based. Our research is partly supported by the project AAP, which is funded by Austrian ministries BMVIT, BMWFJ, the Styrian Business Promotion Agency (SFG), and the departments 3 and 14 of the Styrian Government. The Austrian Research Promotion Agency (FFG) conducted the funding under the Competence Centers for Excellent Technologies (COMET, K-Project), a program of the above-mentioned institutions.

A part of this work was presented at the 59th Open Seminar on Acoustics (Poland, Boszkowo, September 10–14, 2012).

## References

1. BLAU M. (2002), *Difference limens for measures of apparent source width*, [in:] *Forum Acusticum*, Sevilla, Spain.
2. BLAUERT J., LINDEMANN W. (1986), *Spatial mapping of intracranial auditory events for various degrees of interaural coherence*, *The Journal of the Acoustical Society of America*, **79**, 806–813.
3. BOUÉRI M., KYRIAKAKIS C. (2004), *Audio signal decorrelation based on a critical band approach*, [in:] *Preprint 6291, 117th Conv. Audio Eng. Soc.*, San Francisco.
4. FRANK M., MARENTAKIS G., AND SONTACCHI A. (2011), *A simple technical measure for the perceived source width*, [in:] *Fortschritte der Akustik, DAGA*, Düsseldorf.
5. GERZON M. A. (1992), *Signal processing for simulating realistic stereo images*, [in:] *Preprint 3423, 93rd Conv. Audio Eng. Soc.*, San Francisco.
6. GERZON M. A. (1993), *Stereophonic signal processor generating pseudo stereo signals*, Patent, WO 93/25055.
7. HIDAKA T., BERANEK L. L., OKANO T. (1995), *Interaural cross-correlation, lateral fraction, and low- and high-frequency sound levels as measures of acoustical quality in concert halls*, *The Journal of the Acoustical Society of America*, **98**, 2, 988–1007.
8. ISO (2009), *ISO 3382-1:2009: Acoustics – measurement of room acoustic parameters – part 1: Performance spaces*.
9. KARJALAINEN M., PIIRILÄ E., JÄRVINEN A., HUOPANIEMI J. (1999), *Comparison of loudspeaker equalization methods based on DSP techniques*, *J. Audio Eng. Soc.*, **47**, 1/2, 14–31.
10. KENDALL G. S. (1995), *The decorrelation of audio signals and its impact on spatial imagery*, *Computer Music J.*, **19**, 4, 71–87.
11. KIN M. J., PLASKOTA P. (2011), *Comparison of sound attributes of multichannel and mixed-down stereo recordings*, *Archives of Acoustics*, **36**, 2, 333–345.
12. LAITINEN M.-V., PHILAJAMÄKI T., ERKUT C., PULKKI V. (2012), *Parametric time-frequency representation of spatial sound in virtual worlds*, *ACM Trans. Appl. Percept.*, **3**, 2, 8:1–20.
13. ONO K., PULKKI V., KARJALAINEN M. (2001), *Binaural modeling of multiple sound source perception: Methodology and coloration experiments*, [in:] *Preprint 5446, 111th Conv. Audio Eng. Soc.*, New York.
14. OPPENHEIM A. V., SCHAFER R. W., BUCK J. R. (1999), *Discrete-Time Signal Processing*, Prentice Hall, New Jersey.
15. ORBAN R. (1970a), *A rational technique for synthesizing pseudo-stereo from monophonic sources*, *J. Audio Eng. Soc.*, **18**, 2, 157–164.
16. ORBAN R. (1970b), *Letters to the editor: Further thoughts on “a rational technique for synthesizing pseudo-stereo from monophonic sources”*, *J. Audio Eng. Soc.*, **18**, 4, 443–444.
17. PLEWA M., KLECZKOWSKI P. (2011), *Choosing and configuring a stereo microphone technique based on localisation curves*, *Archives of Acoustics*, **36**, 2, 347–363.
18. POTARD G., BURNETT I. (2004), *Decorrelation techniques for the rendering of apparent sound source width in 3D audio displays*, [in:] *Proc. DAFx-04*, 280–284, Napoli.
19. POTARD G. (2006), *3D-audio object oriented coding*, PhD thesis, University of Wollongong.
20. PULKKI V. (1997), *Virtual sound source positioning using vector base amplitude panning*, *J. Audio Eng. Soc.*, **45**, 6, 456–466.
21. SCHRÖDER M. R. (1958), *An artificial stereophonic effect obtained from a single audio signal*, *J. Audio Eng. Soc.*, **6**, 2, 74–79.
22. SZCZERBA M., OOMEN W., THERSSEN D. (2011), *Parametric audio based decoder and music synthesizer for mobile applications*, *Archives of Acoustics*, **36**, 2, 461–478.
23. THURSTONE L. L. (1994), *A law of comparative judgment*, *Psychological Review*, **101**, 2, 266–270.
24. WENDT K. (1963), *Das Richtungshören bei der Überlagerung zweier Schallfelder bei Intensitäts- und Laufzeitstereophonie*, PhD thesis, RWTH Aachen.
25. ZOTTER F., FRANK M., MARENTAKIS G., SONTACCHI A. (2011), *Phantom source widening with deterministic frequency dependent time-delays*, [in:] *DAFx-11 Proceedings*, 307–312, Paris.

# Spectral Mapping Using Kernel Principal Components Regression for Voice Conversion

Peng SONG<sup>(1)</sup>, Li ZHAO<sup>(1)</sup>, Yongqiang BAO<sup>(2)</sup>

<sup>(1)</sup> *Key Laboratory of Underwater Acoustic Signal Processing of Ministry of Education  
Southeast University  
Nanjing, 210096, P.R. China; e-mail: pengsongseu@gmail.com*

<sup>(2)</sup> *School of Communication Engineering, Nanjing Institute of Technology  
Nanjing 211167, P.R. China*

*(received July 2, 2012; accepted December 7, 2012)*

The Gaussian mixture model (GMM) method is popular and efficient for voice conversion (VC), but it is often subject to overfitting. In this paper, the principal component regression (PCR) method is adopted for the spectral mapping between source speech and target speech, and the numbers of principal components are adjusted properly to prevent the overfitting. Then, in order to better model the nonlinear relationships between the source speech and target speech, the kernel principal component regression (KPCR) method is also proposed. Moreover, a KPCR combined with GMM method is further proposed to improve the accuracy of conversion. In addition, the discontinuity and oversmoothing problems of the traditional GMM method are also addressed. On the one hand, in order to solve the discontinuity problem, the adaptive median filter is adopted to smooth the posterior probabilities. On the other hand, the two mixture components with higher posterior probabilities for each frame are chosen for VC to reduce the oversmoothing problem. Finally, the objective and subjective experiments are carried out, and the results demonstrate that the proposed approach shows greatly better performance than the GMM method. In the objective tests, the proposed method shows lower cepstral distances and higher identification rates than the GMM method. While in the subjective tests, the proposed method obtains higher scores of preference and perceptual quality.

**Keywords:** spectral mapping, overfitting, oversmoothing, discontinuity, kernel principal component regression.

## 1. Introduction

In text-to-speech (TTS) synthesis system, the personalized speech generation is one of the biggest challenges. Voice conversion (VC) can be seen as an efficient technique to solve this problem, which refers to modifying the speech spoken by a source speaker to be perceived as that spoken by a target speaker. It also has many other realistic applications, such as identity disguise in secure communications, single channel speech enhancement, and speech-to-speech translation system, etc.

In the last two decades, many spectral mapping approaches have been proposed for VC. The vector quantization (VQ) method is first introduced for spectral transformation (ABE *et al.*, 1988), it can efficiently map the space of the source speaker to that of the

target speaker. However, the transformation is performed in discrete spaces, it will lead to poor perceptual speech quality. The linear multivariate regression method is also proposed to improve the VC performance (VALBRET *et al.*, 1992), although it can obtain better performance than VQ method, it still can not achieve satisfactory results. The Gaussian mixture model (GMM) method is further proposed for VC (STYLIANOU *et al.*, 1998; KAIN, MACON, 1998), the results show that it can significantly outperform the prior methods. There are also some other methods, such as the artificial neural network (ANN) method (DESAI *et al.*, 2010) and the support vector regression (SVR) method (SONG *et al.*, 2011). The experimental results show that compared to the GMM method, these methods can get comparable or even better performance.

From the state-of-the-art references, the GMM method is proven to be most prevalent and efficient, and is chosen as the baseline method in the paper. However, the GMM method has some main shortcomings. Firstly, if the spectral features are sparse and the model is complex, the GMM always tends to overfitting, it has been proven that the GMM with a full covariance matrix is difficult to estimate and will lead to overfitting (MESBAHI *et al.*, 2007), and the partial least squares regression (PLSR) approach (HELANDER *et al.*, 2010) is proposed to avoid this problem efficiently. Secondly, it inevitably introduces the oversmoothing problem, which is caused by the statistical averaging of the model. A hybrid GMM and maximum a posterior (MAP) method (CHEN *et al.*, 2003) has been proposed to solve this problem, and the global variance (GV) is considered to reduce this issue (TODA *et al.*, 2001). Thirdly, the traditional GMM method deals with every frame independently regardless of the adjacent frames, this will lead to discontinuities in the converted spectral features and degrade the perceptual quality. Similar to the hidden Markov model (HMM) based speech synthesis, the relationships between the static and dynamic features are considered to obtain the optimal spectral trajectory (TODA *et al.*, 2005).

In this paper, we first propose a VC method based on the principal component regression (PCR). The numbers of principal components are adjusted to overcome the overfitting problem. In order to describe the nonlinear mapping between the source speech and target speech, the kernel PCR (KPCR) method is proposed, and the combined KPCR and GMM method is further presented to improve the VC performance. Then, in order to solve the discontinuity problem, an adaptive median filtering strategy, which is prevalent in image processing, is adopted to smooth the posterior probabilities. Meanwhile, only the two mixture components with higher posterior probabilities for each frame are chosen to reduce the oversmoothing problem. Finally, the objective and subjective experiments are carried out. Compared to the baseline GMM method, the proposed approach can efficiently avoid the overfitting problem when the number of training utterances is limited. Meanwhile, the converted speech using the proposed method shows better perceptual quality, and is much closer to the target one.

The paper is organized as follows, Sec. 2 describes the baseline GMM based VC method. Section 3 gives the PCR and KPCR based VC methods, respectively, and also presents the combined KPCR and GMM method. The novel post-processing approaches using posterior probability information are proposed to solve the discontinuity and oversmoothing problems in Sec. 4. The experimental results are given and discussed in Sec. 5. Finally, Sec. 6 draws the conclusions of this paper.

## 2. GMM based spectral mapping

In the GMM based VC method, the GMM is employed to model the augmented source and target spectral features. Let  $\mathbf{x}$  and  $\mathbf{y}$  denote the spectral sequences of source and target speakers, respectively, where  $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$  and  $\mathbf{y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N]$  aligned by the dynamic feature warping (DTW) algorithm. The distribution of the augmented features  $\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix}$  can be seen as a mixture of  $M$  Gaussian components, and the probability density function can be written as

$$p\begin{pmatrix} \mathbf{x}_n \\ \mathbf{y}_n \end{pmatrix} = \sum_{m=1}^M \alpha_m N\left[\begin{pmatrix} \mathbf{x}_n \\ \mathbf{y}_n \end{pmatrix}, \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m\right], \quad (1)$$

$$\sum_{m=1}^M \alpha_m = 1,$$

where  $N\left[\begin{pmatrix} \mathbf{x}_n \\ \mathbf{y}_n \end{pmatrix}, \boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m\right]$  denotes a Gaussian distribution,  $\alpha_m$  is a prior probability of each frame belonging to the  $m$ -th component, and  $\boldsymbol{\mu}_m$  and  $\boldsymbol{\Sigma}_m$  are the mean vector and covariance matrix of the  $m$ -th Gaussian component, respectively, which can be represented as

$$\boldsymbol{\mu}_m = \begin{bmatrix} \boldsymbol{\mu}_m^x \\ \boldsymbol{\mu}_m^y \end{bmatrix}, \quad \boldsymbol{\Sigma}_m = \begin{bmatrix} \boldsymbol{\Sigma}_m^{xx} & \boldsymbol{\Sigma}_m^{xy} \\ \boldsymbol{\Sigma}_m^{yx} & \boldsymbol{\Sigma}_m^{yy} \end{bmatrix}, \quad (2)$$

where  $\boldsymbol{\mu}_m^x$  and  $\boldsymbol{\mu}_m^y$  denote the mean vectors of the  $m$ -th component for source and target speakers, respectively.  $\boldsymbol{\Sigma}_m^{xx}$ ,  $\boldsymbol{\Sigma}_m^{yy}$ ,  $\boldsymbol{\Sigma}_m^{xy}$  and  $\boldsymbol{\Sigma}_m^{yx}$  are the blocks of covariance matrix of the  $m$ -th component. The unknown parameters  $\alpha_m$ ,  $\boldsymbol{\mu}_m$  and  $\boldsymbol{\Sigma}_m$  can be estimated by the expectation maximization (EM) algorithm.

The spectral mapping can be seen as a regression problem. Let  $F(\cdot)$  denotes the conversion function, the total squares error can be shown as

$$\varepsilon = \sum_{n=1}^N |\mathbf{y}_n - F(\mathbf{x}_n)|^2. \quad (3)$$

Introducing the least squares estimation (LSE) algorithm, the unknown parameters can be computed, and the conversion function can be represented as

$$F(\mathbf{x}_n) = E(\mathbf{y}_n | \mathbf{x}_n) = \sum_{m=1}^M p_m(\mathbf{x}_n) \left[ \boldsymbol{\mu}_m^y + \frac{\boldsymbol{\Sigma}_m^{yx}}{\boldsymbol{\Sigma}_m^{xx}} (\mathbf{x}_n - \boldsymbol{\mu}_m^x) \right], \quad (4)$$

where  $p_m(\mathbf{x}_n)$  denotes the posterior probability of  $\mathbf{x}_n$  belonging to the  $m$ -th component, and is given by

$$p_m(\mathbf{x}_n) = \frac{\alpha_m N(\mathbf{x}_n, \boldsymbol{\mu}_m^x, \boldsymbol{\Sigma}_m^{xx})}{\sum_{j=1}^M \alpha_j N(\mathbf{x}_n, \boldsymbol{\mu}_j^x, \boldsymbol{\Sigma}_j^{xx})}. \quad (5)$$

### 3. Proposed spectral mapping using KPCR

#### 3.1. PCR based Spectral mapping

When the training utterances are sparse and the number of GMM components is large, the GMM based VC always tends to overfitting. In the paper, a PCR based VC method is proposed to solve this problem. The PCR method is a linear regression, and the principal component analysis (PCA) (JOLLIFFE, 1982) is embedded to work for regression. It can efficiently overcome the multicollinearity and the overfitting problems.

Mathematically, given the spectral sequences of the source and target speakers,  $\mathbf{x}$  and  $\mathbf{y}$ , respectively. By subtracting the mean vectors  $\boldsymbol{\mu}_x$  and  $\boldsymbol{\mu}_y$ , we shall obtain the zero-mean matrices  $\tilde{\mathbf{x}} = [\tilde{\mathbf{x}}_1, \tilde{\mathbf{x}}_2, \dots, \tilde{\mathbf{x}}_N]^T$  and  $\tilde{\mathbf{y}} = [\tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2, \dots, \tilde{\mathbf{y}}_N]^T$ , respectively. The  $n$ -th frames of the spectral features take the forms as

$$\tilde{\mathbf{x}}_n = \mathbf{x}_n - \boldsymbol{\mu}_x, \quad \tilde{\mathbf{y}}_n = \mathbf{y}_n - \boldsymbol{\mu}_y; \quad (6)$$

$$\boldsymbol{\mu}_x = \frac{1}{N} \sum_{n=1}^N \mathbf{x}_n, \quad \boldsymbol{\mu}_y = \frac{1}{N} \sum_{n=1}^N \mathbf{y}_n. \quad (7)$$

In order to perform PCA, the covariance matrix of the source speech is calculated as given by

$$\boldsymbol{\Sigma} = \frac{1}{N} \tilde{\mathbf{x}}^T \tilde{\mathbf{x}}. \quad (8)$$

Performing the eigen-decomposition of the covariance matrix  $\boldsymbol{\Sigma}$ , we shall obtain

$$\boldsymbol{\Lambda} = \mathbf{Q}^T \boldsymbol{\Sigma} \mathbf{Q}, \quad (9)$$

where  $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_K)$  and  $\mathbf{Q}_K = [\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_K]$  with the  $K$  sorted eigenvalues as  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_K$ , and the associated eigenvectors  $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_K$  of the matrix  $\boldsymbol{\Sigma}$ . So the ordinary PCA projection is given by

$$\mathbf{S}_K = \tilde{\mathbf{x}} \mathbf{Q}_K. \quad (10)$$

After choosing a suitable number of principal components, the important features of  $\tilde{\mathbf{x}}$  are retained by  $\mathbf{S}_K$ . Then, we perform a linear regression between  $\mathbf{S}_K$  and the target spectral feature sequence  $\tilde{\mathbf{y}}$ ,

$$\tilde{\mathbf{y}} = \mathbf{S}_K \mathbf{C} + \mathbf{E}, \quad (11)$$

where  $\mathbf{C}$  is the regressor, and  $\mathbf{E}$  is the residual error. Employing the LSE algorithm, the unknown parameter  $\mathbf{C}$  is given by

$$\mathbf{C} = \left( \mathbf{S}_K \mathbf{S}_K^T \right)^{-1} \mathbf{S}_K^T \tilde{\mathbf{y}}, \quad (12)$$

and it is necessary to turn to  $\mathbf{x}$  for prediction, so the conversion function can be written as

$$F(\mathbf{x}_n) = (\mathbf{x}_n - \boldsymbol{\mu}_x) \mathbf{Q}_K \mathbf{C} + \boldsymbol{\mu}_y. \quad (13)$$

#### 3.2. KPCR based spectral mapping

In fact, the relationships between the spectral features of source speech and target speech are nonlinear, so the KPCR algorithm is further proposed for the spectral mapping. It has many advantages, such as it can make a perfect nonlinear regression, shows better mapping performance than the PCR method, and can avoid the overfitting problem efficiently (SCHOLKOPF *et al.*, 1997). Let  $\phi(\tilde{\mathbf{x}})$  be a nonlinear mapping function from the lower feature space to the higher feature space, so that  $\tilde{\mathbf{x}}_n$  is thereby projected to be  $\phi(\tilde{\mathbf{x}}_n)$ . Assuming  $\sum_{n=1}^N \phi(\tilde{\mathbf{x}}_n) = 0$ , the covariance matrix in the feature space is given by

$$\boldsymbol{\Sigma} = \frac{1}{N} \sum_{n=1}^N \phi(\tilde{\mathbf{x}}) \phi(\tilde{\mathbf{x}})^T. \quad (14)$$

The eigenvector  $\mathbf{q}_l$  can be obtained by a linear combination of  $\phi(\tilde{\mathbf{x}}_n)$ , and given by

$$\mathbf{q}_l = \sum_{n=1}^N \beta_{l,n} \phi(\tilde{\mathbf{x}}_n). \quad (15)$$

The projection onto eigenvector  $\mathbf{q}_l$  is given by

$$\begin{aligned} B_l &= \langle \mathbf{q}_l, \phi(\tilde{\mathbf{x}}_j) \rangle = \sum_{n=1}^N \beta_{l,n} \phi(\tilde{\mathbf{x}}_j)^T \phi(\tilde{\mathbf{x}}_n) \\ &= \sum_{n=1}^N \beta_{l,n} k(\tilde{\mathbf{x}}_n, \tilde{\mathbf{x}}_j), \end{aligned} \quad (16)$$

where  $k(\tilde{\mathbf{x}}_n, \tilde{\mathbf{x}}_j)$  is the kernel function, and  $l = 1, 2, \dots, L$  are the numbers of principal components. However, in general, the projected data given by  $\phi(\tilde{\mathbf{x}})$  does not have zero mean, and it is too difficult to calculate the mean values of  $\phi(\tilde{\mathbf{x}})$ . Employing the kernel function can avoid directly working in the feature space. Denote the kernel matrix  $\mathbf{K}$  and the centralized kernel matrix  $\tilde{\mathbf{K}}$ , respectively, and represented as

$$[\mathbf{K}]_{i,j} = [k(x_i, x_j)], \quad [\tilde{\mathbf{K}}]_{i,j} = \left[ \tilde{k}(x_i, x_j) \right]. \quad (17)$$

The  $\tilde{\mathbf{K}}$  is computed after centralizing the kernel function, and given by

$$\tilde{\mathbf{K}} = \mathbf{K} - \mathbf{I}_N \mathbf{K} - \mathbf{K} \mathbf{I}_N + \mathbf{I}_N \mathbf{K} \mathbf{I}_N, \quad (18)$$

where  $\mathbf{I}_N$  is a  $N \times N$  matrix with every element taking the value of  $1/N$ . So the formula (16) can be modified as

$$\mathbf{B}_l = \sum_{n=1}^N \beta_{l,n} \tilde{k}(\tilde{\mathbf{x}}_n, \tilde{\mathbf{x}}_j). \quad (19)$$

The KPCR is based on the analysis of KPCA, and the regression function takes the form as follows

$$F(\mathbf{x}_n) = \mathbf{B}_L \mathbf{w} + \boldsymbol{\mu}_y, \quad (20)$$

where

$$\mathbf{B}_L = \sum_{i=1}^N \beta_{l,i} K(\tilde{\mathbf{x}}_i, \mathbf{x}_n - \boldsymbol{\mu}_x)$$

and

$$\mathbf{w} = (\mathbf{B}_L \mathbf{B}_L^T)^{-1} \mathbf{B}_L^T \tilde{\mathbf{y}}.$$

The optimal number of principal components is critical for the KPCR method, and is chosen by a ten-fold cross-validation in the paper.

### 3.3. Combining with GMM

From the state-of-the-art references, it is not possible to perfectly describe the relationships between source and target speech using single regression function (Helander *et al.*, 2010; SONG *et al.*, 2012). In this paper, the KPCR method is further extended to combine with GMM. Similar to the forms of GMM based VC function mentioned in Sec. 2, the regression function can be seen as a mixture of local regressions. In the  $m$ -th mixture, the regression function is given by

$$F_m(\mathbf{x}_n) = \mathbf{B}_{L,m} \mathbf{w}_m + \boldsymbol{\mu}_m^y, \quad (21)$$

where  $\mathbf{B}_{L,m} = \sum_{i=1}^N \beta_{l,i} k(\mathbf{x}_i - \boldsymbol{\mu}_x, \mathbf{x}_n - \boldsymbol{\mu}_m^x)$ , and  $\mathbf{w}_m$  is the regressor. So the global conversion function can be calculated as shown

$$F(\mathbf{x}_n) = \sum_{m=1}^M p_m(\mathbf{x}_n) F_m(\mathbf{x}_n), \quad (22)$$

where  $p_m(\mathbf{x}_n)$  is the posterior probability as shown in Eq. (5).

The kernel selection is also essential to the accuracy of spectral transformation. There are many kinds of kernel functions, such as linear kernel, polynomial kernel, Gaussian kernel, and wave kernel, etc. Among which, the Gaussian kernel is simple and efficient, and is chosen for spectral mapping, which takes the form as

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right), \quad (23)$$

where  $\sigma$  is the width of the Parzen window, and the decent range is used to find its best value. In the paper, the  $\sigma$  is optimized as 0.6 also by a ten-fold cross-validation.

## 4. Post-processing using posterior probability

The proposed method can be seen as an extension of GMM method, and can efficiently avoid the overfitting problem. But it is well known that except the overfitting problem, GMM method has another two main shortcomings, one is the discontinuity problem because of the frame-by-frame based transformation,

and the other is the oversmoothing problem caused by the statistical averaging of the model. In this section, the two problems are analyzed and efficiently overcome using the posterior probability information.

From the prior studies, it is well known that each frame often obviously dominates one component of the GMM (HELANDER *et al.*, 2010). Table 1 gives the results obtained from the 1500 frames based on a GMM with 8 components, we can easily find that most of the frames have the maximum posterior probabilities higher than 90%, and only a very few with the maximum posterior probabilities lower than 50%. So it can be assumed that each frame belongs to single component of GMM, and it also has been proven that the changes of posterior probabilities from one component to another are rapid (HELANDER *et al.*, 2010), which will lead to the discontinuities in the converted speech, and degrade the perceptual quality.

Table 1. Number of frames in different maximum posterior probabilities.

| Maximum posterior probability [%] | Number of frames |
|-----------------------------------|------------------|
| 90–100                            | 1052             |
| 80–90                             | 213              |
| 70–80                             | 78               |
| 60–70                             | 52               |
| 50–60                             | 38               |
| 0–50                              | 67               |

In this paper, in order to reduce the discontinuity problem, an adaptive median filtering approach is adopted to smooth the posterior probabilities. The adaptive median filtering is a prevalent and efficient algorithm in image processing, and can efficiently smooth the images while preserving the details (HWANG, HADDAD, 1995). Let  $W$  be the rectangle moving window, and the initial length and maximum length are set as 3 and 7, respectively. In each component of GMM,  $p_{\text{cur}}$  is the posterior probability of current frame,  $p_{\text{min}}$ ,  $p_{\text{max}}$  and  $p_{\text{med}}$  are the minimum, maximum and median values of the posterior probability in the window, respectively. The adaptive filtering can be seen as a two level structure:

Level A:

$$A_1 = p_{\text{med}} - p_{\text{min}},$$

$$A_2 = p_{\text{med}} - p_{\text{max}}.$$

In Level A, if  $A_1 > 0$  and  $A_2 < 0$ , then go to level B. Or increase the length of  $W$  and repeat level A. If the length exceeds the maximum, then  $p_{\text{cur}}$  is the output.

Level B:

$$B_1 = p_{\text{cur}} - p_{\text{min}},$$

$$B_2 = p_{\text{cur}} - p_{\text{max}}.$$

In level B, if  $B_1 > 0$  and  $B_2 < 0$ , then  $p_{\text{cur}}$  is the output, or  $p_{\text{med}}$  is chosen as the output.

Meanwhile, the statistical averaging of GMM will introduce the oversmoothing problem, and smoothing the posterior probabilities by adaptive median filtering will also aggravate this issue to some extent, although it can efficiently decrease the discontinuities. As mentioned above, in most cases, each frame dominates one component of GMM. So a new approach is proposed, to each frame, the mixture component with highest posterior probability is chosen, and the local regression in this component is adopted to replace the global transformation. Theoretically, it can efficiently solve the oversmoothing problem. Meanwhile, it can be also observed that after adaptive smoothing, there are always two higher posterior probabilities to each frame in many cases, especially when the highest posterior probabilities of the adjacent frames belong to the different mixture components. So a top-two selection strategy is further proposed, in which, after adaptive smoothing, the mixture components with two higher posterior probabilities are chosen, and the two probabilities are normalized so that they sum to a unity. The transformation function can be modified as

$$F(x_n) = \sum_{m=1}^2 p_m F_m(x_n), \quad (24)$$

where  $p_m$  is the normalized posterior probability, which satisfies  $\sum_{m=1}^2 p_m = 1$ .

## 5. Experiments

We perform the experiments on CMU ARCTIC corpus. Two male and two female speakers are chosen. 25 parallel utterances of each speaker are prepared for the experiments, in which, 15 utterances (about 5 minutes) are used for training, while the others are used for testing. The 30-order Mel-cepstral coefficients (MCEPs) and their  $\Delta$  and  $\Delta^2$  features (totally 90th order) are chosen to represent the spectral features (TODA *et al.*, 2005). Four kinds of VC methods are compared, they are the baseline GMM method (GMM), the PCR method (PCR), the KPCR method (KPCR), and the combined KPCR and GMM method with post-processing (GKPCR). Meanwhile, four types of VC strategies are adopted, including male-to-male transformation (M-M), male-to-female transformation (M-F), female-to-female transformation (F-F), and female-to-male transformation (F-M). The objective and subjective experiments are carried out to evaluate the performance of the proposed method, respectively. The Mel-cepstral distortion (MCD) and speaker identification system are chosen for the objective evaluation, while the ABX test and MOS test are used for the subjective evaluation. The number of GMM is set as 16. The number of principal components is optimized as 40, and 8 experienced people are employed for the listening tests.

### 5.1. Objective evaluation

The MCD is a common method to evaluate the objective performance of VC. It measures the cepstral distance between the converted speech and target speech, and the formula can be computed as follows

$$MCD = 10 / \log 10 \sqrt{2 \sum_{j=1}^D (\mathbf{mc}_j^c - \mathbf{mc}_j^t)^2}, \quad (25)$$

where  $\mathbf{mc}_j^c$  and  $\mathbf{mc}_j^t$  are the  $i$ -th coefficients of converted and target MCEPs, respectively, and  $D$  is the dimension of MCEPs.

Figures 1 to 4 give the average MCD results of different methods, it should be noted that a lower MCD value means a better VC performance. We can observe that when the numbers of principal components are adjusted properly, the proposed method can show greatly better performance than the baseline GMM method.

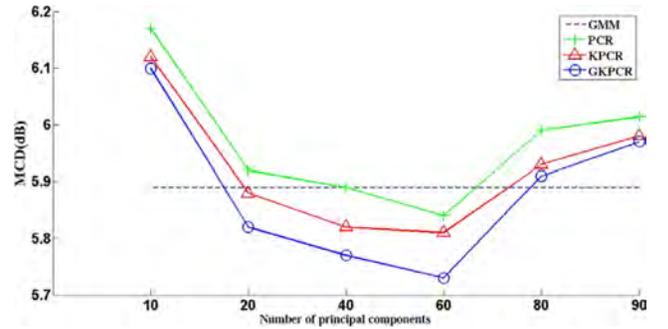


Fig. 1. MCD of M-M.

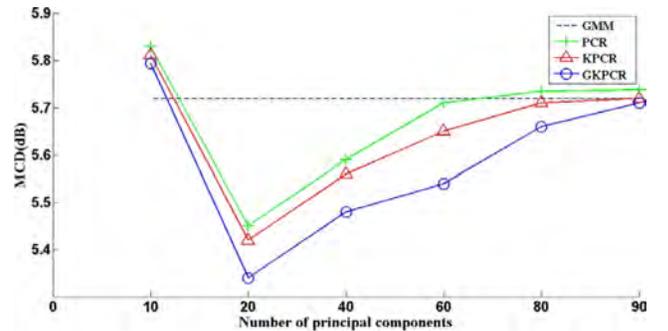


Fig. 2. MCD of M-F.

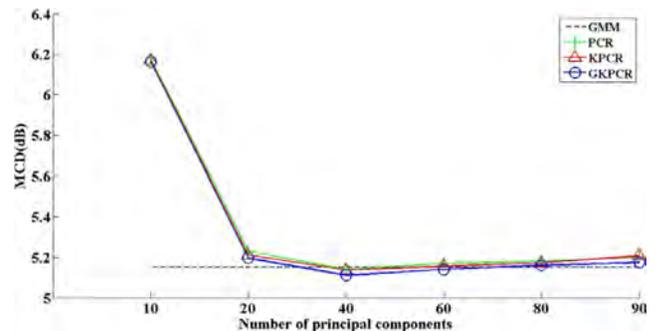


Fig. 3. MCD of F-F.

In the F-F VC, when the numbers of principal components are 40 and 60, respectively, the proposed method shows a little better performance. While in other cases, the proposed method is superior to the other methods.

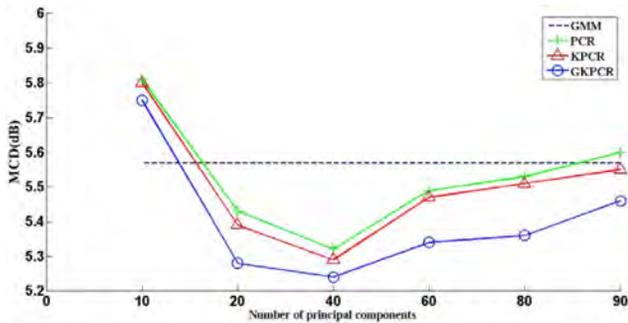


Fig. 4. MCD of F-M.

The GMM based speaker identification system (REYNOLDS *et al.*, 2000) is also adopted for the objective evaluation of VC. Let  $\lambda_S$  and  $\lambda_T$  be the trained speaker models of source and target speakers, respectively, the performance measure  $\theta_{ST}$  is given by

$$\theta_{ST} = \log P(\mathbf{O}|\lambda_T) - \log P(\mathbf{O}|\lambda_S), \quad (26)$$

where  $\mathbf{O}$  is the observed sequence of spectral features, the average recognition results are summarized in Table 2. It can be found that before conversion all the values of  $\theta_{ST}$  are minus, which demonstrates that the speech is recognized as the source speech, and after conversion the values will become plus, which indicates that the converted speech is recognized as the target one. It can be also observed that in all cases, the converted speech is more likely to be recognized as the target one. Compared to other methods, the proposed GKPCR method can efficiently increase the values of  $\theta_{ST}$ .

Table 2. Results of the speaker identification.

| Strategy | $\theta_{ST}$     |                  |       |       |       |
|----------|-------------------|------------------|-------|-------|-------|
|          | Before conversion | After conversion |       |       |       |
|          |                   | GMM              | PCR   | KPCR  | GKPCR |
| M-M      | -4.53             | +3.25            | +3.31 | +3.34 | +3.36 |
| M-F      | -5.16             | +3.01            | +3.08 | +3.09 | +3.12 |
| F-F      | -4.02             | +3.51            | +3.55 | +3.58 | +3.63 |
| F-M      | -5.41             | +3.14            | +3.26 | +3.27 | +3.34 |

### 5.2. Subjective evaluation

The ABX test is carried out for the similarity between converted and target speech. In this method, A and B are the converted speech using GKPCR method and other methods, respectively. X is the target speech. The listeners are asked to choose whether A or B

is closer to X. Figure 5 summarizes the overall performance of different VC methods, and the results confirm the conclusions of objective tests. It can be found that the proposed GKPCR method shows significantly better preference than the baseline GMM method (With probability of 90%). Meanwhile, compared to the PCR method and KPCR method, the proposed method also shows obvious superiority.

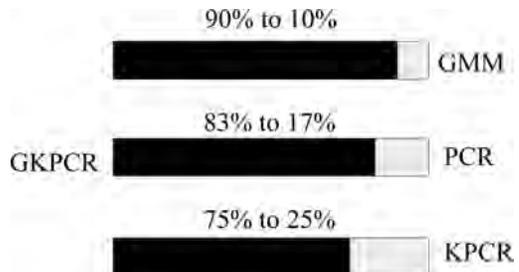


Fig. 5. Results of ABX test.

The MOS test is also performed to evaluate the perceptual quality of the converted speech. The converted utterances using different methods are shown to the listeners, who are asked to rate the converted speech using a 5-point score from 1 “bad” to 5 “excellent”. The average scores of different methods are summarized in Table 3. The mean score and standard deviation (SD) are given, respectively, and the confidence interval is set as 95%. We can easily find that the proposed GKPCR method greatly outperforms the baseline GMM method and other methods, with higher mean scores and lower SD.

Table 3. Results of perceptual quality evaluation.

|            | GMM  | PCR  | KPCR | GKPCR |
|------------|------|------|------|-------|
| Mean score | 3.72 | 3.79 | 3.83 | 4.09  |
| SD         | 0.51 | 0.48 | 0.47 | 0.42  |

## 6. Conclusions

In this paper, a novel spectral mapping method using KPCR is proposed for VC. The proposed KPCR method can efficiently solve the overfitting problem, and also perfectly describe the nonlinear relationships between the source speech and target speech. To further improve the VC performance, the combined KPCR and GMM method is proposed. The discontinuity and oversmoothing problems of traditional GMM method are also analyzed, and can be efficiently overcome by adopting the novel post-processing strategies using posterior probability information. The adaptive median filter is employed to reduce the discontinuities, while a top-two selection strategy is proposed to solve the oversmoothing. Experimental results demonstrate that the proposed method greatly outperforms the traditional GMM method.

## Acknowledgment

The authors acknowledge the work is supported by the National Natural Science Foundation of China (Grant Nos. 60975017 and 51075068), the Natural Science Foundation of Higher Education Institutions of Jiangsu Province (Grant No. 10KJB510005), and the Natural Science Foundation of Guangdong Province (Grant No. 10252800001000001).

## References

1. ABE M., NAKAMURA S., SHIKANO K., KUWABARA H. (1998), *Voice conversion through vector quantization*, Proceedings of the 1998 International Conference on Acoustics, Speech, and Signal Processing, pp. 655–658, New York.
2. CHEN Y., CHU M., CHANG E., LIU J., LIU R. (2003), *Voice conversion with smoothed GMM and MAP adaptation*, Proceedings of Eurospeech 2003, pp. 2413–2416, Geneva.
3. DESAI S., BLACK A. W., YEGNANARAYANA B., PRAHALLAD K. (2010), *Spectral mapping using artificial neural networks for voice conversion*, IEEE Transactions on Audio, Speech, and Language Processing, **18**, 5, 954–964.
4. HELANDER E., VIRTANEN T., NURMINEN J., GABBOUJ M. (2010), *Voice conversion using partial least squares regression*, IEEE Transactions on Audio, Speech, and Language Processing, **18**, 5, 912–921.
5. HWANG H., HADDAD R. A. (1995), *Adaptive median filters: new algorithms and results*, IEEE Transactions on Image Processing, **4**, 4, 499–502.
6. JOLLIFFE I. T. (1982), *A note on the use of principal components in regression*, Applied Statistics, **31**, 3, 300–303.
7. KAIN A., MACON M. W. (1998), *Spectral voice conversion for text-to-speech synthesis*, Proceedings of the 1998 IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 285–288, Seattle.
8. MESBAHI L., BARREAUD V., BOEFFARD O. (2007), *GMM-based speech transformation systems under data reduction*, Proceedings of the 6th ISCA workshop on Speech Synthesis, pp. 119–124, Bonn.
9. REYNOLDS D. A., QUATIERI T. F., DUNN R. B. (2000), *Speaker verification using adapted Gaussian mixture models*, Digital Signal Processing, **10**, 1, 19–41.
10. SONG P., BAO Y. Q., ZHAO L., ZOU C. R. (2011), *Voice conversion using support vector regression*, Electronics Letters, **47**, 18, 1045–1046.
11. SONG P., JIN Y., ZHAO L., ZOU C. R. (2012), *Voice conversion based on hybrid SVR and GMM*, Archives of Acoustics, **37**, 2, 143–149.
12. SCHOLKOPF B., SMOLA A., MULLER K. R. (1997), *Kernel principal component analysis*, Proceedings of the 7th International Conference on Artificial Neural Networks, pp. 583–588, Berlin.
13. STYLIANOU Y., CAPPE O., MOULINES E. (1998), *Continuous probabilistic transform for voice conversion*, IEEE Transactions Speech and Audio Processing, **6**, 2, 131–142.
14. TODA T., SARUWATARI H., SHIKANO K. (2001), *Voice conversion algorithm based on Gaussian mixture model with dynamic frequency warping of STRAIGHT spectrum*, Proceedings of the 2001 International Conference on Acoustics, Speech, and Signal Processing, pp. 841–944, Salt Lake City.
15. TODA T., BLACK A. W., TOKUDA K. (2005), *Spectral conversion based on maximum likelihood estimation considering global variance of converted parameter*, Proceedings of the 2005 International Conference on Acoustics, Speech, and Signal Processing, pp. 9–12, Philadelphia.
16. VALBRET H., MULINES E., TUBACH J. (1992), *Voice transformation using PSOLA techniques*, Speech Communication, **11**, 2-3, 175–187.

# Evaluation of Decay Times from Noisy Room Responses with Pure-Tone Excitation

Mirosław MEISSNER

*Institute of Fundamental Technological Research, Polish Academy of Sciences  
Pawińskiego 5B, 02–106 Warszawa, Poland; e-mail: mmeissn@ippt.pan.pl*

*(received September 18, 2012; accepted December 24, 2012)*

Reverberant responses are widely used to characterize acoustic properties of rooms, such as the early decay time (EDT) and the reverberation times  $T_{20}$  and  $T_{30}$ . However, in real conditions a sound decay is often deformed by background noise, thus a precise evaluation of decay times from noisy room responses is the main problem. In this paper this issue is examined by means of numerical method where the decay times are estimated from the decay function that has been determined by nonlinear polynomial regression from a pressure envelope obtained via the discrete Hilbert transform. In numerical experiment the room responses were obtained from simulations of a sound decay for two-room coupled system. Calculation results have shown that background noise slightly affects the evaluation of reverberation times  $T_{20}$  and  $T_{30}$  as long as the signal-to-noise ratio (SNR) is not smaller than about 25 and 35 dB, respectively. However, when the SNR is close to about 20 and 30 dB, high overestimation of these times may occur as a result of bending up of the decay curve during the late decay.

**Keywords:** room acoustics, reverberation, decay times, room response, background noise, coupled rooms.

## 1. Introduction

One of the most fundamental aims of room acoustics is a prediction of decay times from measurements of a sound pressure decay inside enclosures. An accurate determination of decay times is primary for both absorption measurements in reverberation chambers (BARRON, COLEMAN, 2001; NUTTER *et al.*, 2007) and the evaluation of acoustics of performance spaces (BRADLEY, 2005; GOŁA, SUDERDEBSKA, 2009; ADELMAN-LARSEN *et al.*, 2010; BERANEK, 2011) as well as ordinary rooms (DÍAZ, PEDRERO, 2005, 2007). Decay times are evaluated from a decay curve defined as the graphical representation of the decay of the sound pressure level in a room as a function of time after the cut-off of a continuous sound source (ISO 3382, 2012). The estimation of decay times is achieved by approximation of appropriate parts of the decay curve by fitting lines obtained by a linear least-squares regression and then a calculation of the decay times from the slope of these lines. Another method of determining the decay curve consists in the reverse-time integration of squared impulse response (SCHROEDER, 1965). This method results in exceptionally smooth decay curves, making a determi-

nation of decay times simple and accurate. However, when the room impulse response is contaminated with high level background noise the method's accuracy is substantially reduced because of a distortion of decay curve slope during the late decay. This problem has been extensively studied in the past (CHU, 1978; LUNDEBY *et al.*, 1995; XIANG, 1995; MORGAN, 1997; XIANG, GOGGANS, 2001) and different remedial techniques have been proposed (KARJALAINEN *et al.*, 2002; DRAGONETTI *et al.*, 2009).

A subject of this paper is to study an accuracy of evaluation of decay times from room responses contaminated with background noise. The research is dedicated to low-frequency range where acoustic modes excited inside an enclosure are well separated. The reverberant response of room, which begins just after turning off the continuous sound source, is described theoretically by means of a modal expansion of a sound pressure for room systems with relatively small sound damping. The irregular room, which is considered in the study, has a form of the coupled room system consisting of two rectangular subrooms connected through an acoustically transparent opening. The choice of such a system was dictated by the fact that coupled room systems have the ability to create a nonlinear profile

of pressure level decay (XIANG, GOGGANS, 2001, 2003; MEISSNER, 2007a, 2008a). The decay times were estimated from a decay function computed by nonlinear polynomial regression, corresponding to average long-time changes in a pressure envelope obtained via the discrete Hilbert transform. The practical issues in using the discrete Hilbert transform in signal processing were examined recently by the author (MEISSNER, 2012a, 2012b) and in order to improve an accuracy of this procedure, a numerical algorithm consisted in the appropriate modification of a discrete pressure signal, was proposed.

## 2. Simulation of reverberant response of room

In low-frequency range, the reverberant behaviour of room is strongly frequency dependent (MEISSNER, 2007b, 2008b). Thus, in a theoretical model it was assumed that a room is excited by a pure-tone sound source. When a steady-state is achieved, the source is abruptly switched off and the sound energy accumulated inside the room interior is absorbed on walls and an acoustic reverberation takes place. This reverberation consists of carrier waveforms of the decaying envelope that can be described by exponentially decreasing cosinusoidal functions (MEISSNER, 2008b)

$$P_m(\mathbf{r}, t) = A_m(\mathbf{r})e^{-r_m t} \cos(\Omega_m t - \beta_m), \quad (1)$$

where  $t \geq 0$  is the time,  $\mathbf{r} = (x, y, z)$  represents the receiving position,  $m = 1, 2, 3, \dots$  is the mode number,  $A_m(\mathbf{r})$  describes the space distribution of mode amplitude

$$A_m(\mathbf{r}) = \frac{c^2 \omega_m \Phi_m(\mathbf{r}) \left[ \int_V Q(\mathbf{r}') \Phi_m(\mathbf{r}') dv' \right]}{\Omega_m \sqrt{[(\omega_m^2 - \omega^2)^2 + 4r_m^2 \omega^2]}}, \quad (2)$$

where  $c$  is the sound speed,  $V$  is the room volume,  $\omega$  is the source frequency,  $Q(\mathbf{r})$  is the volume source distribution,  $\Phi_m(\mathbf{r})$  is the eigenfunction,  $\omega_m$  is the natural mode frequency,  $\Omega_m = \sqrt{\omega_m^2 - r_m^2}$  is the mode frequency for damped oscillation,  $\beta_m$  is the initial mode phase

$$\beta_m = \tan^{-1} \left[ \frac{r_m(\omega_m^2 + \omega^2)}{\Omega_m(\omega_m^2 - \omega^2)} \right] \quad (3)$$

and  $r_m$  is the modal damping coefficient

$$r_m = \frac{1}{2} \rho c^2 \int_S \frac{\Phi_m^2(\mathbf{r}') ds'}{Z}, \quad (4)$$

where  $\rho$  is the air density,  $S$  is the surface of room walls and  $Z$  is the wall impedance. In real conditions,

the room response is usually deformed by the background noise, then assuming that the noise has a uniform spectral distribution, the formula for the reverberant response of room in noisy environment can be written as

$$p(\mathbf{r}, t) = \sum_{m=1}^M P_m(\mathbf{r}, t) + A_N \xi(t), \quad (5)$$

where  $A_N$  is the noise amplitude and  $\xi(t)$  is the unity-level random signal. In a computer algorithm, the signal  $\xi(t)$  is created by the function generating random real numbers from the range  $(-1, 1)$ . In Eq. (5) the index  $M$  corresponds to the last mode whose frequency is smaller than the Schroeder frequency

$$f_s = c \sqrt{\frac{6}{\mathcal{A}}}, \quad (6)$$

where  $\mathcal{A}$  is the equivalent absorption area. As was proved by SCHROEDER (1996), below this frequency the modal density is low and particular modes can be decomposed from the room response, thus in multi-mode resonance systems the Schroeder frequency  $f_s$  marks the transition from individual, well-separated resonances to many overlapping modes.

An accuracy of the method of determining the decay times has been investigated for a coupled room system consisting of two connected rectangular subrooms. This was motivated by the fact that such a room is capable of producing a nonlinear profile of pressure level decay which may result in significant differences in decay times in the initial and late stages of sound decay. A horizontal cross-section of the room is shown in Fig. 1. The subrooms have the same height  $h$  of 3 m and their lengths and widths are the following:

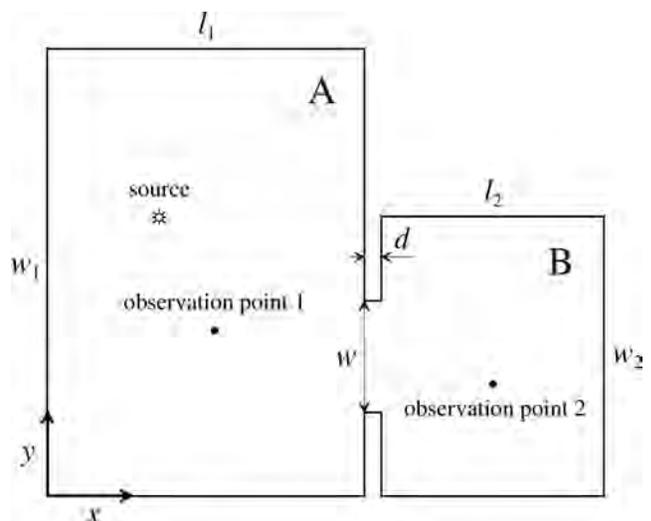


Fig. 1. Horizontal plan view of analysed room system consisting of two connected rectangular subrooms A and B. Symbols show positions of sound source and observation points.

$l_1 = 5.7$  m,  $l_2 = 4$  m,  $w_1 = 8$  m and  $w_2 = 5$  m. The coupling between subrooms is realized by the opening having the height  $h$ , the width  $w$  of 2 m and the thickness  $d$  of 0.3 m. The room system was excited with the power of  $10^{-3}$  W by a harmonic source situated at the point:  $x = 2$  m,  $y = 5$  m,  $z = 1$  m. The room response was received at two observation points:  $x = 3$  m,  $y = 3$  m,  $z = 1.8$  m (subroom A) and  $x = 8$  m,  $y = 2$  m,  $z = 1.8$  m (subroom B). The walls of subrooms A and B were assumed to be covered by materials having the random-absorption coefficients  $\alpha_1$  and  $\alpha_2$ . Thus, the equivalent absorption area  $\mathcal{A}$  is the following

$$\mathcal{A} = \alpha_1 S_1 + \alpha_2 S_2, \quad (7)$$

where  $S_1$  and  $S_2$  are surfaces of walls in subrooms A and B. In a numerical simulation the coefficients  $\alpha_1$  and  $\alpha_2$  were set to 0.04 and 0.185, respectively. Therefore, in the considered case a sound absorption in subroom A was much smaller than in subroom B. Using Eqs. (6) and (7), and the assumed values  $\alpha_1$  and  $\alpha_2$ , it is easy to calculate that the Schroeder frequency is as follows:  $f_s \approx 174$  Hz. Below this frequency 150 eigenmodes were found and for this set of modes the reverberant responses of room in receiving positions

were simulated. The eigenfunctions  $\Phi_m$  were computed by a numerical solution of the wave equation where the finite difference method and the forced oscillator method were employed (MEISSNER, 2007b).

A computer reconstruction of room reverberant response was performed for two source frequencies: 107 and 151 Hz, which correspond approximately to eigenfrequencies of 44th and 107th modes. The room responses computed for these frequencies under noise-free conditions are shown in Fig. 2. Calculation results indicate in a clear way that in the analyzed coupled room system a shape of the room response and an initial amplitude of a sound decay strongly depend on the sound frequency and a position of the observation point. For example, when the source frequency is set to 107 Hz, a smooth decay of the sound pressure is observed inside the subroom A in the first receiving position (Fig. 2a). However, a shift of this position to the subroom B results in a fundamental change of the response consisting in a rapid decrease in a pressure in the initial stage of a sound decay (Fig. 2b). On the other hand, for the source frequency of 151 Hz large wave fluctuations in reverberant responses are noted suggesting that in these cases the beating effect is present (Figs. 2c, d).

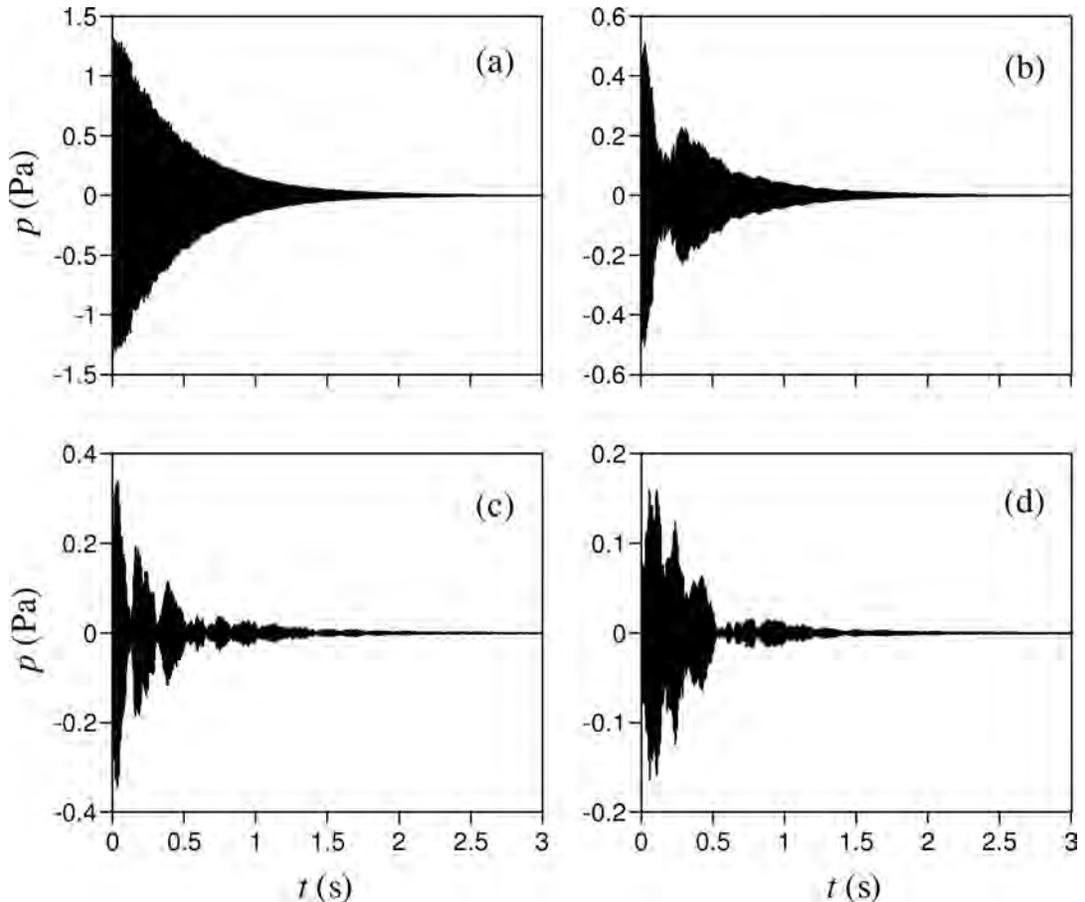


Fig. 2. Temporal changes in sound pressure  $p$  for source frequencies: a), b) 107 Hz and c), d) 151 Hz, and noise-free conditions. Position of observation point: a), c) subroom A, b), d) subroom B.

### 3. Detection of response envelope via discrete Hilbert transform

The Hilbert transform is an important tool for a signal analysis because it can be used in a direct examination of instantaneous properties of the signal such as an envelope and a phase. An application of this method to the reverberant response of rooms enables a detection of the response envelope giving a more accurate prediction of decay times. In continuous time domain a classical definition of the Hilbert transform  $\mathcal{H}$  is as follows (HAHN, 1996)

$$\mathcal{H}[s(t)] = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{s(\tau)}{t - \tau} d\tau, \quad (8)$$

where  $s(t)$  is a real-valued signal and the symbol  $\int$  denotes the principal value integral because of the possible singularity at  $\tau = t$ . Of course, the room response obtained by a numerical simulation is of finite length and digitally sampled, thus if the pressure  $p(\mathbf{r}, t)$  from Eq. (5) is uniformly sampled with the period  $T$ , the discrete-time pressure signal is obtained

$$p[n] = p(\mathbf{r}, t[n]), \quad (9)$$

where  $t[n] = nT$ ,  $n = 0, 1, \dots, N$  and on the left side, the spatial coordinate  $\mathbf{r}$  is omitted for simplicity of notation. In discrete time case, the Hilbert transform  $\mathcal{H}$  is replaced by the discrete Hilbert transform  $\mathcal{H}_d$ . In order to improve an exactness of the discrete Hilbert transform in the prediction of pressure envelope the extended discrete signal  $P[n]$  determined on the basis of the original pressure signal  $p[n]$  is introduced (MEISSNER, 2012a)

$$P[n] = \begin{cases} -p[N - n], & n = 0, 1, \dots, N - 1, \\ p[-N + n], & n = N, N + 1, \dots, 2N. \end{cases} \quad (10)$$

The new discrete signal is determined in a double-extended time interval and represents a discrete function having a rotational symmetry with respect to the origin of coordinate system. This method of signal processing reduces an inaccuracy of the discrete Hilbert transform generated by a limitation of a signal duration (end effect) and caused by the fact that for exponentially decaying harmonic signals the Bedrosian identity (BEDROSIAN, 1963) is not satisfied. Using a definition of the discrete Hilbert transform for non-periodic signals (KAK, 1970), the Hilbert transform of the extended signal  $P[n]$  is determined by

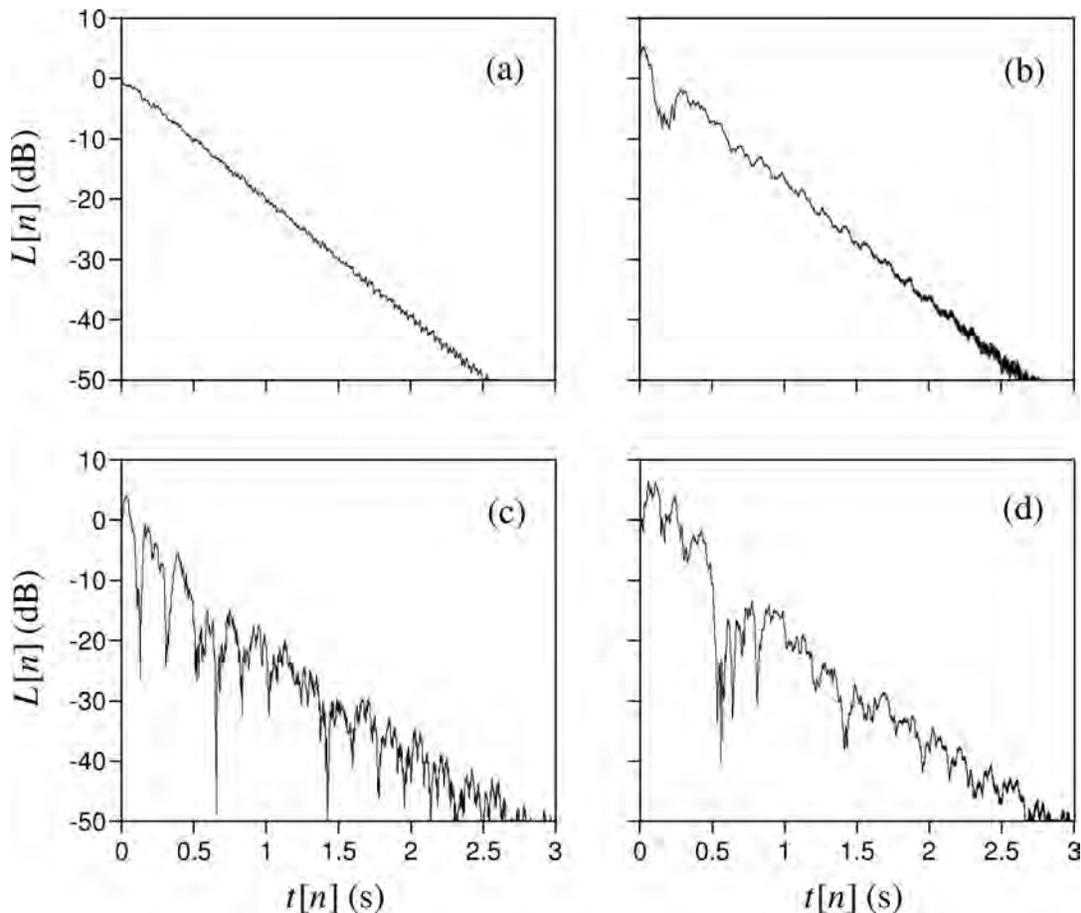


Fig. 3. Temporal decay of level  $L[n]$  of pressure envelope for signals shown in Fig. 2.

$$\mathcal{H}_d\{P[n]\} = \begin{cases} \frac{2}{\pi} \sum_{m=\text{odd}} \frac{P[m]}{n-m}, & n \text{ even,} \\ \frac{2}{\pi} \sum_{m=\text{even}} \frac{P[m]}{n-m}, & n \text{ odd.} \end{cases} \quad (11)$$

Subsequently, a logarithmic decay of a pressure envelope is found from the equation

$$L[n] = 20 \log(E[n]/E[0]), \quad (12)$$

where  $E[n] = \sqrt{P^2[n] + \mathcal{H}_d^2\{P[n]\}}$  and  $E[0]$  is the envelope at a beginning of a sound decay.

Calculation results in Fig. 3 depict temporal changes in the relative level  $L[n]$  of the pressure envelope calculated for the room responses shown in Fig. 2 by applying the method presented above. For the first room response, the level  $L[n]$  in the analysed time domain decreases almost linearly with the time showing that in this case a reverberation process can be described by a single decay time. A characteristic property of the second room response is a considerable difference between decay times in the early and late stages of the decaying sound. In the last two cases significant fluctuations of the level  $L[n]$  are observed and they are due to a presence of two dominant modes of slightly

different frequencies in room responses. This causes the pressure envelope to fluctuate with a frequency equal to the difference between frequencies of these modes (the beating effect).

Numerical data in Fig. 4 show changes in the room responses when the sound decay is deformed by a high level background noise ( $A_N = 2 \times 10^{-2}$  Pa, the noise level  $L_N$  of 60 dB). Under such very noisy conditions, the decrease in the relative level  $L[n]$  is noted in the initial stage of reverberation process until the decaying signal is masked by the noise. A duration of this initial stage depends directly on the signal-to-noise ratio (SNR) and of course, it is visibly smaller for low SNRs as is evident from the comparison of Figs. 2 and 4.

#### 4. Evaluation of decay times

As was shown in the previous section, an application of the discrete Hilbert transform method is a simple way to calculate the relative level  $L[n]$  of a pressure envelope in room responses. However, an accuracy of a direct evaluation of decay times from changes in the level  $L[n]$  is substantially limited by a presence of background noise and large fluctuations of pressure in some responses. Therefore, in a numerical procedure

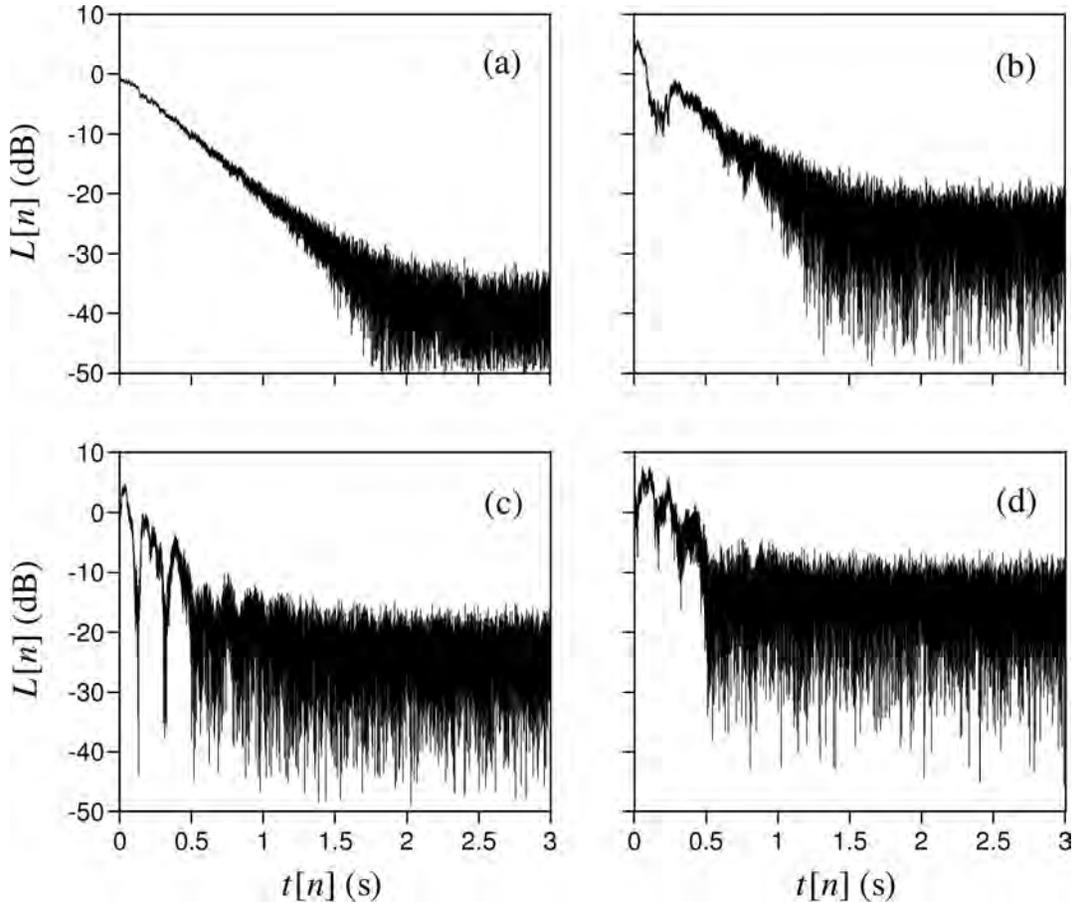


Fig. 4. Temporal changes in level  $L[n]$  for signals from Fig. 2 with added high level background noise ( $L_N = 60$  dB).

the function  $L[n]$  was employed to construct the decay function  $L_{av}$  describing average long-time changes in a sound pressure level. This function was calculated by nonlinear polynomial regression and it was finally used to predict decay times. Examples of functions  $L_{av}$  computed by this method (solid lines) and the level  $L[n]$  corresponding to these functions (gray lines) are shown in Fig. 5. The results were obtained for the sound source of frequency 151 Hz located in the subroom A and the background noise level  $L_N$  of 0 and 60 dB.

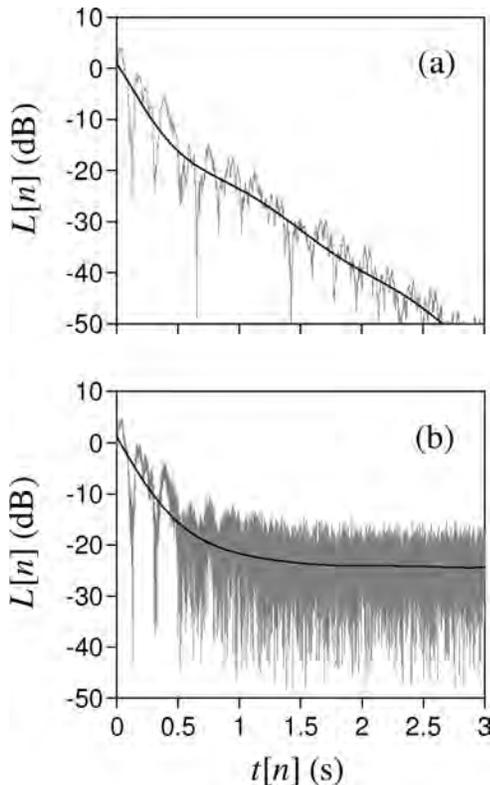


Fig. 5. Level  $L[n]$  of pressure envelope (gray lines) and decay function  $L_{av}$  calculated by polynomial regression of 8th order (solid lines) for source frequency of 151 Hz and noise level  $L_N$ : a) 0 dB, b) 60 dB. Observation point in subroom A.

The estimation of decay times was based on finding a fit line to appropriate parts of the function  $L_{av}$  and it was realized by a linear regression. In order to properly characterize the reverberation process, three decay times were computed: the early decay time (EDT) predicted on the basis of a drop of  $L_{av}$  from 0 to  $-10$  dB and the reverberation times  $T_{20}$  and  $T_{30}$  estimated from a decrease in  $L_{av}$  from  $-5$  to  $-25$  and  $-35$  dB, respectively. The early decay time is always determined from measurements because the initial decay is important from the subjective viewpoint. The decay time  $T_{30}$  is a standard measure of the reverberation time under low-noise conditions. However, when the decay curve does not have sufficient dynamic range due to the presence of high level background noise, the de-

decay time  $T_{20}$  is used as appropriate measures of the reverberation time (ISO 3382, 2012).

Results of evaluation of the decay times EDT,  $T_{20}$  and  $T_{30}$  are summarized in Tables 1 and 2. They illustrate changes in the decay times with the background noise level  $L_N$  increasing from 0 to 60 dB in 10 dB steps and show variations of these times with corresponding SNRs. In addition to these data, in Tables 1 and 2 dimensionless parameters  $\Delta_{EDT}$ ,  $\Delta_{T_{20}}$  and  $\Delta_{T_{30}}$  are also collected. These quantities are defined as

$$\Delta_{EDT} = \frac{|\text{EDT}(L_N) - \text{EDT}(0)|}{\text{EDT}(0)}, \quad (13)$$

$$\Delta_{T_{20}} = \frac{|T_{20}(L_N) - T_{20}(0)|}{T_{20}(0)}, \quad (14)$$

$$\Delta_{T_{30}} = \frac{|T_{30}(L_N) - T_{30}(0)|}{T_{30}(0)}, \quad (15)$$

where  $\text{EDT}(0)$ ,  $T_{20}(0)$  and  $T_{30}(0)$  are values of decay times for the background noise level  $L_N$  equal to zero ( $A_N = 2 \times 10^{-5}$  Pa), thus they represent relative errors in the evaluation of decay times due to the presence of background noise.

Simulation data collected in upper part of Table 1 were obtained in the receiving position located in the subroom A for the source frequency of 107 Hz. In this case the room response is characterized by large and moderate SNRs (37–97 dB) resulting in a good accuracy in determination of decay times. Since the response is dominated by one acoustic mode all decay times are very similar (3.012–3.050 s). For the same source frequency and the observation point located in subroom B, the decay function exhibits a “sagging” appearance in an initial stage of decay (Fig. 3) causing that the early decay time is approximately three times smaller than the reverberation times  $T_{20}$  and  $T_{30}$  evaluated from the late decay. Horizontal lines in Table 1 denote the case when the SNR is below 30 dB making it impossible to determine the reverberation time  $T_{30}$ .

Calculation results in Table 2 were obtained for the source frequency of 151 Hz. They confirm that decay functions created on the basis of signals received at both observation points are highly nonlinear because the decay times considerably differ. For example, if the observation point is located in the subroom A and the background noise level  $L_N$  is zero, the early decay time is approximately 1.9 times smaller than  $T_{20}$  and 2.2 times smaller than  $T_{30}$ . On the other hand, when this point is located in the subroom B, the ratios  $T_{20}/\text{EDT}$  and  $T_{30}/\text{EDT}$  are respectively equal to 2.6 and 2.3 for  $L_N = 0$ . Because of smaller signal dynamics, the effect of background noise on the reverberation process manifests itself through high increase in reverberation times  $T_{20}$  and  $T_{30}$  for the lowest SNRs (observation point in subroom A) or a total impossibility of determining these reverberation times for a number of cases (observation point in subroom B).

Table 1. Decay times EDT,  $T_{20}$ ,  $T_{30}$  and relative errors  $\Delta_{\text{EDT}}$ ,  $\Delta_{T_{20}}$ ,  $\Delta_{T_{30}}$  versus background noise level  $L_N$  and signal-to-noise ratio. Source frequency of 107 Hz. Observation point in subroom A (upper part of table) and subroom B (bottom part of table).

| $L_N$ [dB] | SNR [dB] | EDT [s] | $\Delta_{\text{EDT}}$ [%] | $T_{20}$ [s] | $\Delta_{T_{20}}$ [%] | $T_{30}$ [s] | $\Delta_{T_{30}}$ [%] |
|------------|----------|---------|---------------------------|--------------|-----------------------|--------------|-----------------------|
| 0          | 97       | 3.050   | 0                         | 3.012        | 0                     | 3.028        | 0                     |
| 10         | 87       | 3.050   | 0.006                     | 3.012        | 0.002                 | 3.028        | 0                     |
| 20         | 77       | 3.049   | 0.024                     | 3.012        | 0.009                 | 3.028        | 0.001                 |
| 30         | 67       | 3.047   | 0.080                     | 3.012        | 0.031                 | 3.028        | 0.003                 |
| 40         | 57       | 3.048   | 0.071                     | 3.013        | 0.052                 | 3.028        | 0.007                 |
| 50         | 47       | 3.026   | 0.782                     | 3.015        | 0.107                 | 3.033        | 0.179                 |
| 60         | 37       | 3.035   | 0.497                     | 3.025        | 0.456                 | 3.029        | 0.047                 |
| 0          | 88       | 1.086   | 0                         | 3.498        | 0                     | 3.260        | 0                     |
| 10         | 78       | 1.089   | 0.237                     | 3.498        | 0.013                 | 3.260        | 0.007                 |
| 20         | 68       | 1.097   | 0.959                     | 3.498        | 0.024                 | 3.260        | 0.026                 |
| 30         | 58       | 1.105   | 1.712                     | 3.497        | 0.055                 | 3.258        | 0.057                 |
| 40         | 48       | 1.117   | 2.861                     | 3.495        | 0.109                 | 3.259        | 0.016                 |
| 50         | 38       | 1.186   | 9.179                     | 3.490        | 0.236                 | 3.245        | 0.463                 |
| 60         | 28       | 1.240   | 14.12                     | 3.460        | 1.099                 | –            | –                     |

Table 2. Decay times EDT,  $T_{20}$ ,  $T_{30}$  and relative errors  $\Delta_{\text{EDT}}$ ,  $\Delta_{T_{20}}$ ,  $\Delta_{T_{30}}$  versus background noise level  $L_N$  and signal-to-noise ratio. Source frequency of 151 Hz. Observation point in subroom A (upper part of table) and subroom B (bottom part of table).

| $L_N$ [dB] | SNR [dB] | EDT [s] | $\Delta_{\text{EDT}}$ [%] | $T_{20}$ [s] | $\Delta_{T_{20}}$ [%] | $T_{30}$ [s] | $\Delta_{T_{30}}$ [%] |
|------------|----------|---------|---------------------------|--------------|-----------------------|--------------|-----------------------|
| 0          | 81       | 1.554   | 0                         | 2.889        | 0                     | 3.451        | 0                     |
| 10         | 71       | 1.554   | 0.041                     | 2.888        | 0.055                 | 3.451        | 0.003                 |
| 20         | 61       | 1.554   | 0.008                     | 2.890        | 0.011                 | 3.452        | 0.024                 |
| 30         | 51       | 1.569   | 0.949                     | 2.942        | 1.830                 | 3.463        | 0.352                 |
| 40         | 41       | 1.584   | 1.907                     | 3.002        | 3.909                 | 3.490        | 1.123                 |
| 50         | 31       | 1.604   | 3.175                     | 3.118        | 7.926                 | 6.655        | 92.84                 |
| 60         | 21       | 1.537   | 1.080                     | 5.501        | 90.38                 | –            | –                     |
| 0          | 72       | 1.631   | 0                         | 4.246        | 0                     | 3.818        | 0                     |
| 10         | 62       | 1.631   | 0.041                     | 4.249        | 0.070                 | 3.818        | 0.012                 |
| 20         | 52       | 1.630   | 0.050                     | 4.255        | 0.199                 | 3.823        | 0.153                 |
| 30         | 42       | 1.613   | 1.070                     | 4.276        | 0.693                 | 3.863        | 1.192                 |
| 40         | 32       | 1.589   | 2.527                     | 4.334        | 2.071                 | –            | –                     |
| 50         | 22       | 1.687   | 3.478                     | –            | –                     | –            | –                     |
| 60         | 12       | 2.003   | 22.84                     | –            | –                     | –            | –                     |

## 5. Summary and conclusions

Reverberation is the most basic and easily perceived acoustical property of enclosures, therefore one of the most important objectives of room acoustics is an evaluation of decay times from room reverberant responses. However, in real measurements a late sound decay is usually contaminated with background noise, thus an accurate prediction of reverberation times from noisy room responses is the main concern. The problem was examined using a numerical technique where the decay times are estimated from the best straight line fit to the decay function that has been calculated by non-linear polynomial regression from a pressure envelope obtained via the discrete Hilbert transform. A choice

of such a procedure was a consequence of large fluctuations of the pressure envelope arising from noise disturbances and modal interactions which cause the beating effect.

Room responses exploited in a numerical experiment were adopted from sound decay simulations, performed for a coupled room system consisting of two connected rectangular subrooms. As a result of complex room shape and irregular distribution of absorbing material, a nonlinear behaviour of a pressure level decay was noted. The numerical experiment indicated that the background noise only slightly influences the evaluation of reverberation times  $T_{20}$  and  $T_{30}$  as long as the SNR is not smaller than about 25 and 35 dB, respectively. This implies that the proposed method

tolerates clearly smaller dynamic ranges for evaluation of the reverberation times  $T_{20}$  and  $T_{30}$  than is required by the ISO 3382 standard.

In the past, several methods were applied for estimating decay times. A nonlinear iterative regression method for evaluating reverberation times from Schroeder's decay curves was proposed by XIANG (1995). He found that the SNR of 35 dB is a sufficient dynamic range for accurate prediction of the reverberation time  $T_{30}$  and this result is in accordance with the finding of this study. The method of XIANG was extended by XIANG and GOGGANS (2001) to multirate decay functions using Bayesian probability theory and, as was demonstrated, for two coupled rooms Bayesian decay time estimation yields a reliable value when the SNR is higher than 41 dB.

### References

1. ADELMAN-LARSEN N., THOMPSON E., GADE A. (2010), *Suitable reverberation times for halls for rock and pop music*, Journal of the Acoustical Society of America, **127**, 1, 247–255.
2. BARRON M., COLEMAN S. (2001), *Measurements of the absorption by auditorium seating — a model study*, Journal of Sound and Vibration, **239**, 4, 573–587.
3. BEDROSIAN E. (1963), *A product theorem for Hilbert transform*, Proceedings of the IEEE, **51**, 5, 868–869.
4. BERANEK L. (2011), *Concert hall acoustics*, Architectural Science Review, **54**, 1, 5–14.
5. BRADLEY J. (2005), *Using ISO 3382 measures, and their extensions, to evaluate acoustical conditions in concert halls*, Acoustical Science and Technology, **26**, 2, 170–178.
6. CHU W. (1978), *Comparison of reverberation measurements using Schroeder's impulse method and decay-curve averaging method*, Journal of the Acoustical Society of America, **63**, 5, 1444–1450.
7. DÍAZ C., PEDRERO A. (2005), *The reverberation time of furnished rooms in dwellings*, Applied Acoustics, **66**, 8, 945–956.
8. DÍAZ C., PEDRERO A. (2007), *The reverberation time and equivalent sound absorption area of rooms in dwellings*, Noise and Vibration Worldwide, **38**, 6, 12–20.
9. DRAGONETTI R., IANNIELLO C., ROMANO R. (2009), *Reverberation time measurement by the product of two room impulse responses*, Applied Acoustics, **70**, 1, 231–243.
10. GOŁA A., SUDER-DEBSKA K. (2009), *Analysis of Dome Home Hall theatre acoustic field*, Archives of Acoustics, **34**, 3, 273–293.
11. HAHN S. (1996), *The Hilbert transforms in signal processing*, Artech House Inc., Boston.
12. ISO 3382 (2012), *Acoustics – Measurement of room acoustic parameters. Part 1: Performance spaces. Part 2: Reverberation time in ordinary rooms*, International Organization for Standardization, Genève.
13. KAK S. (1970), *The discrete Hilbert transform*, Proceedings of the IEEE, **58**, 4, 585–586.
14. KARJALAINEN M., ANTSALO P., MÄKIVIRTA A., PELTONEN T., VÄLIMÄKI V. (2002), *Estimation of modal decay parameters from noisy response measurements*, Journal of the Audio Engineering Society, **50**, 11, 867–878.
15. LUNDEBY A., VIGRAN T., BIETZ H., VORLÄNDER M. (1995), *Uncertainties of measurements in room acoustics*, Acustica, **81**, 4, 344–355.
16. MEISSNER M. (2007a), *Analysis of non-exponential sound decay in an enclosure composed of two connected rectangular subrooms*, Archives of Acoustics, **32**, 4S, 213–220.
17. MEISSNER M. (2007b), *Computational studies of steady-state sound field and reverberant sound decay in a system of two coupled rooms*, Central European Journal of Physics, **5**, 3, 293–312.
18. MEISSNER M. (2008a), *Influence of absorbing material distribution on double slope sound decay in L-shaped room*, Archives of Acoustics, **33**, 4S, 159–164.
19. MEISSNER M. (2008b), *Influence of wall absorption on low-frequency dependence of reverberation time in room of irregular shape*, Applied Acoustics, **69**, 7, 583–590.
20. MEISSNER M. (2012a), *The discrete Hilbert transform and its application to the analysis of reverberant decay of modal vibrations in enclosures*, Journal of Vibration and Control, **18**, 11, 1595–1606.
21. MEISSNER M. (2012b), *Accuracy issues of discrete Hilbert transform in identification of instantaneous parameters of vibration signals*, Acta Physica Polonica A, **121**, 1A, 164–167.
22. MORGAN D. (1997), *A parametric error analysis of the backward integration method for reverberation time estimation*, Journal of the Acoustical Society of America, **101**, 5, 2686–2693.
23. NUTTER D., LEISHMAN T., SOMMERFELDT S., BLOTTER J. (2007), *Measurement of sound power and absorption in reverberation chambers using energy density*, Journal of the Acoustical Society of America, **121**, 5, 2700–2710.
24. SCHROEDER M. (1965), *New method of measuring reverberation time*, Journal of the Acoustical Society of America, **37**, 3, 400–412.
25. SCHROEDER M. (1996), *The “Schroeder frequency” revisited*, Journal of the Acoustical Society of America, **99**, 5, 3240–3241.
26. XIANG N. (1995), *Evaluation of reverberation times using a nonlinear regression approach*, Journal of the Acoustical Society of America, **98**, 4, 2112–2121.
27. XIANG N., GOGGANS P. (2001), *Evaluation of decay times in coupled spaces: Bayesian parameter estimation*, Journal of the Acoustical Society of America, **110**, 3, 1415–1424.
28. XIANG N., GOGGANS P. (2003), *Evaluation of decay times in coupled spaces: Bayesian decay model selection*, Journal of the Acoustical Society of America, **113**, 5, 2685–2697.

## HRTF Adjustments with Audio Quality Assessments

Shu-Nung YAO<sup>(1)</sup>, Li Jen CHEN<sup>(2)</sup>

<sup>(1)</sup> *School of Electronic, Electrical and Computer Engineering, University of Birmingham*  
Edgbaston, Birmingham, B15 2TT, UK; e-mail: SXY043@bham.ac.uk

<sup>(2)</sup> *Acoustic and Camera Technology Department, Wistron Corporation*  
21F, 88, Sec.1, Hsin Tai Wu Rd., Hsichih  
New Taipei City 22181, Taiwan, R.O.C.; e-mail: arlen.chen@wistron.com

(received July 15, 2012; accepted January 4, 2013)

There are an increasing number of binaural systems embedded with head-related transfer functions (HRTFs), so listeners can experience virtual environments via conventional stereo loudspeakers or headphones. As HRTFs vary from person to person, it is difficult to select appropriated HRTFs from already existing databases for users. Once the HRTFs in a binaural audio device hardly match the real ones of the users, poor localization happens especially on the cone of confusion. The most accurate way to obtain personalized HRTFs might be doing practical measurements. It is, however, expensive and time consuming. Modifying non-individualized HRTFs may be an effort-saving way, though the modifications are always accompanied by undesired audio distortion. This paper proposes a flexible HRTF adjustment system for users to define their own HRTFs. Also, the system can keep sounds from suffering intolerable distortion based on an objective measurement tool for evaluating the quality of processed audio.

**Keywords:** HRTF, PEAQ, cone of confusion, headphones, surround.

### 1. Introduction

The duplex theory (RAYLEIGH, 1907) provides a model for a listener to distinguish the location of a sound source by feeling interaural time differences (ITDs) and interaural level differences (ILDs). However, the model has a problem specifying a unique three-dimensional position, because the positions on the cone of confusion (CHENG, WAKEFIELD, 2001) share the same ITD and ILD cues. Head-related transfer functions (HRTFs) subsume not only ITD and ILD information but also the spectral characteristics, frequency magnitude, and phase responses for perception of a source in a three-dimensional space. In Fig. 1, where  $s(t)$  is the sound coming from the loudspeaker  $M$ ,  $e_l(t)$  and  $e_r(t)$  are the sounds reaching the listener's left and right eardrums, respectively, which can be represented as

$$e_l(t) = s(t) * h_{Ml}(t) \quad (1)$$

and

$$e_r(t) = s(t) * h_{Mr}(t), \quad (2)$$

where “\*” symbolizes the convolution operator;  $h_{Ml}(t)$  and  $h_{Mr}(t)$  denote a pair of head-related impulse responses (HRIRs), the inverse Fourier Transforms of

HRTFs  $H_{Ml}(f)$  and  $H_{Mr}(f)$ , containing the characteristic of sound affected by the head, pinna, shoulder, and torso.

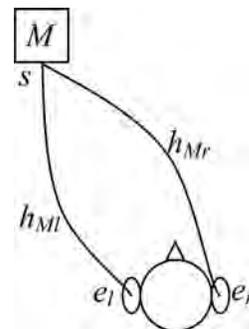


Fig. 1. A pair of HRTFs in the transaural listening:  $s(t)$  – the original mono sound from the speaker  $M$ ,  $h(t)$  – HRIRs,  $e(t)$  – signals at ears.

To date, some labs have measured several kinds of HRTFs by using artificial models (GARDNER, MARTIN, 1994) or real human beings (ALGAZI *et al.*, 2001). However, when users use those databases, they might be unfamiliar with the HRTFs which do not actually belong to them. Then, two negative effects may hap-

pen, the so-called front-back confusions and up-down confusions. Figure 2 illustrates an example of front-back confusions. ITDs are important variables for human beings to locate the positions of sounds, but listeners perceive the same ITD if two sounds are virtually placed in the front hemisphere and symmetrically in the back hemisphere. In this case, listeners tend to confuse the locations of the front sound and the behind sound. On the other hand, Fig. 3 shows an illustration of up-down confusions. When human beings distinguish the elevations of sounds, the peaks and notches of HRTFs in frequency domain are quite important (HEBRANK, WRIGHT, 1974). If the spectrum characteristics of the non-individualized HRTFs stored in 3D audio systems rarely match those of the listener, the listener may have difficulties in locating the correct elevation of the virtual sound, thereby feeling it coming from a position slightly upper or lower than the actual one.

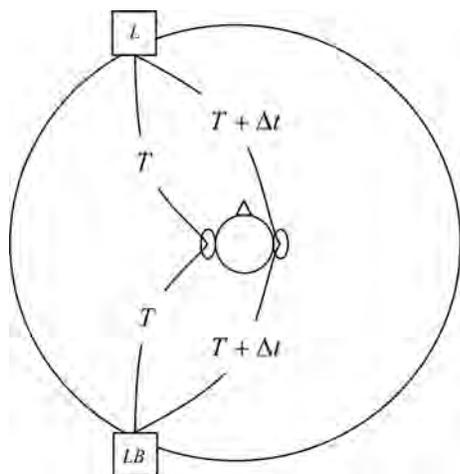


Fig. 2. Front-back confusions:  $T$  – the time delays from speakers to the left ear,  $T + \Delta t$  – the time delays from speakers to the right ear.

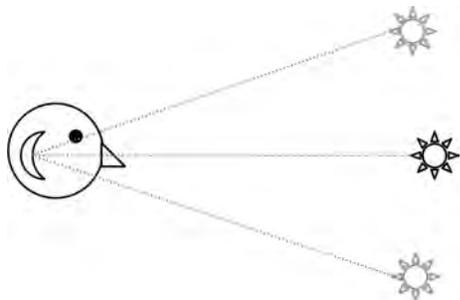


Fig. 3. Up-down confusions.

For medical electronics, DOBRUCKI *et al.* (2010) have introduced an efficient and reliable way to measure HRIRs in the visually impaired. DOBRUCKI and PLASKOTA (2007) have also indicated the measurement of head and ear geometry could provide sufficient

information for modeling a numerical model. For consumer electronics, there have been several kinds of algorithms proposed to help listeners overcome the negative effects without any measurements. TAN and GAN (1998) have set up a 3D sound system by using a high-pass filter, several band-pass filters, and few variable gains, so users can boost or attenuate the frequency components in each filter band which is associated with front, back, up, and down perceptions. GUPTA *et al.* (2002) obtained HRTF modification functions by using a solid sphere with cardboard flaps functioning as a human being's head with ears. ZHANG *et al.* (1998) used weight functions to refine HRTFs, exaggerating the difference between front and back transfer functions. PARK *et al.* (2005) noticed that the weight functions introduced by ZHANG *et al.* (1998) overamplify spectral peaks and notches, and thus proposed more moderate functions. Even if those algorithms enhance the spatial effects, they may also cause distortions in terms of audio quality. Nevertheless, those studies rarely used any objective measurements to evaluate the distortions.

In this paper, we propose an HRTF adjustment system equipped with perceptual evaluation of audio quality (PEAQ), the International Telecommunications Union (ITU) standard for audio quality assessment (ITU-R BS.1387, 1994), so that objective measurements of perceived audio quality can be assessed. Moreover, the proposed system is composed of parametric filters which provide flexibility in the adjustment process.

## 2. Spatial effect enhancement

Based on several psychoacoustic studies (HEBRANK, WRIGHT, 1974; IIDA *et al.*, 2007) and the frequency responses of HRTFs (GARDNER, MARTIN, 1994; ALGAZI *et al.*, 2001), we have observed that the spectrum components in some frequency bands are closely associated with the subjective impression of direction.

HEBRANK and WRIGHT (1974) summarized the following results. Firstly, a sound passing by a 1-octave notch filter with the center frequency located at 7.5 kHz and a high pass filter with the cut-off frequency from 13 kHz to 14 kHz is perceived as a source located directly ahead. Second, a sound filtered by a 1/4-octave peak filter with an 8 kHz center frequency is perceived as a source located straight above. Finally, a peak filter at 11 kHz makes a sound perceived as a source located directly behind.

IIDA *et al.* (2007) found that on the median plane, the peak of an HRTF always happens at about 4 kHz. Therefore, they suggested the peak is the reference information for human beings to analyze other peaks and notches. Moreover, they showed that two notch filters located at 9 and 16 kHz make a source appear as to be behind.

As a result, the adjustment filters for non-individualized HRTFs are designed according to the information in Table 1 which presents the summary of using the special frequency bands related to the subjective impression of direction. Figures 4a, b, and c indicate the magnitude frequency characteristics of the fil-

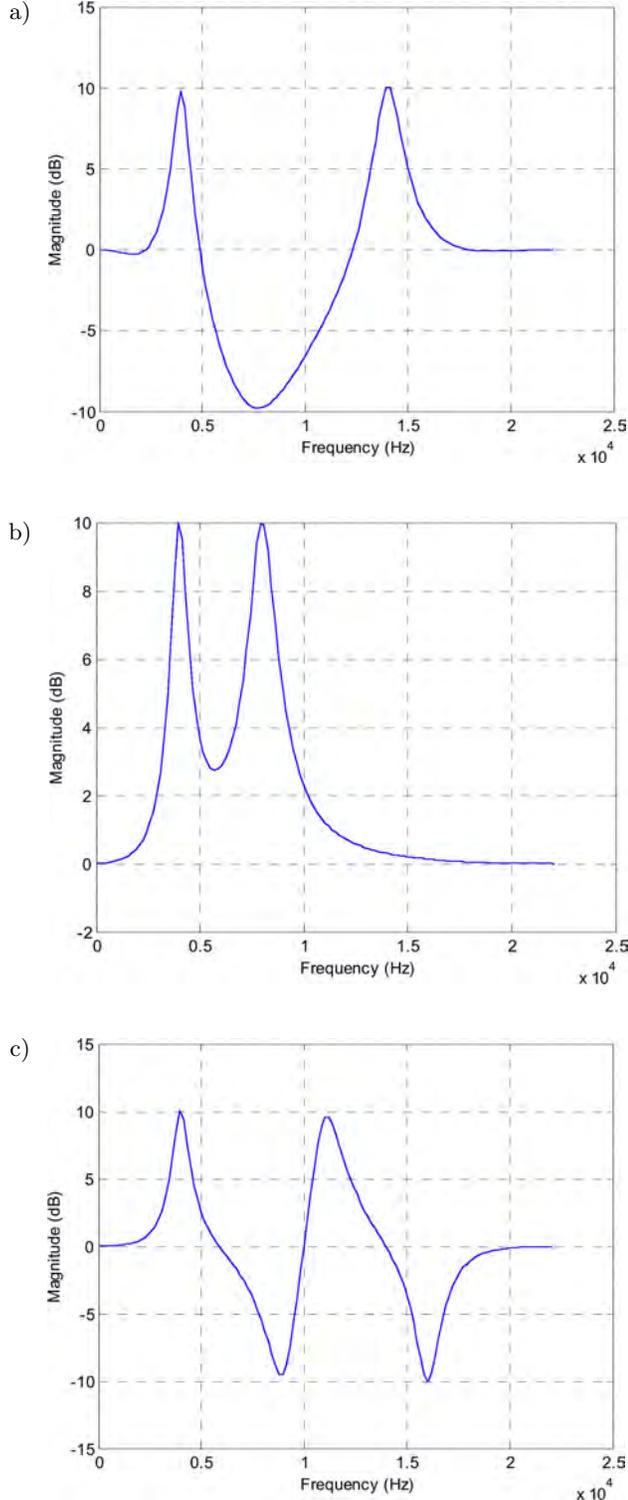


Fig. 4. The magnitude frequency characteristics of the filter structures: a) ahead, b) overhead, and c) behind.

ter structures for making a sound coming from ahead, above, and behind, respectively.

Table 1. The characteristics of filters used in the proposed system.

| “Frontness”  |                  |            |
|--------------|------------------|------------|
| Filter Type  | Center Frequency | Band Width |
| Peak         | 4 kHz            | 1/4 octave |
| Notch        | 7.5 kHz          | 1 octave   |
| Peak         | 14 kHz           | 1/4 octave |
| “Aboveness”  |                  |            |
| Filter Type  | Center Frequency | Band Width |
| Peak         | 4 kHz            | 1/4 octave |
| Peak         | 8 kHz            | 1/4 octave |
| “Behindness” |                  |            |
| Filter Type  | Center Frequency | Band Width |
| Peak         | 4 kHz            | 1/4 octave |
| Notch        | 9 kHz            | 1/4 octave |
| Peak         | 11 kHz           | 1/4 octave |
| Notch        | 16 kHz           | 1/4 octave |

The filter structures are composed of second-order parametric equalizers. The detailed design procedure is described by ZHANG *et al.* (2010), and thus only the results are given in this paper. ZHANG *et al.* (2010) designed three user-defined variables associated with digital audio equalization for octave bands. The first variable is used for determining the radian center frequency  $\omega_c$ . The second variable is for the filter gain,  $G$ . The third variable  $Q$  is used to design the quality factor. Finally, the three parameters form the biquadratic filter transfer function shown as

$$H(z) = \frac{[(1 + M_1) + (M_2 - M_3 - 1)z^{-1} + (M_3 - M_1)z^{-2}]}{[1 + (M_2 - M_3 - 1)z^{-1} + M_3z^{-2}]}. \quad (3)$$

When we design a peak filter, let

$$\begin{aligned} M_1 &= \frac{(G-1)k}{1 + \frac{k}{Q} + k^2}, \\ M_2 &= \frac{4k^2}{1 + \frac{k}{Q} + k^2}, \\ M_3 &= \frac{1 - \frac{k}{Q} + k^2}{1 + \frac{k}{Q} + k^2}, \end{aligned} \quad (4)$$

where  $k = \tan(\omega_c/2)$ .

When designing a notch filter, we choose

$$\begin{aligned} M_1 &= \frac{-(G-1)k}{1 + \frac{kG}{Q} + k^2}, \\ M_2 &= \frac{4k^2}{1 + \frac{kG}{Q} + k^2}, \\ M_3 &= \frac{1 - \frac{kG}{Q} + k^2}{1 + \frac{kG}{Q} + k^2}, \end{aligned} \quad (5)$$

where  $k = \tan(\omega_c/2)$ .

The filter structures are realized by cascading such peak or notch filters. Because of using parametric filters, the magnitude and bandwidth of each peak or notch can be adjusted flexibly. In Fig. 4, we arbitrary tune the magnitude to 10 dB for peaks and  $-10$  dB for notches and the bandwidth, if it cannot be clearly found in references, to  $1/4$  octave.

An example of enhancing non-individualized HRTFs from three directions on the median plane is

shown in Fig. 5. In this example, we assume that the listener's head is perfectly symmetrical, so the left and right ear impulse responses are identical (GARDNER, MARTIN, 1994). The non-individualized HRTFs corresponding to a point ahead, overhead, and behind are respectively filtered by the three filter structures whose frequency responses are shown in Figs. 4a, b, and c. Figure 5 indicates that the effects of HRTFs are exaggerated. That is, by using the proposed adjustment system, the peak values are modified to be higher, while the notch values are lower.

### 3. Audio quality assessment

It is generally accepted that either overamplifying peaks or attenuating notches causes sound distortions (ZHANG *et al.*, 1998; PARK *et al.*, 2005). Therefore, a computer-based objective algorithm, PEAQ, is introduced to analyze the relationship between sound distortions and HRTF enhancements. PEAQ was developed for objective measurement of the perceived audio quality, grading the quality of the audio signals. Unlike traditional objective measurement methods, such as signal-to-noise-ratio (SNR) or total-harmonic-

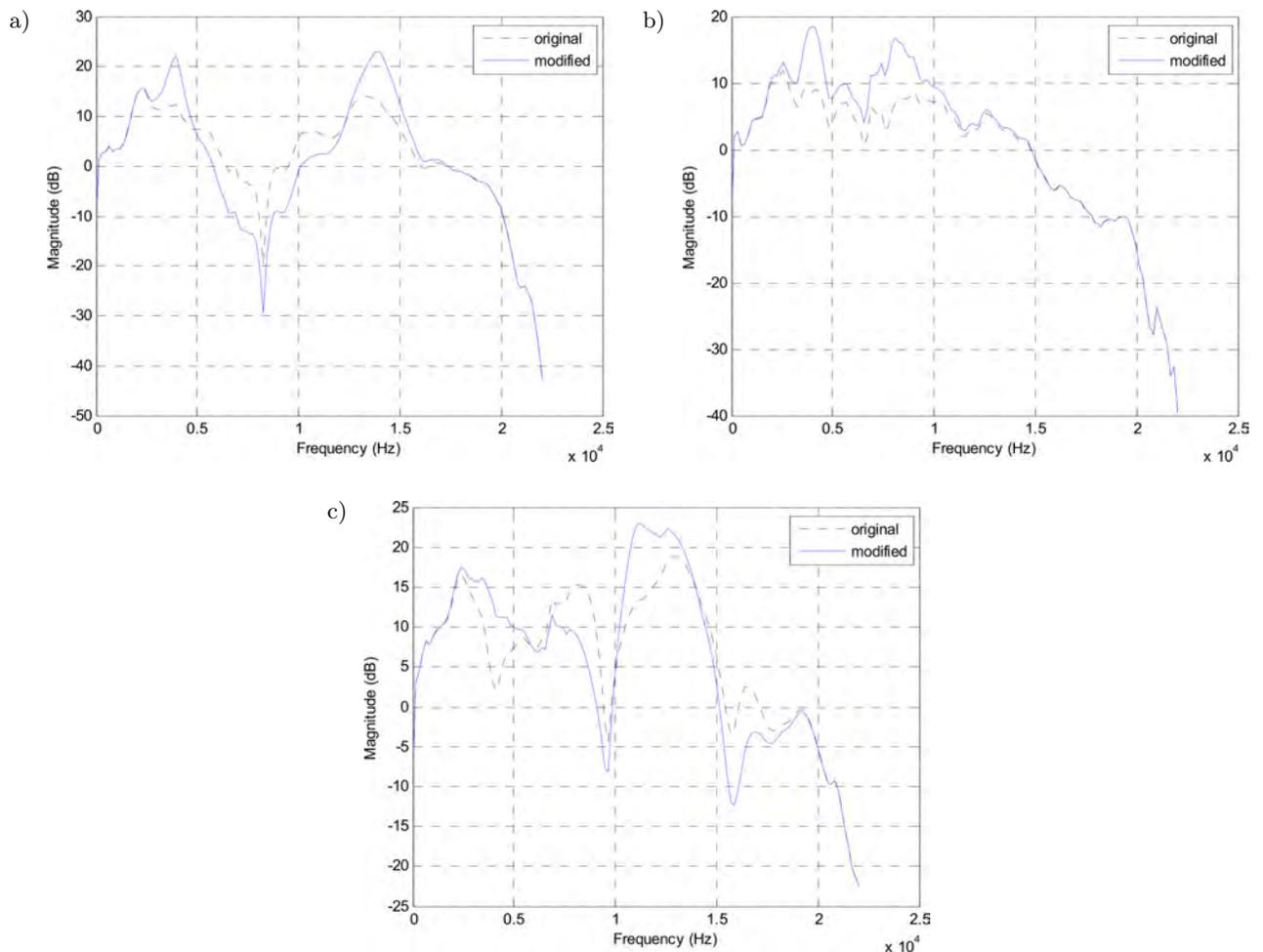


Fig. 5. An example of HRTF adjustment: a) ahead, b) overhead, and c) behind.

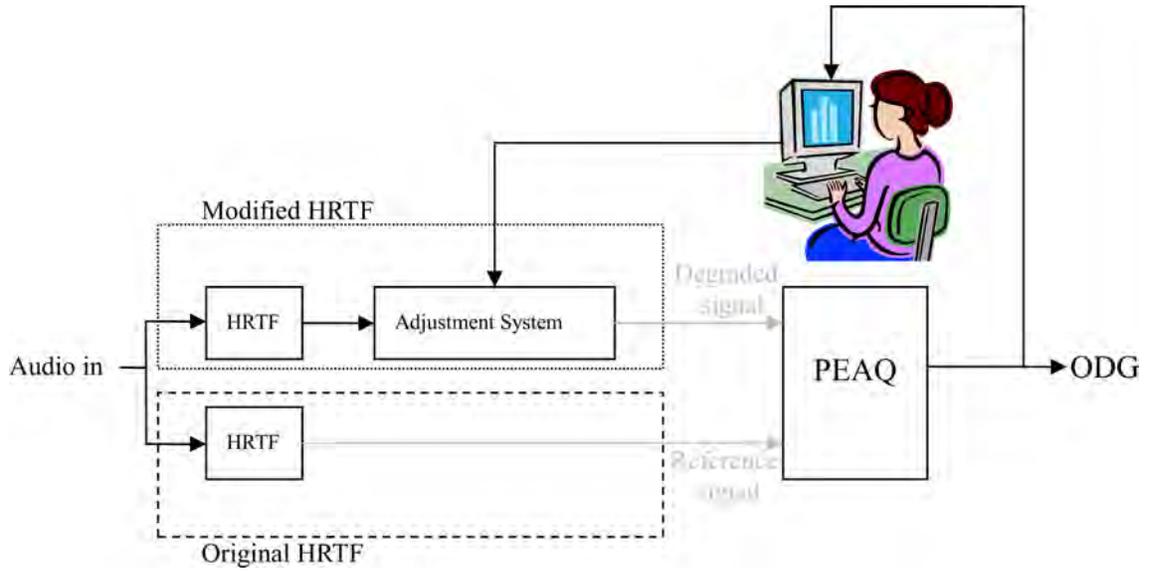


Fig. 6. The measurement scheme by PEAQ.

distortion (THD), PEAQ is designed based on the training of a neural network.

The measurement scheme is shown in Fig. 6. Users can determine the level of enhancement by adjusting the parameters of parametric filters after selecting the non-individualized HRTFs. Then the audio filtered by the modified HRTF will be compared with that filtered by the original one. Finally, overall difference grade (ODG) is generated. Normally, ODG is a value in the range of  $-4$  to  $0$ . As shown in Fig. 7, the more negative the score, the more perceptible the audio distortion. The ODG values provide the information about the tolerance of audio distortion, so HRTFs can be reasonably adjusted by users.

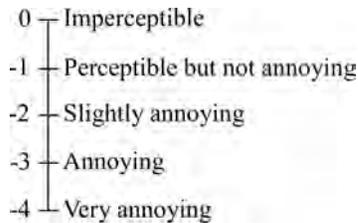


Fig. 7. Five-grade impairment scale.

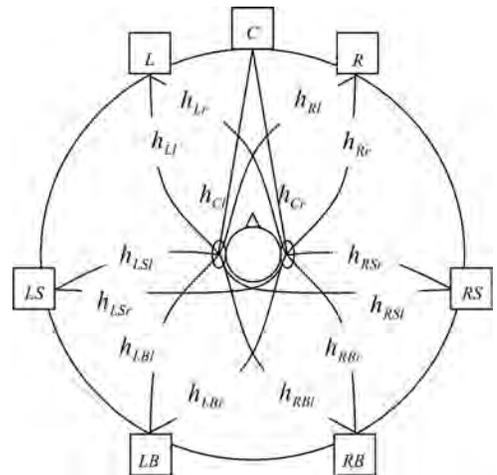
#### 4. Applications

Downmixing is an audio technique used to convert multi-channel surround to stereo format. The performance of downmix equations given by ITU recommendation (1994) had been well discussed (KIN, PLASKOTA, 2011). In addition to using downmix parameters, HEN *et al.* (2008) have exploited HRTFs to reproduce 5-channel audio over headphones. Recently, the conversion from 7-channel to 2-channel format has been proposed (YAO *et al.*, 2011). The main idea is

to extend the situation from Fig. 1 to Fig. 8, so  $e_l(t)$  and  $e_r(t)$  will become (6) and (7), where  $h_{Xy}(t)$  is the HRIR between the loudspeaker  $X$  ( $C, L, R, LS, RS, LB, or RB$ ) and the listener's ear  $y$  ( $l$  or  $r$ ), and  $s_X(t)$  denotes the original audio signal produced by the loudspeaker  $X$ :

$$e_l(t) = s_C(t)*h_{Cl}(t) + s_L(t)*h_{Ll}(t) + s_R(t)*h_{Rl}(t) + s_{LS}(t)*h_{LSl}(t) + s_{RS}(t)*h_{RSl}(t) + s_{LB}(t)*h_{LB l}(t) + s_{RB}(t)*h_{RB l}(t); \quad (6)$$

$$e_r(t) = s_C(t)*h_{Cr}(t) + s_L(t)*h_{Lr}(t) + s_R(t)*h_{Rr}(t) + s_{LS}(t)*h_{LSr}(t) + s_{RS}(t)*h_{RSr}(t) + s_{LB}(t)*h_{LB r}(t) + s_{RB}(t)*h_{RB r}(t). \quad (7)$$

Fig. 8. A 7-channel surround system,  $h$  – HRIRs.

Since the conversion algorithm is based on HRIRs, front-back confusions are expected to happen. To be more specific, the loudspeaker  $L$  and loudspeaker  $LB$

are virtually placed in the front hemisphere and symmetrically in the back hemisphere, respectively. The similar situation can be found by looking into the positions of the loudspeaker  $R$  and loudspeaker  $RB$ . As a result, the proposed HRTF adjustment technique is applied to improve the perceived localization of these four loudspeakers. The experimental results are described in the following section.

## 5. Experiments and results

During the experiments, subjective listening tests are carried out to evaluate spatial effects, while PEAQ is used to assess audio quality. The system is currently being run on a PC with the operating system Windows 7. The functions of HRTF adjustments and audio quality assessments were technically implemented in Matlab programming. The sounds are played via headphones. We compare the experimental results by tuning the magnitude of each peak and notch to three different settings,  $\pm 5$  dB,  $\pm 10$  dB, and then  $\pm 15$  dB.

For subjective listening tests, 15 untrained subjects, 8 males and 7 females, are involved. They are asked to locate white noise filtered through the original HRTFs and the modified ones. Each piece of white noise comes from any of the four positions corresponding to the loudspeaker  $R$ , the loudspeaker  $RB$ , the loudspeaker  $L$ , and the loudspeaker  $LB$  in Fig. 8. The purpose of the subjective tests is to determine the best sound localization of the four settings presented as follows:

Test A – Original HRTFs measured by GARDNER and MARTIN (1994);

Test B – Modified HRTFs by adjusting the peak filter values to 5 dB and the notch filter values to  $-5$  dB;

Test C – Modified HRTFs by adjusting the peak filter values to 10 dB and the notch filter values to  $-10$  dB;

Test D – Modified HRTFs by adjusting the peak filter values to 15 dB and the notch filter values to  $-15$  dB.

There are 24 stimuli in each test. A stimulus is composed of a 1-second pause followed by a 2-second burst of white noise coming from different positions in random order. The same directional white noise was presented 6 times per test. The total duration of each test is 72 seconds. It takes about 5 minutes for each subject to complete the four tests. When a subject correctly recognizes the direction of the sound source, a point is accumulated in this current test. The average scores scaled from 0% to 100% are shown in Fig. 9. The means and standard deviations may appear in connection with the hypothesis that the more exaggeratedly we reshape HRTFs, the more easily listeners can distinguish the locations.

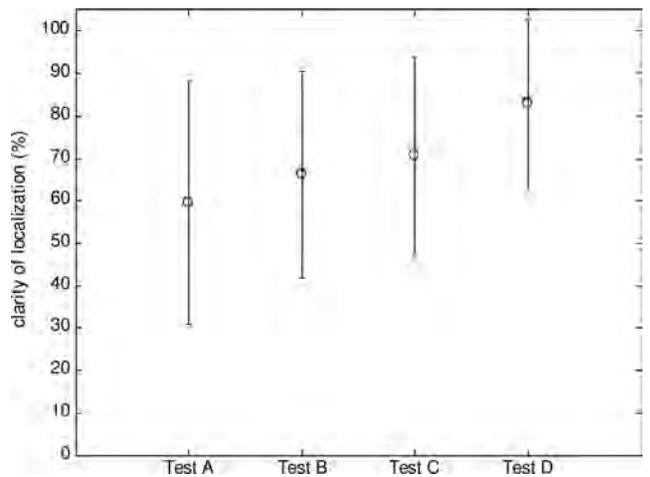


Fig. 9. Perceived clarity of sound locations. The circles are the mean values and the vertical lines symbolize the standard deviation values.

For objective measurement of perceived audio quality, the input signals are several types of audio pieces including string, wind, percussion, and piano music. As shown in Fig. 6, input audio filtered by an original HRTF functions as a reference referring to undistorted signal, while input audio filtered by a modified HRTF is the distorted signal under evaluation. After processing by PEAQ, we summarize the resulting ODGs in Table 2. Through looking into the values of ODGs one can see, as we expect, that the larger the filter gain, the worse the grade.

Table 2. The resulting ODGs by adjusting the magnitude of parametric filters.

| Filter setting | Cello  | Flute  | Cymbals | Piano  |
|----------------|--------|--------|---------|--------|
| $\pm 5$ dB     | -0.522 | -0.357 | -0.909  | -0.288 |
| $\pm 10$ dB    | -1.502 | -1.181 | -1.142  | -1.163 |
| $\pm 15$ dB    | -2.151 | -1.983 | -2.152  | -2.511 |

## 6. Discussion

The proposed filter structures are developed according to spectral cues on the median plane (HEBRANK, WRIGHT, 1974; IIDA *et al.*, 2007), but we find the proposed system still work pretty well when it is applied to the virtual 7-channel surround. This is because positions of loudspeakers  $R$ ,  $RB$ ,  $L$ , and  $LB$  are not far from the median plane, a reason which can be verified by comparing the HRTF characteristics at the positions  $R$ ,  $RB$ ,  $L$ , and  $LB$  and those on the median plane. Figure 10 illustrates an instance of the comparison. Through looking into the magnitude of  $H_{Rl}(f)$ ,  $H_{Rr}(f)$ ,  $H_{Cl}(f)$ , and  $H_{Cr}(f)$  one can see that there is not much difference among the HRTFs.

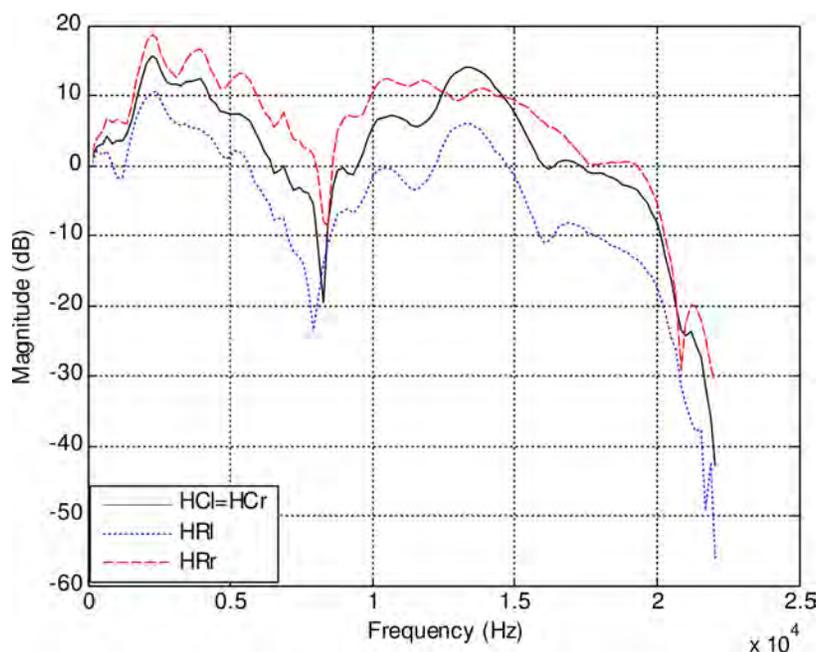


Fig. 10. The HRTFs with elevation  $0^\circ$  and azimuth  $30^\circ$  (the position of the loudspeaker  $R$ ) and the HRTF with elevation  $0^\circ$  and azimuth  $0^\circ$  (the position of the loudspeaker  $C$ ). Detailed description of symbols in text.

Broadly speaking, the experimental results indicate a negative correlation coefficient between the spectral difference in HRTFs and the grade of audio quality, while the correlation coefficient between the spectral difference in HRTFs and the spatial effect of sound is positive. When reinforcing the localization performance, one should be careful about audio distortion. Through looking into the resulting ODGs in Table 2 together with the physical meanings of ODG scale in Fig. 7 it can be seen that the treated audio becomes slightly annoying when HRTF magnitude variations reach  $\pm 15$  dB. On the other hand, the ODGs are in  $\pm 5$  dB setting range between 0 and  $-1$  which means the treatment is perceptible but not annoying.

Although there is a negative relationship between audio quality and localization performance, some of the subjects got a higher score in Test B than in Test D. The possible explanation is that the filter setting used in Test B provides reasonable spectral cues which coincidentally match their real HRTFs, so they are similar to the modified ones. That the unchanged HRTFs match the listener's HRTFs could also happen, but rarely. This can be verified by the high standard deviation in Test A in Fig. 9.

In order to generate the precise spectral cues, the filter setting for the loudspeaker  $L$  and  $LB$  may be different from that for the loudspeaker  $R$  and  $RB$ . This can be illustrated by the following example. Subject No. 7 has no difficulty in identifying the source positions of her left hand side during Test B, as well as the positions of her right hand during Test C, while

the sounds from the right hand side in Test B and the left hand side in Test C lead to poor localization. In this case, we use  $\pm 5$  dB setting to refine the HRTFs,  $H_{Li}(f)$ ,  $H_{Lr}(f)$ ,  $H_{LBl}(f)$ , and  $H_{LBr}(f)$ , and use  $\pm 10$  dB setting to refine  $H_{Rl}(f)$ ,  $H_{Rr}(f)$ ,  $H_{RBl}(f)$ , and  $H_{RBr}(f)$ , a new asymmetric setting which is used for further examination. Upon the extra informal listening tests, the results are quite encouraging, showing better localization performance. Therefore, we hypothesize that asymmetric adjustments sometimes are needed because of the distinct physical characteristics from an individual's left ear to right ear.

## 7. Conclusion

In this paper, we design a system for HRTF customization with an audio quality assessment technique. The HRTF adjustment system is implemented by few parametric filters, so users can flexibly adjust the bandwidth and the magnitude of each filter. Moreover, the computer-based objective algorithm, PEAQ, provides the results of objective measurement for users to evaluate audio quality. As a result, users can clearly assess the trade-off between sound distortion and localization performance. Through subjective localization listening experiments and objective audio quality measurements, the proposed system can improve the sound spatialization in virtual environments without suffering too much annoying sound distortion.

There are some areas of future work concerning the externalization of sounds. The room model chosen for the system and the effects caused by the reflections

and reverberation on sound distortion and localization performance will be investigated. For headphone-based spatial sound, the sense making the sound outside of the head will be of great interest.

### Acknowledgments

The authors would like to thank the anonymous reviewers for their constructive comments. The authors would also like to thank Tim Collins and Peter Jančovič for their guidance and helpful advice in this study.

### References

1. ALGAZI V. R., DUDA R. O., THOMPSON D. M., AVENDANO C. (2001), *The CIPIC HRTF database*, Proceedings of IEEE workshop Applcat. Signal Process. Audio Acoust., 99–102.
2. CHENG C. I., WAKEFIELD G. H. (2001), *Introduction to head-related transfer functions (HRTFs): representations of HRTFs in time, frequency, and space*, J. Audio Eng. Soc., **49**, 4, 231–249.
3. DOBRUCKI A. B., PLASKOTA P. (2007), *Computational modelling of head-related transfer function*, Archives of Acoustics, **32**, 3, 659–682.
4. DOBRUCKI A. B., PLASKOTA P., PRUCHNICKI P., PEC M., BUJACZ M., STRUMILLO P. (2010), *Measurement system for personalized head-related transfer functions and its verification by virtual source localization trials with visually impaired and sighted individuals*, J. Audio Eng. Soc., **58**, 9, 724–738.
5. GARDNER B., MARTIN K. (1994), *HRTF Measurements of a KEMAR Dummy-Head Microphone*, MIT Media Lab.
6. GUPTA N., BARRETO A., ORDONEZ C. (2002), *Spectral modification of head-related transfer functions for improved virtual sound spatialization*, Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, 1953–1956.
7. HEBRANK J., WRIGHT D. (1974), *Spectral cues used in the localization of sound sources on the median plane*, Journal of the Acoustical Society of America, **56**, 1829–1834.
8. HEN P., KIN M. J., PLASKOTA P. (2008), *Conversion of stereo recording to 5.1 format using head-related transfer functions*, Archives of Acoustics, **33**, 1, 7–10.
9. IIDA K., ITOH M., ITAGAKI A., MORIMOTO M. (2007), *Median plane localization using a parametric model of the head-related transfer function based on spectral cues*, Applied Acoustics, **68**, 8, 835–850.
10. ITU-R BS.775-1 (1994), *Multichannel stereophonic sound system with and without accompanying picture*, ITU, Geneva.
11. ITU-R REC. BS.1387 (1998), *Method for objective measurement of perceived audio quality*, ITU, Geneva.
12. KIN M. J., PLASKOTA P. (2011), *Comparison of sound attributes of multichannel and mixed-down stereo recordings*, Archives of Acoustics, **36**, 2, 333–345.
13. PARK M.-H., CHOI S.-I., KIM S.-H., BAE K.-S. (2005), *Improvement of front-back sound localization characteristics in headphone-based 3D sound generation*, Proceedings of IEEE International Conference on Advanced Communication Technology, 273–276.
14. RAYLEIGH L. (1907), *On our perception of sound direction*, Philosoph. Mag., **13**.
15. TAN C.-J., GAN W.-S. (1998), *User-defined spectral manipulation of HRTF for improved localisation in 3D sound systems*, Electronics Letters, **34**, 25, 2387–2389.
16. YAO S. N., COLLINS T., JANCOVIC P. (2011), *A dual-mode architecture for headphones delivering surround sound: low-order IIR filter models approach*, Proceedings of IEEE International Symposium on Consumer Electronics, 62–66.
17. ZHANG M., TAN K.-C., ER M. H. (1998), *A refined algorithm of 3-D sound synthesis*, Proceedings of IEEE International Conference on Signal Processing Proceedings, 1408–1411.
18. ZHANG X., ZHANG R., CHEN W. (2010), *Design of digital parametric equalizer based on second-order function*, Proceedings of IEEE International Conference on Image Analysis and Signal Processing, 182–185.

# Acoustic Features of Filled Pauses in Polish Task-Oriented Dialogues

Maciej KARPIŃSKI

*Institute of Linguistics, Adam Mickiewicz University*  
al. Niepodległości 4, 61-874 Poznań, Poland; e-mail: maciej.karpinski@amu.edu.pl

(received October 8, 2012; accepted January 15, 2013)

Filled pauses (FPs) have proved to be more than valuable cues to speech production processes and important units in discourse analysis. Some aspects of their form and occurrence patterns have been shown to be speaker- and language-specific. In the present study, basic acoustic properties of FPs in Polish task-oriented dialogues are explored. A set of FPs was extracted from a corpus of twenty task-oriented dialogues on the basis of available annotations. After initial scrutiny and selection, a subset of the signals underwent a series of pitch, formant frequency and voice quality analyses. A significant amount of variation found in the realisations of FPs justifies their potential application in speaker recognition systems. Regular monosegmental FPs were confirmed to show relatively stable basic acoustic parameters, which allows for their easy identification and measurements but it may result in less significant differences among the speakers.

**Keywords:** filled pauses, paralinguistics, dialogue, acoustic properties, Polish.

## 1. Filled pauses as paralinguistic components of spoken utterances

Paralinguistic features of utterances (henceforth PLFs) were extensively explored by linguists early in the 1960s (TRAGER, 1960, 1961, 1964; CRYSTAL, 1963, 1966, 1974, 1975). Since then, they have been acknowledged as an indispensable component of spoken communication. They may reveal important facts on the speaker him/herself, including his/her age, gender, origin, social background or education, as well as his/her present emotional state, attitude towards the topic of conversation or towards the conversational partner (GOBL, NI CHASAIDE, 2003; WALLBOTT, SCHERER, 1979; PAKOSZ, 1982; BORTFELD, 2001). Some of them may also provide cues in the analysis of the process of speech production (e.g., disfluencies (FROMKIN, 1971, 1973)). PLFs often contribute to an individual, idiosyncratic speaking style but most people are also able to control many of them consciously.

As PLFs form an extraordinarily heterogeneous group, it is difficult to cover all of them with a single definition. Moreover, they can be defined only as precisely as precise the boundaries of language can be. As a result, researchers tend to define them by enumeration and the inventories proposed in literature are often selective. Among the most frequently mentioned ex-

amples of PLFs, there are prosodic features and voice quality parameters (CRYSTAL, 1966). Some of them can be relatively easily measured as local acoustic parameters of speech signal (energy, pitch frequency). Others can be determined only when longer stretches of speech are analysed. Finally, some of them have more “structural” nature. For example, in an emotional utterance, words may be sequenced in an atypical way, accentuation may be also uncommon. But this can be only noticed once the entire utterance is taken into account and compared to some other typical or common structures.

In recent decades, many studies of paralinguistic features have been focused on potential cues to the emotional or attitudinal value of utterances in psychological and psycholinguistic (e.g., LADD, 1985; PATEL *et al.*, 2011; JOHNSTONE, SCHERER, 2000), communicational (e.g., BENUŠ, 2009; GRAVANO, 2011) or technological (CAMPBELL, 2002; TEN BOSH, 2003) contexts. Social and intercultural studies of paralinguistic features have also brought interesting results (ABELIN, ALWOOD, 2000; DODDINGTON, 2001; ABELIN, 2008; BURKHARDT *et al.*, 2006; SCHERER *et al.*, 2000; CAMPBELL, 2004, 2007), especially in relation to emotionality. But in everyday communicative situations, one rarely faces vividly emotional speech. Subtle background emotions and the mixtures of emotions driv-

ing human everyday behaviour may be extremely difficult to detect and decode. Another area of research related to the form and function of PLFs has been speaker identification and recognition. The variation of the acoustic form of utterances poses additional problems in the area of speech recognition, but it makes automatic speaker recognition and identification possible (MARY, YEGNANARAYANA, 2008; SCHÖTZ, 2002), especially – but not solely – in forensic applications (GONZALES-RODRIGUEZ, 2008; SACKS, KOEHLER, 2005). The importance of paralinguistic features of utterances in language communication fully justifies attempts towards the formulation of spoken language grammar that would encompass paralinguistic information (CAMPBELL, 2002) as well as attempts towards a comprehensive theory of disfluencies (SHRIBERG, 1994).

Silent and filled pauses (SPs and FPs, respectively) seem to be especially frequently explored categories of paralinguistic phenomena. It was found very early in the studies on spontaneous speech that FPs may signal problems in the process of lexical access (MACLAY, OSGOOD, 1959; GOLDMAN-EISLER, 1968) as well as in syntactic processing (BOOMER, 1965). More recent views on the origin and role of FPs are presented, for example, by SHRIBERG (2001) or WARD (2005). Their distribution, length and form say much about the pragmatic aspects of utterances as well as about the speaker him/herself. There is no doubt that pauses have certain communicative value (e.g., SAVILLE-TROIKE, 1985; LOCAL, KELLY, 1986; NISHINUMA, HAYASHI, 2004). They may function as markers of discourse structure (SWERTS, 1998) and can be consciously used as a stylistic device but they may also reveal a peculiar, idiosyncratic speaking style. As opposite to monologue speech, in dialogues, their occurrence may also result from a wide variety of interaction-related factors, e.g. processing input from the conversational partner, formulating a reply, waiting for an appropriate moment for turn transition, as well as from other aspects of alignment tendencies.

FPs may be realised in a variety of ways – not only as centralised vowels or “creaky sounds” but also as full lexical units (sometimes taboo words) that are, however, used for purposes different from exploiting their lexical meaning. In such cases, they may be marked by a peculiar pronunciation, involving atypical lengthenings or pitch changes.

In their study of eight languages, CANDEA *et al.* (2005) found that non-lexical, vocalic fillers were not language-specific in terms of duration or pitch frequency but rather in terms of vocalic quality and segmental structure. Similar findings were reported by VASILESCU *et al.* (2004, 2005) who pointed to some language specific differences in the values of  $f_1$  and  $f_2$ . STEPANOVA (2007) found significant inter-speaker variation in the realisations of FPs in Russian. DUEZ

(2001) reports the  $f_0$  values of filled-pause onsets to be stable within the same speaker as they are “linked to the absolute, physiological aspects of speech”.

These and some other studies show that (1) FPs are relatively easy to detect in the stream of speech on the basis of their acoustic properties; (2) FPs show certain speaker-specific features, both in terms of occurrence patterns and acoustic realisations. As a consequence, FPs are potentially useful in automatic speaker identification or recognition.

Early studies of FPs and SPs in semi-spontaneous Polish utterances (FRANCUZIK *et al.*, 2002; KARPIŃSKI, 2006, 2007) not only show differences in the distributions of their durations (which is intuitively obvious as SPs – unlike FPs – can be arbitrarily long) but also their co-occurrence patterns. While the form of their realisation has been briefly discussed in (KARPIŃSKI, 2008), their acoustic properties have never been studied in depth for the Polish language.

## 2. The aims of the study

In the present study, selected acoustic properties of FPs in Polish are explored in order to find the areas of individual differences that can be potentially useful in speaker identification and recognition or just as markers of individual speaking style.

Speaker recognition and identification are intensively developing areas of research due to their potential applications in security systems, access management systems and forensics (BEIGI, 2011a, 2011b). Each human has an uniquely shaped vocal tract which contributes to the individual acoustic features of voice. Still, it is not easy to define the “vocal print” (voice print) which would be a set of acoustic features that allow for unambiguous identification of the speaker. Most researchers seem to follow the path of gathering possibly numerous measurements of parameters that can be extracted from the acoustic signal and then, with the use of advanced statistical methods, looking for their most efficient combinations and hierarchies that can be used in the procedure of speaker identification or recognition (BEIGI, 2011b). However, as the efficiency of such an approach can be significantly limited by the quality of available voice recordings, another path is to include idiosyncratic structural-linguistic properties of utterances – e.g., the frequency of words or phrases, or the way of building sentences (e.g., DODDINGTON, 2001; SHRIBERG, 2007) but also any potentially speaker-specific paralinguistic features and phenomena, including SPs and FPs.

The analyses are based on dialogue recordings as conversational speech shows peculiar properties and its paralinguistic profile may significantly differ from what is found in monologues. In general, one may expect more disfluencies (including FPs) than in monologues as the cognitive load related to the interactivity

normally seems to be higher than in the case of monologue where the speaker may prepare some portions of her/his talk beforehand and does not face interruptions nor fighting for the floor. Here, instrumental analyses are focused on the pitch frequency, first two formant frequencies as well as on jitter and shimmer measures in simple monosegmental fillers. Detailed occurrence patterns of FPs are not covered by the present study because the data in hand are still too sparse and the number of factors related to interactivity and dialogue-specific environment is too high to build a comprehensive model.

### 3. Material under study: DiaGest2 Corpus

DiaGest2 multimodal corpus (e.g., KARPIŃSKI, JARMOŁOWICZ-NOWIKOW, 2010) consists of twenty task-oriented “origami” dialogue session recordings, each of approximately five minute duration. The task of the participants was to reconstruct a figure made of paper. The figure was fully visible to the instruction giver (IG) and not visible to the instruction follower (IF) who was provided with all the necessary materials for its reconstruction. In ten sessions, IF and IG faced and could see each other (mutual visibility condition, MVC), while in another ten sessions they could not see each other (limited visibility condition, LVC). IGs were gender balanced (both in the MVC and LVC), while most of the IFs were females. Each participant took part in the task only in one of the conditions and only in one role (IF or IG).

FPs were manually tagged in the process of transcription and segmentation using Praat (BOERSMA, WENINK, 2012), both in the word and syllable tiers, and automatically extracted from dialogue recordings. As the number of utterances and FPs produced by IFs was much lower, only the utterances by IGs are analysed here. IFs produced less utterances and most of them were much shorter and built of shorter phrases than in the case of IGs so there were less filler-evoking contexts. The total numbers of IG’s fillers of FPs in the MVC and LVC sessions was almost equal (Table 1), although there were clear individual differences among speakers. The total number of FPs produced by IGs ranged from 7 to 61.

Table 1. The frequencies of FPs in the DiaGest2 corpus.

| Signals                        | MVC | LVC | Sum |
|--------------------------------|-----|-----|-----|
| Extracted                      | 255 | 256 | 511 |
| Female                         | 143 | 142 | 285 |
| Male                           | 112 | 114 | 226 |
| Excluded from further analyses | 82  | 41  | 123 |

A significant proportion of FPs had to be excluded from further instrumental analyses after first audition. Many of them included noises or overlapped

speech, and some were realised as voiceless, breathy sounds. Those realised in creaky voice (ca. 3%) and as nasal(ised) sounds (ca. 5.3%) were also rejected. Finally, some of them were compound sequences of sounds which did not match the analytical framework adopted for this study. As a result, the initial set of signals was significantly reduced but it still allowed for a number of measurements and analyses. This limitation has also its advantages: “Clean” FPs can be more easily traced and extracted automatically so the set under study is more compatible with the potential input data for FPs-based speaker recognition computer systems.

Although only the data from IGs will be analysed in this study, in Fig. 1, the “FPs per syllable” rate is shown for both IGs and IFs in the MVC (calculated by dividing the number of FPs by the number of syllables). On average, the rates are higher for IGs which can be easily justified by a higher complexity of this role. However, the data show that there can be strong individual variation in the FP per syllable rate (e.g., see Session 9 and 10). In any case, the proportions reflect the fact that the cognitive load of this particular task was quite different for IGs and IFs. IGs produced more complex phrases, forming longer sequences while IFs could focus on understanding, guessing and giving feedback.

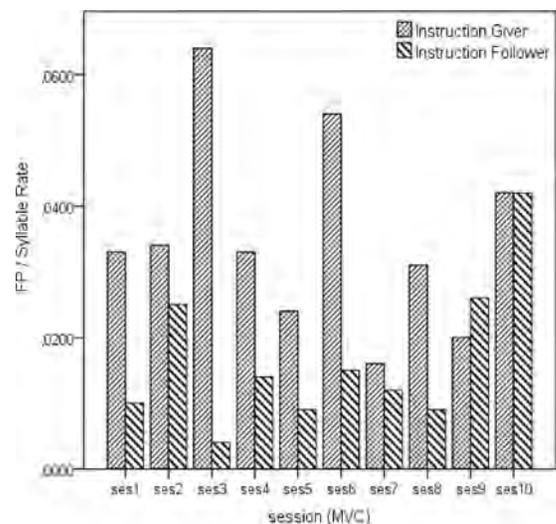


Fig. 1. “Filled pauses per syllable” rates for IGs and IFs in the MV condition.

### 4. Acoustic properties of FPs

Acoustically homogenous FPs are relatively easy to track down automatically in the stream of speech which makes them convenient for automatic sampling and analysis. Although they seem to be relatively stable in terms of pitch frequency and voice quality, they may still have enough of individual variance for speaker identification or recognition. Speakers use idiosyncratic “lexical” fillers so they may also tend to shape their

non-lexical fillers volitionally in a peculiar way. Obviously, the acoustic characteristics of fillers as mostly vocalic sounds must be strongly influenced by the unique shape of each vocal tract.

#### 4.1. Durations of FPs

The durations of fillers vary extremely. The distribution presented in Fig. 2 shows covers only the analysed subset of “clean” signals. It may be slightly different from a distribution of the entire set of FPs as it does not include compound fillers – not very numerous

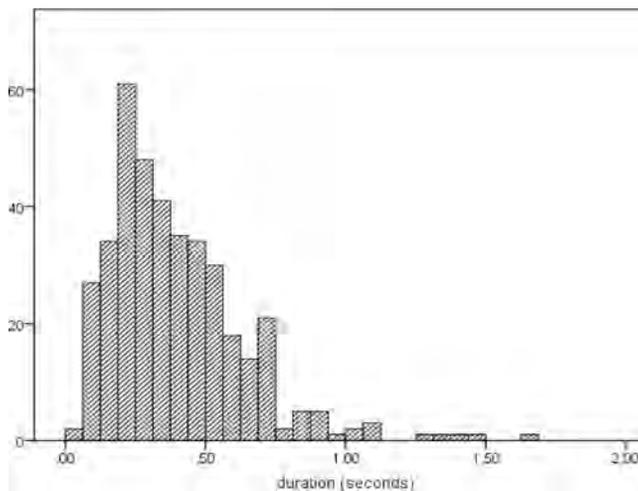


Fig. 2. The distribution of FP's duration in the utterances produced by IGs.

but, as a rule, being much longer. The distribution of FPs duration was similar to that obtained by FRANCUZIK *et al.* (2002). It was also skewed to the left but here the maximum is around 200–300 ms while in the cited work it reached ca. 500 ms.

The difference between the mean values of FP duration in female and male speakers turned out to be statistically insignificant ( $p > 0.05$ ). However, one-way Anova performed on log-transformed data (to compensate for the skew) showed significant differences between the mean durations of FPs for individual speakers ( $F = 3.43$ ,  $p < 0.01$ ;  $df = 18$  as one of the male voices has a very limited number of observations and was excluded from this and further Anova calculations).

#### 4.2. Pitch frequency in FPs

The values of the mean pitch frequency averaged for the entire gender groups and visibility conditions are presented in Table 1. For female speakers, pitch frequency of FPs was significantly higher in the MVC (t-test,  $p = 0.001$ ) and there was no such a difference for males. While the number of subjects does not allow for brave hypothesising, one may understand this result as not necessarily intuitive. One of possible explanations may be that the mutual visibility condition evokes more emotional speech in females (but not in males) and leads to higher pitch values. In Table 2, pitch range data are presented for both female and male speakers in both the conditions.

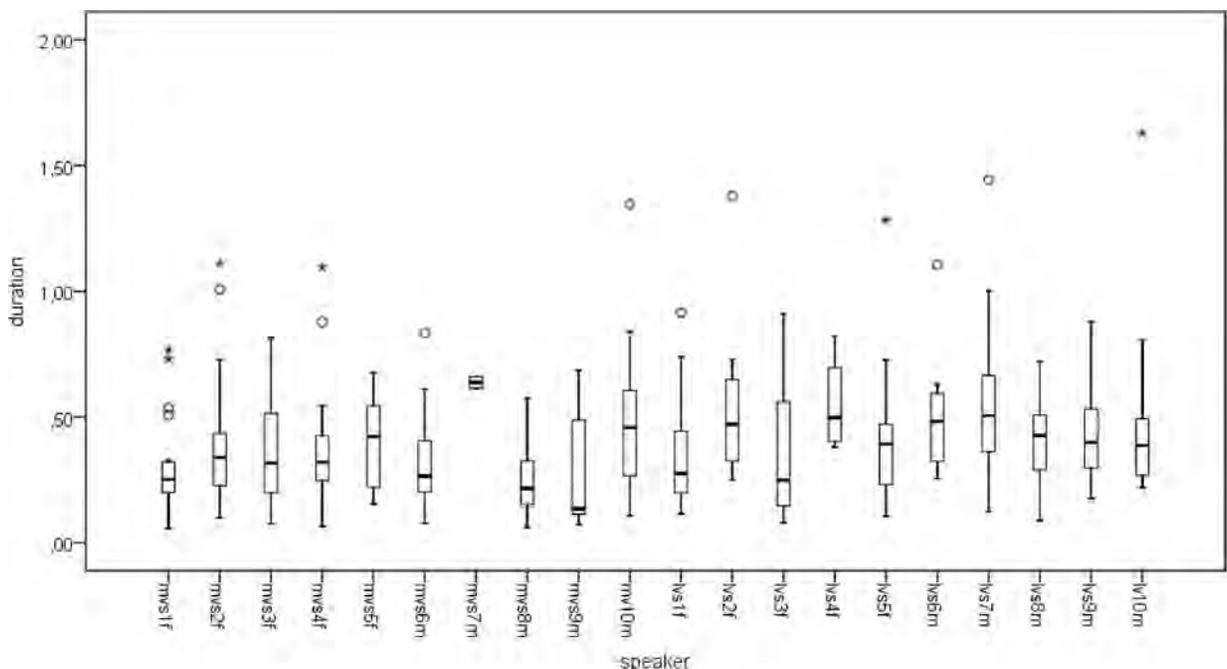


Fig. 3. Individual variation in FP duration. Speakers grouped according to the recording condition (mv – mutual visibility condition, lv – limited visibility condition, f – female, m – male).

Table 2. Mean pitch frequency in female and male voices in both the conditions (minimum and maximum values calculated directly from individual signals).

| condition                        | MVC    |      | LVC    |      |
|----------------------------------|--------|------|--------|------|
| gender                           | female | male | female | male |
| average pitch frequency [Hz]     | 229    | 124  | 219    | 141  |
| standard deviation (pitch freq.) | 22     | 17   | 26     | 22   |
| maximum pitch frequency [Hz]     | 278    | 193  | 120    | 66   |
| minimum pitch frequency [Hz]     | 169    | 84   | 317    | 198  |

As it can be seen in Fig. 5, the mean relative pitch change is also a parameter which may take very individual values (one-way Anova, IGs in two conditions,

Table 3. Relative pitch frequency range for male and female speakers in the two conditions (MVC and LVC) calculated as  $(f_{0\max} - f_{0\min})/f_{0\max}$ .

| condition                      | MV     |      | LV     |      |
|--------------------------------|--------|------|--------|------|
| gender                         | female | male | female | male |
| mean pitch frequency change    | 0.04   | 0.02 | 0.07   | 0.04 |
| standard deviation             | 0.07   | 0.04 | 0.08   | 0.07 |
| maximum pitch frequency change | 0.30   | 0.28 | 0.42   | 0.43 |

$p < 0.001$ ). On the other hand, the majority of mean values are in a limited range. There are more “wide range” speakers among females and there is a signi-

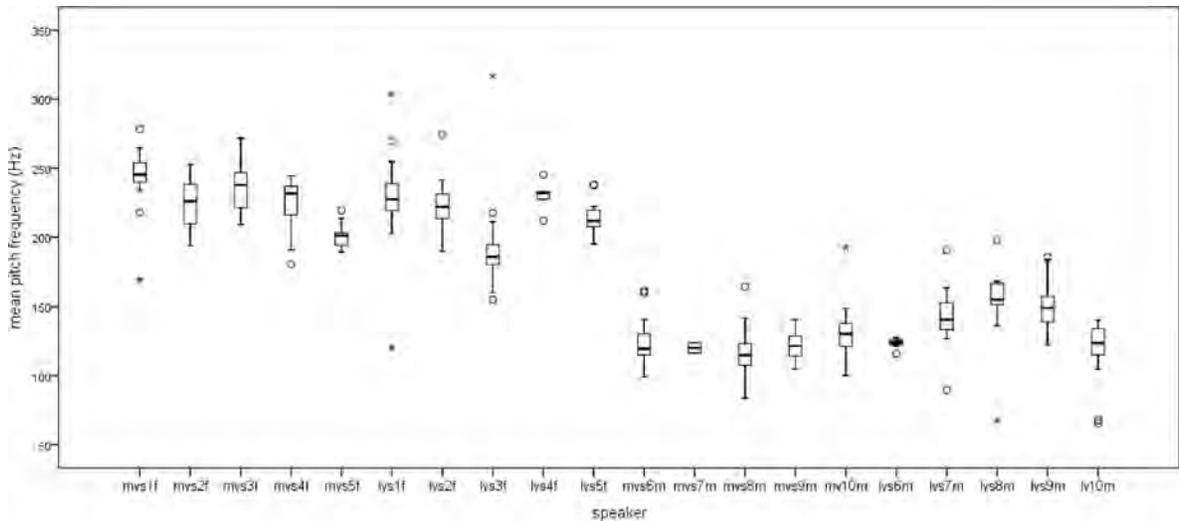


Fig. 4. Mean pitch frequency in FPs realised by twenty speakers (speakers grouped by gender; mv and lv stand for MVC and LVC, respectively).

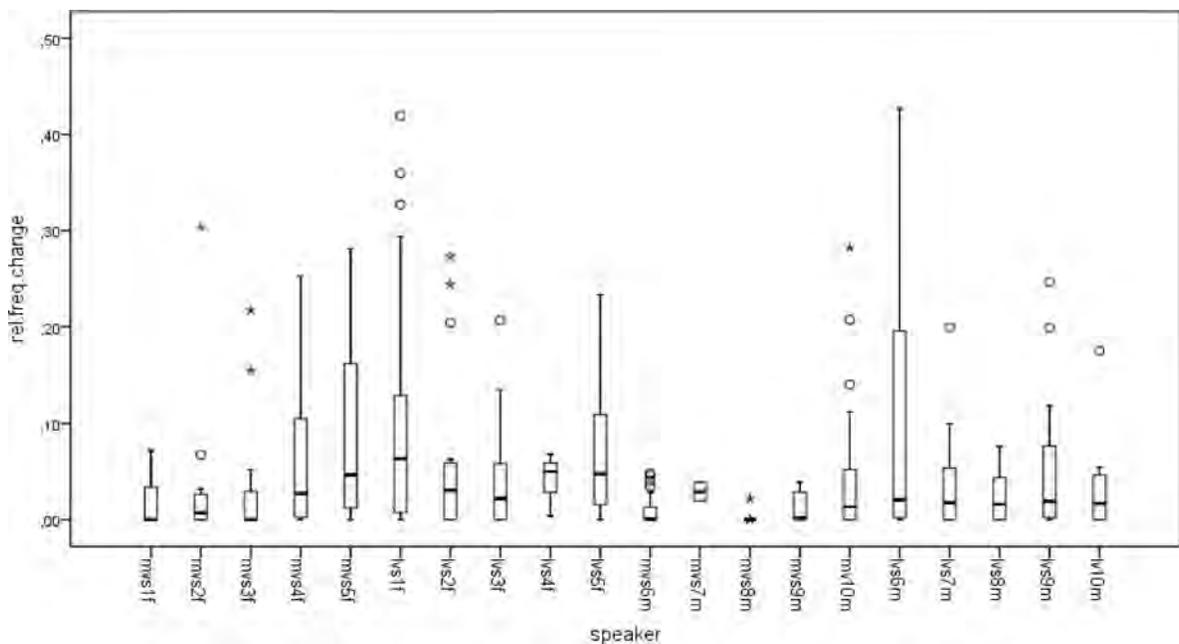


Fig. 5. Mean pitch frequency range within FPs for all the speakers (grouped by gender).

ificant difference between the means for females and males ( $p < 0.001$ ). However, there is no significant difference between the values both for female and male speakers in the two conditions ( $p > 0.05$ ).

#### 4.3. Voice quality in FPs

Voice quality is often understood as a subjective perceptual parameter and, as such, it is difficult to express in terms of acoustic measures of the signal itself. More precisely, it was defined as the auditory coloring of voice, derived from a variety of laryngeal and supralaryngeal features (ABERCROMBIE, 1967, p. 91; LAVER, 1980, p. 1; TRASK, 1996, p. 381; KELLER, 2005). Still, it remains a confusing term as it encompasses phenomena of various origin and character. Voice quality in FPs has been found to be language specific (SHRIBERG, 2001, p. 156). It has been also demonstrated to have some functions in dialogue interaction (LOCAL, KELLY, 1986; OGDEN, 2001). Among the most frequently used acoustic parameters related to voice quality are jitter (corresponding to the variation in the pitch frequency) and shimmer (corresponding to the variation of amplitude). Both are measures of phonation irregularities and their high values may signal pathology. In a number of studies, the usefulness of jitter and shimmer measurements in speaker recognition or identification has been indicated (FARRÚS, HERNANDO, EJARQUE, 2007; FARRÚS, HERNANDO, 2009). They have been shown to be related to the speaker's age, gender, smoking habits, and some other features (LUDLOW *et al.*, 1982).

One can assume that voice quality parameters may change during an utterance, depending on the phonetic fundament and context (segmental, suprasegmental) and as well as possible emotional factors or even changes in the articulatory effort in longer stretches of speech. Still, in the case of monosegmental non-lexical fillers one may expect them to be relatively stable (AUDHKHASI *et al.*, 2009). Therefore, they may be especially precious as a source of data on basic, presumably speaker-dependent voice quality parameters.

In the present study five jitter (local, local absolute, rap, ppq5, DDP) and six shimmer measures (local, local dB, DDA apq3, apq5, apq11) were extracted using Praat. Their basic descriptions can be found, e.g., in (FARRÚS *et al.*, 2007) or (RUSZ *et al.*, 2011), as well as in Praat Manual. Although the measurements were carried out on a reduced set of preselected, “clean” signals, the number of analysis errors was significant, especially for apq5 and apq11 shimmer. As a consequence, only some of the measured variables were taken into account in further steps and, in some cases, two of the speakers were excluded as having too few FPs that could be analysed.

Among jitter measures, only the mean values of local and local absolute jitter were found significantly

different for female and male speakers ( $p < 0.05$ ;  $df = 18$  as one of the voices was excluded). For shimmer, only the mean values of shimmer DDA were significantly different for the two genders ( $p < 0.05$ ). One-way Anova showed significant differences in the mean values for individual speakers in the case of local absolute jitter ( $p < 0.01$ ) as well as for local and PPQ5 jitter ( $p < 0.05$ ). Similar Anova tests were conducted for four shimmer measures but the differences among the means for respective subjects were significant only for DDA and APP3 ( $p < 0.05$ ).

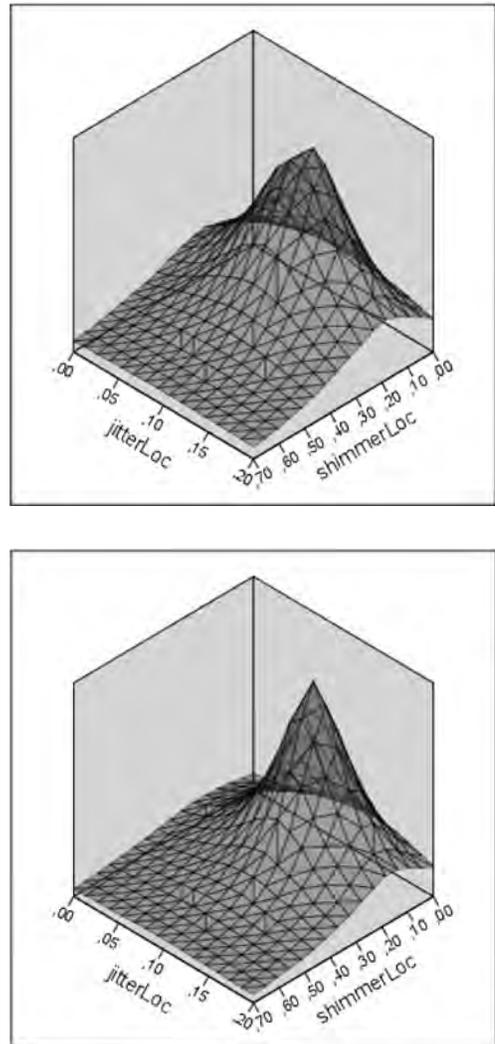


Fig. 6. The distribution of local jitter and shimmer values for female (top panel) and male (bottom panel) speakers.

#### 4.4. Formant frequencies in FPs

Formant frequencies are major cues to the identification of vocalic segments (ROSNER, PICKERING, 1994). For example, ALBALÁ *et al.* (2009) show the variation of formants in vowels in Spanish that can be employed for the purpose of speaker identification. As most of the FPs under study are vocalic, their

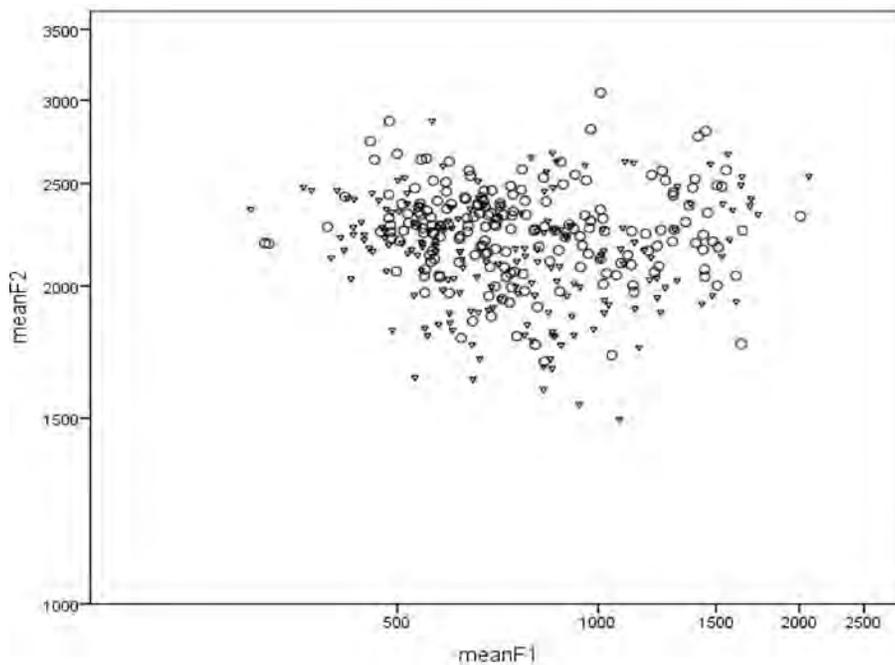


Fig. 7. Formants ( $f_1$  and  $f_2$ ) in FPs produced by female (circles) and male (triangles) speakers.

formant frequencies may bear some important information on their form. VASILESCU *et al.* (2005) argue for potentially language-specific character of  $f_1/f_2$  formant values in FPs. STEPANOVA (2007) finds that FPs in Russian differ in terms of  $f_1$  and  $f_2$  values from vowels in stressed positions. She also points to speaker-specific characteristics of FPs and argues that it makes them useful in speaker recognition. Many methods used for the identification of FPs in the stream of speech are based on the measurements of formant values (AUDHKHASI *et al.* 2009; WU, YAN, 2004).

The material under study was analysed for the first two formant frequencies using the robust formant extraction algorithm in Praat. The signals that gave extreme or clearly erroneous results were excluded from further statistical analyses. Although the mean values of  $f_1$  and  $f_2$  were not significantly different in female and male speakers, Anova test showed significant differences in the mean formant values for individual voices ( $df = 18$ ,  $p < 0.01$ ).

Typical Polish monosegmental vocalic fillers are articulatorily less centralised than their English counterparts. Their formant values tend to group in the upper right area of the vocalic loop (cf. JASSEM, 1973) but they can reach much further in terms of their  $f_1$  values.

Table 4. Mean formant values for female and male speakers.

| gender          | female |       | male  |       |
|-----------------|--------|-------|-------|-------|
|                 | $f_1$  | $f_2$ | $f_1$ | $f_2$ |
| formant freq.   |        |       |       |       |
| mean value [Hz] | 875    | 2265  | 827   | 2156  |
| st. deviation   | 382    | 220   | 362   | 257   |

#### 4.5. Other findings

Some types of FPs were excluded from the above study as they could be hardly analysed for at least some of their acoustic parameters or they were realised as compound units that could not be covered by the present analytic framework. Among them, the following were most frequent:

- *Creaky fillers*. More typical of some speakers in the group but can be found in almost any voice. As in other cases, distorted phonation is more frequent in the final, low-energy stages of fillers but some FPs are realised solely in creaky voice.
- *Voiceless fillers* (sigh-like sounds). Rare but important as it seems that they may have different meaning from their fully voiced counterparts.
- *Compound fillers*. Built of at least two significantly different segments with a perceptually evident transition between the two (as in the case of, e.g., closing lips in the middle of the filler and producing its remaining part as a nasal sound).

These types of FPs would require a modified approach to acoustic measurements, involving different measures. There is no doubt that they deserve further studies. However, the assumption that the proportion of peculiar (creaky, voiceless, compound, etc.) FPs is typical of a given individual and may serve as a basis for speaker recognition or identification remains risky. It may prove dependant on the particular type of communicative situation or the mode of interaction. In any case, a substantially larger corpus, both in terms of the number of speakers and signals under analysis, would be necessary to explore these issues.

## 5. Conclusion and directions for further research

The data analysed in the present study come from a single, peculiar type of communicative situation. As a consequence, they are more coherent, which allows for the control of more variables and facilitates statistical exploration even with a relatively limited number of speakers (from 18 to 20) and signals (ca. 400 used in most analyses). The initial rejection of distorted and imperfectly recorded signals may potentially introduce a bias as it might have reduced the number of samples coming from the most emotional exchanges, with many overlaps and from overloaded recordings or from the stages where the IF was most active and caused most noise manipulating the figure. However, it may actually reflect what can be gathered in real life circumstances where speech samples are recorded for the purpose of speaker identification and where similar quality issues may arise.

The acoustic characteristics of FPs in Polish seem to comply with many of the findings and claims cited above as well with those from some other studies (e.g., STEPANOVA, 2007; WU, YAN, 2004; VASILESCU, ADDA-DECKER, 2007) but a more direct cross-linguistic comparison requires additional work. Polish FPs form a group showing homogeneity in a number of dimensions. While ca. 10% cases were excluded from instrumental analyses (nasalised, creaky, voiceless and compound fillers), the remaining 90% were relatively stable in terms of the analysed parameters, often had flat pitch contours and show little variation in the values of the first two formants.

The range of FPs duration was found to be extremely wide and bear some signs of speaker-dependence, although it was not gender-specific. The mean pitch frequency of the FPs in the mutual visibility condition (MVC) was found significantly higher than in the limited visibility condition (LVC) for female but not for male speakers. The mean values of the pitch range turned out to be significantly different in the set of twenty speakers but it may be due to some extreme individual means, with the remaining group kept within a rather limited range. Accordingly, pitch range alone would be not a clearly speaker-specific measure except for some cases. AUDHKHASI *et al.* (2009) reports that their formant-based technique for filler detection is more accurate than cepstrum and pitch-based methods. Simultaneously,  $f_1$  and  $f_2$  values in the Polish FPs significantly differed in individual speakers. Accordingly, formant frequencies measurements may support both the identification of fillers in the stream of speech and can be expected to support speaker identification. Jitter-related voice parameters showed a limited level of distinctiveness – only the mean values of the local absolute, local and PPQ jitter as well as the DDA and APP3 shimmer turned

out to be significantly different for the voices under study.

A significant proportion of the initial set of signals had to be rejected and new methods of exploration would be certainly needed for a systematic analysis of their acoustic properties different from pitch or formants but referring to some other spectral parameters. Another issue to be addressed in future research is the question of contextual influences on the acoustic realisations of FPs. SHRIBERG (1999) suggests that the pitch pattern of the filler is influenced by the intonational contour of the preceding utterance but this can be only one among many other context-driven phenomena one should take into account. Finally, the structural properties of utterances should be taken into account (e.g., the placement of a given FP in an utterance) as well as the correlation of speech production problems with certain gestural behaviour (CHRISTENFELD, SCHACHTER, BILOUS, 1991; ESPOSITO *et al.*, 2001; JARMOŁOWICZ-NOWIKOW, KARPIŃSKI, 2012). For future studies, the data will be extended with a new corpus of task-oriented dialogues which will allow for a more reliable verification of hypotheses. Various configurations of the features will be explored in order to find optimum sets for the purpose of speaker recognition (cf. DEMENKO, 2000). Alternative, more flexible exploration techniques based on data mining will be also employed.

## Acknowledgment

This work is supported from the financial resources for science in the years 2010–2012 by the National Centre for Research and Development as a development project O R00 0170 12.

## References

1. ABELIN Ā., ALLWOOD J. (2000), *Cross-linguistic Interpretation of Emotional Prosody*, ITR Workshop on Speech and Emotion, Newcastle, Northern Ireland, UK.
2. ABELIN Ā. (2008), *Anger or Fear? – Crosscultural multimodal interpretations of emotional expressions*, [in:] *Emotions in the human voice*, IZDEBSKI K. [Ed.], pp. 65–73, Vol. 1, Plural Publ. Co., San Diego.
3. ABERCROMBIE D. (1967), *Elements of general phonetics*, Edinburgh University Press, Edinburgh.
4. ALBALÁ M. J., BATTANER E., GIL J., LLISTERRI J., MACHUCA M., MARRERO V., RÍOS A. (2009), *Vowel formant structure and speaker identification. A perceptual study*, CIP 2009–3a Conferência Ibérica de Percepção, Guimarães, Portugal, 8–10 Julho 2009.

5. AUDHKHASI K., KANDHWAY K., DESHMUKH O., VERMA A. (2009), *Formant-based technique for automatic filled-pause detection in spontaneous spoken English*, Proc. ICASSP, pp. 4857–4860, Taipei, Taiwan.
6. BEIGI H. (2011a), *Speaker Recognition* [in:] *Biometrics*, YANG J. [Ed.], InTech. ISBN: 978-953-307-618-8, Available from: <http://www.intechopen.com/books/biometrics/speaker-recognition>.
7. BEIGI H. (2011b), *Fundamentals of Speaker Recognition*, Springer, New York.
8. BENUŠ Š. (2009), *Variability and stability in collaborative dialogues: Turn-taking and filled pauses*, Proceedings of Interspeech 2009, pp. 709–799, Brighton.
9. BOERSMA P., WEENINK D. (2013), *Praat: doing phonetics by computer* [Computer program], Version 5.3 (retrieved <http://www.praat.org/>).
10. BOOMER D. S. (1965), *Hesitation and grammatical encoding*, *Language and Speech*, **8**, 148–158.
11. BORTFELD H., LEON S., BLOOM J., SCHOBER M., BRENNAN S. (2001), *Disfluency Rates in Conversation: Effects of Age, Relationship, Topic, Role, and Gender*, *Language and Speech*, **44**, 2, 123–147.
12. TEN BOSCH L. (2003), *Emotions, speech and the ASR framework*, *Speech Communication*, **40**, 1-2, 213–225.
13. BURKHARDT F., AUDIBERT N., MALATESTA L., TÜRK O., ARSLAN L., AUBERGÉ V. (2006), *Emotional Prosody – Does Culture Makes A Difference?*, Proceedings of Speech Prosody 2006 Conference, Dresden, Germany.
14. CAMPBELL N. (2002), *Towards a grammar of spoken language: Incorporating paralinguistic information*, Proceedings ICSLP 2002, Denver, Colorado.
15. CAMPBELL N. (2004), *Listening between the lines. A study of paralinguistic information carried by tone of voice*, pp. 13–16, International Symposium on Total Aspects of Languages TAL2004, Beijing, China.
16. CAMPBELL N. (2007), *Whom we laugh affects how we laugh*, Proc. Workshop on “The Phonetics of Laughter”, pp. 61–65, Saarbrücken, Germany.
17. CANDEA M., VASILESCU I., ADDA-DECKER M. (2005), *Inter- and intra-language acoustic analysis of autonomous fillers*, Proceedings of DISS05, Aix-en-Provence, France.
18. CHRISTENFELD N., SCHACHTER S., BILOUS F. (1991), *Filled pauses and gestures: It's not coincidence*, *Journal of Psycholinguistic Research*, **20**, 1–10.
19. CRYSTAL D. (1963), *A perspective for paralanguage*, *Le Maître Phonétique*, **120**, 25-29.
20. CRYSTAL D. (1966), *The linguistic status of prosodic and paralinguistic features*, Proceedings of the University of Newcastle-upon Tyne Philosophical Society, **1**, 8, 93–108.
21. CRYSTAL D. (1974), *Paralinguistics*, [in:] *Current trends in linguistics*, T. A. SEBEOK [Ed.], **12**, pp. 265–295, Mouton, The Hague.
22. CRYSTAL D. (1975), *Paralinguistics*, [in:] *The body as a medium of expression*, BENTHALL J., POLHEMUS T. (Eds.), pp. 162–174, Institute of Contemporary Arts, London.
23. DEMENKO G. (2000), *Analysis for suprasegmental features for speaker verification*, 8-th Australian International Conference on Speech Science and Technology, pp. 294–299, Canberra.
24. DODDINGTON G. (2001), *Speaker recognition based on idiolectal differences between speakers*, Proceedings of the Eurospeech, **4**, 2521–2524.
25. DUEZ D. (2001), *Acoustic-phonetic Characteristics of Filled Pauses in Spontaneous French*, ITRW on Disfluency in Spontaneous Speech, pp. 41–44, Edinburgh.
26. ESPOSITO A., MCCULLOUGH K. E., QUEK F. (2001), *Disfluencies in gesture: Gestural correlates to filled and unfilled speech pauses*, IEEE Workshop on Cues in Communication, Kauai, Hawaii.
27. FARRÚS M., HERNANDO J. (2009), *Using jitter and shimmer in speaker verification*, IET Signal Processing, **3**, 4, 247–257.
28. FARRÚS M., HERNANDO J., EJARQUE P. (2007), *Jitter and shimmer measurements for speaker recognition*, pp. 778–781, Proceedings of Interspeech 2007 Conference, Antwerp, Belgium.
29. FRANCUZIK K., KARPIŃSKI M., KLEŠTA J. (2002), *A Preliminary Study of the Intonational Phrase, Nuclear Melody and Pauses in Polish Semi-Spontaneous Narration*, Proceedings of Speech Prosody 2002 Conference, Aix-en-Provence, France.
30. FROMKIN V. A. (1971), *The nonanomalous nature of anomalous utterances*, *Language*, **47**, 27–52.
31. FROMKIN V. A. [Ed.] (1973), *Speech errors as linguistic evidence*, Mouton Publishers, The Hague.
32. GARG G., WARD N. (2006), *Detecting Filled Pauses in Tutorial Dialogs*, Technical Report UTEP-CS-06-32.
33. GOBL C., NI CHASAIDE A. (2003), *The role of voice quality in communicating emotion, mood and attitude*, *Speech Communication*, **40**, 1-2, 189–212.
34. GOLDMAN-EISLER F. (1968), *Psycholinguistics. Experiments in spontaneous speech*, The Academic Press, London and New York.
35. GONZALEZ-RODRIGUEZ J. (2008), *Forensic Automatic Speaker Recognition: Fiction or Science?*, Proceedings of Interspeech 2008, Brisbane, Australia.
36. GRAVANO A., LEVITAN R., WILLSON L., BENUŠ Š., HIRSCHBERG J., NENKOVA A. (2011), *Acoustic and Prosodic Correlates of Social Behavior*, Proceedings of Interspeech 2011, Florence, Italy.

37. JARMOŁOWICZ-NOWIKOW E., KARPIŃSKI M. (2012), *The form and function of pointing gestures in task-oriented dialogues*, Conference of the International Society for Gesture Studies (Book of Abstracts), Lund, Sweden.
38. JASSEM W. (1973), *Principles of Acoustic Phonetics*, [in Polish: *Podstawy fonetyki akustycznej*], PWN, Warszawa.
39. JOHNSTONE T., SCHERER K. R. (2000), *Vocal communication of emotion*, [in:] *The Handbook of Emotions*, 2nd Ed., LEWIS M., HAVILAND J. [Eds.], pp. 226–235, Guilford, New York.
40. KARPIŃSKI M. (2006), *Structure and intonation of Polish task-oriented dialogue* [in Polish], Wydawnictwo Naukowe UAM, Poznań.
41. KARPIŃSKI M. (2007), *Selected quasi-lexical and non-lexical units in Polish map task dialogues*, Archives of Acoustics, **32**, 1, 51–65.
42. KARPIŃSKI M., JARMOŁOWICZ-NOWIKOW E. (2010), *Prosodic and Gestural Features of Phrase-internal Disfluencies in Polish Spontaneous Utterances*, Proceedings of Speech Prosody 2010 Conference, Chicago.
43. KELLER E. (2005), *The analysis of voice quality in speech processing*, [in:] *Lecture Notes in Computer Science*, CHOLLET G., ESPOSITO A., FAUNDEZ-ZANUY M. [Eds.], **3445**, pp. 54–73, Springer-Verlag.
44. LADD D., SILVERMAN K. A., TOLKMITT F., BERGMANN G., SCHERER K. R. (1985), *Evidence for the independent function of intonation contour type, voice quality, and f0 range in signalling speaker affect*, Journal of the Acoustical Society of America, **78**, 2, 435–444.
45. LAVER J. (1980), *The Phonetic Description of Voice Quality*, Cambridge University Press, Cambridge.
46. LOCAL J., KELLY J. (1980), *Projection and ‘silences’: Notes on phonetic and conversational structure*, Human Studies, **9**, 185–204.
47. LUDLOW C. L., COULTER D. C., BASSICH C. J. (1982), *Relationships between vocal jitter, age, sex, and smoking*, Journal of Acoustic Society of America, **71**, 55–56.
48. MACLAY H., OSGOOD C. E. (1959), *Hesitation Phenomena in Spontaneous English Speech*, Word, **1**, 19–43.
49. MARY L., YEGNANARAYANA B. (2008), *Extraction and representation of prosodic features for language and speaker recognition*, Speech Communication, **50**, 782–796.
50. NISHINUMA Y., HAYASHI A. (2004), *Silent Pauses in Simulated Request-Refusal Type Dialogues. A Phonetic Analysis of German, Korean, and Japanese*, Symposium of Nordic Association for Japanese and Korean Studies (NAJAKS), Göteborg, Sweden.
51. OGDEN R. (2001), *Turn transition, creak and glottal stop in Finnish talk-in-interaction*, Journal of the International Phonetic Association, **31**, 139–152.
52. PAKOSZ M. (1982), *Intonation and attitude*, Lingua, **56**, 153–178.
53. PATEL S., SCHERER K. R., BJÖRKNER E., SUNDBERG J. (2011), *Mapping emotions into acoustic space: The role of voice production*, Biological Psychology, **87**, 93–98.
54. ROSNER B. S., PICKERING J. B. (1994), *Vowel perception and production*, Oxford University Press, Oxford.
55. RUSZ J., CMEJLA R., RUZICKOVA H., RUZICKA E. (2011), *Quantitative acoustic measurements for characterisation of speech and voice disorders in early untreated Parkinson’s disease*, Journal of Acoustic Society of America, **129**, 1, 350–367.
56. SAKS M. J., KOEHLER J. (2005), *The coming paradigm shift in forensic identification science*, Science, **309**, 5736, 892–895.
57. SAVILLE-TROIKE M. (1985), *The place of silence in an integrated theory of communication*, [in:] *Perspectives on silence*, TANNEN D., SAVILLE-TROIKE M. [Eds.], Norwood, NJ, Ablex.
58. SCHERER K. R., BANSE R., WALLBOTT H. (2001) *Emotion inferences from vocal expression correlate across languages and cultures*, Journal of Cross-Cultural Psychology, **32**, 76–92.
59. SCHÖTZ S. (2006), *Prosodic Cues in Human and Machine Estimation of Female and Male Speaker Age*, [in:] *Nordic Prosody: Proceedings of the IX-th Conference*, BRUCE G., HORNE M. [Eds.], pp. 215–223, Peter Lang Publishing, Lund.
60. SHRIBERG E. (1994), *Preliminaries to a Theory of Speech Disfluencies*, PhD Thesis, Dep. of Psychology, University of California, Berkeley.
61. SHRIBERG E. (1999), *Phonetic consequences of speech disfluency*, Proceedings of International Congress of Phonetic Sciences, pp. 619–622, San Francisco.
62. SHRIBERG E. (2001), *To “Errrr” is Human: Ecology and Acoustics of Speech Disfluencies*, Journal of the International Phonetic Association, **31**, 1, 153–169.
63. SHRIBERG E. (2007), *High-level Features in Speaker Recognition*, [in:] *Speaker Classification I*, MUELLER C. [Ed.], pp. 241–259, Springer-Verlag, Berlin – Heidelberg.
64. STEPANOVA S. (2007), *Some features of filled hesitation pauses in spontaneous Russian*, Proceedings of ICPHS XVI, pp. 1325–1328, Saarbruecken.
65. SWERTS M. (1998), *Filled pauses as markers of discourse structure*, Journal of Pragmatics, **30**, 485–496.
66. TRAGER G. L. (1960), *Taos III, paralinguage*, Anthropological Linguistics, **2**, 2, 24–30.
67. TRAGER G. L. (1961), *The typology of paralinguage*, Anthropological Linguistics, **3**, 1, 17–21.
68. TRAGER G. L. (1964), *Paralinguage: A first approximation*, [in:] *Language in culture and society*, DELL HYMES [Ed.], pp. 274–288, Harper and Row, New York.

- 
69. TRASK R. L. (1996), *A Dictionary of Phonetics and Phonology*, Routledge, London.
70. VASILESCU I., ADDA-DECKER M. (2007), *A cross-language study of acoustic and prosodic characteristics of vocalic hesitations*, [in:] *Fundamentals of Verbal and Nonverbal Communication and the Biometric Issue*, ESPOSITO A., BRATANIĆ M., KELLER E., MARI-NARO M. [Eds.], pp. 140–148, IOS Press.
71. VASILESCU I., CANDEA M., ADDA-DECKER M. (2004), *Hésitations autonomes dans 8 langues: une étude acoustique et perceptive*, Workshop MIDL04, Paris, France.
72. VASILESCU I., CANDEA M., ADDA-DECKER M. (2005), *Perceptual salience of language-specific acoustic differences in autonomous fillers across eight languages*, Proceedings of InterSpeech 2005, pp. 1773–1776, Lisbon, Portugal.
73. WALLBOTT H. G., SCHERER R. S. (1979), *Normal speech – normal people. Speculations on paralinguistic features, arousal, and social competence attribution*, Proceedings of the Social Psychology and Language Conference, Bristol.
74. WARD N. (2004), *Pragmatic Functions of Prosodic Features in Non-Lexical Utterances*, Proceedings of Speech Prosody 2004 Conference, pp. 325–328, Nara, Japan.
75. WU CH.-H., YAN G.-L. (2004), *Acoustic Feature Analysis and Discriminative Modeling of Filled Pauses for Spontaneous Speech Recognition*, Journal of VLSI Signal Processing, **36**, 91–104.

## Modal Parameters of Two Violins with Different Varnish Layers and Subjective Evaluation of Their Sound Quality

Ewa B. SKRODZKA<sup>(1),(2)</sup>, Bogumił B.J. LINDE<sup>(3)</sup>, Antoni KRUPA<sup>(2)</sup>

<sup>(1)</sup> *Institute of Acoustics, Adam Mickiewicz University*  
Umultowska 85, 61-614 Poznań, Poland; e-mail: afa@amu.edu.pl

<sup>(2)</sup> *Faculty of String Instruments, Harp, Guitar and Luthiery*  
*The Ignacy Jan Paderewski Academy of Music*  
Św. Marcin 87, 61-808 Poznań, Poland

<sup>(3)</sup> *Institute of Experimental Physics, University of Gdańsk*  
Wita Stwosza 57, 80-952 Gdańsk, Poland

(received September 18, 2012; accepted January 17, 2013)

Two violins were investigated. The only intentionally introduced difference between them was the type of varnish. One of the instruments was covered with a spirit varnish, the other was oil varnished. Experimental modal analysis was done for unvarnished/varnished violins and a questionnaire inquiry on the instrument's sound quality was performed. The aim of both examinations was to find differences and similarities between the two instruments in the objective (modal parameters) and subjective domain (subjective evaluation of sound quality). In the modal analysis, three strongly radiating signature modes were taken into account. Varnishing did not change the sequence of mode shapes. Modal frequencies A0 and B(1+) were not changed by oil varnishing compared to the unvarnished condition. For the oil varnished instrument, the frequency of mode B(1+) was lower than that of the same mode of the spirit varnished instrument. Our two violins were not excellent instruments, but before varnishing they were practically identical. However, after varnishing it appeared that the oil-varnished violin was better than the spirit-varnished instrument. Therefore, it can be assumed with a fairly high probability that also in general, the oil-varnished violins sound somewhat better than initially identical spirit-varnished ones.

**Keywords:** violins, modal analysis, varnish, subjective sound quality evaluation.

### 1. Introduction

Wooden musical instruments are usually coated with a so-called 'varnish', in order to protect them from moisture and dust and to enhance their visual appearance. These varnishes consist of one or more layers of mixed organic and mineral materials prepared by a luthier by grinding, mixing, solubilising and/or heating the raw materials. The mixtures are subsequently applied with a brush or a piece of cloth on a surface of a violin (ECHARD *et al.*, 2008). Violin-makers distinguish three main varnish types: (a) solutions of resins in alcohol, the so-called spirit varnishes, drying quickly by evaporation of the solvent, (b) solutions of resins in volatile oils, i.e. essential oil varnishes, drying more slowly, (c) mixtures of resins with siccative oils, i.e. oil varnishes, drying slowly. Volatile oils mainly

used are turpentine oil and spike oil. The siccative oils most often used for oil varnishes are linseed and walnut oils. Colophony, Strasburg and Venice turpentines, sandarac, copals, benzoin, elemi, and mastic are vegetal resins typically used in violin varnishes. Shellac is a resin exuded by an insect (ECHARD *et al.*, 2007).

The dynamic behaviour of violins may be investigated by experimental and computational methods. Among those former, the most popular are optical measurements and modal analysis. A detailed modal analysis of a violin was done by MARSHALL (1985) and many others (BISSINGER, 1995, 2003, 2008; BISSINGER, KEIFFER, 2003; SKRODZKA *et al.*, 2009). However, a few reports were published on the effect of structural modifications on vibrational behaviour of violins, among them by WEINREICH *et al.* (2000), examining a violin with holes drilled in the ribs, and the

report by Meinel (1937), who measured the response curves for violins with and without varnish, with thick, normal, and thin plates. Thus, there are many papers about modal analysis of violins and one, rather historical (MEINEL, 1937), on the influence of varnishing on violin's modal behaviour.

Basic information about the nature of vibration of a complete violin and its parts and behaviour of particular elements when strings are excited by different playing techniques can be found in handbooks by CREMER (1981); FLETCHER and ROSSING (1998), and HARTMANN (1997).

Modal analysis technique has been extensively used to investigate a behaviour of particular elements of cellos, violins, and guitars, like a cello tailpiece (FOUILHE *et al.*, 2011), the violin bridge, and a soundpost (BISSINGER, 1995; 2006; SALDNER *et al.*, 2006), symmetric and asymmetric bracing patterns on guitar soundboards (SKRODZKA *et al.*, 2011), isolated guitar soundboards (ELEJABARRIETA *et al.*, 2000; BOULLOSA, 2002), a complete Hutchins-Schelleng violin octet (BISSINGER, 2003), complete violins (MARSHALL, 1985; DÜNNWALD, 1999), and many others.

The aim of the present work was to show differences (if any) in natural vibrations between two complete violins with intentionally introduced differences in varnishes. Parameters of natural vibrations (modal frequencies, modal damping, and modal deformations associated with them), being inherent properties of the structure, are responsible for the sound radiation from any vibrating mechanical system. Additionally, a subjective evaluation of the sound quality of violins was done by a simple auditory test.

## 2. Experiment

### 2.1. Violins

Two replicas of the “Ysaÿe” violin of Giuseppe Guarneri del Gesu (1740) were made by a professional luthier. We named them “1” and “2”. Top plates of both instruments were made of the same piece of spruce. Back plates, ribs, and heads were made of maple (sycamore). Special attention was paid to make both violins as similar as possible, if not identical. The only intentionally introduced difference was the coating varnish. Figure 1 shows the thickness of the front (right) and back (left) plates in millimetres. Violin sizes are listed in Table 1.

Both instruments were equipped with identical sets of strings: G – Pirastro Flexocore, D – Thomastik Infeld Red, A – Pirastro Chromcore Eudoxa, E – Pirastro Gold.

The first modal experiment was performed for both unvarnished instruments. The second modal experiment was done after varnishing. Before varnishing both

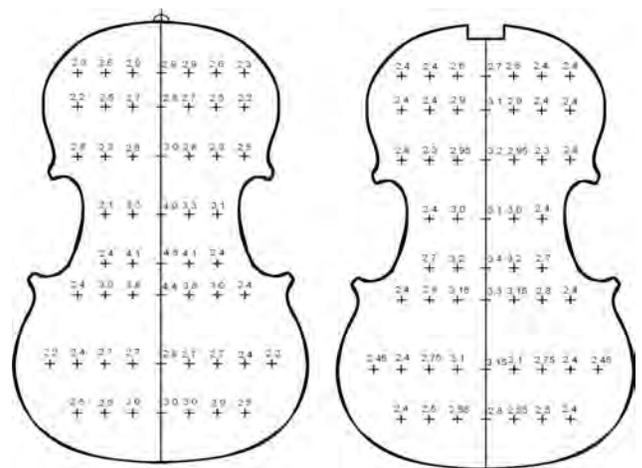


Fig. 1. Thickness of the front (right) and back (left) plates in mm.

Table 1. Violin sizes (mm).

|   | Front plate                         | Back plate       |
|---|-------------------------------------|------------------|
| Body length   | 353                                 | 353              |
| Maximum width in the upper bout                                 | 167.5                               | 167.5            |
| Width at the waist (arch)                                       | 110.5<br>(113.5)                    | 110.5<br>(114.5) |
| Maximum width in the lower bout                                 | 206                                 | 206              |
| Mensur length   | –                                   | 195              |
| Maximum height of arch  | 14                                  | 15.5             |
| Ribs height   | 29–32                               |                  |
| Bass beam length  | 270                                 |                  |
| Maximum bass beam height  | 13                                  |                  |
| Bass beam width   | 3–6                                 |                  |
| Bass beam position according to the instrument axis of symmetry | 18 (upper part)<br>20 (bottom part) |                  |

violins were painted using wood stain of golden-brown colour tone. Next, the instruments were prime painted using linseed oil. Subsequently, instrument 1 was varnished using the spirit varnish; for instrument 2 the oil varnish was applied. The spirit varnish prepared by our luthier consisted of transparent spirit, ruby shellac, sandarac, mastic, gummi elemi, and benzoin. The oil varnish was JOHA<sup>®</sup> Oil Varnish, IA. This varnish contains a mixture of natural and synthetic resins, dissolved in turpentine and linseed oil, as well as essential oils such as lavender or rosemary. Golden yellow, golden brown, and amber colour extracts were used to obtain the desired colour shade. No siccatives were used. A badger painting brush was used. Violin 1 was covered with 20 layers of the spirit varnish and violin 2 with 10 layers of the oil varnish, following the common rules of varnishing, well known among violin makers. Both violins were measured in playing conditions with

undamped strings at tension, without chin, or shoulder rest, similarly to BISSINGER (2008).

## 2.2. Measurements

### 2.2.1. Modal analysis experiment

Modal analysis is a popular experimental method of studying dynamic behaviour of structures, including violins and guitars (MARSHALL, 1985; EWINS, 1995; SKRODZKA *et al.*, 2005, 2006, 2011). The main assumption of the modal analysis is that the system under investigation is linear. In terms of the modal analysis, linearity means that interchanging the positions of the accelerometer and the impact hammer does not change the course of Frequency Response Functions (FRFs) obtained at these two positions. Thus, if the FRFs are the same, the Maxwell’s reciprocity principle is valid and the system is linear. In reality, no mechanical system is linear but the assumption is not very strict (MARSHALL, 1985; SKRODZKA *et al.*, 2011).

In terms of the modal analysis, the FRF for the response  $X_i$  at the point  $i$ , due to an excitation force  $F_k$  applied at the point  $k$ , has the form (EWINS, 1995; MARSHALL, 1985; SKRODZKA, SEK, 1998):

$$\frac{X_i}{F_k} = H_{ik}(s) = \sum_{r=1}^n \left( \frac{r_{ikr}}{s - p_r} + \frac{r_{ikr}^*}{s - p_r^*} \right), \quad (1)$$

where  $r_{ikr}$  is a residue for  $r$ -th mode,  $s$  is Laplace variable,  $p_r = \sigma_r + j\omega_r$ , for  $k = 1, \dots, r$  is the eigenvalue or the pole of the FRF,  $\sigma_r$  is the modal damping of the  $r$ -th mode,  $\omega_r$  is the modal frequency of the  $r$ -th mode.

If all points  $i$  and  $k$  are taken into consideration, then Eq. (1) takes the form:

$$H(\omega) = \sum_{r=1}^n \left( \frac{\{u_r\} \{u_r\}^t}{j\omega - p_r} + \frac{\{u_r^*\} \{u_r^*\}^t}{j\omega - p_r^*} \right). \quad (2)$$

The Laplace variable  $s$  in Eq. (1) has been replaced by the frequency  $\omega$  along the frequency axis;  $r_{ikr}$  and  $\{u_r\}$  are complex quantities. Thus, each mode of vibration is defined by a pair of complex conjugate poles ( $p_r, p_r^*$ ) and a pair of complex conjugate mode shapes ( $\{u_r\}, \{u_r^*\}$ );  $\{u_r\}^t$  and  $\{u_r^*\}^t$  are transpositions of  $\{u_r\}$  and  $\{u_r^*\}$ , respectively. A set of FRFs with indices  $i$  and  $k$  is usually arranged in a matrix called a modal matrix.

The modal analysis method describes the dynamics of any vibrating system in terms of modal parameters: natural frequencies and natural damping, as well as deformation patterns (mode shapes) associated with them. As the measurement setup was similar to that described in our previous works (SKRODZKA *et al.*, 2005, 2009, 2011), only the most crucial details are given below. For measurements, the instruments were mounted in a wooden cage of a significant mass, allowing the soundboard and the back plate to

vibrate without any obstacles. The instruments were excited by an impact hammer to provide a broad-band excitation (PCB Impact Hammer 086C05, sensitivity 2.25 mV/N). As the version of the modal analysis with a fixed response point was used, the response signal was measured at the fixed measuring point marked as a black dot in Fig. 2. An ONO SOKKI NP-2910 accelerometer, with the mass of 2 g and sensitivity 0.3 pC/m/s<sup>2</sup> was used. The mass of the accelerometer was significantly less than 10% of the mass of the top of each instrument (about 120 g) and did not affect the results of the measurements. Both the excitation and the response signals were measured perpendicularly to the top plates, i.e. in the most important direction as regards the vibration of the final instrument. The accelerometer was mounted on a beewax. On the basis of these signals, the frequency response functions (FRFs) were calculated between all 230 excitation points on the front plate, see Fig.2, and the fixed response point where the accelerometer was mounted. The modal parameters extracted from the FRFs were calculated by means of an SMS STAR-Modal<sup>®</sup> package (The STAR system, 1990). The FRFs were measured at 230 points on the front plate. Geometry of the measuring mesh is shown in Fig. 2. A black dot in Fig. 2 denotes the position of the accelerometer. The position was chosen experimentally, avoiding areas of the top plate, where the bass bar was attached, and after a preliminary test, when the FRFs measured between some tested accelerometer positions and some points of excitation were evaluated by the experimenter with respect to a proper course of the coherence function and repeatable frequencies of peaks in FRFs.

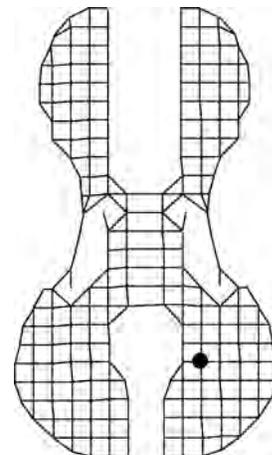


Fig. 2. Geometry of the modal analysis measuring mesh. The black dot denotes the position of the accelerometer.

All FRFs were measured in the frequency range 0–3200 Hz with a 4 Hz spectral resolution, and their quality was controlled by the coherence function. Ten spectral averages were used to improve the signal-to-noise ratio and FRFs.

### 2.2.2. Performance subjective remarks

To judge the sound quality of the varnished instruments, it was decided to evaluate them in natural playing conditions. Thirteen people with some experience in comparison of violin's sound (luthiery students) took part in the experiment. Their task was to fill up a questionnaire shown in Table 2 by assigning the mark “+” to the subjectively better instrument and mark “-” for the worse instrument. In the case when both instruments were evaluated equally they gave “+” (both good) or “-” (both bad). The questionnaire was similar to those used during luthier competitions. Subjects listened to sounds of violins in the concert hall of I.J. Paderewski Academy of Music. Subjects were informed about the symbol of presented instruments (1 or 2) and they were not informed about varnishing. Instructions concerning the questionnaire were given before the test. A professional violinist was the player. Subjects evaluated the sound of individual strings listening to the sounds:  $g$  (196 Hz),  $h$  (245 Hz),  $d^1$  (294 Hz),  $f^1$  (349 Hz),  $a^1$  (440 Hz) on string G;  $d^1$  (294 Hz),  $f^{\#1}$  (370 Hz),  $c^2$  (523 Hz),  $e^2$  (659 Hz) on string D;  $a^1$  (440 Hz),  $c^{\#2}$  (554 Hz),  $e^2$  (659 Hz),  $g^2$  (784 Hz),  $h^2$  (988 Hz) on string A;  $e^2$  (659 Hz),  $g^{\#2}$  (831 Hz),  $h^2$  (988 Hz),  $d^3$  (1175 Hz),  $f^{\#3}$  (1480 Hz) on string E. Then scales were played on all strings and the task of the subjects was to evaluate the level and volume of the sound of all the strings. Finally, the Largo, C-major Sonata, BWV 1005 by J.S. Bach was played and the task of the subjects was to point out the instrument which according to them was better, with a more noble and interesting sound, or, in other words, they were asked about their general opinion about the level, range, and timbre of the sound.

Table 2. Questionnaire for the subjective evaluation of the sound quality.

|  | Sound                                      | Violin 1                      | Violin 2                      |
|--|--|-------------------------------|-------------------------------|
| Sound of individual strings                                    | Scale on string G                          |                               |                               |
|  | Scale on string D                          |                               |                               |
|  | Scale on string A                          |                               |                               |
|  | Scale on string E                          |                               |                               |
| Level and equal volume of the sound of all the strings         | Scale on all the strings                   | Level                         | Level                         |
|  |  | Equal sound volume of strings | Equal sound volume of strings |
| General opinion about the level, range and timbre of the sound | J.S. Bach, Largo, C-major Sonata, BWV 1005 |                               |                               |
| General remarks  |  |                               |                               |

## 3. Results

### 3.1. Modal analysis

A dozen or so modes were found in the frequency range measured. The mode shape sequence was similar to that well described in earlier papers and analyzed in a descriptive manner, being a widely accepted standard in the literature (MARSHALL, 1985; DÜNNWALD, 1999; BISSINGER, 1995; 2003; 2008; BISSINGER, KEIFFER, 2003; SCHLESKE, 2002; SKRODZKA *et al.*, 2005; 2009, 2011). Our discussion of individual modes will be limited to three strongly radiating modes A0, B(1-) and B(1+) (BISSINGER, 2008) of the top plate. These modes fall into the first Dünwald frequency band (190–650 Hz) and are responsible for the sound “richness” (DÜNNWALD, 1999). A0 is a Helmholtz resonance (“air mode”). B(1-) and B(1+) are “plate modes” which arise from the bending and stretching of the front plate. For the top plate, the mode shape B(1-) has two longitudinal nodal curves placed almost symmetrically on both sides of the main axis of symmetry. Mode B(1+) on the top plate has two nodal curves crossing the upper and the lower bouts. In Table 3 the mode shapes, modal frequencies ( $f$ ), and percentage of the critical damping ( $d$ ) are shown for the above modes. Mode B(1-) of the oil varnished violin 2 presented in Table 3 was excluded from the analysis because of a high value of the critical damping ( $> 10\%$ ). The main assumption of the modal analysis is linearity of the system under investigation. Strictly speaking, none of the violins is a linear system but it can be treated as such when critical damping is smaller than 10% (SKRODZKA *et al.*, 2009; EWINS, 1995).

### 3.2. Performance subjective remarks

The results of the subjective evaluation of the sound quality of both varnished instruments are shown in Table 4. The evaluation of the individual strings was as follows: string G – nine persons pointed at violin 2, two persons – violin 1, and two subjects stated that in both instruments it sounded well. String D – eleven subjects preferred instrument 2, one subject – instrument 1, and for one person the string was good for both violins. String A – seven persons pointed at instrument 1, six – instrument 2. String E – eleven subjects preferred instrument 2, one subject – instrument 1, and for one subject the string was good in both instruments. When the sound level was evaluated, none of the instruments was significantly better: six subjects pointed at violin 2 and five persons preferred instrument 1; for one person both instruments were good, and for one both were bad. The strength of the sound was evaluated as better for violin 2 (twelve subjects); only one subject pointed out instrument 1. Instrument 2 covered with the oil varnish was

Table 3. Modal parameters for the top plate modes A0, B(1-) and B(1+). Mode B(1-) for the oil varnished instrument had a damping value exceeding 10% of the critical damping.

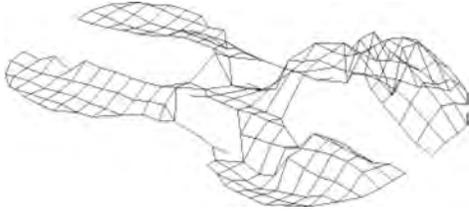
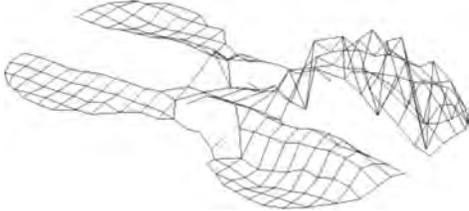
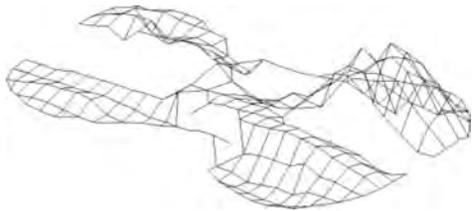
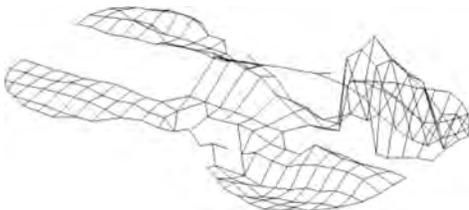
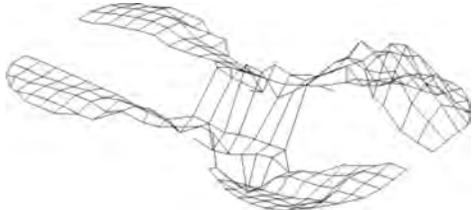
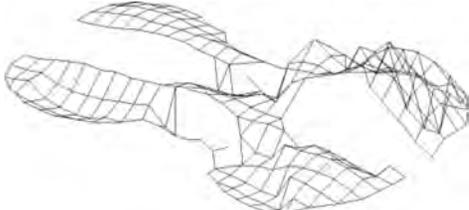
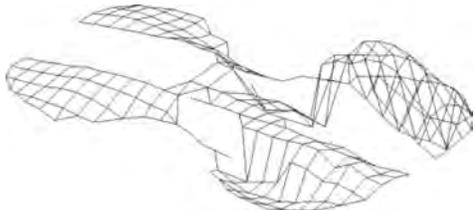
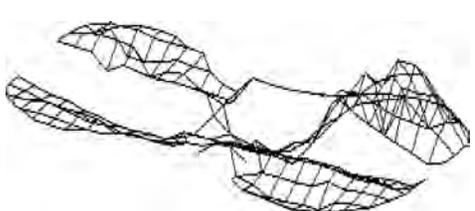
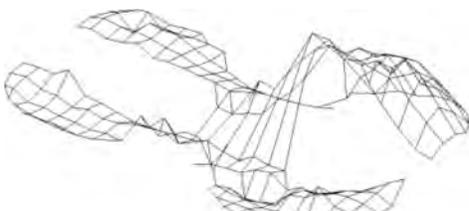
| Mode  | Violin 1 (white)  | Violin 2 (white)  |
|-------|---|---|
| A0    |  <p><math>f = 281 \text{ Hz}, d = 10.0\%</math></p>  |  <p><math>f = 280 \text{ Hz}, d = 9.0\%</math></p>    |
| B(1-) |  <p><math>f = 476 \text{ Hz}, d = 9.6\%</math></p>   |  <p><math>f = 451 \text{ Hz}, d = 7.1\%</math></p>    |
| B(1+) |  <p><math>f = 548 \text{ Hz}, d = 8.1\%</math></p>  |  <p><math>f = 540 \text{ Hz}, d = 5.0\%</math></p>   |
|       | Violin 1 (spirit)   | Violin 2 (oil)  |
| A0    |  <p><math>f = 277 \text{ Hz}, d = 5.2\%</math></p> |  <p><math>f = 276 \text{ Hz}, d = 6.4\%</math></p>  |
| B(1-) |  <p><math>f = 492 \text{ Hz}, d = 4.6\%</math></p> |  <p><math>f = 419 \text{ Hz}, d = 17.0\%</math></p> |
| B(1+) |  <p><math>f = 598 \text{ Hz}, d = 4.1\%</math></p> |  <p><math>f = 534 \text{ Hz}, d = 5.1\%</math></p>  |

Table 4. Results of the subjective evaluation of the sound quality.

|                | Violin 1<br>(spirit) |     | Violin 2<br>(oil) |     | Both<br>good<br>(+) | Both<br>bad<br>(-) |
|----------------|----------------------|-----|-------------------|-----|---------------------|--------------------|
|                | (+)                  | (-) | (+)               | (-) |                     |                    |
| String G       | 9                    | 2   | 2                 | 9   | 2                   | 0                  |
| String D       | 1                    | 11  | 11                | 1   | 1                   | 0                  |
| String A       | 7                    | 6   | 6                 | 7   | 0                   | 0                  |
| String E       | 1                    | 11  | 11                | 1   | 1                   | 0                  |
| Sound level    | 5                    | 6   | 6                 | 5   | 1                   | 1                  |
| Sound strength | 1                    | 12  | 12                | 1   | 0                   | 0                  |
| Total          | 2                    | 10  | 10                | 2   | 1                   | 0                  |

evaluated in total as better (10 persons) than the violin covered with the spirit varnish (2 subjects). For one subject the overall evaluation of both instruments was good.

#### 4. Discussion

For unvarnished (“white”) and varnished instruments, the mode shapes, modal frequencies, and modal damping of the modes under consideration were similar to those described in earlier papers (MARSHALL, 1985; BISSINGER, 2003; 2008; SCHLESKE, 2002). While comparing the modal frequencies of the unvarnished violins the only significant difference was found for mode B(1–), Table 3. Thus, modal parameters of the “white” instruments were very similar. Varnishing did not change the mode shapes but it influenced the modal frequencies, especially of modes B(1–) and B(1+). There was no systematic trend in the modal frequency changes. The modal frequency A0 was the same for both instruments before and after varnishing (measurement accuracy was 4 Hz). For mode B(1–) the modal frequency of “white” instrument 1 was 476 Hz and for the same instrument with the spirit varnish it was 492 Hz. A significant increase in the modal frequency for violin 1 was also observed for mode B(1+): from 548 Hz for the “white” instrument to 598 Hz for the varnished corpus. For unvarnished instrument 2, the frequency of mode B(1–) was 451 Hz; for the same violin with the oil varnish this mode was not observed with the assumed “damping limit” (10% of the critical damping). The modal frequency of mode B(1+) before varnishing was 540 Hz and 534 Hz after it. Thus, it was not changed by oil varnishing of violin 2 within the measurement accuracy of 4 Hz. By comparing the modal frequencies of the varnished instruments it was found that the modal frequency B(1+) was significantly lower for instrument 2 with the oil varnish (534 Hz) than for violin 1 with the spirit varnish (598 Hz). The modal damping was slightly reduced for both varnished corpuses when compared to the unvar-

nished instruments (Table 3), but this fact did not influence the violin’s quality, as damping trends are not robust quality discriminators (BISSINGER, 2008).

The frequency of mode B(1+) is very important for the violin sound and may be a “mechanical” gauge of the sound quality. According to Schleske it acts as a “tonal barometer” of the violin sound quality. When the frequency of mode B(1+) is lower than 510 Hz, an instrument is “soft”, with a rather weak ‘resistance’ to the player, and its sound is dark. The instrument with B(1+) frequency higher than 550 Hz is ‘stubborn’, ‘resistant’ to the player, with a bright sound with a tendency of harshness (BISSINGER, 2008; SCHLESKE, 2002). The sound of our both varnished instruments was evaluated subjectively and the subjects pointed at the oil varnished violin to be better “in total” than the spirit varnished violin. This subjective feeling was objectively confirmed by the frequency of mode B(1+): for the oil varnished violin it was 543 Hz (less than 550 Hz) and for the spirit varnished violin it was much higher (598 Hz). Thus, by comparing the sound quality of the two varnished violins, oil varnished instrument, as evaluated subjectively and in agreement with comparative judgments of other authors (BISSINGER, 2008; FRITZ *et al.*, 2007; SCHLESKE, 2002), was found better. It is worthwhile to note that famous 18th and 19th century instruments were usually oil varnished.

However, there are some reasons that keep us far away from a conclusion that the oil varnish has advantages over the spirit varnish when the sound quality of the violin is concerned. First, only two instruments were investigated. This number is obviously not appropriate for any statistical considerations. Nonetheless, similar situations, when only a few instruments were investigated with the aim of formulating general conclusions, can be found in some reports (MARSHALL, 1985; SKRODZKA *et al.*, 2009; WEINREICH *et al.*, 2000). Next, the “better” violin 2 probably is not acoustically ‘excellent’ because it did not outperformed violin 1 in all categories in the subjective test, and in the objective modal experiment the strongly radiating mode B(1–) was missing. Therefore, a better score obtained by violin 2 can be explained by the lower frequency of the “tonal barometer” B(1+). Finally, our investigations were carried out just after hardening of the varnish. Its drying and ageing induce important changes in its composition and may influence both the visual appearance and sound quality of the instrument.

#### 5. Conclusions

An experimental modal analysis of two unvarnished and varnished violins, as well as a subjective evaluation of the sound quality of the varnished instruments, was done to find differences and similarities in the modal parameters and subjective judgment. Hence, we conclude that:

1. Our violins were not ‘excellent’ instruments, but before varnishing they were practically identical. However, after varnishing the subjects pointed out the oil-varnished violin as better than the spirit-varnished one. Therefore, it can be assumed with a fairly high probability that also in general, the oil-varnished violins sound somewhat better than initially identical spirit-varnished ones.
2. In the case of unvarnished violins the only difference in the modal frequency was observed for mode B(1–).
3. Varnishing did not change the mode shape sequence.
4. For the oil-varnished violin mode B(1–) was not taken into account due to a high value of the critical damping. The modal frequencies A0 and B(1+) were not changed by oil varnishing within a 4 Hz frequency measurement accuracy.
5. The frequency of a “tonal barometer” (mode B(1+)) was lower for the oil varnished instrument than the frequency of the same mode of the spirit varnished instrument. Thus, the objective value of this frequency suggested that the violin 2’s sound quality was better.

### Acknowledgment

We wish to express our deep gratitude to Jan Sebastian Adamczyk, the luthier, for his work. The work was supported by the Polish Ministry of Science (Grant N N105 058437).

### References

1. BISSINGER G. (1995), *Some mechanical and acoustical consequences of the violin soundpost*, J. Acoust. Soc. Am., **97**, 3154–3164.
2. BISSINGER G. (2003), *Modal analysis of a violin octet*, J. Acoust. Soc. Am., **113**, 4, 2105–2113.
3. BISSINGER G. (2006), *The violin bridge as filter*, J. Acoust. Soc. Am., **120**, 1, 482–491.
4. BISSINGER G. (2008), *Structural acoustics of good and bad violins*, J. Acoust. Soc. Am., **124**, 3, 1764–1773.
5. BISSINGER G., KEIFFER J. (2003), *Radiation damping, efficiency, and directivity for violin normal modes below 4 kHz*, Acoust. Res. Lett. Online, **4**, 1, 7–12.
6. BOULLOSA R. R. (2002), *Vibration measurements in the classical guitars*, Appl. Acoust., **62**, 311–322.
7. CREMER L. (1981), *The physics of the violin*, Hirzel Verlag, Stuttgart.
8. DÜNNWALD H. (1999), *Deduction of objective quality parameters on old and new violins*, Catgut Acoust. Soc. J., **2**, 1, 1–5.
9. ECHARD J. P., BENOIT C., PERIS-VICENTE J., MALECKI V., GIMENO-ADELANTO J. V., VAIEDELICH S. (2007), *Gas chromatography/mass spectrometry characterization of historical varnishes of ancient Italian lutes and violin*, Anal. Chim. Acta, **584**, 172–180.
10. ECHARD J. P., COTTE M., DOORYHEE E., BERTRAND L. (2008), *Insights into the varnishes of historical musical instruments using synchrotron micro-analytical methods*, Appl. Phys. A, **92**, 77–81.
11. ELEJABARRIETA M. J., EZCURRA A., SANTAMARIA C. (2000), *Evolution of the vibrational behavior of a guitar soundboard along successive construction phases by means of the modal analysis technique*, J. Acoust. Soc. Am., **108**, 1, 369–378.
12. EWINS D. J. (1995), *Modal Testing: Theory and Practice*, Research Studies Press Ltd., Taunton, Somerset, England.
13. FLETCHER N. H., ROSSING T. D. (1997), *The Physics of Musical Instruments*, Chap. 10, pp. 272–330, Springer-Verlag, New York.
14. FOUILHE E., GOLI G., HOUSSAY A., STOPPANI G. (2011), *Vibration modes of the cello tailpiece*, Arch. Acoust., **36**, 4, 713–726.
15. FRITZ C., CROSS I., MOORE B. C. J., WOODHOUSE J. (2007), *Perceptual thresholds for detecting modifications applied to the acoustical properties of a violin*, J. Acoust. Soc. Am., **122**, 3640–3650.
16. HARTMANN W. M. (1997), *Signals, Sound, and Sensation*, Woodbury, New York.
17. MARSHALL K. D. (1985), *Modal analysis of a violin*, J. Acoust. Soc. Am., **77**, 695–709.
18. MEINEL H. (1937), *On the frequency curves of violin*, Akust. Z., **2**, 22–33.
19. SALDNER H. O., MOLIN N-E., JANSSON E. V. (1996), *Vibration mode of the violin forced via the bridge and action of the soundpost*, J. Acoust. Soc. Am., **100**, 1168–1177.
20. SCHLESKE M. (2002), *Empirical tools in contemporary violin making. 1. Analysis of design, materials, varnish and normal modes*, Catgut Acoust. Soc. J., **4**, 50–65.
21. SKRODZKA E., KRUPA A., ROSENFELD E., LINDE B. B. J. (2009), *Mechanical and optical investigation of dynamic behavior of violins in modal frequencies*, Appl. Opt., **48**, C165–170.
22. SKRODZKA E., ŁAPA A., GORDZIEJ M. (2005), *Modal and spectral frequencies of guitars with differently angled necks*, Arch. Acoust., **30**, 4, 197–201.
23. SKRODZKA E., ŁAPA A., LINDE B. B. J., ROSENFELD E. (2011), *Modal parameters of two incomplete and complete guitars differing in the bracing pattern of the soundboard*, J. Acoust. Soc. Am., **130**, 4, 2186–2194.
24. SKRODZKA E. B., SEK A. P. (1998), *Vibration patterns of the front panel of the loudspeaker system: measurement conditions and results*, J. Acoust. Soc. Jap. (E), **19**, 4, 249–257.
25. *The STAR system version 3.02D. Theory and application*, User’s Manual (1990), SMS.
26. WEINREICH G., HOLMES C., MELLODY M. (2000), *Air-wood coupling and Swiss-cheese violin*, J. Acoust. Soc. Am., **108**, 2389–2402.

# Power Amplification and Selectivity in the Cochlear Amplifier

Piotr KIEŁCZYŃSKI

*Research Group of Acoustoelectronics, Institute of Fundamental Technological Research, Polish Academy of Sciences  
Pawińskiego 5B, 02-106 Warsaw, Poland; e-mail: pkielczy@ippt.pan.pl*

*(received October 30, 2012; accepted January 21, 2013)*

This paper presents a new model that describes the physical phenomena occurring in an individual Outer Hair Cell (OHC) in the human hearing organ (Cochlea). The new model employs the concept of parametric amplification and piezoelectricity. As a consequence, the proposed model may explain in a natural way many as yet unresolved problems about the mechanisms of: 1) power amplification, 2) non-linearity, 3) fine tuning, or 4) high sensitivity that take place in the human hearing organ. Mathematical analysis of the model is performed. The equivalent electrical circuits of an individual OHC are established. The high selectivity of the OHC parametric amplifier is analyzed by solving the resulting Mathieu and Ince differential equations. An analytical formula for the power gain in the OHC's parametric amplifier has been developed. The proposed model has direct physical interpretation and all its elements have their physical counterparts in the actual structure of the cochlea. The numerical values of the individual elements of the electrical equivalent circuits are consistent with the experimental physiological data. It is anticipated that the proposed new model may contribute in future improvements of human cochlear implants as well as in development of new digital audio standards.

**Keywords:** piezoelectricity, parametric amplification, mechanism of hearing, electroacoustic transducers.

## 1. Introduction

Understanding of the human hearing process is of fundamental importance in speech and audio acoustics as well as in medicine and biology. In fact, all modern MPEG digital audio codecs exploit heavily psycho-physical properties of the human hearing organ, such as tone masking, nonlinearity, etc. On the other hand, cochlear implants are often only one alternative for peoples with severely impaired hearing capabilities.

A very important role in the human auditory mechanism play electric and electromechanical phenomena occurring in the cochlea. Electromechanical phenomena taking place in the human hearing organ (the cochlea) are still a fascinating and unsolved mystery (LIAO *et al.*, 2007; MAOILEIDIGH, JULICHER, 2010; ELIOT, SHERA, 2012). In particular, the mechanism of active power amplification and a mechanism that sharpens the output amplitude frequency characteristics, which provides selectivity, are not known.

For a long time it was believed that cochlea is a passive structure. Gold was the first who concluded the necessity of existence of an active process of power amplification in the cochlea (GOLD, 1948). The presence of power gain can explain remarkable properties

of cochlea, such as its frequency selectivity, high sensitivity and nonlinear behavior.

The most important processes determining the selectivity and sensitivity of the human hearing organ occur in the inner ear, namely in the organ of Corti. A key role is played by the cochlear amplifier. It consists of about 20,000 outer hair cells (OHC), located between the basilar membrane (BM) and the tectorial membrane (TM). The outer hair cells (OHCs) are the electro-mechanosensory cells in the cochlea. They are considered to be critically important and responsible for the amplifying and sharp frequency sensitivity of the human ear (DIMITRIADIS, 1999).

In the outer hair cells (OHCs) mechanical signals are transformed into electrical signals, and conversely, electrical signals are converted into mechanical ones (simple and inverse piezoelectric effect).

### *Objective of the paper*

The main scientific objective of the paper is development of a new mathematical and electro-mechanical model of the physical phenomena occurring in a single OHC (Outer Hair Cell). The mechanism of physical phenomena that take place in a single OHC is

still not fully understood and described, even though it was, and is the subject of an intense research in many renowned scientific institutions worldwide. Among others, it is also not known as yet, the role of individual components occurring in a single OHC, such as e.g., the nonlinear capacitance.

The existing models describing the operation of a single OHC are incomplete and do not exactly correspond to physical phenomena occurring in a single OHC (LIU, NEELY, 2009; COHEN, FURST, 2004; MOUNTAIN, HUBBARD, 1994; DE BOER, 1995). In fact, there are a few models describing the phenomenon of power amplification. However, these models are phenomenological in nature and to large extend speculative (SHERA, 2007).

In this paper the author presents a new physical model for outer hair cells (OHCs) in the cochlea, which are represented by hollow piezoelectric cylinders with nonlinear mechanical and electrical properties.

A key element of the proposed model is the parametric amplifier composed of a nonlinear capacitance and other physical elements that are components of a single OHC.

The author proposes a new theory of power amplification in the outer hair cells (OHCs) embedded in the cochlea using (known in electronics and physics) concept of parametric amplification (LOUISELL, 1960). In fact, the nonlinear capacitance that occurs between the inner and outer surfaces of the OHC constitutes a nonlinear reactive element of the proposed parametric amplifier. This nonlinear capacitance is a crucial element of the proposed OHC's parametric amplifier model.

In the present work the author provides the theory of power gain and selectivity (sharpness of frequency characteristics) that occur in a single OHC treated as a parametric amplifier. A quantitative analysis of the processes of power amplification and selectivity, which occur in the parametric amplifier based on a single OHC, was carried out. The proposed new model of operation of an individual OHC, compared to so far existing models, is entirely original.

#### *Properties of the model*

The proposed model has direct physical interpretation and is not only a mathematical formalism. All elements of the OHC's model have their counterparts in the real structure of the cochlea. Physical explanation of the role that play the individual components of a single OHC, including the role that plays nonlinear capacitance, is given.

The developed parametric amplification model may describe in a natural way the following effects and phenomena occurring in the human cochlea:

- 1) how mechanical signals are transformed into electrical signals and *vice versa* (direct and inverse piezoelectric effect),
- 2) how power amplification phenomenon occurs in an input electromechanical amplifier, by the operation of the input electromechanical transistor,
- 3) in which manner, sharpening of the output characteristics and power amplification due to parametric amplifier operation takes place,
- 4) how operates the input electroacoustic transducer (electromechanical transistor or electromechanical control element). Operation of the electromechanical transistor will be the subject of the next publication.

#### *State of the art*

Cochlear modeling has a long tradition, beginning with Helmholtz in 1863, who was the first to elaborate an anatomically motivated cochlear model (the so-called resonance tonotopic model) (HELMHOLTZ, 1954). The essentials of this model were confirmed by von Bekesy's physiological measurements performed some half a century later (1928). He observed traveling acoustic waves along the cochlear basilar membrane (VON BEKESY, 1960). Von Bekesy's discovery gave rise to formulation of a passive hydrodynamic model.

GOLD (1948) hypothesized that cochlear tuning and selectivity is the result of a feedback system consisting of a mechanical-to-electrical transduction process coupled to a piezoelectric-like electric-to-mechanical one. A number of models of cochlear mechanics have been proposed, which incorporate some form of the feedback enhancing sensitivity and frequency selectivity. In these models the details of the electromechanical properties of the OHC are not included, but rather the OHC role is simulated via a force acting on the basilar membrane. Gold further assumed that an active process in the cochlea is similar to that taking place in regenerative or super-regenerative receivers that can provide power amplification.

Recently, in 2000 Eguiluz et al. introduced a Hopf bifurcation concept to describe observed experimental phenomena (EGUILUZ *et al.*, 2000; CAMALET *et al.*, 2000). In this model an active amplifier is represented by a set of mechanosensors, which are maintained at the threshold of an oscillatory (Hopf) instability.

However, this concept is very complicated mathematically and unphysical. In practice, this type of amplifiers (Hopf) can be unstable, generates large amount of noise and is characterized by low fidelity of signal reproduction.

The existing so far models (RAMAMOORTHY *et al.*, 2007; DEO, GROSH, 2005; RATTAY *et al.*, 1998; MAGNASCO, 2003) of the hearing processes are often complex and incomplete, without direct representation in the human auditory organ.

### Assumptions

In the proposed model of an individual OHC it is assumed that:

- 1) the energy for power amplification is taken from the (DC) voltage source (the endolymphatic potentials),
- 2) the input and output circuits of the OHC's amplifier are electro-mechanical circuits,
- 3) the input and output signals are mechanical (acoustic) signals,
- 4) signal processing, power amplification and tuning take place on the electrical side. Input and output mechanical (acoustic) signals are transformed into electrical side and *vice versa* through the piezoelectric effect,
- 5) each individual OHC along with surrounding structures of the TM and the BM can be regarded as an active electromechanical OHC's resonant circuit that has one characteristic frequency (CF), at which it can perform resonance vibrations,
- 6) an individual OHC acts as a sensor (transforms an input acoustic signal into an electrical signal), actuator (transforms an output electrical signal into a mechanical displacement), piezoelectric electroacoustic transducer, power amplifier and the parametric amplifier,
- 7) each individual OHC together with adjacent regions of the TM and the BM can be treated as an electromechanical OHC's parametric receiver.

### OHC's parametric amplifier

Electromechanical parametric amplifier established on the basis of a single OHC performs the following functions:

- 1) amplifies the power of the input acoustical (electrical) signal,
- 2) increases the amplitude of the input acoustical (electrical) signal,
- 3) enhances the selectivity of the OHC's receiver (larger Q-factor),
- 4) improves the sensitivity of the OHC's receiver.

In this article, a physical model of a single OHC and electrical equivalent circuits (series and parallel) of a single OHC, operating as a parametric amplifier, are presented.

The selectivity of the OHC's parametric amplifier is analyzed by solving differential equations of the Mathieu and Ince type.

Analytical formula for power gain that occurs in the OHC's parametric amplifier has also been derived.

Elements of the developed model are uniquely linked to the physical elements that are present in a single OHC and in the whole cochlea, and to the

physical phenomena occurring in a single OHC and in the inner ear (electromotility, the flow of ionic currents, piezoelectricity, nonlinear capacitance, stereocilia movements, movements of the basilar and tectorial membrane, nonlinear effects, metabolic processes in the stria vascularis, etc.).

The proposed model is a physical model, not a phenomenological model. Phenomenological models are constructed so as to match the input and output characteristics. By contrast, in the proposed model, the numerical values of the individual elements of the electrical equivalent circuit are consistent with the physiological data.

The development of mathematical and electronic base of physical model of a single OHC will enable designing of a new generation of hearing aids and cochlear implants, more functional and efficient than those used today. The results obtained in this paper can be used to design various types of electro-acoustic devices, musical instruments and audio coding systems, new in relation to currently existing codecs of the MP-3 type.

The outline of this paper is as follows. Section 2 presents a new electromechanical model for an individual outer hair cell. Section 3 shows electrical equivalent circuits of an OHC's electromechanical resonator operating as a parametric amplifier. Analysis of the selectivity of an OHC's parametric amplifier is included in Sec. 4. Section 5 describes parametric mechanism of power amplification in the cochlea. Conclusions and directions for future research are given in Sec. 6.

## 2. Proposed electromechanical (physical) model of an individual OHC

An individual OHC is a complex electromechanical system. To describe physical phenomena occurring in the cochlea the following concepts are used:

- 1) piezoelectricity,
- 2) the ionic currents,
- 3) the transistor effect,
- 4) nonlinear capacitance,
- 5) parametric amplifier.

Electrical and mechanical phenomena that occur in a single OHC are mutually coupled via the piezoelectric effect (WEITZEL *et al.*, 2003). An individual OHC can be represented by a hollow piezoelectric cylinder with a nonlinear capacitance between inner and outer surface of the OHC, located between the BM and TM in the cochlea, see Fig. 1. OHCs have a diameter of 9 microns and a length that ranges from 20 microns to 90 microns depending of their location in the cochlea.

The OHC's capacitance consists of the linear and nonlinear component (IWASA, 1994). The nonlinear component is due to movement of localized charges and

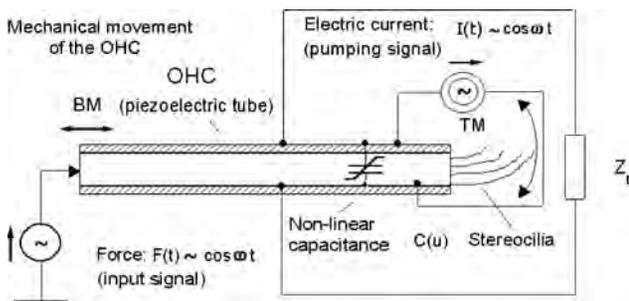


Fig. 1. Hypothetical simplified electro-mechanical diagram of a single OHC operating as a parametric amplifier.  $F(t)$  – represents the input acoustic signal,  $Z_L$  represents the electrical load impedance,  $C(u)$  – is a nonlinear capacitance.

is a prominent feature of the OHC (SPECTOR, 2005). The magnitude of the nonlinear component is comparable to that of the linear component. Non-linear (voltage-dependent) capacitance measured across the inner and outer cylindrical surfaces of the OHC is one of the major features of the electrical behavior of the OHC. From the circuit theory point of view this capacitance is a nonlinear reactive element. The capacitance varies like a bell-shaped function with respect to changes in transmembrane voltage.

Here, the sequence of the physical phenomena that occurs in an individual OHC is presented:

1. The incoming acoustic signal sets in motion the BM and TM.
2. The movement of BM acts on a considered OHC and causes its motion in relation to the TM. Consequently deflection of OHC' stereocilia occurs, what triggers the operation of electro-mechanical input transistor.
3. Subsequently, charging of nonlinear capacitance occurs. Electrical signal which pumps a nonlinear capacitance is represented by an AC current source ( $I(t) \sim \cos \omega t$ ). This source is produced by an input electromechanical amplifier, which is based on the electromechanical transistor. This transistor transforms changes in the mechanical deflection of stereocilia (see right side of Fig. 1) to the channel conductance variations and consequently to changes in the ion channel current ( $K^+$  ions) (OSPECK *et al.*, 2003). Here there is a power gain on a principle similar to the operation of the electronic (unipolar) field effect transistor with a modulated channel conductance. A thorough analysis of an electro-mechanical transistor will be the subject of future work.
4. The input driving signal is an acoustic (mechanical) signal (represented by the force  $F(t) \sim \cos \omega t$  or velocity  $v(t)$ ) acting on the piezoelectric tube from the basilar membrane (BM) side, (see left side of Fig. 1). Through the direct piezoelectric effect this input (acoustic) mechanical signal is transferred to the electrical side as a voltage source  $E(t)$  and subsequently is amplified in the parametric amplifier based on a nonlinear capacitance  $C(u)$ . Power to the nonlinear capacitance is supplied by AC electric pump signal represented by a variable current source  $I(t)$ .
5. Electric signal amplified in a parametric amplifier, as an output electrical signal is dissipated (performs useful work) at the electrical load impedance ( $Z_L$ ). The output useful electrical signal is then transferred to the mechanical side by the inverse piezoelectric effect.
6. The resulting useful output signal is therefore the mechanical signal (force velocity or displacement). The power of this output mechanical signal can surpass many times the power of an input acoustic signal. Subsequently, this output mechanical signal acts on the Basilar Membrane (BM) and Tectorial Membrane (TM), and consequently also on the Inner Hair Cells (IHC). IHCs are sensors which transform mechanical signals into electrical signals (electric pulse trains).

Since the parametric amplifier is a highly selective system, so this amplifier can get a very narrow output characteristics (sharp tuning).

Summing up, the proposed OHC's model implies the following sequence of the physical phenomena occurring in an individual OHC:

- 1) incoming acoustical signal  $\Rightarrow$ ,
- 2) basilar membrane movement  $\Rightarrow$ ,
- 3) mechanical stimulation (deflection) of stereocilia  $\Rightarrow$ ,
- 4a) small "ac" voltage signal generation in the OHC due to the direct piezoelectric effect (at the BM side)  $\Rightarrow$ ,
- 4b) ion channels activation and inactivation (at the TM side)  $\Rightarrow$ ,
- 5) operation of the input electromechanical transistor,
- 5a) "ac" current source generation  $\Rightarrow$ ,
- 6) pumping of the nonlinear capacitance  $\Rightarrow$ ,
- 7) parametric amplification of the power of the small "ac" input voltage signal  $\Rightarrow$ ,
- 8) inverse piezoelectric effect  $\Rightarrow$ ,
- 9) large axial mechanical displacement of the OHC  $\Rightarrow$ ,
- 10) augmentation of the basilar and tectorial membrane mechanical displacement (output useful mechanical signal)  $\Rightarrow$ ,
- 11) IHCs (inner hair cells) movement  $\Rightarrow$ ,
- 12) generation of sequences of electrical impulses in neurons.

### 3. Electrical equivalent circuits of an individual OHC

#### 3.1. Non-linear equivalent circuit of an electromechanical OHC's resonator

Analyzing an electromechanical system from Fig. 1, one can specify the following nonlinear electrical equivalent circuit (series) of a single OHC, see Fig. 2.

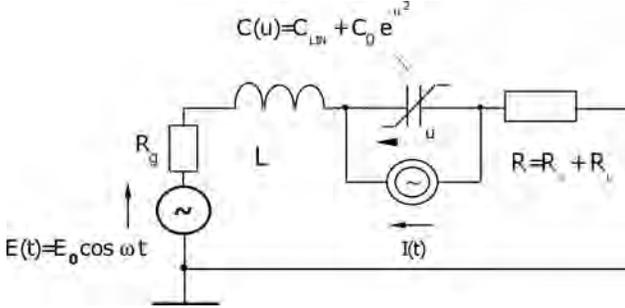


Fig. 2. Hypothetical simplified Thevenin electrical equivalent circuit of the OHC's parametric amplifier treated as a nonlinear oscillator. Circuit represents the operation of a single OHC.

In Fig. 2,  $E(t)$  represents an input electrical signal,  $C(u)$  is the sum of linear capacitance and nonlinear capacitance with a Gaussian bell shape voltage dependence,  $I(t)$  represents the pumping (AC) signal,  $L$  represents the loading mass of a single OHC along with the surrounding TM and BM areas,  $R_g$  is the resistance of the input electrical (source),  $R_s$  represents losses of a nonlinear oscillator circuit, and  $R_L$  is the load resistance.

Using the electrical equivalent circuit of Fig. 2, one can analyze the variations of the electrical variables such as voltage, current and charge flowing in the circuit, and as a consequence the course of the physical phenomena occurring in the individual OHC's resonators.

Elements of the resonant circuit (inductance  $L$ , linear and nonlinear capacitance, load resistance  $R_L$ , loss resistance  $R_s$ , along with an input signal represented by the voltage source  $E(t)$ , and resistance  $R_g$ , and pumping signal  $I(t)$ ), constitute the electrical equivalent circuit of an electromechanical parametric amplifier formed from an individual OHC. In this parametric amplifier, the energy flows from a strong pump signal  $I(t)$  to weak useful signal wave represented by  $E(t)$ .

By analyzing the circuit from Fig. 2, one can investigate the linear and nonlinear behavior of the OHC's parametric amplifier.

In this work a quantitative analysis of physical phenomena occurring in an individual OHC is performed, using the linearized model of the parametric amplifier based on a single OHC. Analysis of nonlinear effects occurring in a single OHC will be the subject of future work.

#### 3.2. Linearized equivalent circuits of an OHC's resonator

##### 3.2.1. Series (Thevenin) equivalent circuit

For moderate level of input driving signal, assuming that the pumping signal (represented by the current source  $I(t)$ ) is much larger than the input signal (represented by the voltage source  $E(t)$ ), the system from Fig. 2 can be linearized.

Consequently, one arrives at the equivalent circuit such as in Fig. 3. In this circuit,  $C'_0 = C_{Lin} + C_0$  is the static component of an overall capacitance of the OHC's resonator.  $C(t)$  is a time varying capacitance resulting from the pumping of the nonlinear capacitance by the electrical current  $I(t)$ , see Fig. 2. Inductance  $L$  represents the inertial effects in the OHC's resonator. Resistance  $R_s$  represents losses in the OHC's resonator.  $E(t)$  is the driving voltage generated by an input force  $F(t)$  and transformed into the electrical side through the direct piezoelectric effect.  $R_L$  is the load resistance, and  $R_g$  is the resistance of an input voltage source.

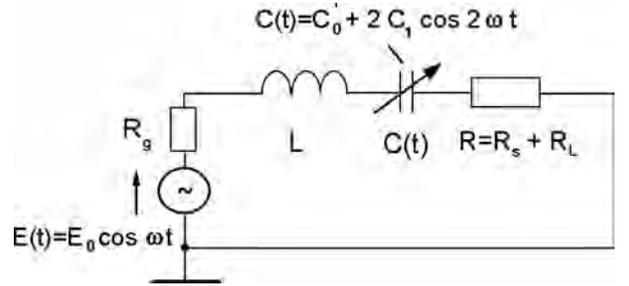


Fig. 3. Linearized (lumped-element) Thevenin electrical equivalent circuit of an individual OHC operating as an electromechanical resonator.

Differential equation that describes the flow of charge in the circuit of Figure 3 is a linear differential equation with variable coefficients:

$$\frac{d^2q}{dt^2} + \frac{\omega_0}{Q_{\text{pass}}} \frac{dq}{dt} + \omega_0^2 \left( 1 - \frac{2C_1}{C'_0} \cos 2\omega t \right) q = \frac{1}{L} E(t), \quad (1)$$

where  $q$  is the electric charge flowing in the circuit,  $\omega_0$  is the angular resonant frequency ( $\omega_0^2 = 1/LC'_0$ ), and  $Q_{\text{pass}}$  is the quality factor of the passive resonant circuit ( $Q_{\text{pass}} = \omega_0 L / (R + R_g)$ ).

Equation (1) is an inhomogeneous differential equation of the Mathieu type. This equation describes the operation of the parametric OHC's amplifier for moderate levels of driving input signals.

By solving Eq. (1) one can determine the electric charge, current and voltage distribution in the circuit and describe the electrical and electromechanical properties of the OHC's resonator (e.g., obtain the output

frequency amplitude characteristics, power amplification etc.).

### 3.2.2. Parallel (Norton) equivalent circuit

A parallel (Norton) equivalent circuit of an individual OHC (see Fig. 4), similar to the series equivalent circuit from Fig. 3, can also be specified.

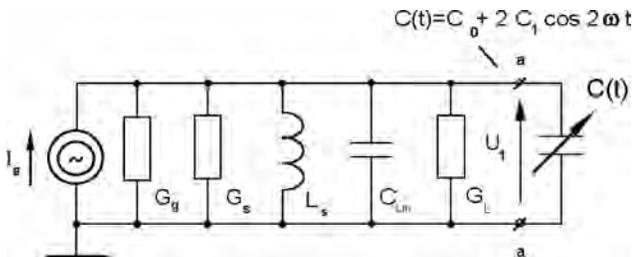


Fig. 4. Norton linear equivalent circuit of a degenerate OHC's parametric amplifier.

$I_g$  is the source current,  $G_g$  is the source conductance,  $G_L$  is the load conductance,  $G_s$  is the loss conductance,  $L_s$  is the inductance of the OHC's resonator,  $C(t)$  is a time varying capacitance.

As shown in Fig. 4, the input electric signal represented by the current source  $I_g$  of admittance  $G_g$  is amplified in the parametric amplifier (formed with a time-variable capacitance). After amplification, this signal is dissipated as the useful signal at the output load conductance  $G_L$ . The useful output power is then transformed (through the inverse piezoelectric effect) to the mechanical side, where it performs useful work on the TM and BM, which in turn will stimulate further IHCs.

Applying Kirchhoff's laws to the circuit from Fig. 4, the following differential equation that describes voltage ( $u_1$ ) changes in the circuit is obtained:

$$\frac{1}{\omega_0^2} \left( 1 + \frac{2C_1}{C'_0} \cos 2\omega t \right) \frac{d^2 u_1}{dt^2} + L_s (G_w - 4C_1 \omega \sin 2\omega t) \frac{du_1}{dt} + u_1 = L_s \frac{dI_g}{dt}, \quad (2)$$

where  $G_w = G_g + G_L + G_s$ ,  $C'_0 = C_{Lin} + C_0$ ,  $\omega_0^2 = 1/L_s C'_0$ .

Equation (2) is a nonhomogeneous differential equation of the Ince type. Solving this differential equation Eq. (2), one can also determine the time course of the voltage  $u_1$  on the load conductance, output frequency characteristics, and power gain.

## 4. Selectivity (sharpening of frequency characteristics)

Single OHC with adjacent areas of the BM and the TM acts as a selective parametric power amplifier of an input acoustic signal. In addition to power, the

amplitude of the input signal (deflection, velocity or voltage) is also amplified. Consequently, amplification of the input signal amplitude is also accompanied by an increase in the quality factor of the OHC's resonance vibrating circuit. This makes it also narrower resonance curve of the oscillating OHC's resonant circuit.

The result of the input signal power gain is:

- 1) increasing the level of the input signal (displacement, velocity, amplitude or current),
- 2) enhancing the sensitivity of the OHC's receiver,
- 3) improving the selectivity of the OHC's receiver (higher quality factor).

Damped Mathieu differential equation Eq. (1) describes the behavior of the resonant OHC's parametric amplifier. By solving this equation one can determine the selectivity of the resonant circuit from Fig. 3. For this purpose, Eq. (1) was solved numerically (for various values of frequency about the resonant frequency  $f_0$ ) using Mathcad software package. In this way, the resonant curve of Fig. 5 was obtained. Also one gets a similar curve by solving the differential equation of the Ince type Eq. (2).

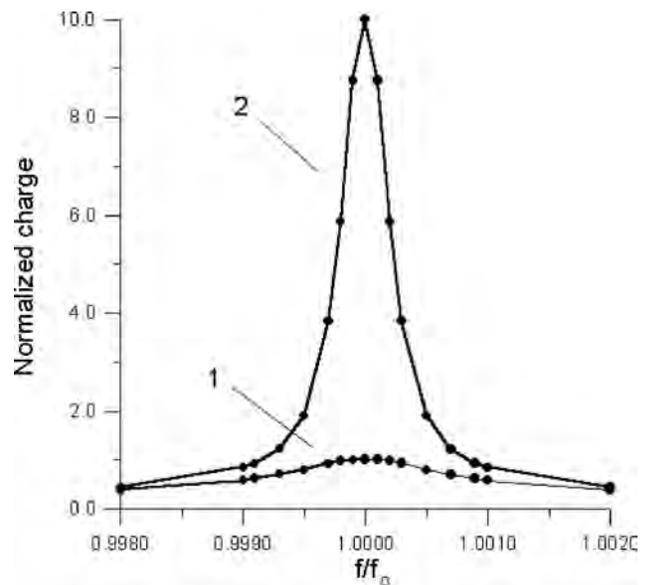


Fig. 5. Sharpening of the resonant curve of the OHC's electromechanical resonator due to the parametric amplification. Curve 1 – passive vibrations, curve 2 – active vibrations.  $f_0$  is a resonant frequency.

The following parameters of the damped vibrating parametric system from Fig. 3 were assumed in the numerical calculations:

$$f_0 = \omega_0 / 2\pi = 1000 \text{ Hz}, \quad Q_{\text{pass}} = 12,$$

$$C_1 / C'_0 = 0.075.$$

The above data were chosen according to the physiological data given in (WEITZEL *et al.*, 2003; OSPECK

*et al.*, 2003; LIU *et al.*, 2006; REINCHENBACH *et al.*, 2010).

Consequently, the resonant curves representing the vibrating OHC's structure were derived in two cases a) passive vibrations of the OHC's structure without parametric effect (curve 1), and b) active vibrations of the OHC's resonator with the parametric amplification (curve 2), see Fig. 5.

It can be seen in Fig. 5, that the resonant curve of the circuit with the variable capacitance and parametric amplification (curve 2) has a higher selectivity (bandwidth is about 10 times narrower), and higher quality factor ( $Q_{\text{act}} = 10 \cdot Q_{\text{pass}}$ ) compared to the resonance curve without the parametric amplification (curve 1).

The amplitude of the amplified signal (active vibration) at the resonant frequency  $f_0$  is 10 times larger than the amplitude of the signal that is not subject to parametric amplification (passive vibration).

Numerical calculations show that the level of the amplitude near the resonant frequency and the quality factor augment with the increase in capacitance modulation ratio  $C_1/C'_0$ .

## 5. Power gain in the OHC's parametric amplifier

The physical mechanism of power amplification in the cochlea is not yet known. As a result, there is no a satisfactory quantitative theory of the power gain occurring in the cochlea. The existing attempts to explain this phenomenon are rather qualitative, descriptive and nonphysical.

In the model presented in this paper the phenomenon of power amplification in an individual OHC resonator is explained using the concept of the parametric amplification. In order to calculate the power gain of the OHC's parametric amplifier, a parallel Norton equivalent circuit from Fig. 4 is used.

The input power amplification in the electromechanical OHC's amplifier occurs in two stages 1 and 2, described below.

The incoming input acoustic signal acts on an individual OHC from the BM side and TM side simultaneously. As a results, the following sequence of events will take place in the OHC:

1. BM movement sets in motion a considered OHC. The movement of an OHC in relation to the TM causes deflection of its stereocilia. The movement of the stereocilia initializes the influx of  $K^+$  ions into the OHC and consequently triggers the operation of an electromechanical input transistor. This transistor in turn amplifies the power of the input signal, and as a consequence loads the nonlinear OHC's capacitance [ $C(u)$  from Figs. 1, 2, and also  $C(t)$  from Figs. 3, 4]. The amplifier formed from the electromechanical transistor is of aperiodic (nonselective) type. The electrical signal, amplified in the

electromechanical transistor amplifier, powers the nonlinear OHC's capacitance. In this way, the modulation of the nonlinear capacitance is produced. For moderate levels of input signals, the nonlinear capacitance can be approximated by the time-variable capacitance  $C(t) = C_0 + 2C_1 \cdot \cos(2\omega_s t)$ , see Figs. 3 and 4. Frequency of the capacitance modulation  $2\omega_s$  is equal to exactly twice the frequency of the input signal,  $\omega_s$ .

2. Input acoustic signal (acting from the BM side) is transformed to the electrical signal through a direct piezoelectric effect. The mass, compliance and mechanical losses of the adjacent areas of the TM and BM are also transformed on the electrical side through a direct piezoelectric effect. These electrical elements, including existing linear and nonlinear (time varying) capacitance, form the electrical equivalent circuit of the parametric amplifier constructed from a single OHC, see Figs. 2, 3 and 4. Due to the parametric amplifier operation, the power gain of the electrical (AC) input signal occurs. Subsequently, by an inverse piezoelectric effect this amplified input electrical signal is transformed into the mechanical side, where it performs a useful work on the BM and TM.

At the same time, due to the parametric amplifier operation, the output frequency characteristics of the OHC are also sharpened, what results in a higher frequency selectivity.

### 5.1. Analytical formula for power amplification

The effective resonant power gain  $k_p$  of the parametric amplifier, represented in Fig. 4, is defined as the ratio of the output power ( $|U_s|^2 \cdot G_L$ ) in the load  $G_L$ , to the available input power ( $|I_g|^2/4G_g$ ) from the source, and can be expressed as follows:

$$k_p = \frac{|U_s|^2 G_L}{|I_g|^2/4G_g} = \frac{4G_g G_L}{(G_g + G_L + G_s + G_{zs})^2}, \quad (3)$$

where  $G_{zs} = -\omega_s C_1$  is the negative conductance introduced to the circuit by the parametric effect, see Eq. (17). This negative conductance is due to the delivery of the additional energy to the circuit by ionic currents  $K^+$ , triggered by stereocilia movement.

The derivation of the analytical expression for the conductance  $G_{zs}$  is given in the Appendix.

Formula (3) can be rewritten in the form:

$$k_p = \frac{4G_g G_L}{G_w^2 (1 - \gamma)^2}, \quad (4)$$

where  $G_w = G_g + G_L + G_s$  and  $\gamma = -\frac{G_{zs}}{G_w} = \frac{\omega_s C_1}{G_g + G_s + G_L}$  is the un-damping coefficient.

The maximum power amplification is achieved when the load conductance  $G_L$  is matched to the source conductance, i.e.,

$$G_{Lopt} = G_g + G_s. \quad (5)$$

Substituting Eq. (5) into Eq. (4) one has:

$$k_{pmax} = \frac{G_g}{G_g + G_s} \frac{1}{(1 - \gamma)^2}. \quad (6)$$

If  $G_s \ll G_g$  then Eq. (6) reduces to

$$k_{pmax} = \frac{1}{(1 - \gamma)^2}, \quad (7)$$

where  $\gamma = \frac{\omega_s C_1}{2G_g}$  is the resulting un-damping coefficient.

Equation (7) shows that the power gain  $k_{pmax}$  depends, among others, on the capacitance coefficient  $C_1$ , which characterizes the level of the pumping signal. Formula (7) was used in the sequel in calculations of the power gain occurring in the OHC's parametric amplifier.

The power gain was calculated for the following set of electrical parameters for the vibrating OHC's parametric resonator (from Fig. 4):

$$P = 10^{-14} \text{ W}, \quad C'_0 = 30 \text{ pF}, \quad Q_{pass} = 12,$$

where  $P$  is the available power of an input electrical signal,  $C'_0$  is the static component of an overall capacitance of the OHC's resonator, and  $Q_{pass}$  is the quality factor of the passive resonant circuit (without the parametric amplification).

The estimated value of the input power ( $P = 10^{-14} \text{ W}$ ) corresponds to a moderate level of acoustic driving signal i.e., 50 dB. The results of calculations of the power gain and other electrical parameters characterizing the operation of the OHC's parametric amplifier from Fig. 4 are presented in Table 1.

Table 1. Coefficient of power amplification  $k_{pmax}$ , source conductance  $G_g$ , negative un-damping conductance  $G_{zs}$  introduced to the resonant circuit, input source current  $I_g$ , capacitance modulation ratio  $C_1/C'_0$ , and quality factor  $Q_{act}$  of the active resonant circuit with the parametric amplification.

| Resonant frequency $f_0$ [Hz] | $k_{pmax}$ | $2G_g$ [nS] | $G_{zs}$ [nS] | $I_g$ [pA] | $\frac{C_1}{C'_0}$ | $Q_{act}$ |
|-------------------------------|------------|-------------|---------------|------------|--------------------|-----------|
| 1000                          | 100        | 15.70       | -14.12        | 17.72      | 0.075              | 120       |
|                               | 18         | 15.70       | -11.93        | 17.72      | 0.063              | 50.9      |
| 3000                          | 100        | 47.10       | -42.39        | 30.69      | 0.075              | 120       |
|                               | 18         | 47.10       | -35.80        | 30.69      | 0.063              | 50.9      |

Values of the parameters used in the numerical calculations are in accordance with the experimental data

obtained for mammalian cochleas (WEITZEL *et al.*, 2003; OSPECK *et al.*, 2003; LIU *et al.*, 2006; REINCHENBACH, HUDSPETH, 2010).

As it is seen from the Table 1, the OHC's parametric amplifier can effectively provide power gain of the useful acoustic input signal. Table 1 shows also that the negative conductance  $G_{zs}$ , introduced to the resonant circuit by the parametric amplification, un-damps the circuit, and enhances the effective quality factor,  $Q_{act}$ .

For low levels of input signals the governing equation for the OHC element is a nonlinear differential equation. The solution of this equation provides an augmented value of the quality factor and therefore higher value of amplification of the signal amplitude. Numerical solutions of this nonlinear differential equation will be performed in future works.

## 6. Conclusions

The theory developed in this study, explains quantitatively in a natural way the physical phenomena occurring in an individual OHC. It was shown that the phenomena of power gain and sharpening of the output amplitude frequency characteristics are particularly important. The developed theory also may explain the phenomenon of otoacoustic emission.

It was demonstrated that an individual OHC can operate as an electromechanical parametric amplifier, that can amplify the power of an input acoustic signal and improve selectivity.

Series and parallel electrical equivalent circuits of a single OHC that operates as a parametric amplifier have been established.

Physical and mathematical models describing the course of electrical signals in the developed series and parallel electric equivalent circuits have been formulated.

The derived inhomogeneous Mathieu and Ince differential equations that describe the operation of a single OHC as a parametric amplifier, have been solved numerically. Output amplitude frequency characteristics of a single OHC working as a parametric amplifier are evaluated.

Sharpening of frequency response amplitude (enhancement in quality factor and selectivity), as a result of operation of the parametric amplifier based on a variable in time capacitance, was stated.

Analytical formula for the input signal power gain occurring in the OHC's parametric amplifier was also derived. Depending on the assumed values of the parameters of equivalent circuits and driving level, the opportunity to amplify the input signal power in a wide range (from several to several hundred times) has been stated.

The developed electromechanical model of a single OHC is a physical model, not a phenomenological model. Each element of the model has its physical

counterpart in the actual structure of the cochlea. The proposed in this work a new parametric-piezoelectric model of a single OHC, along with its mathematical and electronic description, is an original contribution to the state-of-the-art in modeling of the human hearing organ.

The presence of the parametric amplification provides a sharp resonant curve of the oscillating circuit (representing a vibrating OHC), and large output signal to noise ratio.

The material presented in this paper is the first stage of the research. In the next step the author aims to develop a complete electromechanical model of the cochlea and description of the physical phenomena occurring in the whole cochlea's structure. Future analysis will be conducted on nonlinear effects occurring in the considered OHC's parametric amplifier.

The developed model conforms qualitatively to the existing experimental data. New refined theory of the hearing processes based on the piezoelectricity and parametric amplification can open new horizons for research, e.g., deeper understanding of the physical phenomena underlying auditory processes, new phenomena in musical and speech acoustics. This will allow the design of new musical instruments and audio codecs, that will be new in relation to the existing MP3 codecs.

## Appendix

The nonlinear OHC's capacitance has a bell shape voltage dependence (IWASA, 1994), therefore it can be approximated by the following Gaussian function:

$$C(u) = C_0 \cdot \exp(-u^2), \quad (8)$$

where  $u$  is a normalized voltage.

Expanding the above formula into the Taylor series gives rise to:

$$C(u) = C_0 \left( 1 - u^2 + \frac{1}{2}u^4 - \dots \right). \quad (9)$$

The incoming pumping signal, that is applied across the nonlinear capacitance, can be expressed as:

$$v = v_1 \cos(\omega t). \quad (10)$$

The voltage given by formula (10) is next substituted into formula (9). As a consequence, the resulting time variable (time-dependent) capacitance is produced:

$$C = C(t) = C_0 + 2C_1 \cos(2\omega_s t) + 2C_2 \cos(4\omega_s t) + \dots + \quad (11)$$

For the useful input electrical signal  $u_s$  of frequency  $\omega_s$ , the electric charge  $\delta Q$  resulting from the parametric effect equals (GOLDE, 1976):

$$\delta Q = C(t)U_s. \quad (12)$$

For a time-harmonic input signal:

$$U_s = U_{sm} \cos(\omega_s t + \varphi_s), \quad (13)$$

where  $\varphi_s$  is a phase shift between the pump and useful input signal, the current flowing through the time varying capacitance is given by:

$$i(t) = \frac{d(C(t)U_s)}{dt}. \quad (14)$$

Inserting Eqs. (11)–(13) into Eq. (14) and neglecting in Eq. (14) all components of frequency other than  $\omega_s$  one obtains:

$$i(\omega_s t) = -C_0 U_{sm} \omega_s \sin(\omega_s t + \varphi_s) - C_1 U_{sm} \omega_s \sin(\omega_s t - \varphi_s). \quad (15)$$

The second term in Eq. (15) is due to the parametric effect and can be represented by the effective admittance  $Y_{zs}$  introduced across the  $a - a$  terminals in Fig. 4.

According to Eq. (13) and Eq. (15) the admittance  $Y_{zs}$  equals:

$$Y_{zs} = j\omega_s C_1 \exp(-2j\varphi_s) = -\omega_s C_1 \sin(2\varphi_s) + j\omega_s C_1 \cos(2\varphi_s) = G_{zs} + jB_{zs}. \quad (16)$$

For  $\varphi_s = -\pi/4$  one has:

$$Y_{zs} = G_{zs} = -\omega_s C_1 < 0. \quad (17)$$

Therefore, the presence of a time varying capacitance  $C_1$  gives rise to a negative conductance  $G_{zs}$  in the circuit from Fig. 4. This negative conductance can be responsible for the power gain in the OHC's parametric amplifier.

## References

1. VON BEKESY G. (1960), *Experiments in Hearing*, McGraw-Hill, New York.
2. DE BOER E. (1995), *On equivalence of locally active models of the cochlea*, Journal of the Acoustical Society of America, **98**, 3, 1400–1409.
3. CAMALET S., DUKE T., JULICHER F., PROST J. (2000), *Auditory sensitivity provided by self-tuned critical oscillations of hair cells*, Proc. Natl. Acad. Sci. U.S.A., **97**, 3183–3188.
4. COHEN A., FURST M. (2004), *Integration of outer hair cell activity in a one-dimensional cochlear model*, Journal of the Acoustical Society of America, **115**, 5, 2185–2192.
5. DEO N. V., GROSH K. (2005), *Simplified nonlinear outer hair cell models*, Journal of the Acoustical Society of America, **117**, 4, 2141–2146.
6. DIMITRIADIS E. K. (1999), *Solution of the inverse problem for a linear cochlear model: A tonotopic cochlear amplifier*, Journal of the Acoustical Society of America, **106**, 4, 1880–1892.

7. EGUILUZ V. M., OSPECK M., CHOE Y., HUDSPETH A. J., MAGNASCO M. O. (2000), *Essential nonlinearities in hearing*, Physical Review Letters, **84**, 5232–5235.
8. ELIOT S. J., SHERA C. A. (2012), *The cochlea as a smart structure*, Smart Materials and Structures, **21**, 064011.
9. GOLD T. (1948), *Hearing. II. The physical basis of the action of the cochlea*, Proc. R. Soc. London B. Biol. Sci., **135**, 492–498.
10. GOLDE W. (1976), *Electronic systems* [in Polish], Vol. II, Chap. 5, WNT, Warsaw.
11. HELMHOLTZ H. L. F. (1954), *On the sensations of Tones as a Physiological Basis for the Theory of Music*, Dover Publications Inc., New York.
12. IWASA K. H. (1994), *A membrane motor model for the fast motility of the outer hair cell*, Journal of the Acoustical Society of America, **96**, 4, 2216–2224.
13. LIAO Z., FENG S., POPEL A. S. (2007), *Outer hair cell active force generation in the cochlear environment*, Journal of the Acoustical Society of America, **121**, 5, 2758–2773.
14. LIU T. K., ZHAK S., DALLOS P., SARPESHKAR R. (2006), *Fast cochlear amplification with slow outer hair cells*, Hearing Research, **214**, 45–67.
15. LIU Y. W., NEELY S. T. (2009), *Outer hair cell electromechanical properties in a nonlinear piezoelectric model*, Journal of the Acoustical Society of America, **126**, 2, 751–761.
16. LOUISELL W. H. (1960), *Coupled Mode and Parametric Electronics*, Chap. 4, Wiley, New York.
17. MAGNASCO M. O. (2003), *A wave travelling over a Hopf instability shapes the cochlear tuning curve*, Physical Review Letters, **90**, 058101-4.
18. MAOILEIDIGH D. O., JULICHER F. (2010), *The interplay between active hair bundle motility and electromotility in the cochlea*, Journal of the Acoustical Society of America, **128**, 3, 1175–1190.
19. MOUNTAIN D. C., HUBBARD A. E. (1994), *A piezoelectric model of outer hair cell function*, Journal of the Acoustical Society of America, **95**, 1, 350–354.
20. OSPECK M., DONG X., IWASA K. H. (2003), *Limiting frequency of the cochlear amplifier based on electromotility of outer hair cells*, Biophys. J., **84**, 739–749.
21. RAMAMOORTHY S., DEO N. V., GROSH K. (2007), *A mechano-electro-acoustical model for the cochlea: Response to acoustic stimuli*, Journal of the Acoustical Society of America, **121**, 5, 2758–2773.
22. RATTAY F., GEBESHUBER I. C., GITTER A. H. (1998), *The mammalian auditory hair cell: A simple electric circuit model*, Journal of the Acoustical Society of America, **103**, 3, 1558–1565.
23. REINCHENBACH T., HUDSPETH A. J. (2010), *A ratchet mechanism for amplification in low-frequency mammalian hearing*, Proc. Natl. Acad. Sci. U.S.A., Vol. 107, No 11, 4973–4978, Supporting Information.
24. SHERA C. A. (2007), *Laser amplification with a twist: Travelling-wave propagation and gain functions from throughout the cochlea*, Journal of the Acoustical Society of America, **122**, 5, 2738–2758.
25. SPECTOR A. A. (2005), *Effectiveness, active energy produced by molecular motors, and nonlinear capacitance of the cochlear outer hair cell*, J. Biomech. Eng., **127**, 391–399.
26. WEITZEL E. K., TASKER R., BROWNELL W. E. (2003), *Outer hair cell piezoelectricity: Frequency response enhancement and resonant behavior*, Journal of the Acoustical Society of America, **114**, 3, 1462–1466.

# Ultrasonic Properties of Magnetic Nanoparticles with an Additional Biocompatible Dextrane Layer

Leszek DĄBEK, Tomasz HORNOWSKI, Arkadiusz JÓZEF CZAK, Andrzej SKUMIEL

*Institute of Acoustics, Faculty of Physics, Adam Mickiewicz University  
Umultowska 85, 61-614 Poznań, Poland; e-mail: hornaku@amu.edu.pl*

*(received October 2, 2012; accepted January 22, 2013)*

The temperature dependence of the particle size distribution (PSD) of the magnetic fluid with an additional biocompatible dextran layer was studied using an ultrasonic method. The measurements of the ultrasound velocity and attenuation were carried out as a function of the volume concentration of magnetite particles at temperatures ranging from 15°C to 40°C. In order to extract the PSD from ultrasonic measurements, the theoretical model of Vinogradov-Isakovich was used. The extraction of PSD from the ultrasonic data requires also the measurements the density and viscosity of the ferrofluid samples. The calculated PSD of the magnetic fluid with an additional biocompatible layer shows a greater thermal stability than that of a magnetic fluid with a single surfactant layer.

**Keywords:** magnetic nanoparticles, ultrasound, particle size distribution.

## 1. Introduction

According to the usual definition, magnetic nanoparticles can attain a size between 1 nm and 100 nm. Their importance relates to the fact that the characteristics of the magnetic nanoparticles are different from those of bulk materials of the same composition, which is due mainly to the size effects, magnetic and electronic properties, and the role played by surface phenomena as the size is reduced. Magnetic nanoparticles dispersed in a carrier fluid are called magnetic liquids or ferrofluids. Due to the possibility of remote controlling the parameters of the fluid containing magnetic nanoparticles with the aid of an external magnetic field, magnetic liquids have attracted considerable interest because of potential applications in technological, biological and medical domains, such as seals, bearings, sensors, drug delivery or magnetic hyperthermia.

The parameter that control the distinctive features of magnetic nanoparticles in magnetic fluids is the particle size which determines the scale of interparticles interactions and the concentrations. Real magnetic fluids are always polydisperse systems with a log-normal particle size distribution and a mean magnetic core diameter of about 10 nm (REGULSKA *et al.*, 2007; RAŞA, 2000).

Since the magnetic moments of the particles are proportional to the volume of their magnetic cores, larger particles can be oriented more easily at low field strengths, and the smaller ones can be oriented only under rather strong magnetic fields applied. As a result, the magnetization properties of a polydisperse system differ from that of a monodisperse system even in the case of very diluted magnetic fluids. More importantly, a small fraction of the large particles in the real system has a typical size of 15–20 nm. Since their dipole-dipole interaction energy is considerably larger than the thermal energy at room temperature, these particles can aggregate into bigger structures. Experiments have demonstrated that this small fraction of large particles can play a major role in determining the physical properties of polydisperse magnetic fluids (BACRI *et al.*, 1990). Aggregation of magnetic particles due to magnetic interactions could block vessels and capillaries when magnetic fluids are used in drug-delivering therapies. Thus prior to the use of a the magnetic fluid in medical therapies it is necessary to study the particle (or aggregate) size distribution (PSD) and its dependence on temperature.

Among the methods commonly used to measure the PSD function and study its properties are ultrasonic techniques, which are recognized as very promising from an industrial point of view (DUKHIN,

GOETZ, 1998). In the so-called ultrasonic spectroscopy (HORNOWSKI *et al.*, 2008), the particle size distribution and concentration of a colloidal dispersion is usually determined by measuring the ultrasonic velocity and/or attenuation coefficient as a function of frequency and using a suitable mathematical model to interpret the spectra. However, it was shown (JÓZEF CZAK *et al.*, 2011) that in case of diluted magnetic fluids it is also possible to extract the parameters of the PSD function from the dependence of ultrasonic data measured only at one frequency on the concentration of magnetic nanoparticles.

The purpose of this paper is to compare the PSDs obtained by the ultrasonic method for two types of magnetic fluids with different structures of the stabilizing layer. In order to achieve the repulsive mechanism preventing agglomeration, the particles were coated by a surfactant that produces an entropic repulsion. This constitutes one stabilizing layer surrounding the magnetic core. However, some biomedical applications require an encapsulation of the magnetic core by an additional inorganic or a polymeric coating that renders the particles biocompatible and may serve as a support for biomolecules. The first type of magnetic fluids studied in this work was composed of a single surfactant stabilizing layer, the second one possesses an additional biocompatible layer.

## 2. Materials and methods

The superparamagnetic nanoparticles were prepared by co-precipitation using ferric and ferrous salts and ammonium hydroxide. The freshly prepared magnetic nanoparticles (magnetite) were sterically stabilized by sodium oleate (*ca.* 1:1 to  $\text{Fe}_3\text{O}_4$ ) to prevent their agglomeration. By centrifuging 30 min at 9000 rpm an initial magnetic fluid (sample 1 – MF), arose. Then, dextran (average molecular weight = 64000 Daltons) was added to achieve a dextran to  $\text{Fe}_3\text{O}_4$  weight ratio equal to 0.5 (sample 2 – MF Dextran) (KONERACKÁ *et al.*, 2010). Dextran is a polymer  $(\text{C}_6\text{H}_{10}\text{O}_5)_n$ , of anhydroglucose having mainly alpha-D(1-6) linkages with some unusual 1,3 glucosidic linkages at branching points (BAUTISTA *et al.*, 2005).

From the magnetization curve obtained using SQUID, the saturation magnetization and volume concentration,  $\phi_0$ , of the magnetite particles in both samples were determined (RAŞA, 2000, ROZYNEK *et al.*, 2011). Samples with six different volume concentrations of magnetite particles have been prepared by diluting the initial ferrofluid with water. The volume concentration of the samples studied can be found in Table 1. This method of preparation of the samples should not affect the particle sizes as these kinds of ferrofluids have been shown to be stable for long periods. In other words, the measured PSD for the diluted

magnetic fluids is independent of the volume concentration of the samples (YOON, 2009).

Table 1. Volume concentration of magnetic material,  $\phi_m$  for the samples studied.

| Samples    | $\phi_0$<br>[%] | $\phi_1$<br>[%] | $\phi_2$<br>[%] | $\phi_3$<br>[%] | $\phi_4$<br>[%] | $\phi_5$<br>[%] |
|------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| MF         | 1.38            | 0.69            | 0.35            | 0.17            | 0.09            | 0.04            |
| MF Dextran | 1.40            | 0.70            | 0.35            | 0.17            | 0.09            | 0.04            |

The ultrasonic measurements were carried out using the ResoScan (Germany) ultrasonic device, which measures the ultrasonic velocity,  $c$ , and attenuation,  $\alpha f^{-2}$ , of the sample, where  $f$  is the frequency of the ultrasonic wave. The resolution of the ultrasonic velocity is  $0.01 \text{ ms}^{-1}$ . The relative error of the ultrasonic attenuation coefficient was calculated from the series of repeated measurements and is better than 5%.

The shear viscosity coefficient was measured using a cone-and-plate Brookfield DV II+ viscometer within  $10^\circ\text{C}$  to  $40^\circ\text{C}$  temperature range with a temperature control uncertainty of  $0.2^\circ\text{C}$ .

The density was measured using a DMA-38 oscillating U-tube Anton Paar density meter that measures sample density values accurately to  $1 \text{ kg m}^{-3}$  in the temperature range of  $15^\circ\text{C}$  to  $40^\circ\text{C}$ .

## 3. Results and discussion

In order to extract the PSD from the ultrasonic data, it was necessary to measure densities and shear viscosities for both the types of magnetic fluids studied in the experiment. The densities of the samples were measured within the temperature range  $15^\circ\text{C}$  to  $40^\circ\text{C}$ ; the results of measurements are shown in Figs. 1 and 2. The variation of the density with temperature for all the samples studied obeys the quadratic relation.

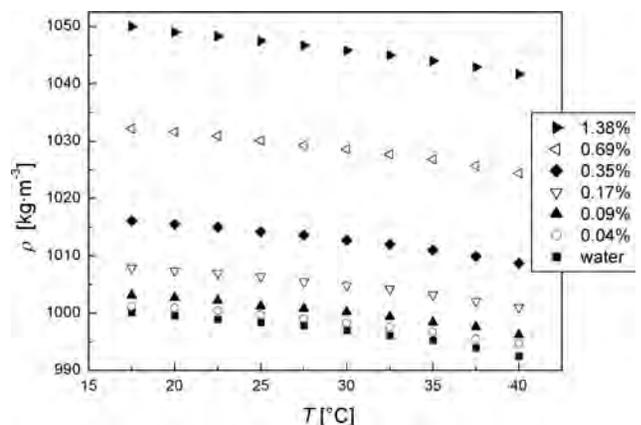


Fig. 1. Temperature dependencies of density of the magnetic fluid with a single surfactant layer (MF) for different concentrations of magnetic nanoparticles.

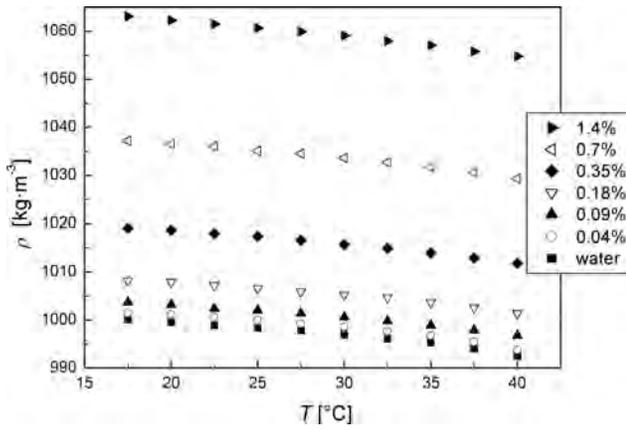


Fig. 2. Temperature dependencies of density of the magnetic fluid with both the surfactant and dextran layers for different concentrations of magnetic nanoparticles.

Figures 3 and 4 show the temperature dependence of the viscosity for both the types of the magnetic fluids

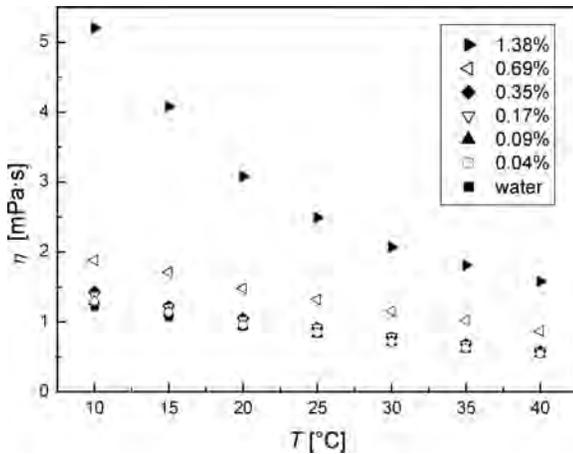


Fig. 3. Temperature dependence of the shear viscosity in magnetic fluids with a single surfactant layer for different concentrations of magnetic nanoparticles.

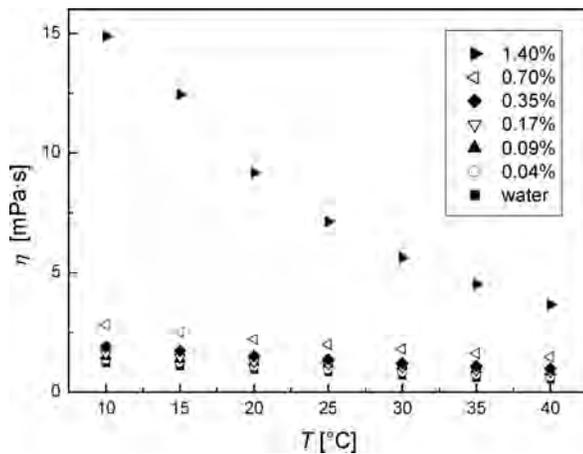


Fig. 4. Temperature dependence of the shear viscosity in magnetic fluids with both surfactant and biocompatible layers for different concentrations of magnetic nanoparticles.

studied. The viscosity increases after the addition of dextran and decreases with temperature according to the Arrhenius law (JÓZEF CZAK *et al.*, 2013).

To calculate the PSD function from ultrasonic attenuation data, information about the various physical parameters of the solid particles and the liquid dispersion medium is required. The studied ferrofluids were considered as aggregates composed of magnetite/oleic acid in the case of the magnetic fluid with a single surfactant layer and as magnetite/oleic acid+dextran in the case of the magnetic fluid with both surfactant and biocompatible shells. It was assumed that these aggregates are spherical and dispersed in water. The aggregates densities, the oleic acid and dextran volume concentrations can be calculated from the physicochemical model described in details in (HORNOWSKI *et al.*, 2008; JÓZEF CZAK *et al.*, 2011).

The results of the ultrasonic velocity measurements are shown in Figs. 5 and 6. The results obtained are in

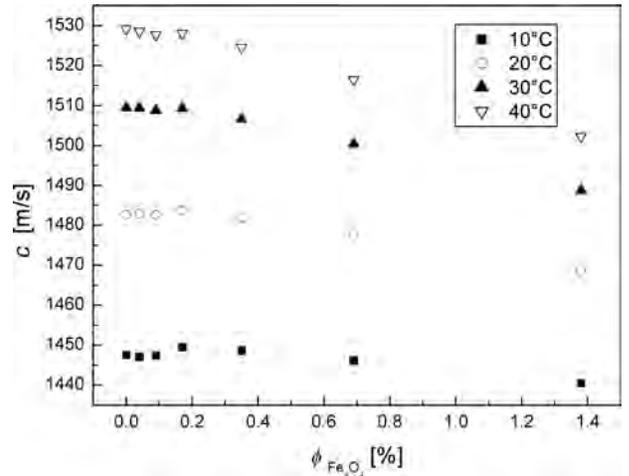


Fig. 5. Ultrasonic wave velocity as a function of the magnetite volume concentration in magnetic fluids with a single surfactant layer for different temperatures.

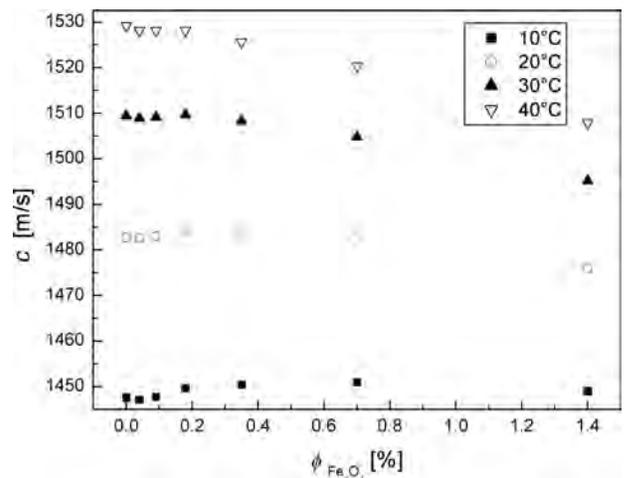


Fig. 6. Ultrasonic wave velocity as a function of magnetite volume concentration in magnetic fluids with both surfactant and biocompatible layers for different temperatures.

a qualitative agreement with simple macroscopic theories that predict a parabolic relationship between the ultrasonic velocity in a colloidal mixture and the volume concentration of the continuous phase (CHALLIS *et al.*, 2005).

Figures 7 and 8 shows the experimental results of ultrasonic attenuation,  $\alpha f^{-2}$ , as a function of the magnetite particle concentration for temperature range of 10°C to 40°C. For all the temperatures the ultrasonic attenuation increases with increasing volume concentration of the magnetite particles.

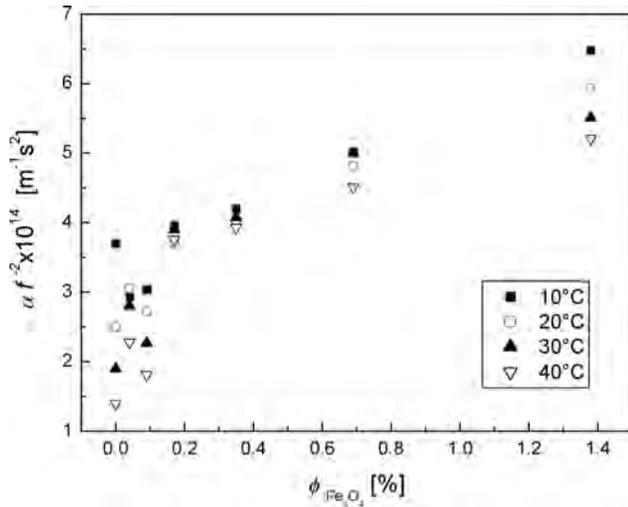


Fig. 7. Ultrasonic wave attenuation as a function of the magnetite volume concentration in magnetic fluids with a single surfactant layer for different temperatures.

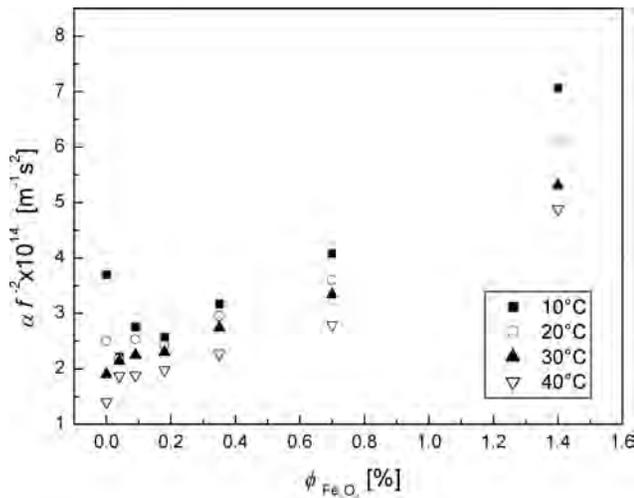


Fig. 8. Ultrasonic wave attenuation as a function of magnetite volume concentration in magnetic fluids with both surfactant and biocompatible layers for different temperatures.

The interaction between the acoustic wave and the suspended particles of nanometer size leads to an additional attenuation of the sound compared to that in the carrier liquid,  $\alpha_0$ . This additional term  $\Delta\alpha$  describes the attenuation that can be attributed to

visco-inertial, thermal, and scattering losses (DUKHIN, GOETZ, 2001). The additional attenuation depends on the particle size distribution that can be assumed to obey a log-normal distribution characterized by the density as a sum of the absorption coefficient of dispersion:

$$p(R) = \frac{1}{\sqrt{2\pi}\sigma(R-R_0)} \cdot \exp\left(-\frac{[\ln(R-R_0)-m]^2}{2\sigma^2}\right). \quad (1)$$

Here  $m$  is the logarithmic mean,  $\sigma$  is the logarithmic standard deviation, and  $R_0$  is a minimal radius. The mean  $m$  and standard deviation  $\sigma$  which give the best agreement between the measured and predicted ultrasonic attenuation can be found by a least squares analysis.

The mean value  $M_\phi$ , and root-mean-square deviation  $D_\phi$ , of the particle radius  $R$  can be determined from the expressions:

$$M = \exp(m + 0.5\sigma^2) + R_0, \quad (2)$$

$$D = \exp(2m + \sigma^2)(\exp \sigma - 1). \quad (3)$$

To extract the parameters of the particle size distribution from ultrasonic attenuation data, the theoretical model proposed by Vinogradov and Isakovitch (VINOGRADOV, 2003) was used.

In this model three mechanisms are responsible for the additional ultrasound attenuation in ferrofluids with respect to that of the carrier liquid: visco-inertial absorption,  $\alpha_\eta$ , associated with the difference in density between the magnetic nanoparticles and the carrier liquid, thermal absorption,  $\alpha_T$ , associated with the difference in the thermal properties between the magnetic particles and the carrier liquid and scattering losses,  $\alpha_S$ , associated with the difference in the acoustic impedance. Neglecting the scattering effect which is very small for the wave with frequency equal to 8 MHz, the overall ultrasonic attenuation in the magnetic liquid per frequency squared can be conveniently expressed as the sum of the first two contributions:

$$\frac{\alpha}{f^2} = \frac{\alpha_\eta}{f^2} + \frac{\alpha_T}{f^2}. \quad (4)$$

The contributions to the attenuation of sound due to viscoinertial and thermal processes are given by

$$\frac{\alpha_\eta}{f^2} = \frac{K\pi\phi}{c} \int_{R_{\min}}^{R_{\max}} f(R)R^2 \times \frac{Q(R, f) + S(R, f)q_3\tau_V f}{Q(R, f)^2 + W(R, f)} dR + \frac{2\pi^2\tau_V}{c}, \quad (5)$$

where

$$K = \frac{4\pi(1-\phi)(\rho_a - \rho_f)^2}{9\rho\eta_f}, \quad Q(R, f) = 1 + R\sqrt{q_1f},$$

$$S(R, f) = R\sqrt{q_1 f} + \frac{2}{9}q_1 R^2 f + q_2 R^2 f,$$

$$W(R, f) = q_1 R^2 f + \frac{4}{9}R^3 \sqrt{(q_1 f)^3} + 2q_2 R^3 \sqrt{q_1 f},$$

$$q_1 = \frac{\pi \rho_f}{\eta_f}, \quad q_2 = \frac{4\pi(1-\phi)\rho_a \rho_f}{9\rho \eta_f},$$

$$q_3 = \frac{4\pi \rho_a^2}{(1-\phi)(\rho_a - \rho_f)^2}, \quad \tau_V = \frac{\alpha_f}{f^2} \frac{\rho_f c_f^3}{2\pi^2 \rho c^2},$$

$$\frac{\alpha_T}{f^2} = -\frac{\phi}{f^2} \int_{R_{\min}}^{R_{\max}} f(R) \cdot \text{Im} \left[ \frac{j\lambda_a \lambda_f [Z_a - \tanh(Z_a)] (Z_f + 1)}{\lambda_a [Z_a - \tanh(Z_a)] + \lambda_f (Z_f + 1) \tanh(Z_a)} \right] dR,$$

$$Z = \frac{3\rho T c}{2R^2} \left( \frac{\gamma_a}{\rho_a C_a} - \frac{\gamma_f}{\rho_f C_f} \right)^2,$$

$$Z_a = (1+j)R\sqrt{\pi f \rho_a C_a \lambda_a^{-1}}, \quad j = \sqrt{-1},$$

$$Z_f = (1+j)R\sqrt{\pi f \rho_f C_f \lambda_f^{-1}}.$$

Here  $\rho$ ,  $\beta$ ,  $\eta$ ,  $c$  and  $\alpha$  are the density, adiabatic compressibility, viscosity, velocity and sound absorption,  $T$  is the temperature of the medium, and  $C$ ,  $\lambda$ , and  $\gamma$  are the heat capacity at constant pressure, the heat conductivity coefficient and the thermal expansion coefficient. The subscripts  $a$  and  $f$  refer to the properties of the magnetite particles and the continuous phase, respectively.

Figures 9 and 10 show the log-normal PSDs obtained by the ultrasonic method for various temperatures in the range of 10°C to 40°C.

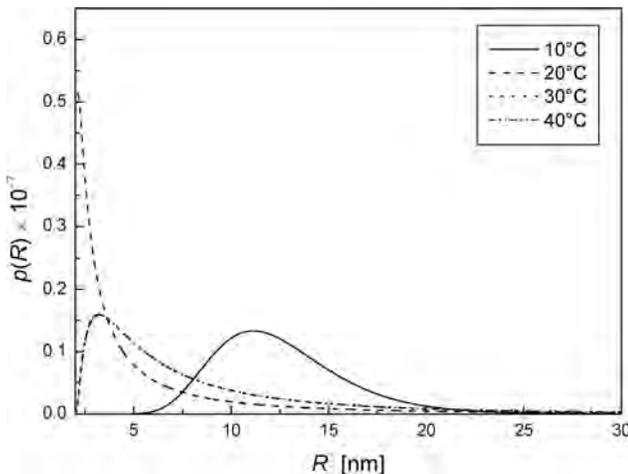


Fig. 9. Log-normal PSD functions in magnetic fluids with a single surfactant layer obtained from ultrasonic measurements for different temperatures.

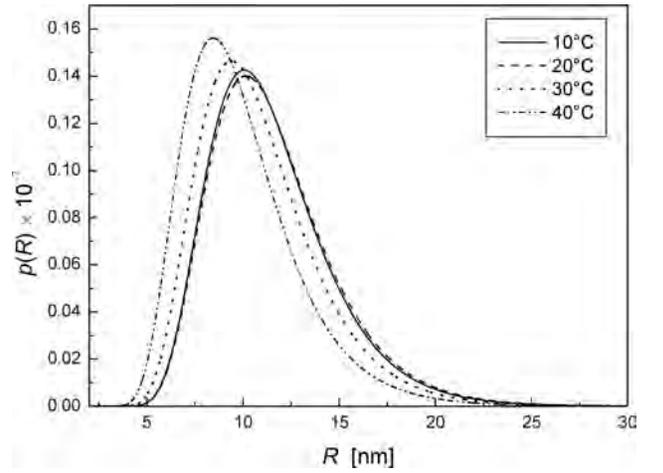


Fig. 10. Log-normal PSD functions in magnetic fluids with both the surfactant and biocompatible layers obtained from ultrasonic measurements for different temperatures.

Figures 9 and 10 present PSD functions extracted from the ultrasonic attenuation data for both the types of magnetic fluids studied in the experiment. The values of the PSD function parameters,  $m$  and  $\sigma$ , as well as the values of  $M_\phi$  and  $D_\phi$  for all temperatures are listed in Table 2. The mean particle radius  $M_\phi$ , evaluated from ultrasonic measurements is bigger by twice as much as the average magnetic nanoparticle size. This is because the ultrasonic granulometry gives the value of the so-called “hydrodynamic” radius which is greater than the size of the magnetic core by a magnitude,  $\delta_s + \delta_m$ , where  $\delta_s$  denotes the thickness of the protective surfactant layer/layers and  $\delta_m$  is the thickness of the magnetically inactive layer on the surface of the particles (PSHENICHNIKOV *et al.*, 1996).

Table 2. Parameters  $m$  and  $\sigma$  of the log-normal particle size distribution, mean value  $M_\phi$ , and root-mean-square deviation  $D_\phi$  of particle radius  $R$  calculated from ultrasonic data for different temperatures.

| Sample     | $T$ [°C] | $-m$   | $\sigma$ | $M_\phi \times 10^9$ [m] | $D_\phi \times 10^{18}$ [m <sup>2</sup> ] |
|------------|----------|--------|----------|--------------------------|---|
| MF         | 10       | 18.414 | 0.312    | 12.57                    | 40.90                                     |
|            | 20       | 20.194 | 1.590    | 8.01                     | 140.10                                    |
|            | 30       | 19.292 | 1.118    | 9.82                     | 125.78                                    |
|            | 40       | 18.947 | 0.857    | 10.53                    | 125.78                                    |
| MF+Dextran | 10       | 18.529 | 0.329    | 11.47                    | 34.95                                     |
|            | 20       | 18.517 | 0.331    | 11.59                    | 36.11                                     |
|            | 30       | 18.603 | 0.347    | 10.85                    | 32.49                                     |
|            | 40       | 18.723 | 0.370    | 9.91                     | 28.05                                     |

The PSDs of magnetic fluid with a single surfactant layer shown in Fig. 9 change substantially with temperature. This instability can be related to the creation and destruction of magnetic particle aggregates and is confirmed by rather large values of the root-

mean-square deviation  $D_\phi$  of the particle radius. However, at human body temperature this process seems to run slower. The addition of the dextran layer has a dramatic impact on the magnetic fluid PSD as can be seen in Fig. 10. The PSDs show now an excellent stability for all temperatures. This proves that apart from providing additional biocompatibility, the dextran layer can also prevent agglomeration of magnetic nanoparticles.

#### 4. Conclusions

The effect of a dextran additional layer on the particle size distribution was studied. The ultrasonically determined PSD functions show that the surface modification of magnetic nanoparticles with dextran enhances the stability of biocompatible magnetic fluids. This finding is in agreement with our studies of the rheological properties of dextran modified magnetic nanoparticles (JÓZEF CZAK *et al.*, 2013).

#### Acknowledgment

This work was supported by the Polish National Science Centre grant number DEC-2011/03/B-ST7/00194. The authors gratefully acknowledge Professor Milan Timko from the Slovak Academy of Science (Košice) for samples of the magnetic fluids.

#### References

1. BACRI J. C., PERZYNSKI R., SALIN D. (1990), *Ionic ferrofluids: A crossing of chemistry and physics*, J. Magn. Magn. Mater., **85**, 27–32.
2. BAUTISTA M. C., BOMATI-MIGUEL O., MORALES M. P., SERNA C. J., VEINTEMILLAS-VERDAGUER S. (2005), *Surface characterisation of dextran-coated iron oxide nanoparticles prepared by laser pyrolysis and coprecipitation*, J. Magn. Magn. Matter., **293**, 20–27.
3. CHALLIS R. E., POVEY M. J., MATHER M. L., HOLMES A. K. (2005), *Ultrasound techniques for characterizing colloidal dispersions*, Rep. Prog. Phys., **68**, 1541–1637.
4. DUKHIN A. S., GOETZ P. J. (1998), *Characterization of aggregation phenomena by means of acoustic and electroacoustic spectroscopy*, Colloids Surf. A, **144**, 49–58.
5. DUKHIN A. S., GOETZ P. J. (2001), *Acoustic and electroacoustic spectroscopy for characterizing concentrated dispersions and emulsions*, Adv. Colloid Interface Sci., **92**, 73–132.
6. HORNOWSKI T., JÓZEF CZAK A., ŁABOWSKI M., SKUMIEL A. (2008), *Ultrasonic determination of the particle size distribution in water-based magnetic liquid*, Ultrasonics, **48**, 594–597.
7. JÓZEF CZAK A., HORNOWSKI T., ROZYNEK Z., SKUMIEL A., FOSSUM J. O. (2013), *Rheological study of dextran modified magnetite nanoparticles water suspension*, Int. J. Thermophys., to be published.
8. JÓZEF CZAK A., HORNOWSKI T., SKUMIEL A. (2011), *Temperature dependence of particle size distribution in transformer oil-based ferrofluid*, Int. J. Thermophys., **32**, 795–806.
9. KONERACKÁ M., ANTOŠOVÁ A., ZÁVIŠOVÁ V., LANČA G., GAŽOVÁ Z., ŠIPOŠOVÁ K., JURÍKOVÁ A., CSACH K., KOVÁČ J., TOMAŠOVICOVÁ N., FABIÁN M., KOPČANSKÝ P. (2010), *Characterization of Fe<sub>3</sub>O<sub>4</sub> Magnetic Nanoparticles Modified with Dextran and Investigation of Their Interaction with Protein Amyloid Aggregates*, Acta Phys. Pol. A, **118**, 983–985.
10. PSHENICHNIKOV A. F., MEKHONOSHIN V. V., LEBEDEV A. V. (1996), *Magneto-granulometric analysis of concentrated ferrocolloids*, J. Magn. Magn. Mater., **161**, 94–102.
11. RAŠA M. (2000), *Magnetic properties and magneto-birefringence of magnetic fluids*, Eur. Phys. J. E, **2**, 265–275.
12. REGULSKA P., SKUMIEL A., HORNOWSKI T., JÓZEF CZAK A. (2007), *Magnetic properties and anisotropy of ultrasound attenuation in APG-832 magnetic liquid*, Archives of Acoustics, **32**, 4 (Supplement), 95–100.
13. ROZYNEK Z., JÓZEF CZAK A., KNUDSEN K. D., SKUMIEL A., HORNOWSKI T., FOSSUM O., TIMKO M., KOPČANSKÝ P., KONERACKÁ M. (2011), *Structuring from nanoparticles in oil-based ferrofluids*, Eur. Phys. J. E, **34**, 28.
14. VINOGRADOV A. N. (2003), *Application of Acoustic Spectroscopy to Investigation of Microinhomogeneous Media*, Colloid J., **65**, 539–544.
15. YOON S. (2009), *Magnetic Methods for Determining the Monodispersity of Ferrofluids*, J. Korean Phys. Soc., **54**, 163–168.

## Detection of Objects Buried in the Sea Bottom with the Use of Parametric Echosounder

Eugeniusz KOZACZKA<sup>(1),(2)</sup>, Grażyna GRELOWSKA<sup>(1)</sup>, Sławomir KOZACZKA<sup>(1)</sup>,  
Wojciech SZYMCZAK<sup>(1)</sup>

<sup>(1)</sup> *Polish Naval Academy*

Śmidowicza 69, 81-103 Gdynia, Poland; e-mail: g.grelowska@amw.gdynia.pl

<sup>(2)</sup> *Gdańsk University of Technology*

Narutowicza 11/12, 80-233 Gdańsk, Poland; e-mail: kozaczka@pg.gda.pl

(received October 30, 2012; accepted January 24, 2013)

The paper contains results of a *in situ* research main task of which was to detect objects buried, partially or completely, in the sea bottom. Object detecting technologies employing acoustic wave sources based on nonlinear interaction of elastic waves require application of parametric sound sources. Detection of objects buried in the sea bottom with the use of classic hydroacoustic devices such as the sidescan sonar or multibeam echosounder proves ineffective. Wave frequencies used in such devices are generally larger than tens of kHz. This results in the fact that almost the whole acoustic energy is reflected from the bottom. On the other hand, parametric echosounders radiate waves with low frequency and narrow beam patterns which ensure high spatial resolution and allows to penetrate the sea bottom to depths of the order of tens of meters. This allows to detect objects that can be interesting, among other things, from archaeological or military point of view.

**Keywords:** sea bottom, sea bottom acoustics, buried objects.

### 1. Introduction

Detection of objects occurring on or under the seabed surface presents a challenge for researchers interested in exploration of the sea bottom. The problem relates to objects buried at depths of up to several tens of meters from the seabed surface. Finding such objects is the subject of interest for a wide group of professionals, including archaeologists, marine safety specialists, and the military responsible for defending coastal waters. One of the currently developing non-invasive remote sensing methods consists in the use of phenomena accompanying nonlinear propagation of elastic waves.

Hydroacoustic examination of the seabed upper layer requires systems with high directivity beams in order to minimize sediment reverberation (GALLOWAY, COLLINS, 1998; HAMILTON *et al.*, 1999). Parametric sonar systems meet this requirement and generate low-frequency/wide-band beams with extremely concentrated main lobes (GRELOWSKA *et al.*, 2010; KOZACZKA *et al.*, 2012a). Due to their compar-

atively small dimensions and weight, parametric systems can be easily mounted on ROVs (remotely operated underwater vehicles) or AUVs (autonomous underwater vehicles); for this reason, they can be used as e.g. relocalization sensors for one-shot disposal systems combating mines buried in the seabed.

There are various types of high-resolution sub-bottom profiling systems, differing mainly in energy sources and receiving elements, their specific merits and demerits, and fields of application (WALTER *et al.*, 1997; TURGUT, 1998; STERNLICHT, MOUSTIER, 2003). One of the most popular and widely used sub-bottom profiling systems is that utilizing air gun(s) as the energy source and a separate receiving cable for recording reflected acoustic signals. A much more accurate system, called the parametric echosounder, is based on parametric sound generation. Probably the best known solution of this type is TOPAS (Topographic Parametric Sonar manufactured by Kongsberg) that allows to penetrate sea floor up to thousands of meters and is a superior sub-bottom profiling system as far resolution is concerned but is less

popular due to its high cost. There are also other mobile parametric sediment echosounder systems available that allow to carry out surveys in shallow waters. The ultimate objective of this technique is to provide a spatially detailed and resolved picture of the seafloor and the subsurface sediment structures. High resolution seismic surveys are primarily confined to the uppermost 80 meters of sediments. This is the area used by majority of typical engineering applications. It is estimated that about 80% of sub-seabed technical infrastructure is located in the upper 15 to 20 meters. Some typical major applications include reconnaissance geological surveys, minerals exploration, foundation studies for offshore platforms, detailed site surveys for engineering projects, cable and pipeline route investigations, harbor development, and environmental studies.

However, single-beam echosounders, even parametric ones, provide information on the seabed only immediately below the surveying vessel. The footprint on the seabed varies in size, depending on the water depth and the local slopes, but is generally large. Seafloor coverage will therefore be variable and rather small.

The main feature of the parametric echosounder consists in generating a sounding pulse of frequency that can be set between 4 kHz and 15 kHz and occurs as a consequence of interaction of two main sounding pulses of higher frequencies, e.g. 100 kHz and 115 kHz. In the device used in the course of experiments described in this survey, the sea bottom was sounded with low frequencies sounding pulse using small-size antenna and additionally gaining narrow main lobe without side lobes.

The fields of most extensive commercial use of this technique include the oil/gas industry and the subsea engineering. High-resolution seismic/sub-bottom profiling surveys provide essential information necessary to make decisions concerning oil rig/platform site selection. Cable and pipeline route investigations need a very detailed picture of the top few tens of meters of the sediment, and sub-bottom reflection profiling method is the primary source of such information.

The technique of precise sub-bottom survey has one more application important for safety at sea. Nowadays, mass destruction weapons are frequently placed in shallow waters in a way making them very difficult to find. Detecting such objects in the sea requires the use of devices that offer possibility to penetrate sediment covering them.

This paper presents results of experimental research aimed at detection of underwater objects with the use of a device called the sub-bottom profiler. The research area was the Southern Baltic, with particular interest focused on the Gulf of Gdańsk.

Some examples of actual acoustics images obtained during the sea trials will be shown in the following together with physical interpretation.

## 2. *In situ* measurements

A bathymetrical measurement system with sub-bottom profiler was installed on a 10.5 m long small survey vessel. The parametric echosounder antenna was installed on the starboard and the multibeam EM3002 transducer on the port side 100 cm below water surface, with both devices fixed to special mounting arms. Additional navigation devices were tested in different locations and finally mounted in places optimal from the point of view of their functions. GPS receiver was installed in the center line of the vessel, close to the deck in order to minimize speed and position errors. The motion sensor MRU-Z was fitted near the vessel's center of gravity. After mounting the devices, measurements were made to define lever arms for each bathymetrical unit. GPS position, heading, and motion speed sensor signals were distributed to different devices used for the sea bottom investigation. Small measurement vessel with calibrated measuring units was used during trials on Gdańsk Bay. Some interesting results of sounding and processing methods will be brought up.

The sub-bottom structure was investigated with the use of parametric echosounder SES-2000 manufactured by Innomar. This is a nonlinear transducer source which simultaneously transmits two signals with slightly different high frequencies at high sound pressures. Nonlinear interactions generate new frequencies in water, one of them being the difference frequency that has a bandwidth similar to the primary frequency. Both the primary HF signal (100 kHz) and the secondary LF signal (6 to 12 kHz) were recorded. Penetration occurred up to a few tens of meters in soft sediments. Advantages of the parametric acoustic system include:

- 1) small beam width at low frequencies;
- 2) deep penetration with high resolution of sediment layers and objects;
- 3) accurate depth measurements with the high frequency signal.

Data processing was carried out with the processing software ISE 2.9 which allows to edit and export layers to ASCII data, extend signal processing, convert and export data, and correct results for tide, water sound velocity, and GPS  $z$ -level.

Conditions of elastic wave propagation in the Baltic Sea depend strongly on hydrological parameters. For that reason, vertical distributions of sound speed as well as temperature and salinity were determined before measurements using STD/CTD sounder (SAIV A/S).

Echograms of the investigated area were taken during the research project devoted to detection of objects buried in the sea bottom. The purpose of this study was to observe the sea bottom structure and compare the obtained results to data given on the geological

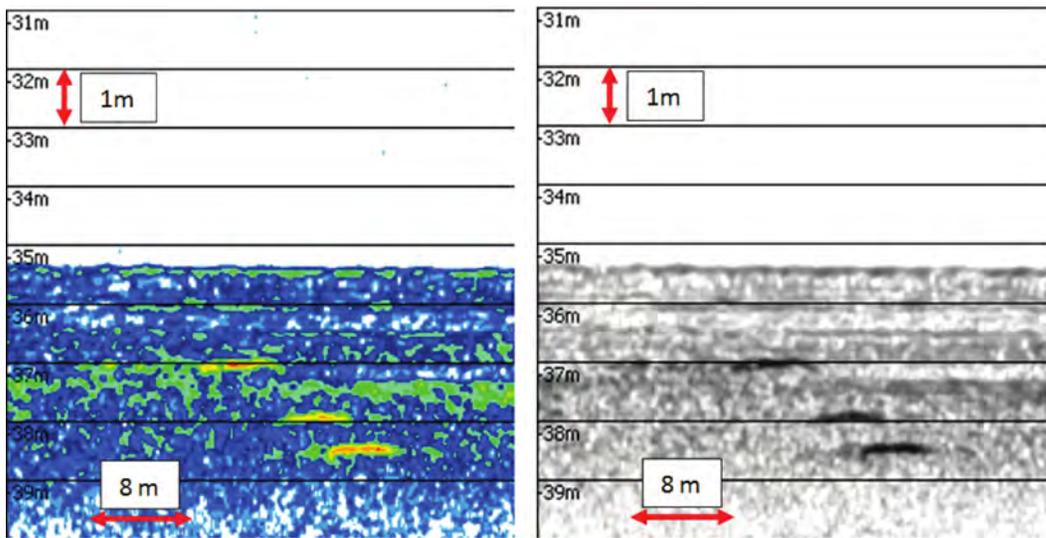


Fig. 1. Echogram with buried objects (the same echogram, but in different color scales).

map of the region (USCINOWICZ, ZACHOWICZ, 1994). Measurements were carried out along the paths crossing the Gulf of Gdańsk.

Some examples of data collected during sea bottom measurements are given in the following figures. They allow to assess the penetration properties of the equipment and determine the presence of buried objects as shown in Fig. 1.

High-resolution sediment echo-sounding allows to differentiate between sediment layers with different impedances (WUNDERLICH, MÜLLER, 2003). Typically, an image is characterized by presence numerous distinct closely spaced continuous parallel horizontal reflectors. There are particularly strong major reflectors within such vertical sequence. Parametric echosounders allow to obtain information based only on perpendicular reflection from layer boundaries

(KOZACZKA *et al.*, 2012b). Investigations of wave propagation within individual layers have not been carried out to date. Acoustic penetration of the Gulf of Gdańsk ranged from about 5 to 40 meters depending on the seabed geoacoustic parameters.

To extract more information on the seabed structure, special post-processing software was developed. Data converted from the software dedicated to the parametric echosounder were imported to Matlab programming scripts where raw recorded signals (reflected from sediments) were processed. It allowed to present data collected during the measurements in different ways (as an envelope or typical signal). Some of them were helpful in obtaining more detailed information about the objects, especially relative strengths of the targets and their longitudinal dimensions, as in the example shown in Fig. 2.

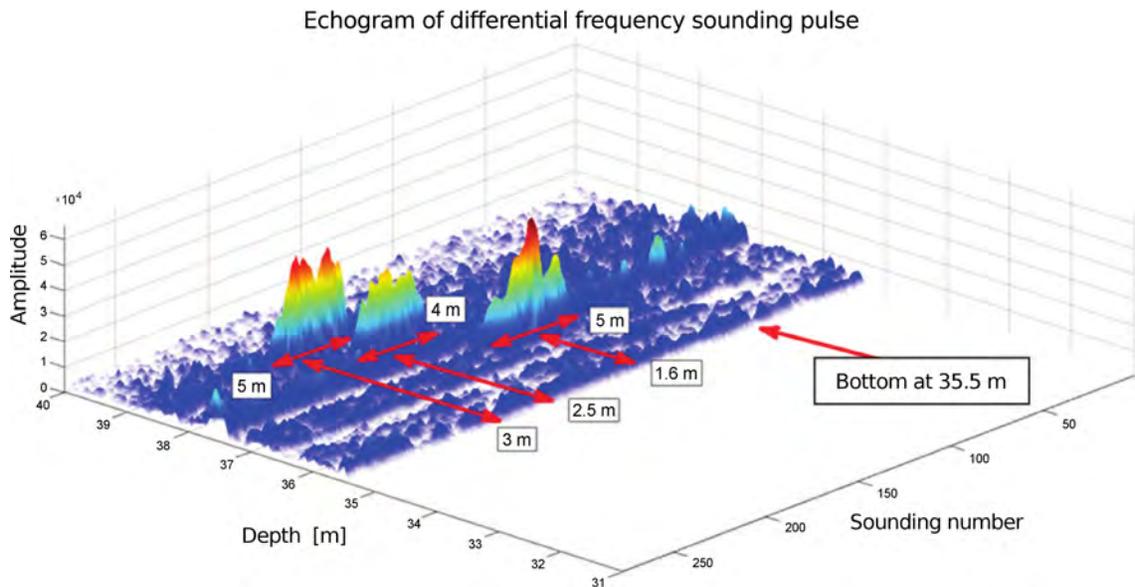


Fig. 2. Data from Fig. 1 after post-processing.

However, the problem how to differentiate remotely between objects with various shapes remains still unsolved. For this reason, measurements on a stationary range were carried out. Different types of objects of known shapes shown in Fig. 4 were used in the experiment (an object with the form close to Manta mine and steel canisters with different types of material inside).

The equipment used in this experiment was the same as in the measurements carried out in the sea. The configuration of the measurement range is shown in Fig. 3. Antenna was mounted on a stable, aluminium holder. Objects were located directly under the transducer on the bottom surface and buried at different depths. All equipment used for the trials was placed

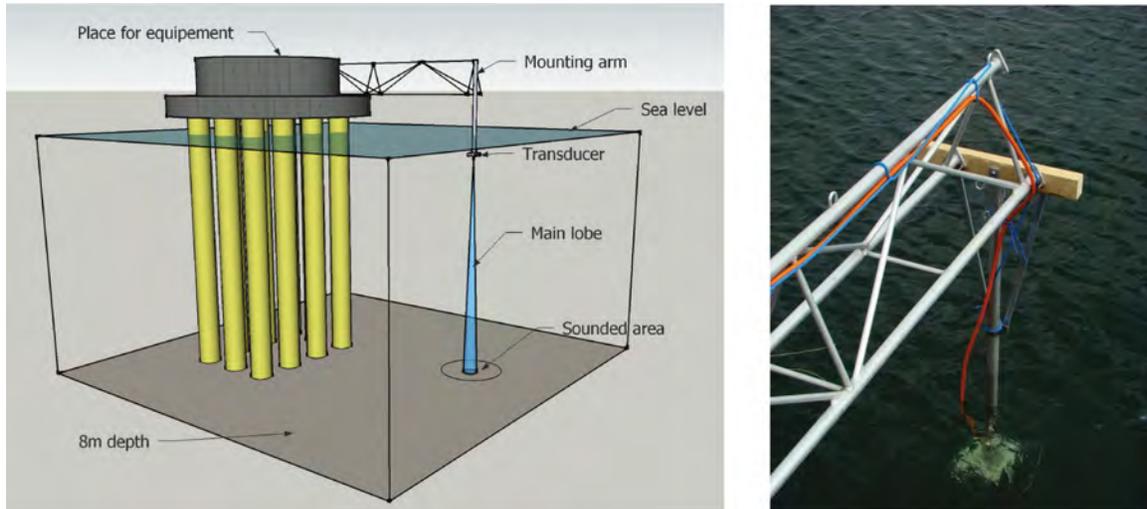


Fig. 3. Stationary range for *in situ* measurements.

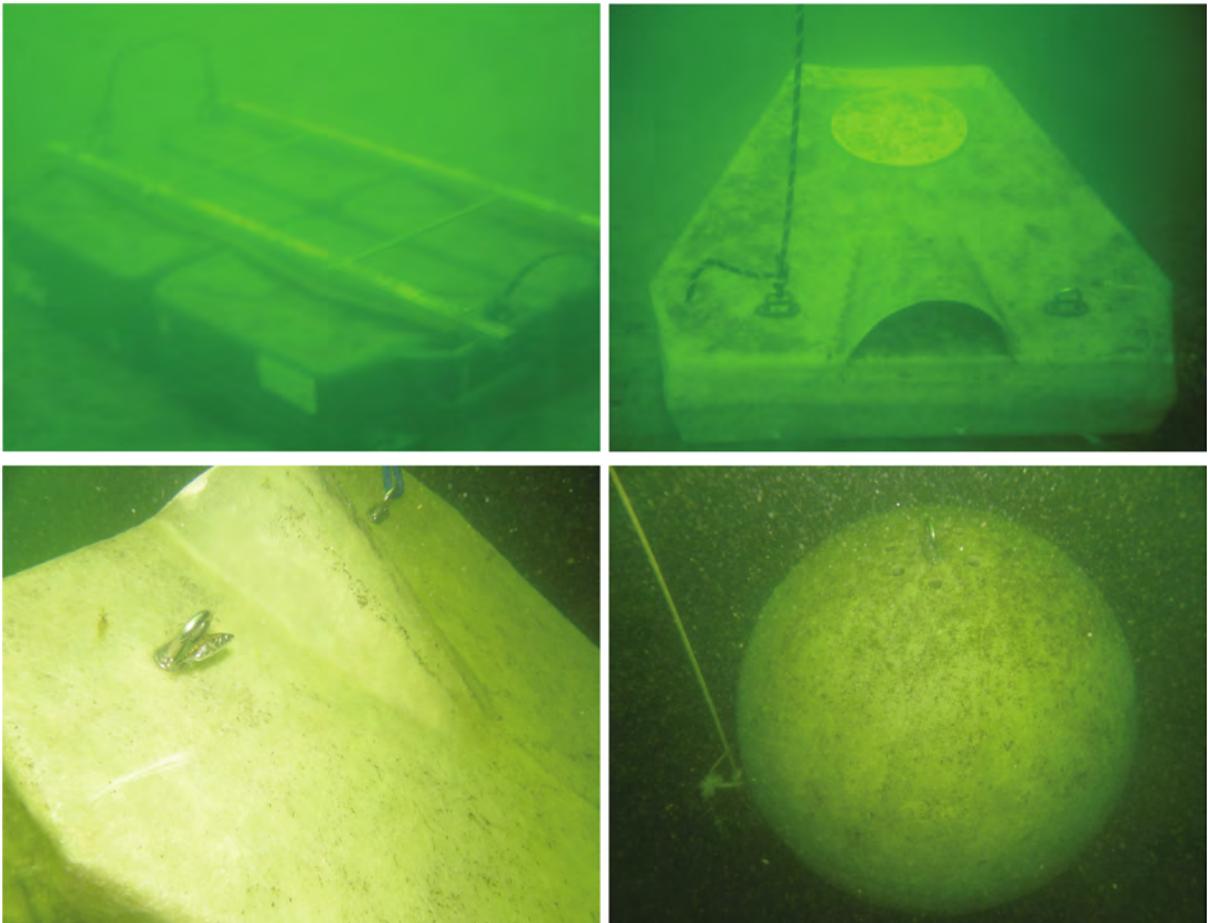


Fig. 4. Objects used as targets.

on a concrete base and powered via a long power cable plugged in ashore.

In the course of the experiment, echoes obtained from the bottom free of any objects were compared to data with an object or objects covered with the sediment and buried. Examples of the results are shown in Figs. 5 and 6. Results of sounding are shown as a comparison between the clear bottom image and the same area with objects placed on or under the seabed surface. In Fig. 5, first 440 pings correspond to the clear bottom, and the second part of visualization (pings numbered 441 to 1000) represents the object (a group of canisters filled with dry sand) placed on the bottom. Figure 6 can be interpreted the same way but there is a difference in location of canisters that are buried and covered with a 5 cm deep layer of seabed material.

### 3. Conclusions

Echograms of Southern Baltic bottom and sub-bottom obtained by means of parametric echosounder confirm usefulness of the device that allows to determine remotely the structure of the upper part of sea bottom and locate object buried in the seabed.

However, we cannot differentiate remotely between several different materials of sediments. This task requires a lot of experimental work in the course of which the acoustically obtained echograms should be attributed to sedimentary structures determined from sediment cores.

An important advantage of the presented high-resolution sub-seabed survey method would be the possibility to identify objects posing a threat in some cir-

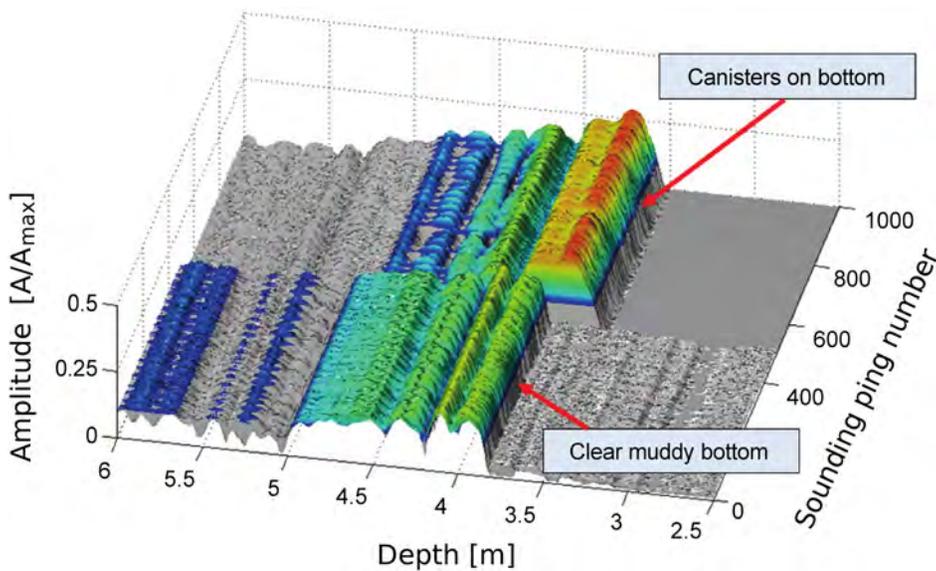


Fig. 5. A comparison of echograms: bottom clear and canisters on the seabed.

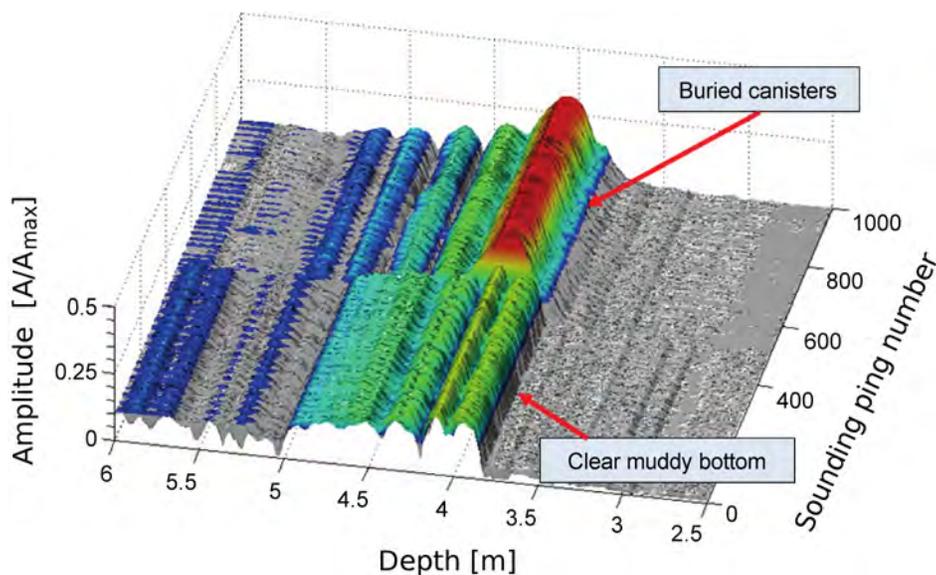


Fig. 6. A comparison of echograms: bottom clear and canisters buried in the seabed.

cumstances. Remote assessment of the type of buried objects needs a lot of experimental investigation and creation of a database of acoustical characteristics corresponding to different targets.

### Acknowledgment

The investigation was partially supported by the Ministry of Science and Higher Education in the framework of the fund for statutory activities of the Polish Naval Academy.

### References

1. GALLOWAY J., COLLINS W. (1998), *Dual frequency acoustic classification of seafloor habitat using the QTC view*, Oceans'98, Nice, France.
2. GRELOWSKA G., KOZACZKA E., SZYMCZAK W. (2010), *Methods of data extraction from sub-bottom profiler's signal*, Hydroacoustics, **13**, 109–118.
3. HAMILTON L., MULHEARN P., POECKERT R. (1999), *Comparison of RoxAnn and QTC-view acoustic bottom classification system performance for the Cairns area, Great Barrier Reef, Australia*, Continental Shelf Research **19**, 1577–1597.
4. KOZACZKA E., GRELOWSKA G., KOZACZKA S. (2012a), *Searching of the buried objects in the sea bottom by means of noninvasive methods*, Proceedings of the 11th European Conference on Underwater Acoustics.
5. KOZACZKA E., GRELOWSKA G., KOZACZKA S., SZYMCZAK W. (2012b), *Processing data on sea bottom structure obtained by means of the parametric sounding*, Polish Maritime Research, **19**, 4(76), 3–10.
6. STERNLICHT D. D., MOUSTIER CH. P. (2003), *Time-dependent seafloor acoustic backscatter (10–100) kHz*, J. Acoust. Soc. Am., **114**, 2709–2723.
7. TURGUT A. (1998), *Method and apparatus of classifying marine sediment*, Technical Report U.S.
8. USCINOWICZ S., ZACHOWICZ J. (1994), *Geological map of the bottom of the Baltic Sea, 8 sheets and pamphlet, scale 1:200,000*, Polish Geological Institute, Warsaw.
9. WALTER D., LAMBERT D., YOUNG D., STEPHENS K. (1997), *Mapping sediment acoustic impedance using remote sensing acoustic techniques in a shallow-water carbonate environment*, Geo-Marine Letters, **17**, 260–267.
10. WUNDERLICH J., MÜLLER S. (2003), *High-resolution sub-bottom profiling using parametric acoustics*, International Ocean Systems, **7**, 4, 6–11.

# Matched-field Source Localization with a Mobile Short Horizontal Linear Array in Offshore Shallow Water

Dexin ZHAO, Zhiping HUANG, Shaojing SU, Ting LI

*Department of Instrument Science and Technology, College of Mechatronics and Automation  
National University of Defense Technology*

47, Yanwachi Street, Changsha City, Hunan Province, 410073, P. R. China; e-mail: derekzhao27@yahoo.com

*(received July 12, 2012; accepted January 25, 2013)*

Passive source localization in shallow water has always been an important and challenging problem. Implementing scientific research, surveying, and monitoring using a short, less than ten meter long, horizontal linear array has received considerable attention in the recent years. The short array can be conveniently placed on autonomous underwater vehicles and deployed for adaptive spatial sampling. However, it is usually difficult to obtain a sufficient spatial gain for localizing long-range sources due to its limited physical size. To address this problem, a localization approach is proposed which is based on matched-field processing of the likelihood of the passive source localization in shallow water, as well as inter-position processing for the improved localization performance and the enhanced stability of the estimation process. The ability of the proposed approach is examined through the two-dimensional synthetic test cases which involves ocean environmental mismatch and position errors of the short array. The presented results illustrate the localization performance for various source locations at different signal-to-noise ratios and demonstrate the build up over time of the positional parameters of the estimated source as the short array moves at a low speed along a straight line at a certain depth.

**Keywords:** passive source localization, matched-field processing, inter-position processing, short horizontal array, shallow water.

## 1. Introduction

Passive sonar is a method for detecting acoustic signals in an underwater environment and is typically deployed in the form of long towed arrays measuring up to a thousand feet or longer to enhance the ability of detecting acoustic sources (BERNECKY, KRZYCH, 2008). The emphasis changes from detection to parameter estimation when the localization stage of sonar processing arrives (HAVELOCK *et al.*, 2008; HODGES, 2010). A number of approaches to beamforming and other acoustic source localization techniques exist. These generally include plane-wave beamforming, wavefront curvature ranging, target motion analysis, multipath ranging, and matched-field processing (MFP).

Plane-wave beamforming is the most mature of these schemes and the easiest to implement. It is assumed that the source is in the far field of the array. However, it only provides bearing information for a single linear array, and its accuracy diminishes as the source moves from the broadside to the end-fire

(HODGES, 2010; BERNECKY, KRZYCH, 2008). Wavefront curvature ranging is only accurate for near-range sources under the assumption of a spherical spread for the acoustic wavefront (HAVELOCK *et al.*, 2008).

Target motion analysis is capable of determining the source's trajectory (i.e. range, course, and speed) in the open ocean where the ray assumption is valid. It is a method that builds up the source positional parameters over time through the motion (constant velocity or maneuvering) of the source and/or the receiver. Clearly, its estimation process strongly depends on the type of measurement data and the model for the source motion and other specific problem features (HAVELOCK *et al.*, 2008; WILSON, VEENHUIS, 1997; INCE *et al.*, 2009; HODGES, 2010; BERNECKY, KRZYCH, 2008).

Multipath ranging incorporates the structure of sound propagation in the ocean into its estimation process. It is assumed that the sound radiated from a distant source arrives at the receiver via multiple paths having simple geometric interpretation (e.g. surface and bottom reflections). The process is tedious,

prone to error, and only works in specific environments (HAVELOCK *et al.*, 2008; BERNECKY, KRZYCH, 2008). It should be noted that all of the source localization techniques described thus far also do not work well in shallow water.

The shallow water environment is extremely complex, and the assumption of plane waves or other relatively simple sound propagation model in the processing scheme can lead to severe degradation of the estimation (DEBEVER, 2009; EHLERS *et al.*, 2010). As an alternative, the technique based on MFP exploits the complexity of the ocean's structure to improve source localization. Thus, the complex shallow water environment actually aids the estimation process (TOLSTOY, 1993; BAGGEROER *et al.*, 1993). This process constitutes a range-depth ambiguity surface, as it spatially correlates the actual field (measured at an array of sensors emitted from a point source at a particular location) with the replica field (computed by a numerical propagation model over a grid of hypothetical source locations). The maximum match on the ambiguity surface is regarded as the estimated source location. Given the sufficient ocean environmental information, MFP has been shown to be a promising signal processing technique. It has been used in practical applications from the stage of scientific experiments with the advent of powerful microprocessors and various optimal MFP algorithms (DEBEVER, KUPERMAN, 2007; TOLLEFSEN, DOSSO, 2009; WILSON, VEENHUIS, 1997; KIM *et al.*, 2010; FIALKOWSKI *et al.*, 1997; BERNECKY, KRZYCH, 2008).

The intent of this study is to apply MFP to a short horizontal linear array for passive source localization in a shallow water environment. The short horizontal linear array specifically refers to a passive sonar system equipped on a relatively small sensor platform, such as an autonomous underwater vehicle (MILLARD, 2003). This type of sensor platform can be conveniently deployed at any desired site in the ocean, for scientific research, surveys, and industry. It has excellent characteristics – such as low self-noise and vibration coupled with a high stability – all of which are desirable for many acoustic measurements. However, the platform also limits the array length to less than ten meters. In MFP theory, for the vertical linear array, increasing the array length to span more of the water column can significantly improve MFP sidelobe reduction and peak resolution, and the horizontal linear array usually requires a much longer array length than the vertical array for the same localization performance (TOLSTOY, 1993; TANTUM, NOLTE, 2000). The short arrays can hardly obtain sufficient spatial gain. The direct application of MFP to such arrays would result in a presumed failure in estimating the source positional parameters. The technique of synthetic aperture processing is usually used to synthesize a longer array when dealing with the problem of this

kind of mobile short horizontal linear array. The corresponding goal is to increase the bearing estimation by combing data from widely separated sampling points (AUTREY, 1988; FERNANDEZ *et al.*, 2004; WILLIAMS, HARRIS, 1992; XUDONG *et al.*, 2007). Inspired by the synthetic aperture technique, a localization approach that combines matched-field processing and inter-position processing is expected to allow for the short horizontal linear array capable of passively localizing long-range acoustic sources in range and depth rather than only bearing for the two-dimensional scenario. MFP primarily provides the likelihood of passive source localization in shallow water, and inter-position processing is used to improve the localization performance and stabilize the estimation process. Thus, the estimated source position is built up over time through combining the source localization ambiguity surfaces generated at widely separated sampling positions, as the short horizontal linear array moves at a low speed along a straight line at a certain depth.

The paper is organized as follows: Section 2 introduces the implementation of the proposed localization approach. Section 3 describes the synthetic test cases used to evaluate the capacity of the proposed approach to localize the acoustic sources in offshore shallow water. Section 4 presents and discusses the results of passive source localization using the proposed approach for various source locations at different signal-to-noise ratios (SNRs) in the presence of environmental mismatch and position errors of the short horizontal linear array. Section 5 provides a short conclusion.

## 2. Localization approach

This section describes the implementation of the proposed localization approach. In our study, the source localization ambiguity surface is generated using the incoherent broadband minimum variance (*MV*) processor as the short horizontal linear array moves along a straight line to each sampling position. Then, the source localization output at each sampling position is formed by averaging all of the generated ambiguity surfaces. Hence the issue of temporal and spatial coherence among the different sampling positions will not be taken into account due to an incoherent combining of the ambiguity surfaces.

In order to successfully localize the acoustic source in our problem, we have to overcome the shortcoming of the short array. We employ a high resolution *MV* processor to exploit as much as possible the unique information arising from the source of interest at each sampling position. The incoherent broadband processing is also used at each sampling position to increase the amount of available data and to stabilize the estimation process. It is a widely used approach to take advantage of the temporal complexity of the signal for an additional gain over narrow pro-

cessing (TOLSTOY, 1993; SOARES, JESUS, 2003; JESUS, SOARES, 2001). The inter-position processing further exploits the spatial complexity of the signal that is emitted by the source of interest and sampled by the short array as it moves to a sampling position.

The *MV* processor (also known as the minimum variance distortionless response processor and the maximum likelihood method) is a high resolution adaptive MFP method, whose essence is “optimum in the sense that the output noise power be minimized subject to the constraint of unity undistorted signal response from the desired source location” (TOLSTOY, 1993; BAGGEROER *et al.*, 1993; COX *et al.*, 1987; STRYCZNIOWICZ, 2006; LEE *et al.*, 1993). Thus, its weight vector  $w$  is determined by solving

$$\min_w \mathbf{w}^* \mathbf{R} \mathbf{w} \text{ subject to } \mathbf{w}^* \mathbf{d} = 1, \quad (1)$$

where  $\mathbf{R} = E\{\mathbf{x}\mathbf{x}^*\}$  is the cross-spectral density matrix (CSDM) at the frequency of interest,  $E\{\}$  denotes the expectation value operation, and  $\mathbf{x}$  represents the measured data vector. The superscript  $*$  denotes the conjugate transpose operation, and  $\mathbf{d}$  represents the replica vector.

The well-known solution of this optimization problem is

$$\mathbf{w}_{MV} = \frac{\mathbf{R}^{-1} \mathbf{d}}{\mathbf{d}^* \mathbf{R}^{-1} \mathbf{d}}. \quad (2)$$

The output of the *MV* processor is the square of the magnitude of the correlation between the weight vector and the measured data vector in the frequency domain, as expressed by

$$P_{MV} = \mathbf{w}_{MV}^* \mathbf{R} \mathbf{w}_{MV} = \frac{1}{\mathbf{d}^* \mathbf{R}^{-1} \mathbf{d}}. \quad (3)$$

The *MV* processor is constrained to pass the signal from the hypothetical source location (a single point on the ambiguity surface), while minimizing the response from all other locations. For high SNR cases, and in the absence of an ocean environmental mismatch, the *MV* processor ambiguity surface provides a very sharp peak where the replica vector corresponds to the data vector for the true source location, whereas it flattens the background level and suppresses sidelobes elsewhere. However, this exceptional resolution capability comes with an increased sensitivity to a slight mismatch between the modeled and actual environments. Also, the adaptive MFP methods usually require received source levels to exceed a threshold SNR (TOLSTOY, 1993; BAGGEROER *et al.*, 1993; DEL BALZO *et al.*, 1988; SMITH *et al.*, 1993; HAMSON, HEITMEYER, 1989).

Consider the noisy data vector  $\mathbf{x}_{f_j} = \mathbf{s}_{f_j} + \mathbf{n}$  received by  $N$  hydrophones in the array at the  $j$ -th frequency component. The signal vector  $\mathbf{s}_{f_j}$  and the replica vector  $\mathbf{d}_{f_j}$  are both calculated by RAM codes based on the parabolic equations (PE) theory (JENSEN *et al.*, 2011; COLLINS, 1995), and then normalized such that  $\|\mathbf{s}_{f_j}\| = 1$  and  $\|\mathbf{d}_{f_j}\| = 1$ , where  $\|\mathbf{s}_{f_j}\|$  and  $\|\mathbf{d}_{f_j}\|$

are the  $L_2$  norm of  $\mathbf{s}_{f_j}$  and  $\mathbf{d}_{f_j}$ . A common assumption is that the additive noise vector  $\mathbf{n}$  is white, Gaussian, zero-mean, and uncorrelated with the signal vector  $\mathbf{s}_{f_j}$ . Thus the components  $n_i$  in  $\mathbf{n}$  are independent random complex Gaussian variables with a zero mean. That is

$$f(n_i) = \frac{1}{\sqrt{2\pi}\sigma_n} e^{-n_i^2/2\sigma_n^2}, \quad (4)$$

where the strength of noise  $\sigma_n^2 = 1/(Nr)$ , and  $r$  denotes the input SNR averaged across the array. Since the amplitude of the signal vector varies across the array, the actual SNR on any individual hydrophone may be higher or lower. From a computational point of view, the components  $n_i$  are generated using the Box-Muller formula (PORTER, TOLSTOY, 1994):

$$n_i = \sigma_n \sqrt{-\log X_i} e^{i2\pi Y_i}, \quad (5)$$

where  $X_i$  and  $Y_i$  are random variables with a uniform distribution on the interval  $(0, 1]$ .

For each sampling position and each frequency component we generate  $L = 300$  realizations (snapshots):  $\mathbf{x}_{f_j}^l$ . The CSDM at the  $j$ -th frequency component is constructed as follows:

$$\mathbf{R}_{f_j} = E\{\mathbf{x}_{f_j} \mathbf{x}_{f_j}^*\} = \sum_{l=1}^L \mathbf{x}_{f_j}^l \mathbf{x}_{f_j}^{l*}. \quad (6)$$

Then, the normalized CSDM at the  $v$ -th sampling position is expressed by

$$\mathbf{K}_{f_j}^v = \frac{N \mathbf{R}_{f_j}^v}{(1 + \sigma_n^2)L}. \quad (7)$$

Thus, the corresponding weight vector  $\mathbf{w}_{f_j}^v$  can be calculated using Eq. (2), which yields that the source localization ambiguity surface generated by the *MV* processor at the  $j$ -th frequency component and the  $v$ -th sampling position is

$$P_{f_j}^v = \mathbf{w}_{f_j}^{v*} \mathbf{K}_{f_j}^v \mathbf{w}_{f_j}^v. \quad (8)$$

If  $M$  is the number of discrete frequencies considered, then the source localization output at the  $v$ -th sampling position is

$$P_{\text{output}} = (P_{f_1}^1 + P_{f_2}^1 + \cdots + P_{f_M}^1 + \cdots + P_{f_1}^v + P_{f_2}^v + \cdots + P_{f_M}^v)/v. \quad (9)$$

### 3. The synthetic test cases

This section introduces the synthetic test cases for evaluating the performance of the proposed localization approach. We employ a two-dimensional coordinate system, where the range  $r$  is the horizontal distance and  $z$  is the depth below the ocean surface. The test scenario is described as follows. A single static acoustic source emits multitone signals at 75 Hz, 100 Hz, 150 Hz, and 250 Hz at a fixed location. A short horizontal linear array comprised of

6 hydrophones that are evenly spaced at 1 m intervals is used to localize the source. The source is further assumed to be at the endfire of the array, and the array moves towards the source at a low speed of 2 m/s along a straight line at a 50 m depth. It should be noted that the multitone signal would arrive at the endfire of the array according to the above assumption. If the source is assumed to be at the other bearing, then the effective length of the array is shorter, and elements are more closely spaced. The relationship between the actual and effective array lengths and respective sensor positions is a function of the source range and it may be determined through trigonometric relationships (TANTUM, NOLTE, 2000). Moreover, the depth of the horizontal linear array does impact on the source localization performance attained by MFP techniques (TOLSTOY, 1993; TANTUM, NOLTE, 2000). However, this study aims to apply MFP to a short horizontal linear array for passive source localization in shallow water. In this case, we are looking forward to having a good chance of estimating the source positional parameters in depth and range rather than only bearing, as in the previous studies. The issue of examining the effects of the array depth on the localization performance is left for future research. Hence, the above test scenario assumption is justifiable.

A range-dependent ocean environmental model representative of offshore shallow water (PORTER, TOLSTOY, 1994) was chosen for this study. The model consists of a water column and a seabed with a sediment layer over a semi-infinite basement. It has a sloping bottom denoted by two parameters ( $D_1$ ) and ( $D_2$ ) as shown in Fig. 1. The seabed geoacoustic parameters include the sound speed at the top ( $c_{BD1}$ ) with respect to ( $D_1$ ) and bottom ( $c_{200}$ ) of the sediment layer, the density ( $\rho_1$ ) and attenuation ( $\alpha_1$ ) in the sediment,

and the constant sound speed equal to ( $c_{200}$ ) in the basement. The sound speed profile (SSP) in the water column is described by two parameters ( $c_0$  and  $c_{WD1}$ ) at the ocean surface and with respect to ( $D_1$ ). The sound propagation is modeled based on the PE theory (JENSEN *et al.*, 2011), in which the usual approach to the problem of simulating the basement condition is to add an artificial absorption layer of several wavelengths thickness associated with a relatively large attenuation value ( $\alpha_2 = 10$  dB/ $\lambda$ ). Table 1 shows the detailed values of the ocean environmental parameters in the simulation study.

The environmental mismatch is unavoidable and is likely the most outstanding obstacle to the general experimental application of MFP techniques. At times, the output of MFP deteriorates greatly even with low levels of the mismatch, especially for high resolution MFP methods. We use different environmental parameter sets for calculating the data vector and replica vector as shown in Table 1, thus, SSP mismatch, water depth mismatch, and mismatch in the bottom parameters are all involved in our study. Additionally, the position errors of the short horizontal linear array are introduced. In reality, we usually cannot guarantee that the positions of the array can be estimated exactly every time when the replica vectors are calculated. Hence, there always exist errors of the estimated array position in range and depth, and both are assumed to follow the uniform distribution in  $[-2$  m,  $2$  m] in our study.

We next examine the ability of the proposed approach to localize the acoustic source at three selected source locations. Source A is located at a 5.7 km range and 70 m depth, source B is located at a 6 km range and 72 m depth, and source C is located at a 8.1 km range and 58 m depth, each measured from the start

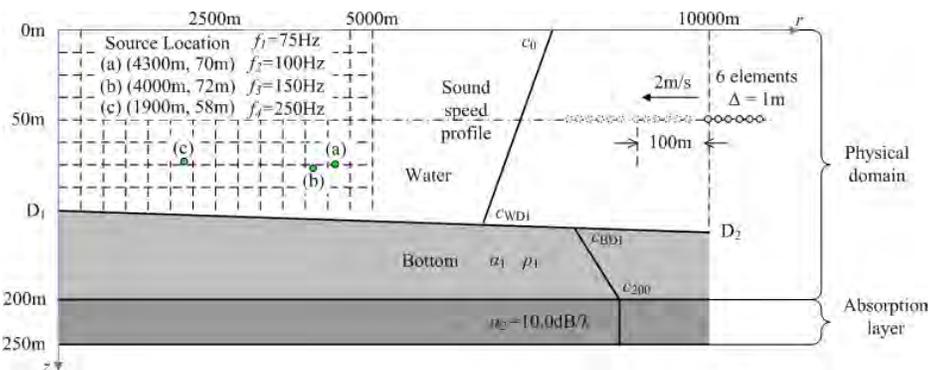


Fig. 1. Schematic of the ocean environment and the geometric configuration of the source and the array.

Table 1. Ocean environmental parameters used in the simulation study.

| Parameter (unit) | $c_0$<br>[m/s] | $c_{WD1}$<br>[m/s] | $c_{BD1}$<br>[m/s] | $c_{200}$<br>[m/s] | $\alpha_1$<br>[dB/ $\lambda$ ] | $\rho_1$<br>[g/cm <sup>3</sup> ] | $D_1$<br>[m] | $D_2$<br>[m] |
|------------------|----------------|--------------------|--------------------|--------------------|--------------------------------|----------------------------------|--------------|--------------|
| Data vector      | 1497.9         | 1478.3             | 1604               | 1798               | 0.11                           | 1.65                             | 101.2        | 104.6        |
| Replica vector   | 1500           | 1480               | 1600               | 1750               | 0.35                           | 1.75                             | 102.5        | 102.5        |

position of the array. The estimation performance is evaluated at three SNRs, 40 dB, 10 dB, and  $-5$  dB. The search region for localizing the source extends from 0 m to 100 m in depth (spanning the entire water column), and from 5 km to 10 km in range (as measured from the start position of the array). The corresponding search grid spacing is 50 m in range and 1 m in depth. The horizontal distance between the adjacent uniformly separated sampling positions along the straight trajectory of the array is 100 m. The geometric configuration of the source and the array is also shown in Fig. 1, using relative position relationships for the source and the array based on calculation convenience.

#### 4. Results and discussion

This section presents the results of the proposed approach applied to passive source localization with a mobile short horizontal linear array. The data are simulated using the synthetic test scenarios discussed in

Sec. 3. We intend to find out how the convergence over time of the estimated source location to the true source location, and the ocean environmental mismatch and the position errors of the array affect the localization performance in our problem. To address this question, 30 sampling positions are processed. Thus, the source localization output is updated every 50 s, the interval for the updated outputs between the first and the last sampling positions is 1450 s, and the array travels up to 2.9 km until the last localization output is updated according to the above assumption. For each updated localization output, we just use the simple “peak picker” algorithm to estimate the source location with respect to the current position of the array. The localization performance in range and depth for various source locations at different SNRs is then illustrated.

Figure 2 displays the source localization outputs for source A in the high SNR case ( $\text{SNR} = 40$  dB). The six plots in this figure are selected from 30 localiza-

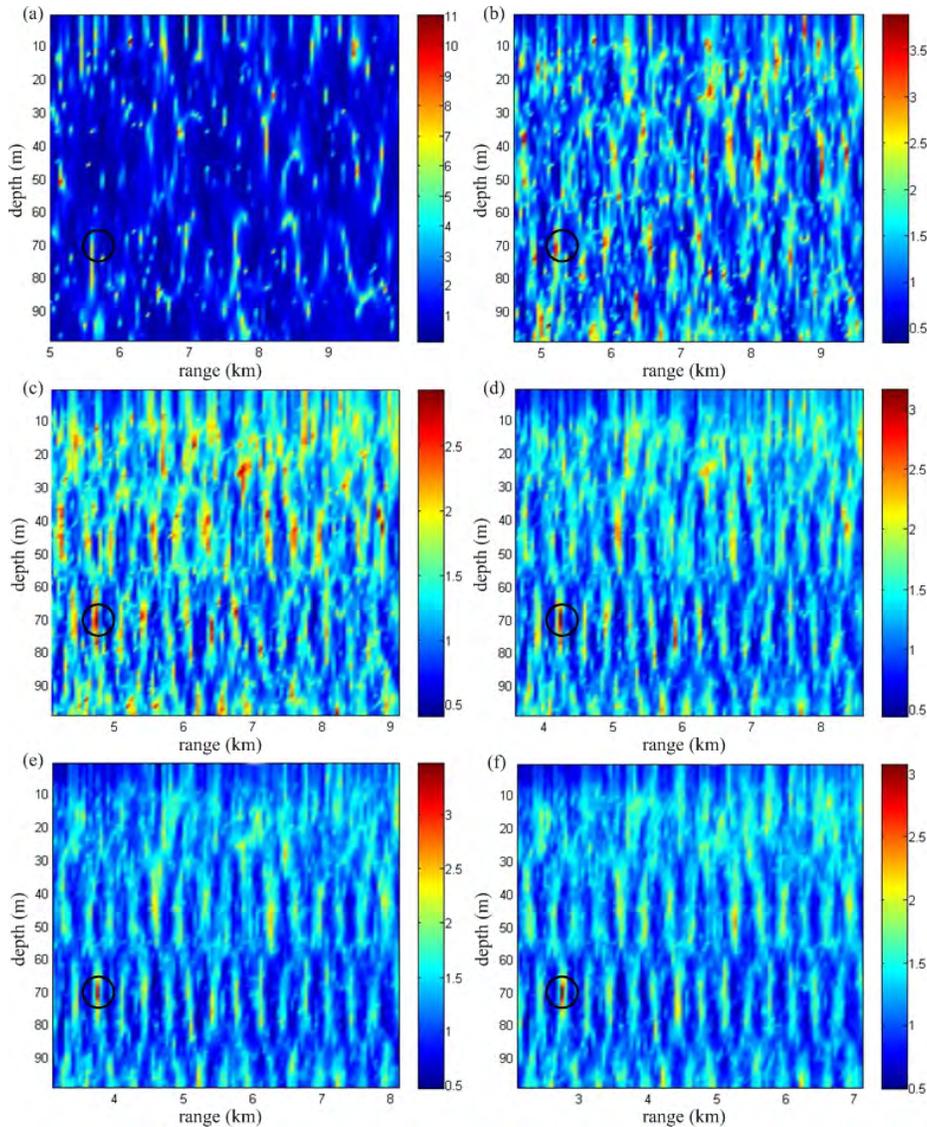


Fig. 2. Localization outputs for source A at  $\text{SNR} = 40$  dB after: a) 0 s, b) 200 s, c) 450 s, d) 700 s, e) 950 s, f) 1450 s.

Table 2. Depth and range errors of the array positions used in the simulation study.

|                 |      |      |      |     |     |      |      |      |      |      |      |      |      |      |      |
|-----------------|------|------|------|-----|-----|------|------|------|------|------|------|------|------|------|------|
|                 | 1    | 2    | 3    | 4   | 5   | 6    | 7    | 8    | 9    | 10   | 11   | 12   | 13   | 14   | 15   |
| Depth error [m] | 1.2  | -0.8 | 0    | 0.8 | 1.6 | 2    | 0    | -1.6 | -1.6 | -0.8 | 1.2  | -0.8 | 1.2  | -1.2 | 1.6  |
| Range error [m] | -1.7 | -1.8 | 0.1  | 1.1 | 1.7 | -1.5 | 0.3  | -0.1 | -2   | -0.7 | -1.4 | 1.2  | -0.8 | 0.1  | -1.3 |
|                 | 16   | 17   | 18   | 19  | 20  | 21   | 22   | 23   | 24   | 25   | 26   | 27   | 28   | 29   | 30   |
| Depth error [m] | -0.8 | -1.2 | -0.8 | 0.4 | 0   | -0.4 | 1.2  | 0.4  | 0    | 1.6  | -0.8 | 1.2  | 1.2  | -0.4 | 0.4  |
| Range error [m] | 0.4  | -0.9 | 0.6  | 0.8 | 1   | -0.2 | -1.7 | -1   | 1.7  | -1.4 | 1.3  | 0.2  | 2    | -1.7 | -0.2 |

tion outputs generated at 30 sampling positions. The circles denote the true source location in these plots. Figure 2 shows that the source of interest gradually appears at the correct location and that the sidelobes are better suppressed over time even in the presence of various ocean environmental mismatches and position errors of the array. Table 2 displays the depth and range errors of the array positions at each sampling position used in our simulation. It means that the incoherent broadband *MV* processor exploits the infor-

mation arising from the source of interest effectively at each sampling position at a high SNR and that its shortcoming of sensitivity to the ocean environmental mismatch and position errors of the array can be restrained through inter-position processing which not only increases the data snapshots but also exploits the spatial characteristics of the acoustic field due to the source of interest.

Figures 3 and 4 display the source localization outputs for source B at a SNR of 10 dB and source C in the

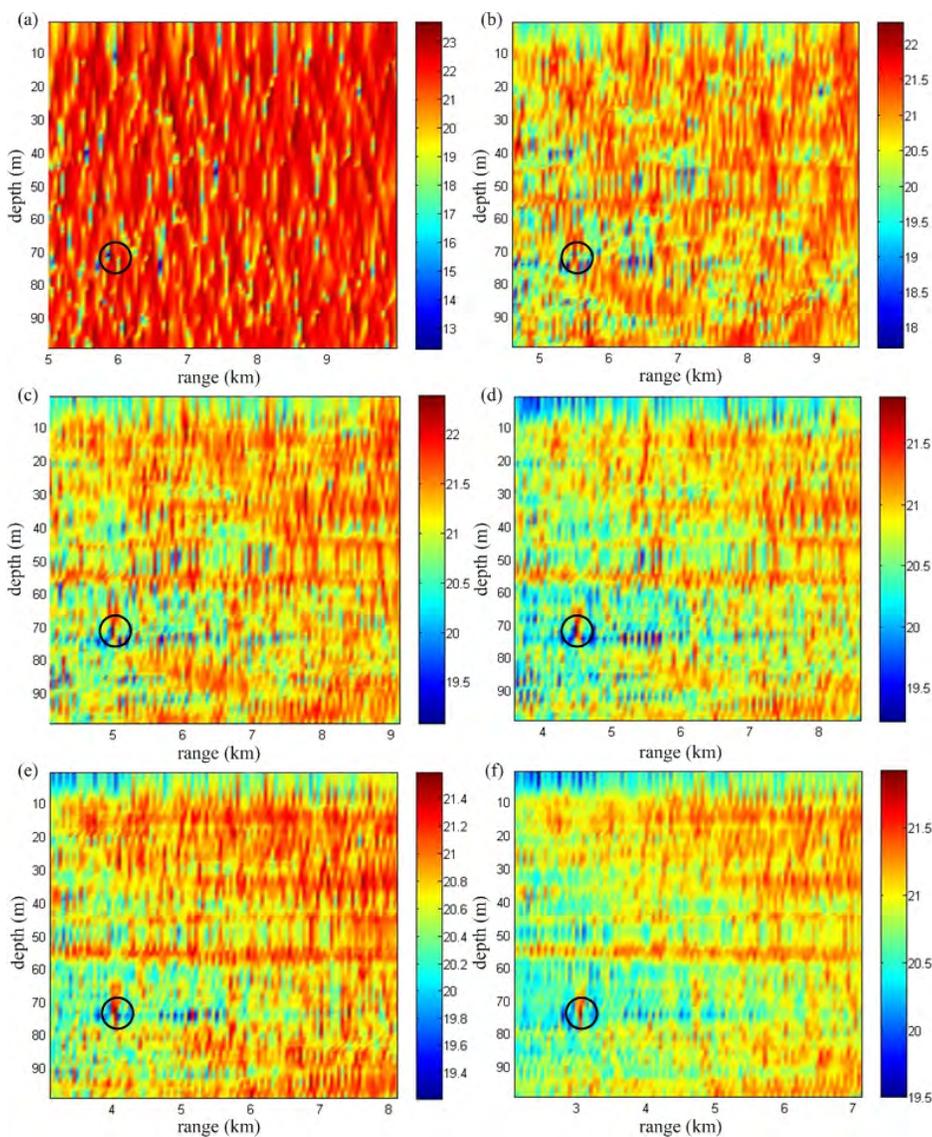


Fig. 3. Localization outputs for source B at SNR = 10 dB after: a) 0 s, b) 200 s, c) 450 s, d) 700 s, e) 950 s, f) 1450 s.

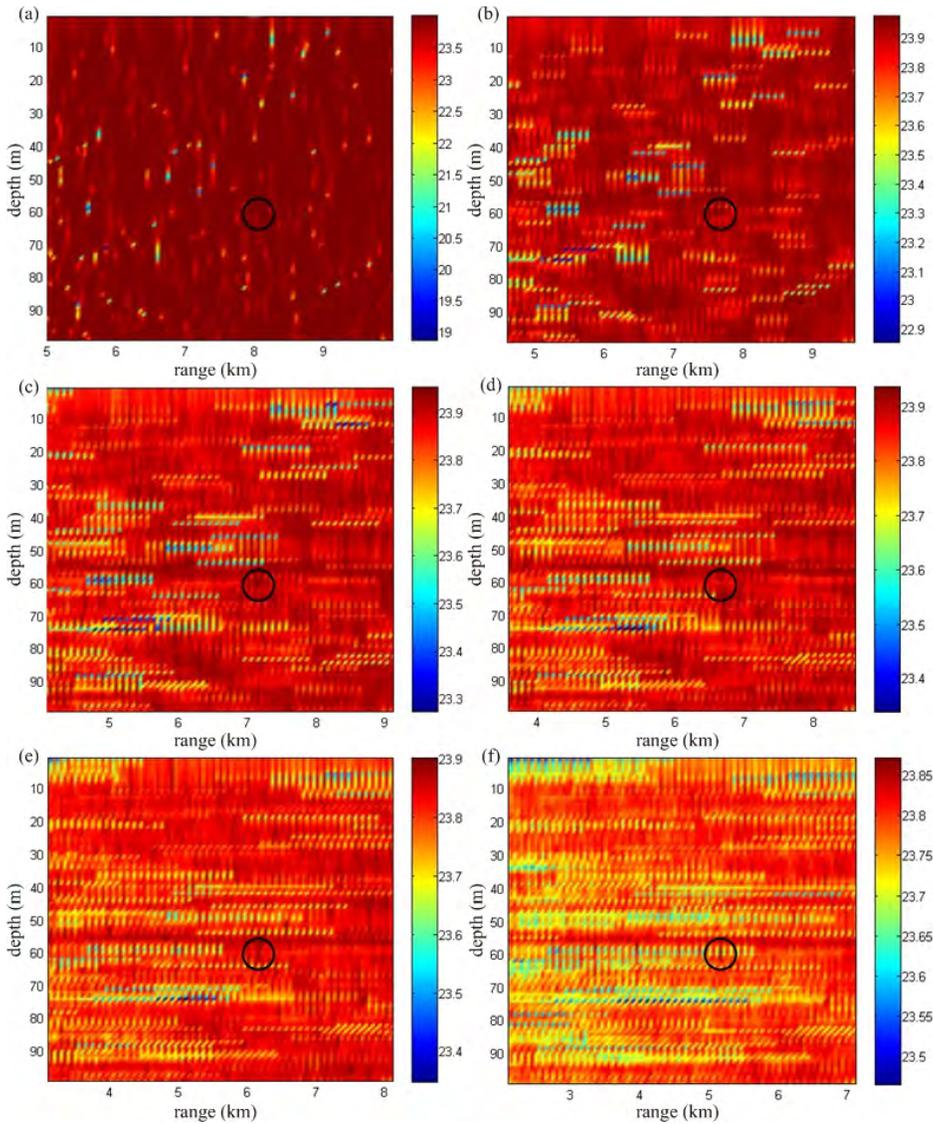


Fig. 4. Localization outputs for source C at SNR = -5 dB after: a) 0 s, b) 200 s, c) 450 s, d) 700 s, e) 950 s, f) 1450 s.

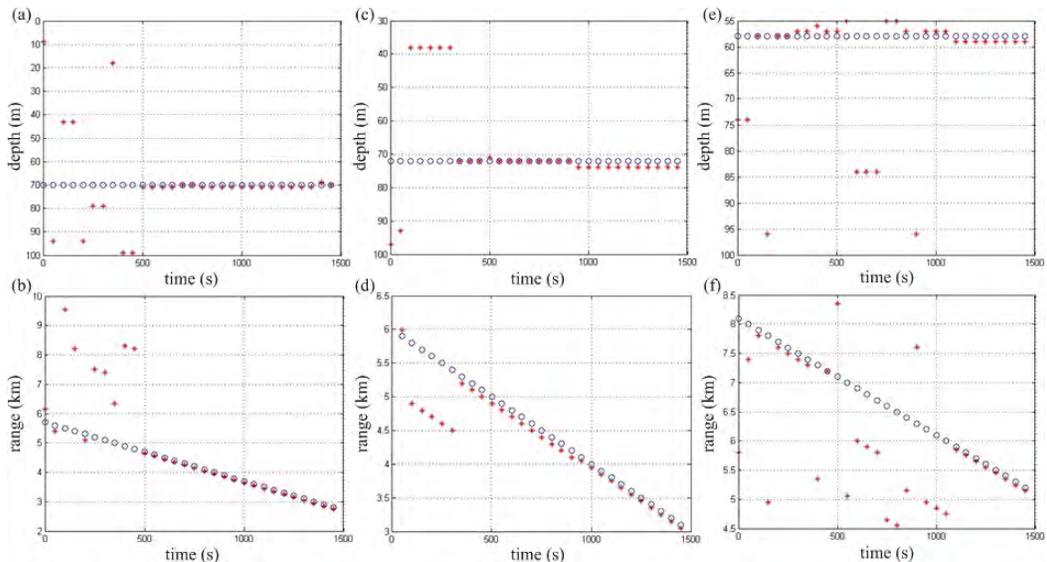


Fig. 5. a) Depth and b) range estimation *versus* time for source A at SNR = 40 dB, c) depth and d) range estimation *versus* time for source B at SNR = 10 dB, and e) depth and f) range estimation *versus* time for source C at SNR = -5 dB.

case of the source of interest submerged in the white noise (SNR =  $-5$  dB), respectively. Figure 3 shows that the source can also be localized at the correct location but with much higher sidelobes compared to the results shown in Fig. 2. However, Fig. 4 shows that the dynamic range displayed on the right side of each plot gets smaller over time and that it is already not easy to discriminate the source of interest from the background in the search region. It is seen that the localization performance degrades rapidly as the white noise level increases.

Figure 5 presents the source depth and range estimation versus the time for source A, B, and C at SNRs of 40 dB, 10 dB, and  $-5$  dB, respectively. The circles denote the true source positional parameters, and the asterisks denote the estimated source positional parameters. The parameter estimation is considered to be acceptable if the absolute range error is less than 600m and the absolute depth error is less than 6m, as defined in (TOLLEFSEN, DOSSO, 2009). It is seen that satisfactory results are attained over time. However, the estimation process becomes extremely time-consuming when the source of interest is submerged in the noise, and the high sidelobes on the ambiguity surfaces usually make the estimated results unstable, as shown in Figs. 5e and 5f. In this case, the proposed localization approach is not realistic in practical situations. We need more resolution MFP methods to exploit more information arising from the source at each sampling position. The coherent broadband MFP that can offer additional gain (DEBEVER, KUPERMAN, 2007) may be required.

## 5. Conclusions

This paper presents an approach to passive source localization using a mobile short horizontal linear array in shallow water. The proposed approach is based on high resolution MFP methods which exploit as much as possible the information arising from the source of interest at each sampling position, as well as inter-position processing for further localization performance improvement in terms of the stability in the estimation process and the robustness to the ocean environmental mismatch and the position errors of the array. The proposed approach was applied to synthetic data in a simulated environment for three source locations at different SNRs. The results show that the source positional parameters can be built up over time, as the short horizontal linear array moves at a low speed along a straight line at a constant depth.

## Acknowledgment

This work is supported by China Shipbuilding Industry Corporation and China Scholarship Council No. 2011611091. The authors would also like to thank

Prof. Woojae Seong and Keunhwa Lee at the Underwater Acoustic Laboratory, Seoul National University, for helpful comments and discussions.

## References

1. AUTREY S. W. (1988), *Passive synthetic arrays*, Journal of the Acoustical Society of America, **84**, 592–598.
2. BAGGEROER A. B., KUPERMAN W. A., MIKHALEVSKY P. N. (1993), *An overview of matched field methods in ocean acoustics*, IEEE Journal of Oceanic Engineering, **18**, 401–424.
3. DEL BALZO D. R., FEUILLADE C., ROWE M. M. (1988), *Effects of water-depth mismatch on matched-field localization in shallow water*, Journal of the Acoustical Society of America, **83**, 2180–2185.
4. BERNECKY W. R., KRZYCH M. J. (2008), *Point source localization sonar system and method*, Patent Application Publication, United States.
5. COLLINS M. D. (1995), *User's guide for RAM versions 1.0 and 1.0p*, Naval Research Lab, Washington, DC.
6. COX H., ZESKIND R. M., OWEN M. M. (1987), *Robust adaptive beamforming*, IEEE Transactions on Acoustics, **35**, 1365–1376.
7. DEBEVER C. (2009), *Study of how environmental fluctuations influence the coherence of acoustic signals*, DTIC Document.
8. DEBEVER C., KUPERMAN W. A. (2007), *Robust matched-field processing using a coherent broadband white noise constraint processor*, Journal of the Acoustical Society of America, **122**, 1979–1986.
9. EHLERS F., FOX W., MAIWALD D., ULMKE M., WOOD G. (2010), *Advances in signal processing for maritime applications*, EURASIP Journal on Advances in Signal Processing, **2010**: 1–5.
10. FERNANDEZ J. E., MATTHEW A. D., COOK D. A., STROUD J. S. (2004), *Synthetic aperture sonar development for autonomous underwater vehicles*, OCEANS, **4**, 1927–1933.
11. FIALKOWSKI L. T., COLLINS M. D., PERKINS J. S. (1997), *Source localization in noisy and uncertain ocean environments*, Journal of the Acoustical Society of America, **101**, 3539–3545.
12. HAMSON R. M., HEITMEYER R. M. (1989), *Environmental and system effects on source localization in shallow water by the matched-field processing of a vertical array*, Journal of the Acoustical Society of America, **86**, 1950–1959.
13. MARANDA B. H. (2008), *Passive sonar*, [in:] *Handbook of signal processing in acoustics*, Havelock D., Kuwano S., Vorländer M. [Eds.], Springer, New York.
14. HODGES R. P. (2010), *Underwater acoustics: Analysis, design and performance of sonar*, Wiley, West Sussex.
15. INCE L., SEZEN B., SARIDOGAN E., INCE H. (2009), *An evolutionary computing approach for the target mo-*

- tion analysis (TMA) problem for underwater tracks*, Expert Systems with Applications, **36**, 3866–3879.
16. JENSEN F. B., KUPERMAN W. A., PORTER M. B., SCHMIDT H. (2011), *Computational ocean acoustics*, Springer, New York.
  17. JESUS S. M., SOARES C. (2001), *Broadband MFP: coherent vs. incoherent*, OCEANS, **2**, 776–781.
  18. KIM K., SEONG W., LEE K. (2010), *Adaptive surface interference suppression for matched-mode source localization*, IEEE Journal of Oceanic Engineering; **35**, 120–130.
  19. LEE Y. P., MIKHALEVSKY P., FREESE H., HANNA J. (1993), *Robust adaptive matched-field processing*, OCEANS, **3**, 387–392.
  20. MILLARD N. W. (2003), *Multidisciplinary ocean science applications of an AUV: the autosub science missions programme*, [in:] *Technology and Applications of Autonomous Underwater Vehicles*, Griffiths G. [Ed.], Taylor & Francis, London and New York.
  21. PORTER M. B., TOLSTOY A. (1994), *The matched-field processing benchmark problems*, Journal of the Acoustical Society of America, **2**, 161–185.
  22. SMITH G. B., CHANDLER H. A., FEULLADE C. (1993), *Performance stability of high-resolution matched-field processors to sound-speed mismatch in a shallow-water environment*, Journal of the Acoustical Society of America, **93**, 2617–2626.
  23. SOARES C., JESUS S. M. (2003), *Broadband matched-field processing: coherent and incoherent approaches*, Journal of the Acoustical Society of America, **113**, 2587–2598.
  24. STRYCZNIEWICZ L. (2006), *The methods of inversion and the maximum likelihood estimation in acoustic tests of industrial sources in the environment*, Archives of Acoustics, **31**, 4S, 295–301.
  25. TANTUM S. L., NOLTE L. W. (2000), *On array design for matched-field processing*, Journal of the Acoustical Society of America, **107**, 2101–2111.
  26. TOLLEFSEN D., DOSSO S. E. (2009), *Three-dimensional source tracking in an uncertain environment*, Journal of the Acoustical Society of America, **125**, 2909–2917.
  27. TOLSTOY A. (1993), *Matched field processing for underwater acoustics*, World Scientific, Singapore.
  28. WILLIAMS R., HARRIS B. (1992), *Passive acoustic synthetic aperture processing techniques*, IEEE Journal of Oceanic Engineering, **17**, 8–15.
  29. WILSON J. H., VEENHUIS R. S. (1997), *Shallow water beamforming with small aperture, horizontal, towed arrays*, Journal of the Acoustical Society of America, **101**, 384–394.
  30. XUDONG Y., JIANGUO H., QUNFEI Z., QI T. (2007), *Study on the long-distance target apperception techniques for underwater vehicles*, Journal of Systems Engineering and Electronics, **18**, 484–490.

## Technical Notes

Practical Concerns Associated with Single-Number Ratings in Measuring  
Sound Transmission Loss Properties of Partition PanelsNaveen GARG<sup>(1),(2)</sup>, Anil KUMAR<sup>(1)</sup>, Sagar MAJI<sup>(2)</sup><sup>(1)</sup> Apex Level Standards and Industrial Metrology Division  
CSIR – National Physical Laboratory  
New Delhi-110 012, India<sup>(2)</sup> Department of Mechanical and Production Engineering, Delhi Technological University  
Delhi – 110 042, India; e-mail: ngarg@mail.nplindia.ernet.in

(received November 13, 2011; accepted January 28, 2013)

The paper presents an extensive review investigating the practical aspects related to the use of single-number ratings used in describing the sound insulation performance of partition wall panels and practical complications encountered in precise measurements in extensive frequency range from 50 Hz to 5 kHz. SWOT analysis of various single number ratings is described. A laboratory investigation on a double wall partition panel combination revealed the significant dependence of STC rating on transmission loss at 125 Hz attributed to 8 dB rule. An investigation conducted on devising alternative spectrums of aircraft noise, traffic noise, vehicular horn noise and elevated metro train noise as an extension to ISO 717-1  $C_{tr}$  for ascertaining the sound insulation properties of materials exclusively towards these noise sources revealed that the single-number rating  $R_w + C_{tr}$  calculated using ISO 717-1  $C_{tr}$  gives the minimum sound insulation, when compared with  $R_w + C_x$  calculated using the alternative spectrums of aircraft noise, traffic noise, etc., which means that material provides a higher sound insulation to the other noise sources. It is also observed that spectrum adaptation term  $C_x$  calculated using the spectrum of noise sources having high sound pressure levels in lower frequencies decreases as compared to ISO 717-1  $C_{tr}$  owing to significant dependence of  $C_{tr}$  at lower frequencies.

**Keywords:** Sound Transmission Loss (TL), sound transmission class (STC), spectrum adaptation terms ( $C$ ,  $C_{tr}$ ), ISO 717-1, weighted sound reduction index ( $R_w$ ), spectrum adaptation term corresponding to noise source,  $C_x$ .

## 1. Introduction

The harmful effects of traffic noise and annoyance caused necessitates the concept of providing better sound insulation in dwellings. Strengthening of the exterior facades of buildings is thus very essential for combating the accentuated ambient noise levels due to vehicular noise. The effective sound insulation has been thus persistently a focus of acoustical engineers towards developing sandwich configurations that provide better sound insulation and are cost effective as well. Transmission Loss (TL) is a performance of sound insulation measured in reverberation chambers. There are varied single-number ratings viz., Sound Transmission Class (STC), Weighted sound reduction index ( $R_w$ ), Outdoor-Indoor Transmission Class (OITC) and adaptation terms  $C$ ,  $C_{tr}$  used for describing the

sound insulation properties of partition wall panels used in dwellings, offices, exterior facades, etc. Sound Transmission Class (STC) is an integer rating of how well a building partition attenuates airborne sound used widely to rate interior partitions, ceilings/floors, doors, windows and exterior wall configurations. The STC value is derived from sound attenuation values tested at sixteen standard frequencies from 125 Hz to 4000 Hz. These sound transmission-loss values are plotted versus frequency and the resulting curve is compared to a standard reference contour subject to that the sum of deficiencies at all frequency cannot exceed 32 dB and TL value at any one frequency cannot be more than 8 dB below the STC contour. The STC value is defined as TL value where the STC contour intersects the 500 Hz line (ASTM E413-87, 1999). There are various other ratings used in the similar context

viz., weighted sound reduction index,  $R_w$  and Outdoor-Indoor Transmission Class, OITC.  $R_w$  is used to facilitate the comparison of sound insulation performance of different materials in European continent. STC rating has been described in standard (ASTM E413-87, 1999) to correlate in a general way with subjective impressions of sound transmission for speech, radio, television, and similar sources of noise in offices and buildings, etc., but is inadequate for sound sources such as machinery, industrial processes, bowling allies, power transformers, musical instruments, and transportation noises such as motor vehicles, aircraft and trains, etc. Thus, it is imperative that for sources like transportation noise, machinery noise, etc., a scientific analysis of individual frequency bands is required for characterizing the sound transmission associated with acoustical materials.

The single-number ratings are very crucial in not only describing the acoustical properties of materials but also in deciding the sound insulation regulations required in dwellings. It may be noted that although the sound insulation characteristics as a function of frequency is a true parameter to judge the sound insulation provided by any material, yet the adoption of single number ratings provide an easy guide for comparison and thus finds to be more popular particularly for manufacturers, architects and layman. The single number rating used in laboratory and field measurements are very crucial in describing the sound regulation requirements in building elements. Thus, devising the single number rating based on scientific principles and fulfilling the characteristics listed had been always a major challenge before acousticians:

- Easily understandable, well defined with no pitfalls;
- Address entire frequency range from 50 Hz to 5 kHz;
- Correlate well with subjective perception;
- Shouldn't have high influence of any particular frequency band either low or high.

The recent studies pertaining to recommendations on a new system of single-number quantities proposed (SCHOLL *et al.*, 2011; SCHOLL, WITTSTOCK, 2012) viz., traffic noise sound reduction index,  $R_{\text{traffic}}$ ; living noise sound reduction index,  $R_{\text{living}}$  and speech sound reduction index,  $R_{\text{speech}}$  is simpler than the existing one and facilitates a clear identification and suitability w.r.t usage of single number quantities for rating the sound insulation in building and of building elements and also harmonizes airborne sound insulation using sound reduction index,  $R$  as a common descriptor.

The STC is a precise rating with well defined rules commonly used, but suffers from limitations in case of partition panels with poor low frequency sound insulation. A subjective study of Sound Transmission Class system carried out four decades ago (CLARK, 1970) for rating building partitions concluded that the present STC system is overconservative in rating changes in

a TL curve and that narrow coincidence type dips are not very important. The limitations in STC rating were cited in literature (Green Glue Company) by illustrating an practical example of two hypothetical poor walls with very bad low frequency performance, but one is STC 32, the other is STC 42. The 125 Hz cut-off leads to some very misleading results. Researchers have tried with various new proposals for a single number rating based on the subjective response in terms of psychoacoustics parameters. VIAN *et al.*, 1983 related the subjective ratings of sound insulation to frequency limited (125 Hz to 4 kHz) A-weighted level differences. The subjective judgments of loudness of transmitted sounds were also correlated with simple arithmetic average transmission loss over frequency (TACHIBANA *et al.*, 1988). Recent research (GOVER, BRADLEY, 2004) had shown the intelligibility of speech from meeting rooms to be well related to frequency weighted signal to noise ratio suggesting possible new wall transmission loss ratings. The two most accurate predictors of the intelligibility of transmitted speech were an arithmetic average transmission loss over the frequencies from 200 Hz to 2.5 kHz and addition of a new spectrum weighting term to  $R_w$  that included frequencies from 400 Hz to 2.5 kHz (PARK *et al.*, 2008a). An STC measure without an 8-dB rule and an  $R_w$  rating with a new spectrum adaptation term were better predictors of annoyance and loudness ratings of speech sounds (PARK, BRADLEY, 2009). The low frequency noise annoyance has been a motivating factor in development of spectrum adaptation terms  $C$  and  $C_{\text{tr}}$  in ISO 717-1 standard (ISO 717-1, 1996). The spectrum adaptation terms have been included to take into account the different spectra of noise sources:  $C$  and  $C_{\text{tr}}$  (corresponding to pink noise and road traffic noise) for airborne sound insulation. The standard covers the spectrum adaptation term  $C_{\text{tr}}$  which is to be applied when a representative urban traffic noise is assumed as the loading noise. There are various other metrics viz., acoustic insulation factor,  $R_{av}$ , etc., proposed by researchers to quantify the sound insulation in terms of single number ratings. In case of a simple approach of using  $R_{av}$  (KOYASU, TACHIBANA, 1990) in 100–3150 Hz, the metrics has been found to be related satisfactorily to loudness effect between 63 Hz to 125 Hz and 4 kHz although this index doesn't differentiate between high and low frequency insulation. The representative spectrum chosen for traffic noise in  $C_{\text{tr}}$  rating has high variability associated owing to the dependence of traffic noise levels on site and situation specific, heterogeneous mix traffic with horn noise component included, vehicular density and percentage of heavy vehicles. It is envisaged that the spectrum adaptation term shall be better correlated in Indian environment if the representative traffic noise spectrum is modified strictly as per the Indian conditions.

## 2. Practical implications of 8 dB rule

Sound transmission loss measurements in the present work are conducted in Reverberation chambers at National Physical laboratory (PANCHOLY *et al.*, 1977; GARG *et al.*, 2011). The test specimen was mounted in an opening of 1 m<sup>2</sup> between the source and receiving room. An 100 mm thick partition consisting of two layers of 12.5 mm thick Gypsum board on either side of a 50 mm thick metal frame spaced to get an overall thickness of 100 mm with an air cavity of 50 mm was tested and found to have poor transmission loss characteristics at lower frequencies. Another modification in the same partition panel with attaching the metal partition to Gypsum board via a steel C-stud was tested and found to have better performance in range from 400 Hz to 4 kHz. The experimental results reveal that dip in transmission loss observed in the original sample at high frequencies was significantly arrested with modified C-stud combination. Although the measurement conducted in an opening size of 1 m<sup>2</sup> in present work is very less as compared to that prescribed in ISO 140-3 standard (1995), yet the relative comparison of two material configurations tested is major point of consideration here for evaluation of the single number rating.

The STC of the original sample was observed to be 34 while that for the new sample, it is calculated to be 35, which creates an ambiguity as there is an appreciable improvement in the TL characteristics as shown in Fig. 1. However, without conforming to the 8 dB rule, the STC is calculated to be 40, which is more practical considering the TL characteristics of both the configurations. The poor transmission loss at 125 Hz thus creates an ambiguity with respect to the characterizing the sound insulation characteristics of partition panels in terms of sound transmission class. It can be observed that with a modified C-stud, although the dip is significantly arrested, yet the STC value has one to one correspondence with the TL at 125 Hz. In case of partition panels having low frequency perfor-

mance, the STC value has one to one correspondence with the corresponding TL at 125 Hz. The inconsistency attributed due to 8 dB rule in STC calculation is resolved in case of  $R_w$  calculation, wherein for the original partition panels,  $R_w$  value comes out to be 34 and for the improved configuration, it comes out to be 38.

Another experimentation performed in reverberation chamber for measuring the transmission loss of double glazed window of size 920 mm × 620 mm and aperture size 930 mm × 630 mm showed a strange behavior of STC directly dependent on the transmission loss at lower frequency as shown in Fig. 2 (GARG *et al.*, 2011). This double glazing configuration used clear float glass of various thickness and size 832 mm × 532 mm with edges damped in window frame. It can be observed that the STC value is decreased with pronounced resonance dip observed in case of air and vacuum as compared to argon although the transmission loss curve shows similar behavior in entire frequency range. It is also observed that the STC value strongly depends upon TL at 160 Hz attributed to the 8 dB rule adopted in calculating the STC value. The above ambiguity is resolved in case of  $R_w$  value, which comes to be 35 for argon, 34 for air and 33 for vacuum in the gap. The ambiguous behavior of pronounced resonance dip observed in case of vacuum is however beyond the scope of present work.

It is thus evident that STC may create an ambiguity in judgment of the sound insulative characteristics of partition wall panels having poor low frequency performance. The low frequency insulation plays a vital role as most of the noise radiated due transportation systems dominates the lower frequency region. In such cases, it is observed that weighted sound reduction coefficient tries to resolve these issues.

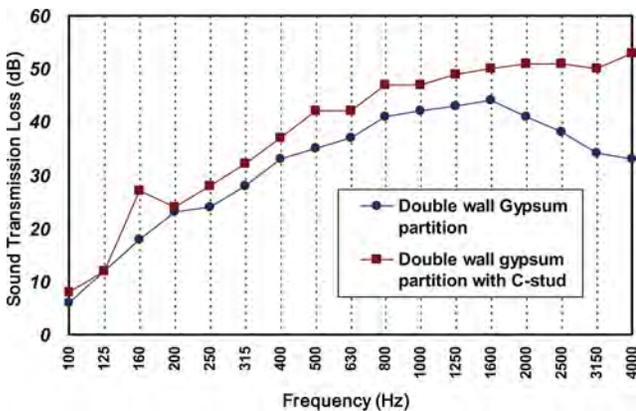


Fig. 1. Sound transmission loss characteristics of Double wall Gypsum partition panel.

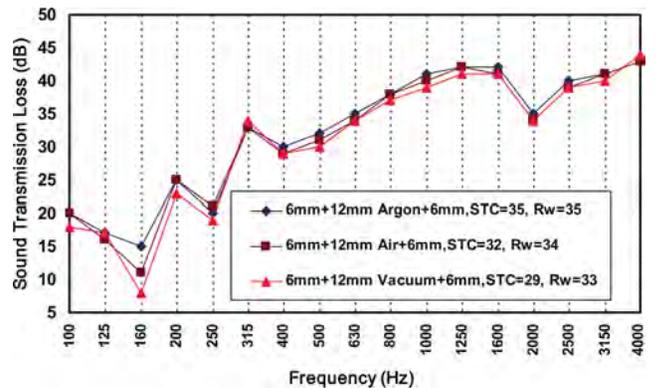


Fig. 2. Sound transmission loss of sandwich construction of clear float glass with Argon and vacuum in air gap (fix design).

There are numerous such practical examples particularly wherein low frequency resonances are encountered and the STC value may create confusion in the

assessment of sound insulation characteristics. This confusion has been observed to be resolved by consideration of  $R_w$  rating. Figure 3 shows the sound transmission loss of sandwich gypsum drywall constructions with 90 mm wood studs at 406 mm on centre and 90 mm blown cellulose fiber insulation in cavity and incremented gypsum layers on each side (HALLIWELL *et al.*, 1998). It can be observed from Fig. 3 that  $R_w$  rating better correlates with the improvement in transmission loss properties associated with the addition of gypsum layers on each side rather than STC rating. So, the 8 dB rule followed to compensate for the poor transmission loss at some frequencies may create confusion in the overall judgement of the actual sound insulation provided by the material and also in relative comparison of sound transmission loss properties of the acoustical materials. A recent subjective survey (PARK *et al.*, 2008b) however substantiates the usefulness of 8 dB rule and provides a different subjective perception towards music and speech wrt 8 dB rule followed. The study reveals that 8 dB rule is useful as it influences low frequency dips in the transmission loss versus frequency characteristics resulting in better prediction of subjective response to sounds with significant low frequency content such as music. The subjective response to speech sounds were observed to be better predicted without 8 dB rule.

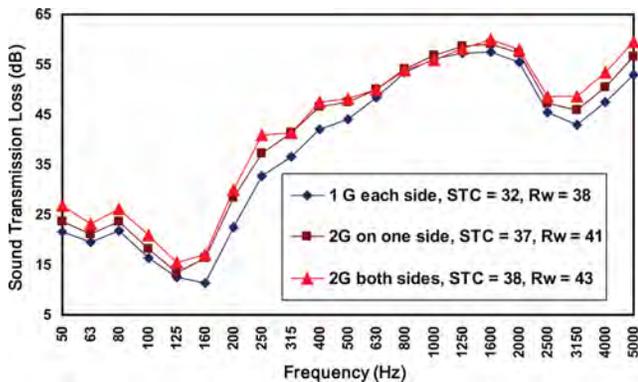


Fig. 3. Sound transmission loss of sandwich gypsum drywall constructions with incremented gypsum layers on each side (HALLIWELL *et al.*, 1998).

### 3. Comparison of single-number ratings

The varied single-number ratings used for describing the sound transmission loss properties of the partition wall panels have common feature of calculation except for the Outdoor Indoor Transmission Class (OITC). The 8 dB rule is skipped in the  $R_w$  method which makes it more reliable and unambiguous for reporting the sound insulation in terms of a single specific number directly proportional to the amount of insulation provide by panel. The spectrum adaptation terms introduced in ISO 717-1 viz.,  $C$  and  $C_{tr}$  value are calculated either from 100 Hz to 3150 Hz

or from 50 Hz to 5 kHz. The spectrum adaptation term  $C$  pertains to living activities, children playing, railway traffic, highway road traffic, jet aircrafts and factories emitting mainly medium and high frequency noise; while the spectrum adaptation term  $C_{tr}$  considers urban traffic noise, railway traffic at low speeds, aircraft, propeller driven, jet aircraft, disco music, etc. (ISO 717-1, 1996). The frequency range used traditionally is 100 to 3150 Hz. For light weight buildings, it is especially important that low frequency spectrum adaptation terms down to 50 Hz are included implying a significantly improved correlation between subjective and objective evaluation (RASMUSSEN, 2010).  $C_{tr}$  significantly concentrates performance outcomes on basis of results at 100 Hz to 160 Hz. The TL at 100 Hz could be often decisive for the final result owing to a high measurement uncertainty attributed to strong  $C_{tr}$  emphasis on lower frequencies (SMITH *et al.*, 2007). The spectrum adaptation terms are adversely affected for light weight constructions and a high variability of around 9 dB average for  $C_{50-3150}$  is observed caused by 50 Hz adaptation term (RASMUSSEN, 2010). LANG (1997) and GOYDKE *et al.*, (2003) also point out uncertainty value associated with  $C_{tr}$  to be much higher. It is thus imperative that wide usage of the single-number rating along with spectrum adaptation terms also imply the need for calculation of associated uncertainties. WITTSTOCK (2007) investigations in this regard reveals that the calculation of the uncertainty of single number ratings from third-octave band sound insulation is possible. The recent study at PTB Germany (SCHOLL *et al.*, 2011) shows that uncertainties are no general obstacle for including third-octave bands with centre frequencies from 50 to 80 Hz into the single-number rating. Another aspect regarding the variability of results after interchanging the source and receiving rooms was investigated by WARNOCK (2004). The investigations reveal that STC and OITC rating are largely affected by changing the test direction i.e. interchanging the source and receiving room, while  $R_w$  is largely unaffected. Table 1 shows the SWOT (Strengths, Weaknesses, Opportunities and Threats) analysis of these standard matrixes with reference to their standards published and findings of various studies (SMITH *et al.*, 2003; PATTERSON, 2004; FITZELL, FRICKE, 2004; RASMUSSEN, RINDEL, 2010).

An investigation carried out to correlate the STC and  $R_w$  rating of sound transmission loss of 25 gypsum board walls (HALLIWELL *et al.*, 1998) leads to a very interesting conclusion on linear relationship between the two ratings. The linear relationship for exclusively gypsum partition panels is observed to be best fit as:

$$R_w = 0.8596 \times STC + 7.7962, \quad (1)$$

$$r^2 = 0.97,$$

Table 1. SWOT analysis of different single number ratings for sound transmission loss measurement.

| Single number ratings                    | Strengths   | Weakness  | Opportunities   | Threats   |
|--|---|---|---|---|
| Sound Transmission Class (STC)           | Simple and easy to calculate, widely used   | 8 dB rule sometimes gives misleading results, low frequency below 125 Hz not addressed                            | Widely used amongst manufacturers and architects  | 8 dB rule sometimes results in confusion especially in cases wherein low frequency resonances are encountered   |
| Outdoor Indoor Transmission Class (OITC) | Suitable for walls, doors, windows; low frequency upto 80 Hz is included  | Low frequency below 80 Hz not included  | Used in walls, doors and windows  | –   |
| Weighted Sound Reduction Index ( $R_w$ ) | Simple and easy to calculate, widely used   | Low frequency below 100 Hz not included   | Widely used amongst manufacturers and architects. $R_w$ in conjunction with spectrum adaptation terms is used in building regulations | –   |
| Spectrum adaptation term, $R_w + C$      | Spectrum adaptation term $C$ is analogous to A-weighting as it is calculated from A-weighting spectrum  | Adversely affected for lightweight constructions and variations are large   | Used in sound regulation requirements in some countries   | Practical problems in measurements down to 50 Hz. $C_{50-3150}$ is highly influenced by 50 Hz spectrum adaptation term  |
| Spectrum adaptation term, $R_w + C_{tr}$ | It is applicable for urban road traffic, railway traffic at low speeds, aircraft propeller driven, Jet aircraft, Disco music and factories emitting mainly low and medium frequency noise | It is not effective in dealing with normal living noise issues and generates too much emphasis at low frequencies | Used in sound regulation requirements in building codes of some countries like Australia, UK, etc.                                    | Practical problem in measurements down to 50 Hz. $C_{tr}$ significantly concentrates performance outcomes on result at 100 Hz to 160 Hz. Variation in measurements of 2–3 dB at lower frequencies can result a significant negative $C_{tr}$ correction value change from –5 to –12 dB. |

A further analysis on correlating  $R_w$  and  $R_w + C_{tr}$  term with STC was done using the sound transmission loss data of 25 gypsum constructions (HALLIWELL *et al.*, 1998) and 34 facade constructions (BRADLEY, BIRTA, 2000) as shown in Fig. 4.

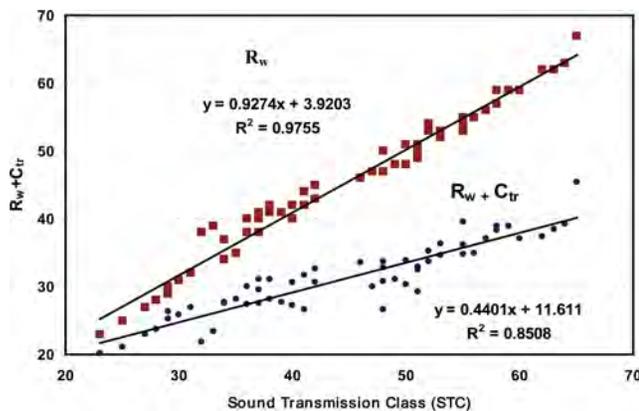


Fig. 4. Linear correlation between  $R_w + C_{tr}$  (dB) and STC rating.

#### 4. Implications of Spectrum Adaptation terms

The spectrum used to calculate  $R_w + C_{tr}$  is an average of eighteen road traffic noise spectra from Copenhagen and Gothenburg (NTACOU 061-1987) with mixed urban road traffic at 50 km/h and about 10% of heavy vehicles. So, an investigation was conducted to devise new reference spectrums for calculating the spectrum adaptation terms denoted by  $C_x$  as an extension to the existing ISO 717-1  $C_{tr}$  for assessing the sound insulation characteristic of materials exclusively towards different noise sources viz., road traffic, aircraft, metro trains, etc. The measurements for traffic noise spectrum were conducted on specific site with average vehicle density between 4000 to 4500 vehicles per hour and dominant horn noise component included. The difference of this spectra with that proposed in ISO 717-1 lies in the fact that considerable sound energy emanated in form of horn noise accentuates the high frequency bands (2.5 kHz to 3.15 kHz) as shown in Fig. 5. Further investigations were conducted

to record the spectra of aircraft landing while undergoing reversed thrust at a distance of 200 m from runway, spectrum of Delhi metro trains running on elevated track and average spectra of horn noise emanated from various vehicles. The noise spectrums were recorded on Norsonic, Nor 118 sound level analyzer and further analyzed in Nor Xfer software and Nor Review software. These spectrums were then normalized such that their sum is zero in compliance with ISO 717-1 and EN 1793-3 standard (1997). The normalized spectra pertaining to highway traffic noise at 90 km/h and 10% heavy vehicles, aircraft starting and propeller aircraft was taken from the DAVY (2004) work and Nord test method. DAVY (2004) conducted an extensive investigations on evaluating the mean, standard deviation, maximum and minimum values of A-weighted sound level attenuation relative to weighted sound reduction index  $R_w$  across 104 sound insulation spectra for different transportation noise spectra. Nord test method (NT ACOU 061-1987) also prescribes six representative spectra for evaluation of adaptation term. Figure 5 shows the normalized spectra of high density traffic with horn noise component included; highway traffic with average speed of 90 km/h and 10% heavy vehicles; metro train noise running on elevated track in Delhi and vehicular horn noise exclusively.

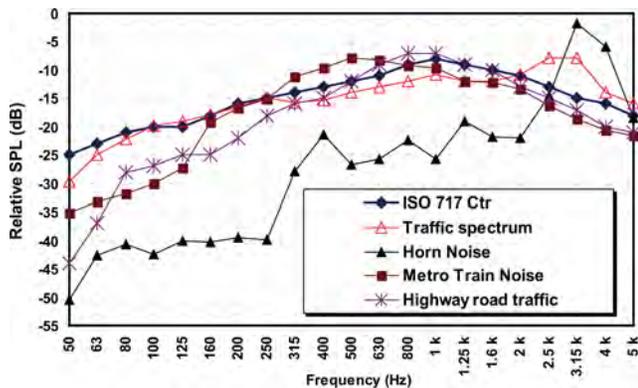


Fig. 5. Normalized spectra for ascertaining sound insulation performance towards traffic, horn noise, Metro train noise on elevated track.

The aircraft starting spectrum (Fig. 6) is extracted from NT ACOU 061-1987 standard, which represents a mean value of 59 starts at Kastrup airport 500 m from runway, while the propeller aircraft spectra is evolved from the mean value of 10 different types of aircrafts starting (DAVY, 2004). It may be noted that these spectrums have been derived from the experimental observations and thus the implications of these spectra in finally evaluating the sound transmission characteristics in terms of a single number rating is a challenging issue rather than their validation. Recent studies (BURATTI *et al.*, 2010; BURATTI, MORETTI, 2010) have confirmed the validity of the proposed revised spectrums by measuring facade sound insulation

index,  $D_{2m,nT_w} + C_{tr}$  value for windows and comparing it with A-weighted level abatements. However, this aspect requires further investigations for correlating the single number descriptors with A-weighting although it devaluates the low frequency noise.

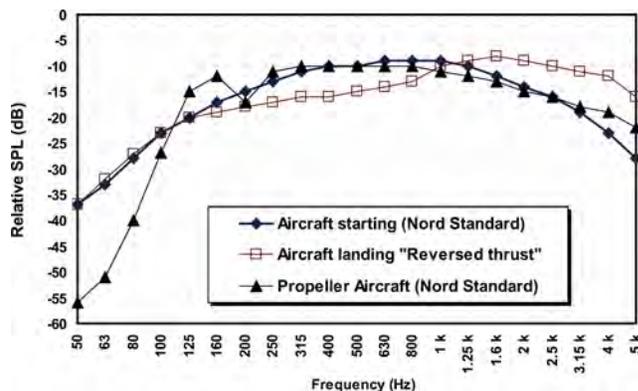


Fig. 6. Normalized spectra for ascertaining sound insulation performance towards aircraft noise.

Sound transmission loss data (HALLIWELL *et al.*, 1998) of 20 gypsum sandwich partition panels was used to evaluate the ISO 717-1  $C_{tr}$  and  $C_x$  term corresponding to noise sources discussed. Table 2 shows the comparison of average value of  $R_w + C_{tr}$  calculated using ISO  $C_{tr}$  and  $R_w + C_x$  calculated using normalized spectra of other noise sources using the sound transmission loss data from 20 gypsum sandwich construction. These observations reveal that  $R_w + C_{tr}$  value is minimum when calculated using the ISO  $C_{tr}$  as compared to other noise sources, which implies that the same material provides a higher sound reduction to other noise sources in comparison to traffic noise. It can be inferred from Table 2 that positive deviation of  $R_w + C_x$  value is observed for each of these noise sources. These investigations when extended to ascertaining the  $R_w + C_{tr}$  value using the energy domain averaged spectrum (Fig. 7) of Rail denoted by Rail Diesel E and decibel domain averaged spectrum denoted by Rail diesel D (DAVY, 2004) reveal an interesting fact that spectrum adaptation term,  $C_x$  calculated using the noise sources having high sound pressure level in lower frequencies decrements as compared to ISO 717-1  $C_{tr}$  owing to its significant dependence on the lower frequencies. This fact was validated from average  $R_w + C_{tr}$  value calculated for 20 gypsum constructions and an average negative deviation of  $-8.2$  dB and  $-4.7$  dB observed for Rail Diesel E and Rail Diesel D spectra w.r.t ISO 717-1  $C_{tr}$ . These investigations reveal that for noise sources having high sound pressure level in lower frequencies, the  $R_w + C_x$  value is least and even lesser than  $R_w + C_{tr}$  of ISO 717-1. The minimum value of  $R_w + C_x$  observed for noise sources dominated by high sound pressure levels in low frequency bands in a way justifies its subjective correlation as low frequency noise is perceived as a source of annoyance (RINDEL,

Table 2. Comparison of  $R_w + C_{tr}$  value calculated using ISO 717-1  $C_{tr}$  spectrum and  $R_w + C_x$  calculated using normalized spectra of various other noise sources.

|  | ISO 717-1 $C_{tr}$ | Horn Noise       | Aircraft landing “Reversed thrust” | Traffic noise       | Highway Traffic Noise      | Elevated Metro Train Noise | Aircraft starting spectrum (Nord) | Propeller Aircraft (Nord) |
|--|--------------------|------------------|------------------------------------|---------------------|----------------------------|----------------------------|-----------------------------------|---------------------------|
| $C_{tr}$ & $C_x$ [dB] computed with corresponding spectrum | $C_{tr}$           | $C_{Horn}$       | $C_{Aircraft}$                     | $C_{traffic}$       | $C_{Highwaytraffic}$       | $C_{Metro\ train}$         | $C_{Aircraft}$                    | $C_{Propeller}$           |
|  | -16.2              | 1.0              | -12.1                              | -15.5               | -8.2                       | -13.4                      | -12.3                             | -13.6                     |
| $R_w + C_{tr}$ & $R_w + C_x$ [dB]                          | $R_w + C_{tr}$     | $R_w + C_{Horn}$ | $R_w + C_{Aircraft}$               | $R_w + C_{traffic}$ | $R_w + C_{Highwaytraffic}$ | $R_w + C_{Metro\ train}$   | $R_w + C_{Aircraft}$              | $R_w + C_{Propeller}$     |
|  | 32.2               | 49.4             | 36.3                               | 32.9                | 40.2                       | 35.0                       | 36.1                              | 34.8                      |
| Difference w.r.t to ISO 717-1 $C_{tr}$ [dB]                | 0.0                | +17.2            | +4.1                               | +0.7                | +7.9                       | +2.8                       | +3.9                              | +2.5                      |

2003) and thus requires special sound insulation measures for its abatement.

These observations thus invite further discussion on applicability of ISO 717-1  $C_{tr}$  to other noise sources as well. However, being a general shaped curve, the use of ISO  $C_{tr}$  is justified to avoid practical and administrative complications associated with implementation of normalized spectra of each individual noise source and representing the minimum sound insulation that material shall provide amongst all the noise sources.

3 (1995) to increase the distance between microphone position and room boundaries and sampling of sound field, increasing the number of loudspeaker positions, the averaging time and use of absorbing materials to decrease the reverberation time still becomes inadequate to enhance the reproducibility of results below 100 Hz (BRAVO, ELLIOTT, 2004; ROLAND, 1995; PEDERSEN *et al.*, 2000). If room volume differs by about 40%, the predicted sound insulation could differ by atleast 3 dB (MALUSKI, GIBBS, 2000). Some studies (BRAVO, ELLIOTT, 2004) have tried to investigate about reducing the effect of source room on measured sound reduction index at low frequencies by using a number of suitable driven loudspeakers close to the panel to stimulate a diffuse incident field. The low frequency diffusion is a cumbersome task achieved by scientifically selecting the volume, surface area of reverberation chamber and enhancing the state of diffusion for reducing the spatial variance in the value of sound pressure level and reverberation time observed at various positions in the room. The volume of the reverberation chamber at Acoustics and Vibration Standard of National Physical Laboratory is 260 m<sup>3</sup> with dimensions 6 × 6.5 × 7 m. The walls, floor and ceilings are non parallel, the average inclination between walls being 6° and between floor and ceiling 2° to 3° (PANCHOLY *et al.*, 1977). Additional diffusing plates have been suspended from ceiling oriented at random to ensure better diffusion. The extent of diffusion can be judged by uniformity of reverberation time within the volume of room, linearity of sound decay at different points in the room and uniformity of sound intensity distribution within the room. The distribution within the room of sound level of filtered band of white noise is within ±0.5 dB at high frequencies and within ±1 dB at low frequencies. The standard deviation of correlation coefficient  $\left(\frac{\sin kr}{kr}\right)$  was measured to be within

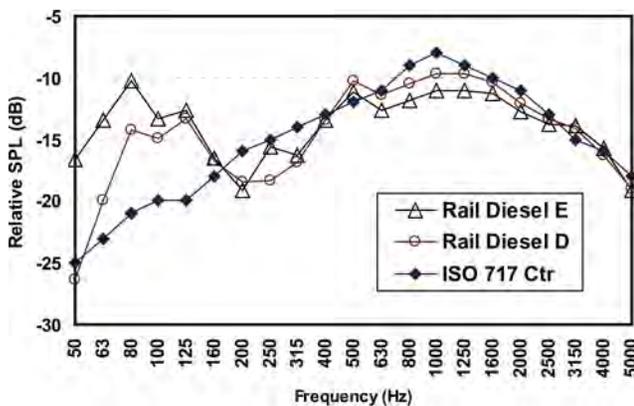


Fig. 7. Normalized spectrums for ascertaining sound insulation performance towards Rail Diesel spectrums (DAVY, 2004).

### 5. Low frequency diffusion issue

The low frequency sound insulation is not only affected by the properties of test wall but also by geometry and dimensions of room-wall-room system (OSIPOV *et al.*, 1997). The diffuse field assumption is only valid in medium and high frequency ranges, as the sound field at low frequencies is dominated by few normal modes in reverberation chamber (SCHROEDER, 1996). The recommendation included in annex in ISO 140-

$\pm 0.06$  (PANCHOLY *et al.*, 1977; BALACHANDRAN, 1959) in frequency range 125–140 Hz. The cross-correlation coefficient for sound pressure at any two points in a room is a means of determining the degree of randomness of sound field with an assumption that in a diffuse field the excitation of two microphones are independent of each other, as soon as certain distance is exceeded, correlation function becomes zero (COOK *et al.*, 1955). A diffuse field can be established in a rectangular room if there is at least 20–30 modes in the measurement bandwidth (NÉLISSE, NICOLAS, 1997), and there is at least one mode per Hz. In the present case, the number of normal modes  $\Delta N$  has the value 21 for  $f = 100$  Hz and  $\Delta f = 13$  Hz (1/6 octave bandwidth). The Schroeder frequency which denotes the boundary between reverberant room behavior above and discrete room modes is calculated as (SCHROEDER, 1962; 1996).

$$f_s \approx \sqrt[3]{\frac{\alpha c^3}{4\pi\eta V}}, \quad (2)$$

where  $\alpha$  is the model overlap. Schroeder has proposed a model overlap  $\alpha = 3$ . For a damping of  $\eta = 5 \times 10^{-3}$  (NÉLISSE, NICOLAS, 1997), the  $f_s$  is calculated to be 192 Hz in present case. The above formulation reveals that for achieving a  $f_s$  value of 50 Hz, the volume of the room should be of the order 15 000 m<sup>3</sup> which is practically impossible. The diffusion of the room increases when the room dimensions are carefully chosen to separate room modes and equalize the frequency response of the room. However, the use of larger reverberation room is restricted by a limit determined by the maximum usable frequency with increasing volume attributed to the increase in dissipation of sound energy during transit between reflections (*Principles and Application of Room Acoustics*, 1982, p. 327). Thus, it can be inferred that measurements down to 50 Hz requires a systematic approach with optimization of the room dimensions as well as augmenting the state of diffuse field by use of rotator diffusers (ISO 3741, 2010). The inclusion of spectrum adaptation terms in range 50 Hz to 3150 Hz in building sound regulations is an effective measure to resolve this issue and avoid practical complications while testing the laboratory or field transmission loss properties of partition panels.

## 6. Conclusions

The present work shows a case study of the limitations associated with use of 8 dB rule in calculation of the STC value. The work also points out the practical limitations associated with the measurement environment for precision measurements down to 50 Hz and use of the adaptation terms in extended frequency range of 50 Hz to 5 kHz. Although the performance of the test specimen to pink noise and traffic is exclusively ascertained in ( $C$ ,  $C_{tr}$ ) matrix, yet there has

to be a trade-off in selecting the reverberation chamber volume for catering to larger wavelengths at low frequencies and energy dissipation at higher frequencies. The present work also discusses the suitability of different representative normalized spectrum for ascertaining the sound insulation performance towards aircraft and traffic noise exclusively. It is known that all different kinds of traffic noise have a different spectral content which will further vary with percentage of heavy vehicles particularly for road traffic noise. Thus it is essentially required to devise a general shape curve for derivation of a single number rating representative of all other sources. The normalized spectra for traffic noise including horn noise component shown in fig 5 can serve as substitute for ISO 717-1  $C_{tr}$  for adjudging the sound insulation properties of material towards road traffic noise in Indian context. The investigations conducted in this paper justifies the use of ISO 717-1  $C_{tr}$  term for adjudging the sound insulation property of material as it represents the minimum sound insulation that a material will provide when exposed to all kinds of traffic noise viz., air traffic, road noise, railway noise, etc. Besides it facilitates a harmonization in the description of a single-number rating for sound insulation properties of acoustical materials rather than following a country specific spectrum adaptation term  $C_x$  so as to avoid any confusion or ambiguity amongst manufacturers and users. The adoption of single number ratings including the spectrum adaptation terms in sound regulation requirements in building elements in Europe necessitates the similar principles and methodology to be followed in Indian perspectives also for harmonization of the sound descriptors for global perspectives and tackling the adverse effects of noise pollution.

## Acknowledgments

Authors would also specially acknowledge the reports published by experts of NRC Canada, *IRC-IR-761* and *IRC-IR-818* and J. L. DAVY, 2004 paper, whose valuable data has been used in various places of paper. Authors thank Mr. Omkar Sharma for his valuable guidance and conducting reverberation chamber measurements and the anonymous reviewers for taking pains in making the paper more interesting and error-free.

## References

1. American Society of Testing Materials, ASTM E413-87 (1999), *Classification for Rating Sound Insulation*.
2. BALACHANDRAN C. G. (1959), *Random sound field in reverberation chambers*, J. Acoustical Society of America, **31**, 10, 1319–1321.
3. BRADLEY J. S., BIRTA J. A. (2000), *Laboratory measurements of the sound insulation of building facade elements*, IRC Internal Report, IRC IR-818.

4. BRAVO T., ELLIOTT S. J. (2004), *Variability of low frequency sound transmission measurements*, J. Acoustical Society of America, **115**, 6, 2986–2997.
5. BURATTI C., MORETTI E., VERGONI M. (2010), *Sound insulation performances of windows: Evaluation of the influence of different traffic noise spectra in laboratory and field measurement*, Proceedings of 17th International Congress on Sound and Vibration, Cairo.
6. BURATTI C., MORETTI E. (2010), *Traffic noise pollution: spectra characteristics and windows sound insulation in laboratory and field measurements*, J. of Environmental Science and Engineering, **4**, 12, 28–36.
7. CLARK D. M. (1970), *Subjective study of sound-transmission class system for rating building partitions*, J. Acoustical Society of America, **47**, 3, 676–682.
8. COOK R. K., WATERHOUSE R. V., BERENDT R. D., EDELMAN S., THOMPSON M. C. (1955), *Measurement of correlation coefficients in reverberant sound fields*, J. Acoustical Society of America, **27**, 1072–1077.
9. CREMER L., MÜLLER H. A., SCHULTZ T. J. (1982), *Principles and application of Room Acoustics*, Vol. 1, Applied Science Publishers.
10. DAVY J. L. (2004), *Insulating buildings against transportation noise*, Proceedings of annual conference of Australian Acoustical Society, 3rd to 5th November, Gold Coast, Australia.
11. European Standard, EN 1793-3 (1997), *Road traffic noise reducing devices-Test method for determining the acoustic performance, Part 3: Normalized traffic noise spectrum*.
12. FITZELL R. J., FRICKE F. R. (2004), *2004 changes to the BCA – Are they a step forward*, Proc. of Acoustics 2004, pp. 151–156, 3rd to 5th November, 2004, Gold Coast, Australia.
13. GARG N., SHARMA O., MAJI S. (2011), *Experimental investigations on sound insulation through single, double and triple window glazing for traffic noise abatement*, J. of Scientific and Industrial Research, **78**, 471–478.
14. GOVER B. N., BRADLEY J. S. (2004), *Measures for assessing architectural speech security (privacy) of closed offices and meeting rooms*, J. Acoustical Society of America, **116**, 6, 3480–3490.
15. GOYDKE H., SIEBERT B. R. L., SCHOLL W. (2003), *Considerations on the evaluation of uncertainty values of building acoustic single number quantities*, EuroNoise, Naples.
16. Green Glue Company (n.d), Retrieved from <http://www.greengluecompany.com/understandingSTC.php>.
17. HALLIWELL R. E., NIGHTINGALE T. R. T., WARNOCK A. C. C., BIRTA J. A. (1998), *Gypsum board walls: Transmission loss data*, NRC Report No. IRC-IR-761.
18. International Organization for Standardization, ISO 717-1 (1996), *Acoustics – Rating of sound insulation in buildings and of building elements, Part 1: Airborne sound insulation*, Switzerland.
19. International Organization for Standardization, ISO 140-3 (1995), *Acoustics, Measurement of sound insulation in buildings and of building elements, Part 3: Laboratory measurements of airborne sound insulation of building elements*, Switzerland.
20. International Organization for Standardization, ISO 3741 (2010), *Acoustics, Determination of sound power levels and sound energy levels of noise sources using sound pressure, Precision methods for reverberation test rooms*, Switzerland.
21. KOYASU M., TACHIBANA H. (1990), *Subjective tests on assessment of sound insulation efficiencies of building facades*, *Internoise*, 155–158.
22. LANG J. (1997), *A round robin on sound insulation in buildings*, *Applied Acoustics*, **52**, 225–238.
23. MALUSKI P. S., GIBBS B. (2000), *Application of a finite-element model to low-frequency sound insulation in dwellings*, J. Acoustical Society of America, **108**, 1741–1751.
24. NÉLISSE H., NICOLAS J. (1997), *Characterization of a diffuse field in a reverberant room*, J. Acoustical Society of America, **101**, 6, 3517–3524.
25. NT ACOU 061-1987, traffic noise reduction indices, <http://www.nordicinnovation.net/nordtestfiles/acou061.pdf>.
26. OSIPOV A., MEES P., VERMEIR G. (1997), *Low-frequency airborne sound transmission through single partitions in buildings*, *Applied Acoustics*, **52**, 273–288.
27. PARK H. K., BRADLEY J. S. (2009), *Evaluating standard airborne sound insulation measures in terms of annoyance, loudness and audibility ratings*, J. Acoustical Society of America, **126**, 1, 208–219.
28. PARK H. K., BRADLEY J. S., GROVER B. N. (2008a), *Evaluating airborne sound insulation in terms of speech intelligibility*, J. Acoustical Society of America, **123**, 3, 1458–1471.
29. PARK H. K., BRADLEY J. S., GOVER B. N. (2008b), *Rating airborne sound insulation in terms of the annoyance and loudness of transmitted speech and music sounds*, IRC RR-242, <http://www.irc.nrc-cnrc.gc.ca>.
30. PANCHOLY M., CHHAPGAR A. F., SHARMA O. (1977), *Construction of a reverberation chamber at National Physical Laboratory of India*, J. Acoustical Society of India, **V**, 27–33.
31. PATTERSON M. J. (2004), *Recent changes to the sound insulation provisions of the Building code of Australia*, Proc. of Acoustics, 3rd to 5th November, Gold Coast, Australia.
32. PEDERSEN D. B., ROLAND J., RAABE G., MAYSENHÖLDER W. (2000), *Measurement of low-frequency sound insulation of building components*, *Acta Acustica united with Acustica*, **86**, 495–505.
33. RASMUSSEN B. (2010), *Sound insulation between dwellings – Requirements in building regulations in Europe*, *Applied Acoustics*, **71**, 373–385.

34. RASMUSSEN B., RINDEL J. H. (2010), *Sound insulation between dwellings – Descriptors applied in building regulations in Europe*, Applied Acoustics, **71**, 171–180.
35. RINDEL J. H. (2003), *On the influence of low frequencies on the annoyance of noise from neighbours*, Proceedings Inter-noise, Seogwipo, Korea.
36. ROLAND J. (1995), *Adaptation of existing test facilities to low frequency measurements*, Proceedings of Inter-noise 95, pp. 1113–1116, Newport Beach, CA.
37. SCHOLL W., LANG J., WITTSTOCK V. (2011), *Rating of Sound Insulation at Present and in Future. The Revision of ISO 717*, Acta Acustica united with Acustica, **97**, 686–698.
38. SCHOLL W., WITTSTOCK V. (2012), *Does it matter whether single-number values of sound reduction indices are evaluated from third-octave band values or from octave band values?*, Acta Acustica united with Acustica, **98**, 354–360.
39. SCHROEDER M. R. (1962), *Frequency-correlation functions of frequency responses in rooms*, J. Acoustical Society of America, **34**, 1819–1823.
40. SCHROEDER M. R. (1996), *The “Schroeder frequency” revisited*, J. Acoustical Society of America, **99**, 5, 3240–3241.
41. SMITH R. S., MACKENZIE R., MACKENZIE R., WATERS-FULLER T. (2003), *The implications of ISO 717 spectrum adaptation terms for residential buildings*, Proceedings of the Institute of Acoustics, **25**, Pt. 5.
42. SMITH R. S., MACDONALD R., LURCOCK D., MACKENZIE R. K. (2007), *Sensitivity analysis of ISO 717-1 spectrum adaptation terms*, Proceedings of Institute of Acoustics, **29**, Pt. 3.
43. TACHIBANA H., HAMADO, Y., SATO F. (1988), *Loudness evaluation of sounds transmitted through walls – Basic experiments with artificial sounds*, J. Sound Vibration, **127**, 3, 499–506.
44. VIAN J.-P., DANNER W. F., BAUER J. W. (1983), *Assessment of significant acoustical parameters for rating sound insulation of party walls*, J. Acoustical Society of America, **73**, 1236–1243.
45. WARNOCK A. C. C. (2004), *Forward and reverse transmission loss measurements*, The 33rd International Congress and Exposition on Noise Control Engineering, Internoise, Prague, August, 22–25.
46. WITTSTOCK V. (2007), *On the uncertainty of Single-Number Quantities for Rating Airborne Sound Insulation*, Acta Acustica united with Acustica, **93**, 375–386.

# Ultrasonic Method for Monitoring Environmental Risks Associated with Precipitation

Tadeusz GUDRA<sup>(1)</sup>, Dariusz BANASIAK<sup>(2)</sup>, Krzysztof HERMAN<sup>(1)</sup>,  
Krzysztof OPIELIŃSKI<sup>(1)</sup>

<sup>(1)</sup> *Institute of Telecommunications, Teleinformatics and Acoustics*

<sup>(2)</sup> *Institute of Computer Engineering, Control and Robotics  
Wrocław University of Technology*

Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland; e-mail: Tadeusz.Gudra@pwr.wroc.pl

(received May 10, 2012; accepted February 4, 2013)

This paper presents a solution that utilises ultrasonic technology to allow monitoring snow layer thickness or water level based on measurement from air. It describes the principle of operation of a measurement device using three methods of compensating for changing external factors affecting appliance's precision. Block diagram of the device is also provided. In order to verify the proposed solutions, the research team tested the device in laboratory and operating conditions. The results obtained this way make it possible to select a configuration of device operation depending on the required measurement precision and limitations associated with installing the system for actual operation.

**Keywords:** ultrasonic methods, snow layer, monitoring of environmental risk.

## 1. Introduction

Ultrasonic technology has many practical applications in science and technology. It can, for instance, be used for monitoring precipitation amounts (e.g. rain, snow). Measuring snow layer thickness is especially important as excessive load can compromise the structural integrity of e.g. flat roofs. Water level measurements are associated with assessing flood hazards. Among many methods that can be used to perform such measurements, those utilising ultrasonic transducers operating in air environment are especially interesting. They are based on the phenomenon of reflection of ultrasonic waves from the air-snow surface or air-water surface boundary. If the ultrasonic wave travel time along transducer-impedance discontinuity path is measured, it is possible to determine the changes in snow or water level.

## 2. Physical parameters of snow and ice

Solid and liquid precipitation forms at altitudes at which water vapour becomes saturated. Crystal nuclei are formed around freezing nuclei. This results in the formation of water layer with droplets or snowflakes

if sublimation or coagulation processes are involved. After the snowflakes fall, they become snow cover the specific properties of which depend on internal and external factors (mostly temperature and humidity, cf. GUDRA, 2005). The character of snow cover changes continuously as a result of physical processes occurring both inside and on its surface. Density is the most important snow parameter for risk assessment. Actual automatic monitoring of snow density is very difficult. It is, however, possible to measure snow layer depth using the ultrasonic method. Acoustic impedance of snow is related to snow density which means that it is possible to determine the latter parameter based on reflection coefficient measurement at air-snow layer surface boundary.

Ice is formed in the process of water freezing; water has maximum density at 4°C and further cooling results in decreasing density. During freezing, the density of water decreases by about 10%. As in the case of snow cover, there are various types of ice. Its density, however, is similar in all cases.

Snow cover measurements which are a part of standard meteorological observation, are performed everyday by weather stations. Automation of measurement of snow layer parameters with the use of ultra-

sound allows more detailed and frequent observations of changes which can help assess risks in advance.

Non-contact monitoring of water level in rivers and water basins can be performed in a similar fashion as in the case of snow layer depth measurement.

### 3. Methods of compensating for parameters affecting measurement precision

In order to perform precise measurement of snow layer thickness or water level for the purposes of monitoring changes occurring in time, it is necessary to compensate for external factors affecting value  $c$  – velocity of sound propagation in a given medium (temperature, humidity, pressure, wind, etc.). Depending on the required measurement precision, it is possible to use three methods of compensation (GUDRA *et al.*, 2011):

- temperature compensation (CANALI *et al.*, 1982);
- parametric compensation with the use of a reflector (CHANDE *et al.*, 1984);
- parametric compensation with the use of two measurement probes.

Figure 1 shows measurement principles with one of the above mentioned compensation methods utilised.

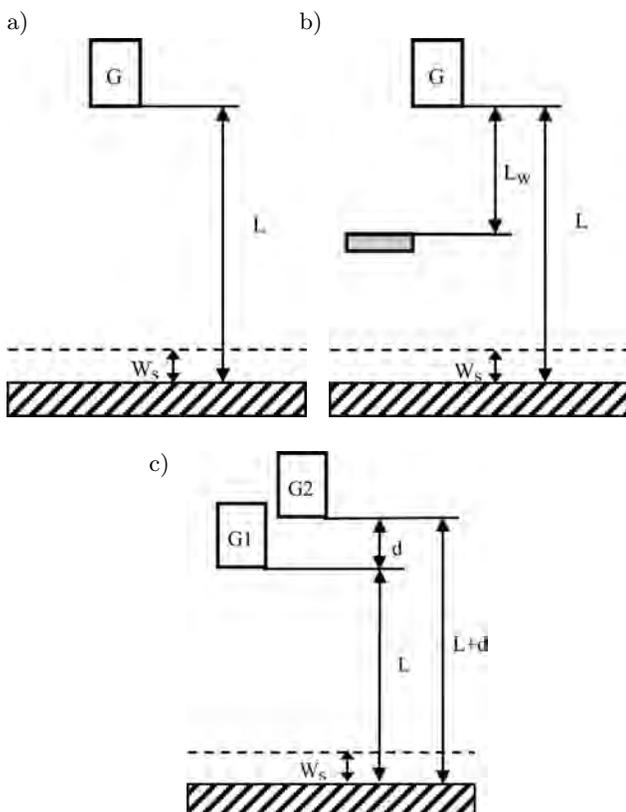


Fig. 1. Methods of compensating for external factors affecting measurement precision: a) a system with temperature compensation; b) a system with parametric compensation with the use of a reflector; c) a system with parametric compensation with the use of two probes.

In the system with temperature compensation, the device measures  $t_1$  – ultrasonic wave travel time from the measurement probe G to the measured surface and back, as well as ambient temperature  $T_c$ . Snow layer depth  $W_s$  can then be determined using the following formula:

$$W_s = L - \frac{t_1 \cdot c}{2}, \quad (1)$$

where  $L$  is the reference distance (no snow layer) and  $c$  is the sound velocity equal to  $c = 331.5(1 + T_c/273.15)^{1/2}$  m/s.

In the system with a reflector, the device measures  $t_1$  – the ultrasonic wave travel time from probe G to the reflector and back and  $t_2$  – the ultrasonic wave travel time from probe G to the measured surface and back. Snow layer depth  $W_s$  can then be determined using the following formula:

$$W_s = L - L_w \frac{t_2}{t_1}, \quad (2)$$

where  $L$  is the reference distance (no snow/ice layer) and  $L_w$  is the distance of the measurement probe from the reflector.

In the system with two measurement probes, the probes are located at a different altitude in relation to the measured surface (assuming uniform snow surface below the probes) and the distance between the probes is known and is  $d$ . The device measures two times of flight:  $t_1$  – the ultrasonic wave travel time from probe  $G_1$  to the measured surface and back and  $t_2$  – the ultrasonic wave travel time from probe  $G_2$  to the measured surface and back. Snow layer depth is determined using the following formula:

$$W_s = L - d \frac{t_2}{t_2 - t_1}, \quad (3)$$

where  $L$  is the reference distance (no snow/ice layer) measured for probe  $G_2$ .

Figure 2 shows block diagram of the measurement device for all compensation methods. According to the measurement principles presented in Fig. 1, temperature measurement is not required in systems with parametric compensation and only one probe is sufficient in systems with a reflector.

Figure 2b shows the structure of a measurement ultrasonic probe. The probe consists of a piezoelectric ultrasonic transducer intended to work in air environment at the frequency of  $f = 50$  kHz (ATK 50, Airmar, USA) and all the electronic systems necessary for the measurement device operation (BANASIAK *et al.*, 2009; BEDNAREK *et al.*, 2010). A parabolic acoustic horn is used to guarantee a required ultrasonic beam shape (width at  $-3$  dB below  $20^\circ$ ). Theoretical measurement range of snow layer thickness is 10 m.

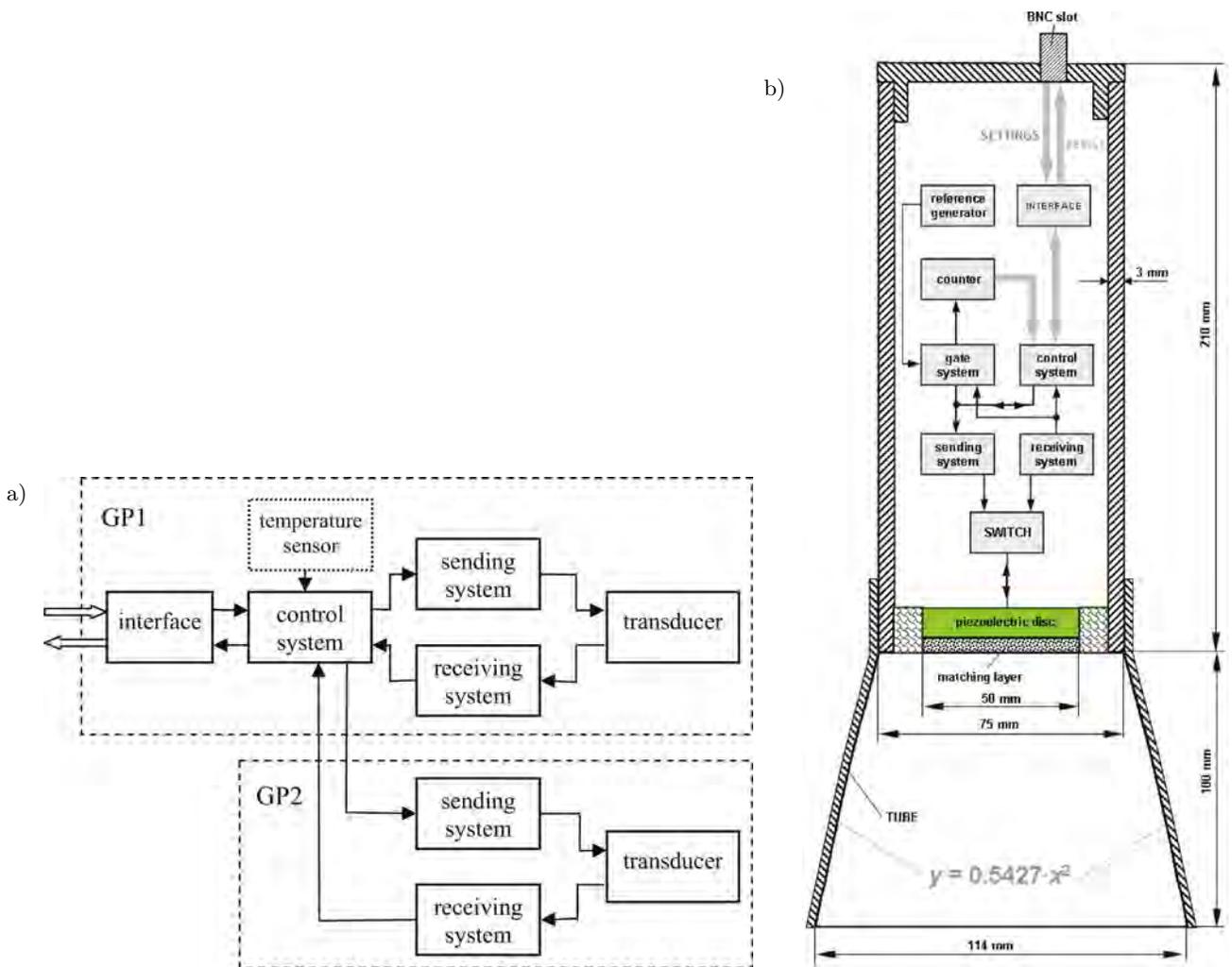


Fig. 2. Diagram of the measurement device.

#### 4. Measurements

The developed measurement device was tested in various operating configurations and measurement conditions (also during snow precipitation). All the measurements were intended to check whether the device operates properly and the obtained results are consistent as well as to determine uncertainty of distance measurement (snow layer depth measurement). Table 1 shows the results of analysis of the measuring inaccuracy for individual measurement methods.

Using the exact differential method to analyse measurement inaccuracy shows that total measurement uncertainty includes components related to device configuration and direct error components related to the actual measurement. The values related to measuring instrument configuration are a result of measurement geometry; their effect can be minimised through calibration.

As the device is intended to work in harsh conditions (low temperatures, humidity, icing), it had to be properly tested. Initial measurements were performed

Table 1. Analytical relations for snow layer thickness and measurement uncertainty with various measurement methods.

|                        | Single-probe measurement  | Two-probe measurement   | Single-probe measurement with a reflector  |
|------------------------|---|---|--|
| Snow layer             | $W_s = L - \frac{2 \cdot c}{t}$   | $W_s = L - d \frac{t_1}{t_2 - t_1}$   | $W_s = L - L_w \frac{t_2}{t_1}$  |
| Measurement inaccuracy | $\Delta W_s = \Delta L + \left  \frac{2 \cdot c}{t^2} \right  \Delta t$ | $\Delta W_s = \Delta L + \frac{t_1}{t_2 - t_1} \Delta d + \frac{d}{(t_2 - t_1)^2} \Delta t_1 + \frac{dt_1}{(t_2 - t_1)^2} \Delta t_2$ | $\Delta W_s = \Delta L + \frac{t_2}{t_1} \Delta L_w + \frac{L_w}{t_1} \Delta t_2 + L_w \frac{t_2}{t_1^2} \Delta t_1$ |

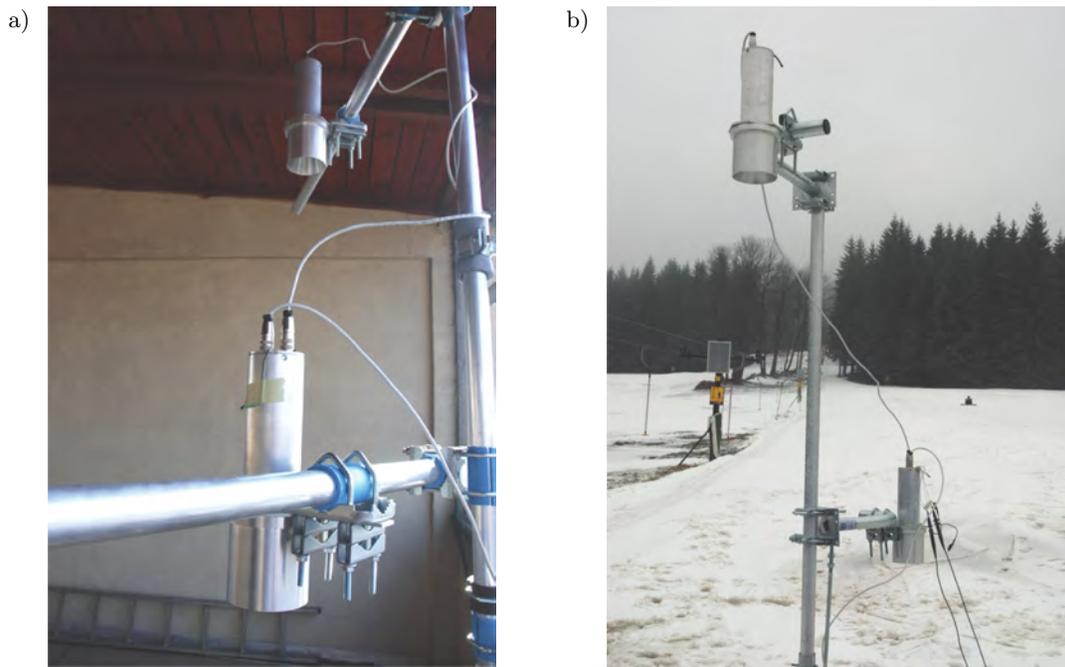


Fig. 3. Probe configuration during measurement: a) in laboratory conditions; b) in operating conditions.

on the roof of building C-5 of the Wrocław University of Technology where the measurement instrument was mounted. A view of the appliance is shown in Fig. 3a. With the presented configuration it is possible to study the effect of weather conditions on the operation of the device for methods I and II. During the measurements, the team observed result consistency and studied the effect a given compensation method has on result spread. As there was no snow, the device was tested in the above described study arrangement using mineral wool with appropriate parameters to simulate snow (IWASE *et al.*, 2001).

According to recommendations related to snow layer measurements presented by RYAN *et al.* (2007), such a measurement should be performed periodically with a 5 minute interval. To test result repeatability, the distance was measured 4 times for each channel. This allowed the team to discard measured values with gross error and average the results.

Measurements in operating conditions were performed in Czarna Góra ski resort. Device mount is shown in Fig. 3b.

## 5. Measurement results

Measurements performed at the measurement position provided results that can be further analysed. As the device is characterised by fractional result inconsistency which is manifested as drastic result differences inside a single measurement series, the team used a simple calibration algorithm that made it possible to discard such results. In the algorithm, a certain level of calibration is assumed for a given channel (base

level with no snow), and then the measured values are compared to previous results in accordance with the following relation:

$$x_n = \begin{cases} x_n & \text{for } x_n \geq k \cdot x_{n-1}, \\ x_{n-1} & \text{for } x_n < k \cdot x_{n-1}, \end{cases} \quad (4)$$

where  $k$  is a coefficient from 0–1 range that discriminates dynamic properties of measurement result changes (snow cover increase). Values of  $k$  close to one describe a situation in which even very small changes in the measurement interval will be discarded by principle as an error result. It is important to note that this policy is justified because snow cover is not likely to change significantly between 5 minute intervals. Initial device tests included laboratory measurements the purpose of which was to determine device precision and result repeatability. Examples of laboratory measurement results were presented as time series diagrams and bar charts shown in Fig. 4 and Fig. 5.

Figure 6 shows the result of a measurement in configuration III which can also be treated as two independent measurement results in configuration I. The measurement was performed outdoors without temperature compensation, as it is clearly visible in the chart. If temperature compensation is applied according to relation dependence (1), it is possible to achieve the result visible in Fig. 7. Analysis of the chart suggests that distance measurement results are not correlated with temperature, which in turn decreases the mean value spread. It can however be observed that both histories are to some extent correlated as a result of environmental conditions other than tempera-

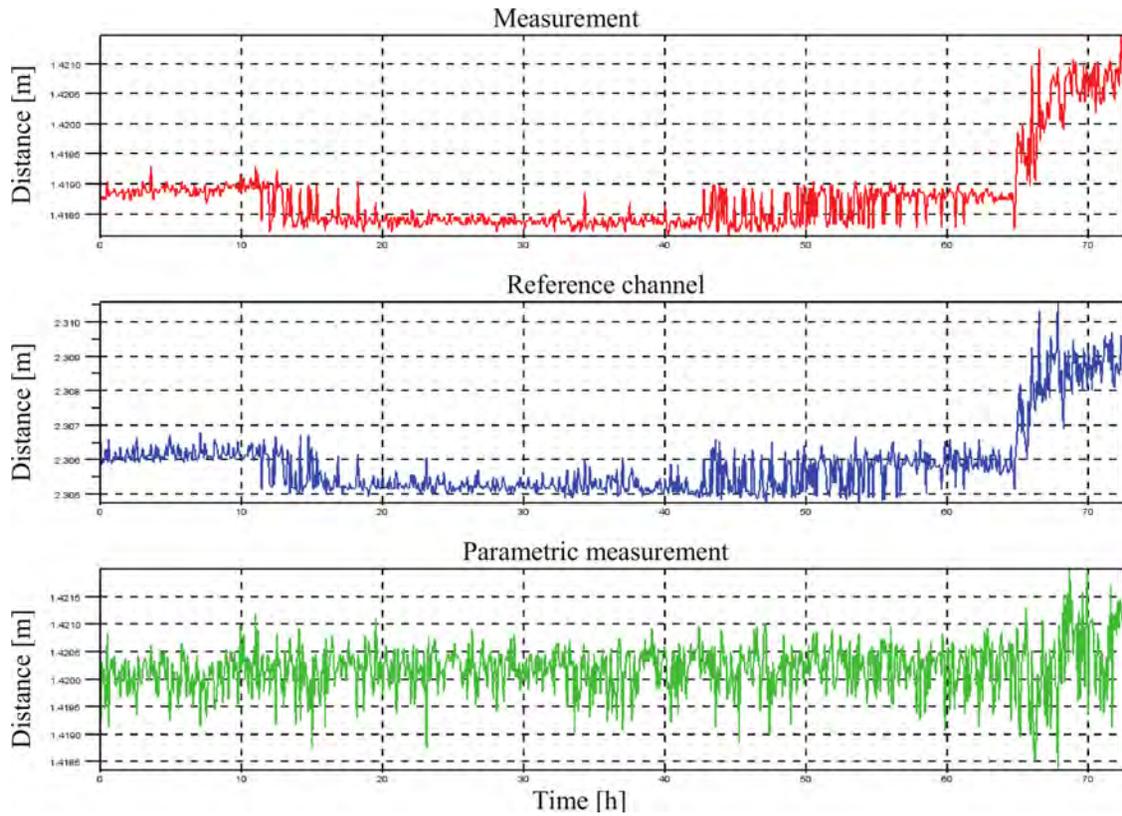


Fig. 4. Device laboratory tests – results for distance measurement in configuration III with temperature compensation. Parametric measurement calculated for distance  $d = 0.888$  m.

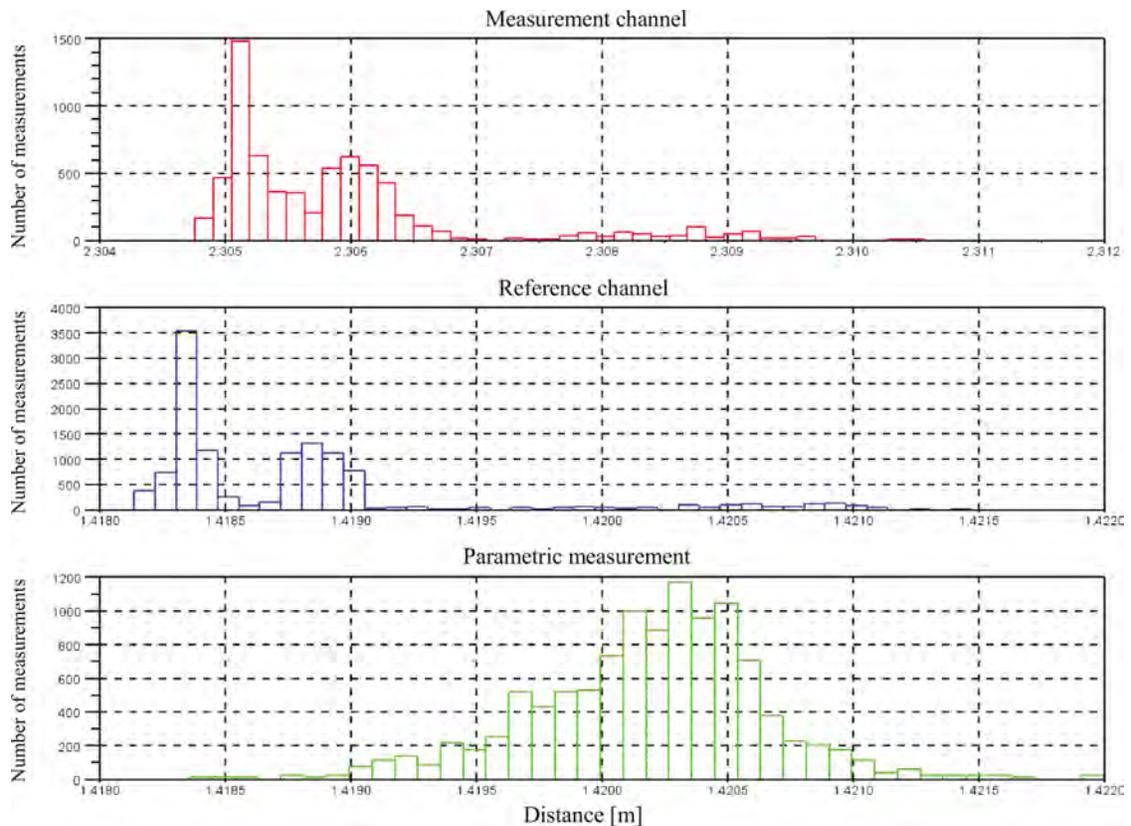


Fig. 5. Device laboratory tests – histograms of measurement results in configuration III with temperature compensation. Parametric measurement calculated for distance  $d = 0.888$  m.

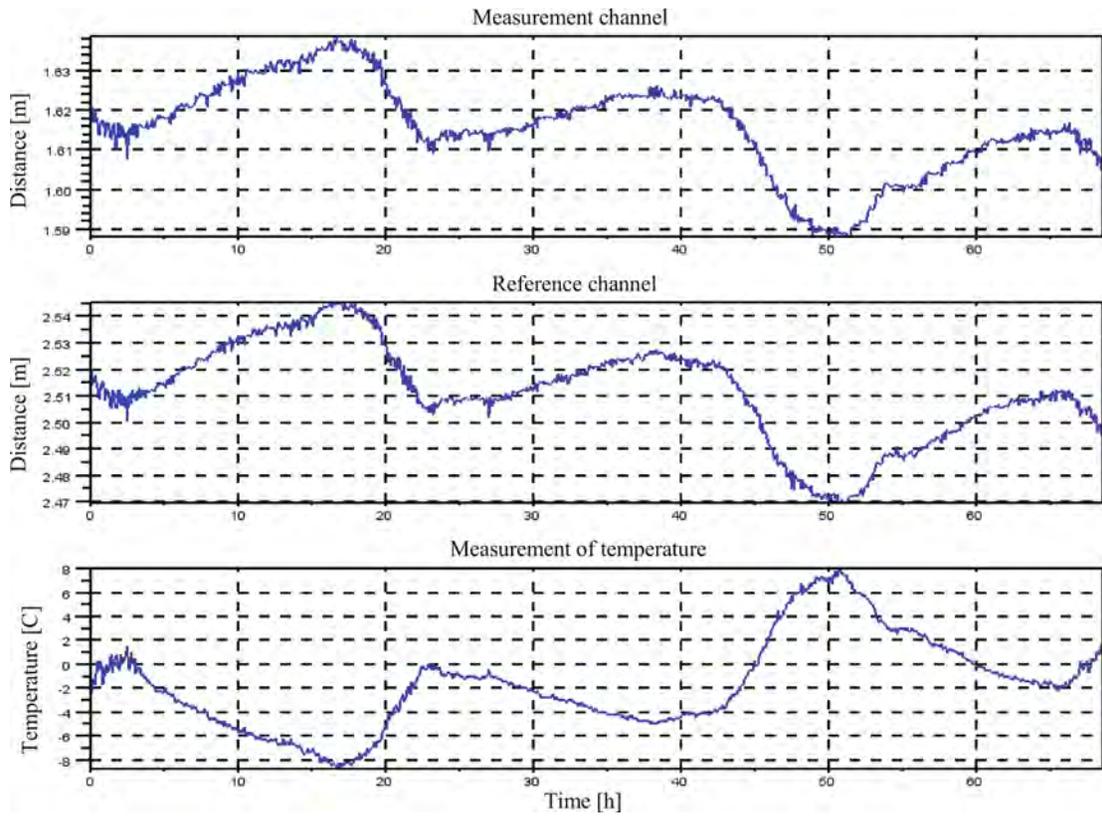


Fig. 6. Distance measurement using two probes without temperature compensation.

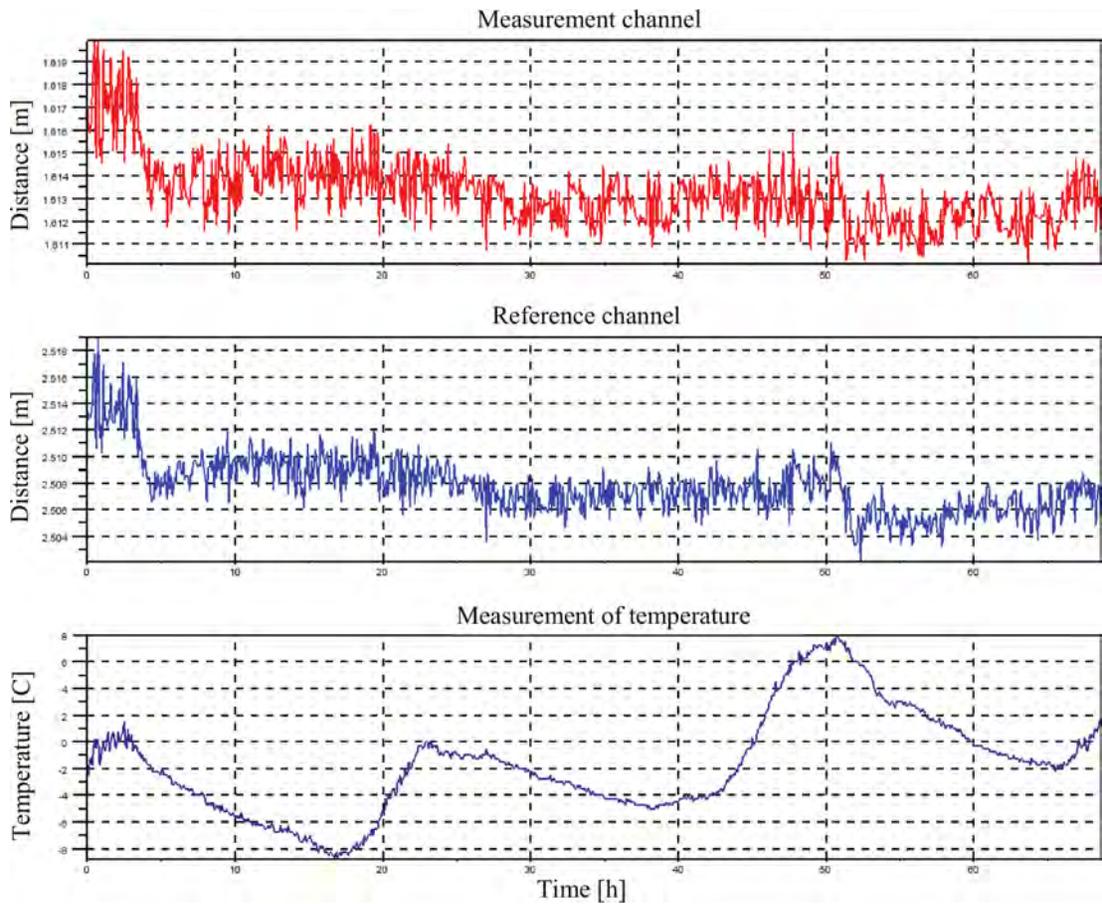


Fig. 7. Distance measurement using two probes with temperature compensation.

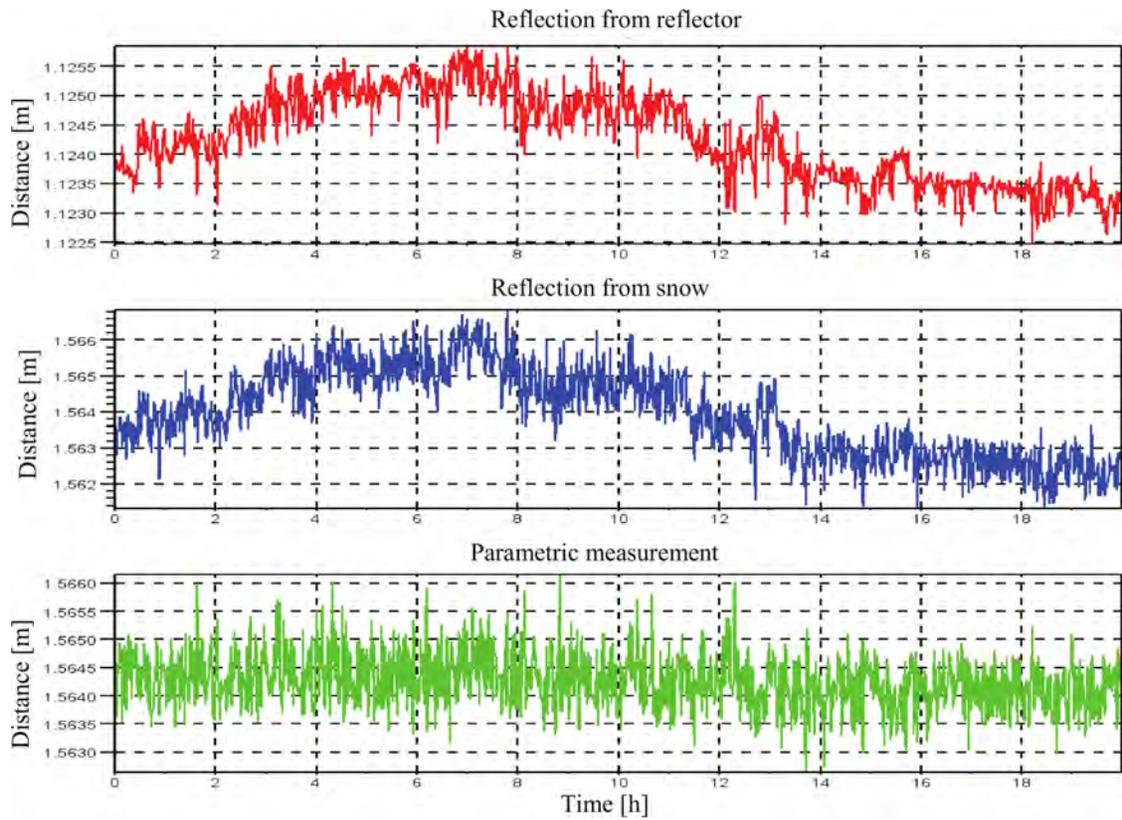


Fig. 8. Measured distance values for a measurement with parametric compensation and the use of a reflector.

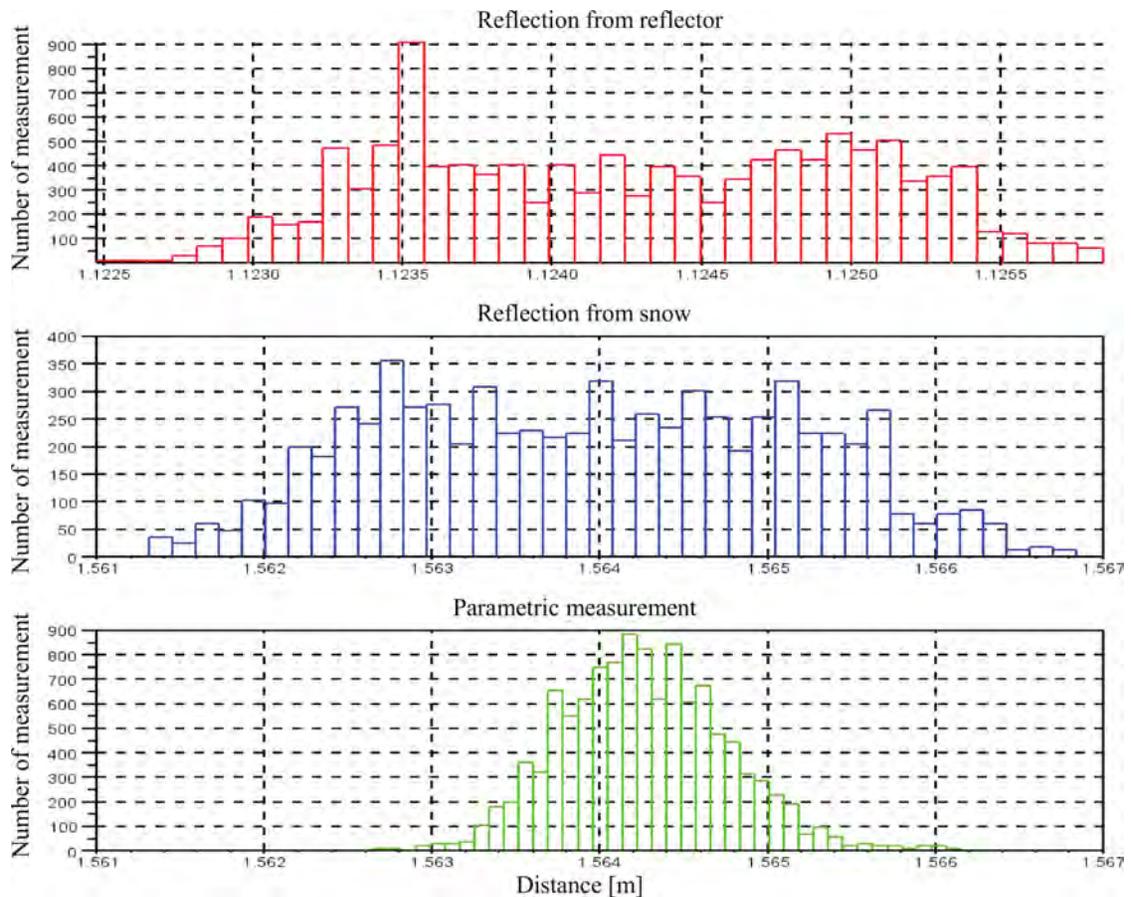


Fig. 9. Histograms for a parametric measurement with the use of a reflector and simulated snow layer.

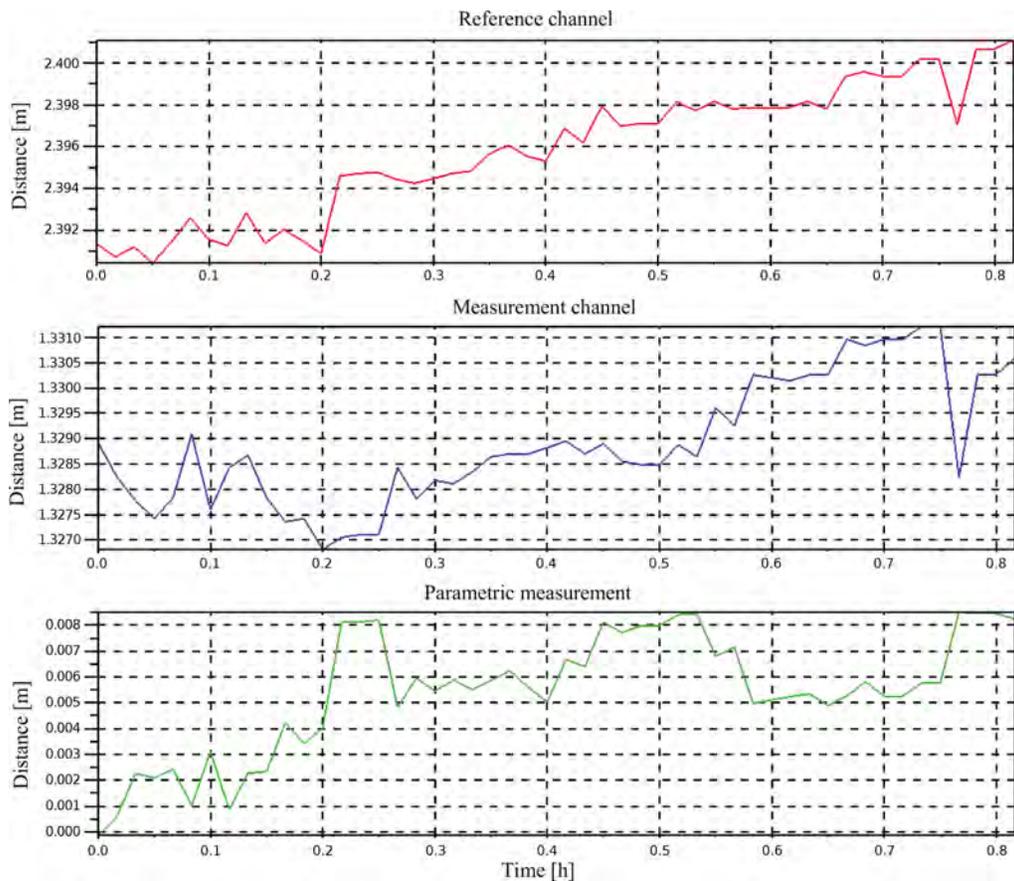


Fig. 10. Parametric measurement of snow layer (configuration III) in actual operating conditions. The lower figure shows snow cover thickness.

ture. The effect of humidity and pressure can be corrected using parametric compensation. Figures 8 and 9 show measured distances and their corresponding histograms (visualising spread of mean value obtained from measurement results) for a measurement with parametric compensation and the use of a reflector (copper rod with the diameter of  $D = 5$  mm was used as the reflector). The parametric measurement histogram is the most compact which should be interpreted as an evidence of low measurement uncertainty. Figure 10 shows an example parametric measurement of snow layer in configuration III in actual operating conditions.

Currently, the measurement device's accuracy is being verified in extreme conditions during 34th Polar Scientific Expedition to Spitsbergen organised by the Department of Polar Research of the Institute of Geophysics of the Polish Academy of Sciences in Warsaw. The snow layer thickness measuring instrument operates only in sub-zero temperature conditions. Harsh weather conditions (low temperature, wind) make it possible to verify the selected measurement methods.

Figure 11 shows the measurement instrument mounted on a mast near the Polish Polar Station and operated with temperature compensation.



Fig. 11. Ultrasonic instrument measuring snow layer thickness mounted on a meteo mast during research on Spitsbergen.

Figure 12 shows distance measurement results (related to snow layer thickness) in a system with temperature compensation, and Fig. 13 shows measurement results separately for snow layer thickness and temperature without temperature compensation (no precipitation).

The presented relations clearly show that temperature alteration results in distance measurement errors. The characteristics of the histories of both values (distance and temperature) are very similar. It is

not possible to obtain a valid distance measurement without compensating for temperature changes. Resolution of measurement in the laboratory condition is 0.2%, whereas resolution in outdoor condition is approximately 1%. The uncertainty of thickness measurement is comparable with other commercial devices using temperature compensation (e.g. SR 50, USH). By applying methods with parametric compensation with the use of a reflector or two measurement probes, the measurement uncertainty is reduced.

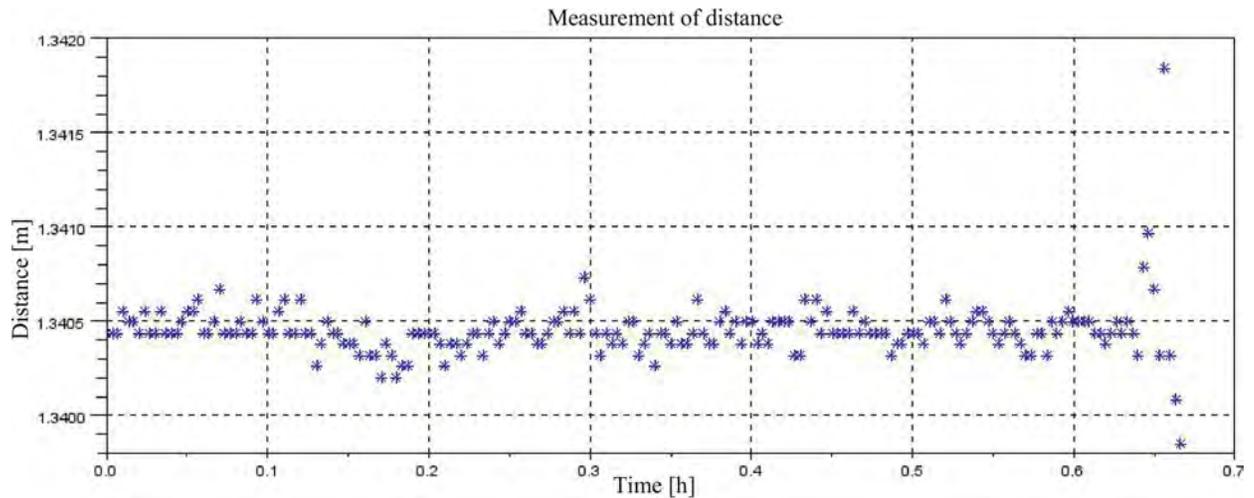


Fig. 12. Measured distance values for a measurement with one probe and temperature compensation.

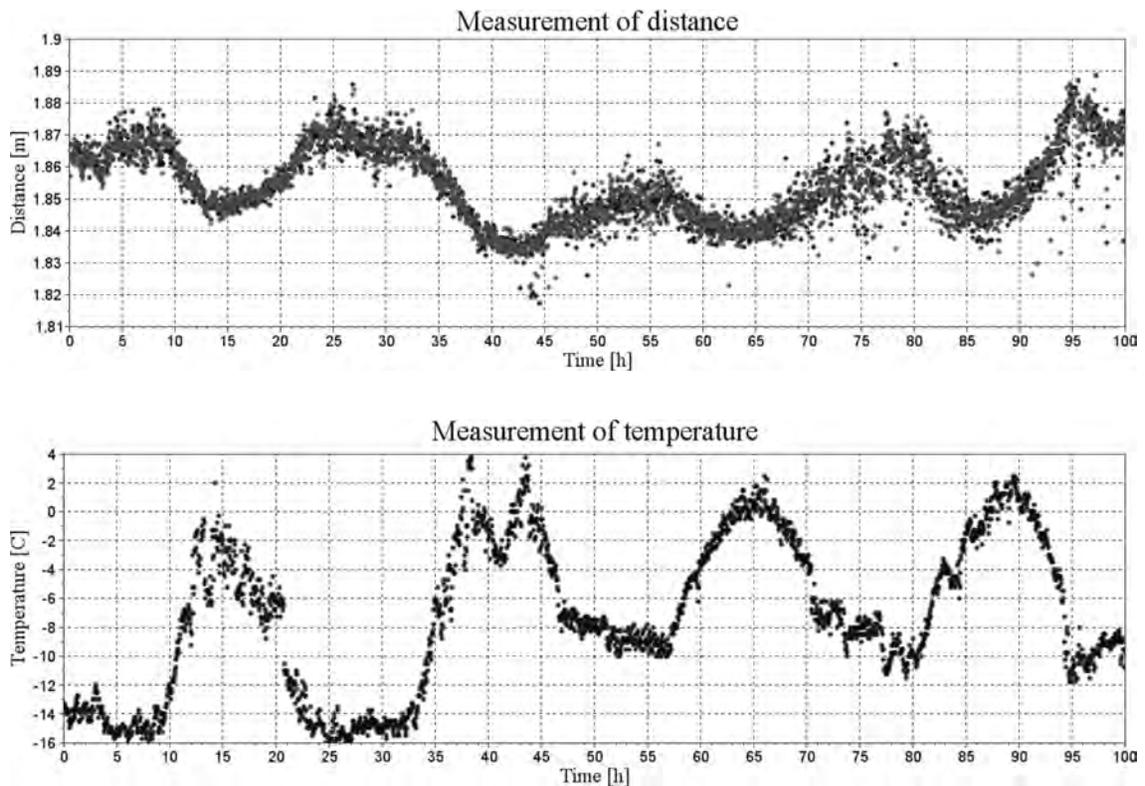


Fig. 13. Measurement results: a) for distance without temperature compensation; b) temperature.

## 6. Prospects for using signal reflected from a snow layer for assessment of the density of various snow types

Figure 14 shows the dependence of pressure coefficient of reflection from snow surface for the studied snow density values presented in Table 2. It is easy to notice that the coefficient value increases linearly with increasing snow density. In order to visualise the linearity better, a trend line was added to the chart. Minor deviations of the obtained results in relation to the trend line can be caused by measurement errors. They are unavoidable in case of snow – a very unstable medium.

Table 2. Coefficient of reflection of the surface of snow with various density values (GUDRA, NAJWER, 2011).

| Snow density [kg/m <sup>3</sup> ] | Pressure reflection coefficient |
|-----------------------------------|---------------------------------|
| 120                               | 20.30%                          |
| 124                               | 20.60%                          |
| 132                               | 22.19%                          |
| 170                               | 25.77%                          |
| 232                               | 31.58%                          |
| 325                               | 37.24%                          |
| 495                               | 53.29%                          |
| 591                               | 63.19%                          |
| 696                               | 70.24%                          |

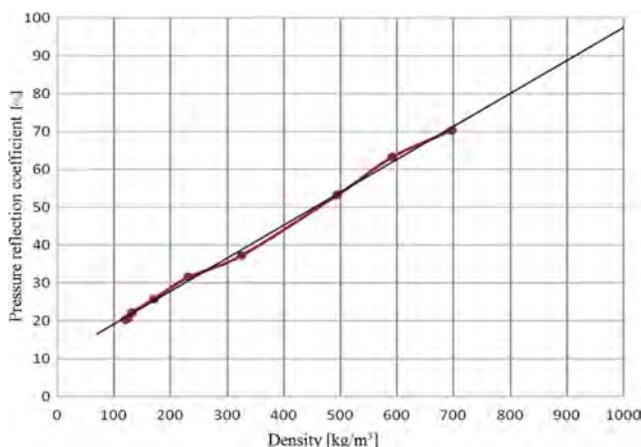


Fig. 14. The relation of the coefficient of reflection from the surface of snow to snow density (GUDRA, NAJWER, 2011).

Detailed analysis of the direction of trend line increase shows that for 1000 kg/m<sup>3</sup> density (which is close to water density at 4°C), the value of the pressure reflection coefficient would be 97%. Theoretical pressure reflection coefficient  $R'$  on the boundary between two media is described by the following relation:

$$R' = \frac{1 - m}{1 + m}, \quad (5)$$

where  $m = (\rho_1 c_1)/(\rho_2 c_2)$  is the ratio of acoustic impedances of the two media, air and snow.

Acoustic impedance of air is 429 kg/(m<sup>2</sup>·s), of water – 1.48·10<sup>6</sup> kg/(m<sup>2</sup>·s) (SZCZENIOWSKI, 1967). The value of the theoretical pressure reflection coefficient calculated with formula (5) is 99.94%. Comparison of the obtained values suggests that the measurement results are biased with little error. Precise measurements of reflection coefficient can be used to assess the density of various snow types (GUDRA, NAJWER, 2011).

## 7. Conclusions

The developed device can operate in three different modes, depending on required precision and mounting conditions. It should be noted that the results were obtained for a prototypical set of devices which can be further optimised. The choice of interface for the device includes RS-422 with MODBUS/ASCII protocol or standard RS-232 port operating in ASCII mode. It is also possible to use more sophisticated methods of result interpretation in order to reduce measurement uncertainty. Such calculations can be performed by a sensor network control centre.

The developed measurement device can almost immediately be used for measuring distances in air (e.g. to measure water level in rivers and other water basins).

## Acknowledgments

Development of this measurement device was possible as a result of funds obtained in relation to a project called “Detectors and sensors for measuring factors hazardous to environment – modelling and monitoring of threats” – POIG.01.03.01-02-002/08 financed by the European Union from European Regional Development Fund and the national budget.

The authors would like to express their deepest gratitude for the opportunity to test the measurement device during 34th Polar Scientific Expedition to Spitsbergen organised by the Department of Polar Research of the Institute of Geophysics of the Polish Academy of Sciences in Warsaw.

This article is an extended version of a paper presented during 58th Open Seminar on Acoustics, 13–16 September 2011, Gdańsk-Jurata.

## References

- BANASIAK D., GUDRA T., OPIELIŃSKI K., BEDNAREK J. (2009), *Elaboration of ultrasonic transducers to monitoring of snow or ice layer level on flat roofs* [in Polish: *Opracowanie przetworników ultradźwiękowych do monitorowania poziomu warstwy śniegu lub lodu na płaskich dachach*], Report of Wrocław University of Technology, No. I-28/2010/S-007.

2. BEDNAREK J., BANASIAK D., GUDRA T., OPIELIŃSKI K. (2010), *Measurement of parameters and characteristics of elaborated transducers* [in Polish: *Pomiar parametrów i charakterystyk opracowanych przetworników*], Report of Wrocław University of Technology, No. I-28/2010/S-087.
3. CANALI C., DE CICCIO G., MORTEN B., PRUDENZIA-TI M., TARONI A. (1982), *A Temperature Compensated Ultrasonic Sensor Operating in Air for Distance and Proximity Measurements*, IEEE Trans. on Industrial Electronics, **29**, 4, 336–341.
4. CHANDE P. K., SHARMA P. C. (1984), *A Fully Compensated Digital Ultrasonic Sensor for Distance Measurement*, IEEE Trans. on Instrumentation and Measurement, **33**, 2, 128–129.
5. GUDRA T. (2005), *Properties and application of ultrasonic transducers working in gaseous media* [in Polish: *Właściwości i zastosowanie przetworników ultradźwiękowych do pracy w ośrodkach gazowych*], Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław.
6. GUDRA T., BEDNAREK J., HERMAN K., OPIELIŃSKI K., BANASIAK D. (2011), *Ultrasonic sensor to monitoring of snow or ice level on flat roofs – monitoring from air side* [in:] *Detectors and sensors for measuring of factors hazardous in environment*, Part 1 [in Polish: *Czujnik ultradźwiękowy do monitorowania poziomu warstwy śniegu lub lodu na płaskich dachach (monitoring od strony powietrza)*], [w:] *Czujniki i sensory do pomiarów czynników stanowiących zagrożenia w środowisku*, Grzebyk W.E. [Ed.], Monografia projektu POIG.01.03.01-02-002/08, pp. 103–129.
7. GUDRA T., NAJWER L. (2011), *Ultrasonic Investigation of Snow and Ice Parameters*, Acta Physica Polonica A, **120**, 4, 625–629.
8. IWASE T., SAKUMA T., YOSHIHISA K. (2001), *Measurements on Sound Propagation Characteristics in Snow Layer*, 17th International Congress on Acoustics, Roma, Italy.
9. RYAN W. A., DOESKEN N. J. (2007), *Ultrasonic snow depth sensors for national weather service. Snow measurements in the US: evaluation of operational readiness*, 14th Symposium on Meteorological Observation and Instrumentation, 17 January.
10. SZCZENIOWSKI S. (1967), *Experimental Physics* [in Polish: *Fizyka doświadczalna*], Part I, PWN, Warszawa.

## Chronicle

**42nd Winter School on Wave and Quantum Acoustics  
41th Winter School on Vibroacoustical Hazards Suppressions****Szczyrk, Poland, February 25 – March 1, 2013**

Traditionally you are invited to acquaint yourselves with some information connected with *42nd Winter School on Wave and Quantum Acoustics, 41st Winter School on Vibroacoustical Hazards Suppressions*.

For some years together Winter Schools are also organized the international optical conference – *Integrated Optics – Sensors, Sensing Structures and Methods*.

The Conferences traditionally held at the turn of February and March and now they are again organized in Szczyrk, in beautiful winter scenery.

The Conferences, as in the previous years, have some organizers, i.e. Upper Silesian Division of the Polish Acoustical Society (the main organizer), the Department of Optoelectronics and the Institute of Physics – Science-Didactic Center at the Silesian University of Technology and the Committee on Acoustics of the Polish Academy of Sciences as well as the Committee of Electronics and Telecommunication of the Polish Academy of Sciences.

The Conferences are traditionally sponsored by the Ministry of Science and Higher Education.

The Conferences have the splendid honorary patron – Prof. Dr. Hab. Eng. Andrzej Karbownik, His Magnificence Rector of the Silesian University of Technology.

On behalf of Organizers  
*prof. Marian Urbanczyk*

Chairman Upper Silesian Division of PAN

**Abstracts****Applications of precision measurements of electrical conductivity of the solutions**

BALD A.

Modern, precise studies of electrical conductivity of electrolyte solutions are one of the main tools of research in physical chemistry of electrolyte solutions. These studies allow to get in a fairly simple way the values of molar limiting

conductance, values of association (dissociation) constant as well as the values of the distance parameters of ions. After using an appropriate methods can also determine the values of ion limiting molar conductance. Values of limiting molar conductance are used mainly to describe the interactions of ions with solvent. These values of limiting molar conductance are also important from the application point of view. Extremely important is the role of conductivity research to determining of equilibrium constants of different types. Examples of applications of modern conductivity measurements to solve various problems and determining the value of various physicochemical parameters has been presented. It also discusses the difficulties associated with the interpretation of some of the results of conductivity measurements.

\* \* \*

**Some remarks on the experimental aspects of determining apparent and partial molar volumes**

BALD A., KLIMASZEWSKI K., TRZCIŃSKA I.

Apparent and partial molar volumes are the fundamental properties describing solute and its interaction with the solvent. For the determination of these parameters density measurements are used most frequently. However, studies of density of solutions must meet the relevant standards related to the methodology used for research and measurement apparatus. Otherwise, the partial molar volume values can not be determined or these values are unreliable. Specific examples are shown to present difficulties in determining the value of the apparent and partial molar volume and ways to overcome these obstacles and also, what are the restrictions on the use of density measurement. Mainly electrolyte solutions were discussed as creating more research problems.

\* \* \*

**Forced pendulum as a low frequency mechanical spectroscopy to probe physical phenomena in semi-crystalline polymers**

BOYER S. A. E., RIVIERE A.

Originally, a sub-resonant forced pendulum suspension is proposed to probe reliable source of information related

to relaxation processes in semi-crystalline polymer materials. The forced pendulum is a prototype permitting accurate damping measurements for a wide frequency scale  $10^{-4}$ –50 Hz under 298–873 K [Woïgard, Mazot, Rivière, *J. Phys.*, 42 C5 (1981)]. Mechanical loss measurements are carried out as a function of vibration frequency at fixed temperatures.

Internal friction and shear-modulus are obtained from frequency characteristics, as illustrated with the position of the maxima of mechanical losses for the alpha relaxation processes in a polyamide 11.

\* \* \*

### Ultrasound research on DHP

BRÓŻYNA A., KACZMAREK-KLINOWSKA M., SKUMIEL A.

Cyclodextrin molecule can form an inclusion complex with a guest molecule. For our study water and ethyl alcohol solutions of nimodipine (derivative of 1,4-dihydropyridine) and methyl- $\beta$ -cyclodextrin were prepared. Solutions with different guest and cyclodextrin ratio were fixed in both environments. Measurement of velocity (using the resonance method at the frequency of 8 MHz), thermal capacity and density of the samples were conducted. Finally the redundant functions of the molar volume and adiabatic compressibility were calculated. Based on the results of the research the existence of inclusion complex between nimodipine and methyl- $\beta$ -cyclodextrin molecules can be determined.

\* \* \*

### Thermodynamics of dihaloalkanes at high pressures

CHORAŻEWSKI M., GROLIER J.-P.

The knowledge of the thermophysical properties of liquid halogenoalkanes is of high interest on account of their wide usage in science and industrial processes. High temperature – high pressure properties are directly applicable to the designing of chemical processes as well as for the progress of thermodynamic theories. pVT behavior will be useful later in analyzing the effect of pressure on a number of physical and thermal properties, on chemical equilibrium, and to some extent on chemical reaction kinetics.

\* \* \*

### 1-alkoxymethylimidazolium salicylates: correlation between ionic structure and physicochemical properties

CZECH B., FEDER-KUBIS J., ZOREBSKI M., CHORAŻEWSKI M., DZIDA M., GEPPERT-RYBCZYŃSKA M., HENSEL-BIELÓWKA S., ZOREBSKI E., ŻARSKA M.

A new family of ionic liquids, composed of salicylate anion and 1-alkoxymethylimidazole cation, were synthesized and characterized. The acoustic, volumetric, calorimetric and refractometric investigation for these salts was performed. The properties under study are sensitive to the structure of the cation. The sound velocity, density and refractive index decreases linearly with increasing alkoxy chain length substituted in imidazole ring whereas molar heat capacity is changing in opposite way. Elongation of alkoxy chain leads to an expansion of the nonpolar region

what influences the interactions and consequently the basic physicochemical properties.

\* \* \*

### Photopyroelectric Calorimetry of Magnetic Nanofluids. Effect of Type of Surfactant and Magnetic Field

DADARLAT D., LONGUEMART S., TURCU R., STREZA M., VEKAS L., SAHRAOUI A. H.

Five types of magnetic nanofluids, based on  $\text{Fe}_3\text{O}_4$  nanoparticles with water as carrier liquid, were investigated by using the two PPE detection configurations (back and front), together with the TWRC technique as scanning procedure. The difference between the nanofluids was the type of surfactant: double layers of Lauric (LA-LA), Oleic (OA-OA) and Miristic (MA-MA) acids and also double layers of Lauric-Miristic (LA-MA) and Palmitic-Oleic (PA-OA) fatty acids were used. In both detection configurations, the information was contained in the phase of the PPE signal. In the BPPE configuration, the thermal diffusivity of the nanofluids was obtained from the slope of the phase of the signal as a function of liquid's thickness. In the FPPE configuration the thermal effusivity was directly measured. The influence of a 0.12 kG magnetic field on the thermal effusivity and diffusivity was also investigated. Due to different surfactants the value of the thermal effusivity of the investigated nanofluids ranges from  $1530 \text{ W s}^{1/2} \text{ m}^{-2} \text{ K}^{-1}$  to  $1790 \text{ W s}^{1/2} \text{ m}^{-2} \text{ K}^{-1}$ , and the value of thermal diffusivity, from  $14.54 \times 10^{-8} \text{ m}^2/\text{s}$  to  $14.79 \times 10^{-8} \text{ m}^2/\text{s}$ . The magnetic field has practically no influence on the thermal effusivity, and produces a maximum increase of the value of the thermal diffusivity (LA-LA surfactant) of about 4%.

\* \* \*

### A New Photothermal Calorimetry: the Photothermoelectric (PTE) Technique

DADARLAT D., STREZA M., KURIAKOSE M., DEPRIESTER M., SAHRAOUI A. H.

The recently introduced photothermoelectric (PTE) effect is proposed as an alternative for measuring dynamic thermal parameters of solid samples. The front PTE configuration, together with the thermal-wave resonator cavity (TWRC) method as scanning procedure, was used to measure the value of thermal effusivity. The back PTE configuration, together with the chopping frequency of incident radiation as scanning parameter, leads to the direct measurement of thermal diffusivity. The theory of the two detection configurations was developed and applications on some solids, covering a large range of typical values of thermal parameters (aluminum, glass, teflon, polyethylene,  $\text{LiTaO}_3$ ), were described, in order to demonstrate the suitability of the method. Some comparison with the photopyroelectric (PPE) method has been done.

\* \* \*

### Impact of infrasound noise emitted by wind turbines for human

DOBRUCKI A., BOGUSZ B.

The state of knowledge concerning wind turbine noise in infrasound and low-frequency range as well as its in-

fluence into humans are presented in the paper. The infrasound is defined as noise in frequency range below 16 Hz. Low-frequency noise is defined by various authors or standards as the noise with frequency range above 16 Hz but lower than 100, 160 or 200 Hz. Both kinds of noise are produced aerodynamically by rotating wind turbine blades. This noise should be measured in 1/3-octave frequency bands or with using frequency weighting curve G. Infrasound influences the auditory system, internal organs, respiratory system, nervous system and coronary vascular of humans. High levels of infrasound can be perceived by auditory system. Typical values of G-level are equal to 60–70 dB at the distance 300–600 m from wind turbines. The permitted values of infrasound levels are 102 dBG at workplaces and 86 dBG for conceptual works requiring special attention span. This means that the wind turbines are not oppressive sources of infrasound noise.

\* \* \*

### Hydration of Na<sup>+</sup> and K<sup>+</sup> with hydrogen maleate and hydrogen carbonate. Car-Parrinello Molecular Dynamics vs. Speed of Sound Experiment

DOPIERALSKI P., BURAKOWSKI A., GLIŃSKI J., LATAJKA Z.

Ion hydration plays an important role in many chemical processes and thus significant number of work have been done concerning this phenomena. The importance of metal ions, commonly found in nature, and their numerous applications in the field of chemistry is undeniable, as well as their role in the life processes.

In this work we study the hydration of Na<sup>+</sup> and K<sup>+</sup> cations alone and with hydrogen maleate and hydrogen carbonate anions by means of Car-Parrinello Molecular Dynamics simulations. The influence of hydrogen maleate or hydrogen carbonate ion on hydration numbers of Na<sup>+</sup> and K<sup>+</sup> will be discussed as well as the opposite effect. Obtained from theoretical investigations hydration numbers will be compared with experimental data from speed of sound measurements.

\* \* \*

### Pulse photothermoacoustics of small-volume liquid probes

EGEREV S.

Laser sound generation in small-volume liquid probes is considered under the conditions of small laser fluence thus giving rise to the linear thermoacoustic conversion. The results obtained are of great demand by analytical chemistry. The outgoing acoustic signal has some peculiarities if the probe layer has thickness within 10–450 μm. With a laser pulse width of about 20 ns the absorbing probe layer exhibits acoustically thin or acoustically thick properties. This depends on a specific task. Physics of the photoacoustic conversion in the small-volume probes is revealed for both acoustically thin and acoustically thick cases under the general conditions of thermally and optically thin probes. We developed an optoacoustic measurement cell of a layered-prism (LP) optimized type. Using this LP cell we

provided precise concentration measurements for a model system.

\* \* \*

### Ultrasonic absorption spectra of paramagnetic 1-alkyl-3-methylimidazolium ionic liquids: [EMIm]<sub>2</sub>[Co(NCS)<sub>4</sub>] and [BMIm]<sub>2</sub>[Co(NCS)<sub>4</sub>]

GEPPERT-RYBCZYŃSKA M., DZIDA M., ZOREBSKI E., ZOREBSKI M., CHORAŻEWSKI M., ŻARSKA M., CZECH B., HEINTZ A., PEPPEL T., KÖCKERLING M.

The ultrasound absorption measurements were performed in two paramagnetic 1-alkyl-3-methylimidazolium ionic liquids, i.e., [EMIm]<sub>2</sub>[Co(NCS)<sub>4</sub>] and [BMIm]<sub>2</sub>[Co(NCS)<sub>4</sub>] within the frequency range from 10 to 300 MHz at the temperatures 293.15 and 298.15 K and at atmospheric pressure. The measurements were made by the use of the measuring set operates on the standard pulse technique with a variable path length. The dependence of the quotient  $\alpha \cdot f^{-2}$  on frequency changes strongly with temperature. It appears also that within the investigated frequency range the quotient  $\alpha \cdot f^{-2}$  is clearly dependent on frequency as early as above 10 MHz both in the case of [EMIm]<sub>2</sub>[Co(NCS)<sub>4</sub>] and [BMIm]<sub>2</sub>[Co(NCS)<sub>4</sub>]. Thus, in both cases the dispersion characteristics  $d\alpha \cdot f^{-2}/df < 0$  is observed. Moreover, the both liquids are rather highly absorbing, i.e., the frequency normalized attenuation (at  $T = 298.15$  K and  $f = 100$  MHz) are  $837 \cdot 10^{-15}$  and  $1321 \cdot 10^{-15}$  s<sup>2</sup>·m<sup>-1</sup>, respectively. Above some frequency, the experimental values of  $\alpha \cdot f^{-2}$  are smaller than those predicted by the Navier-Stokes relation. The absorption curves indicate that  $\alpha > \alpha_{cl}$  at lower frequencies and  $\alpha < \alpha_{cl}$  at higher frequencies, and that the frequency for which  $\alpha = \alpha_{cl}$  decreases with the increasing temperature. Most probably this kind of behaviour would result from a relaxation mechanism of the viscous type. Similar behaviour has been reported previously for 1-dodecanol and castor oil.

\* \* \*

### Acoustic properties of magnetosomes

JÓZEF CZAK A., HASHIM A., MOLCAN M., HORNOWSKI T., SKUMIEL A., KOPČANSKÝ P., TIMKO M.

The objective of the work is to study the biological magnetic nanoparticles (magnetosomes) as a product of the biomineralization process of magnetotactic bacteria *Magnetospirillum* sp. AMB-1. This paper presents an ultrasound method based on the measurements of compression wave velocities and determination of the phase velocity of a wave. The obtained data allow the determination of the mechanical characteristics of the suspension. The study of elastic properties shows, that bulk modulus of a bacterial magnetosomes suspension increased with the increase of temperature like in chemically synthesized magnetite nanoparticles suspension

This work was supported by a Polish National Science Centre grant, no DEC-2011/03/B/ST7/00194 and by Slovak Academy of Sciences, in the framework of Centre of Excellence NANOFLUID, projects VEGA No. 2/0043/2012, APVV 0171-10 and Ministry of Education

Agency for structural funds of EU in frame of projects Nos. 26220220061 and 26220120033.

\* \* \*

### **Non-Critical Fluctuations of Liquids – Cinderella of Ultrasonic Spectroscopy?**

KAATZE U.

In basic research ultrasonic broadband spectroscopy of liquids is mainly focused on chemical relaxations and critical fluctuations in local concentrations. Chemical relaxations follow Debye-type relaxation behaviour. In contrast, near a critical demixing point binary and ternary liquid mixtures display spectra with broad distribution of relaxation times. In applications small-band spectra related to the scattering of ultrasonic waves from suspensions and emulsions are often used.

In this paper attention is directed to a less popular field in the broadband ultrasonic spectroscopy of liquids. Many liquid systems reveal spectra which are only slightly broader than a Debye relaxation term. They can be favourably represented by relaxation functions following from models of non-critical fluctuations in the local concentrations. The so-called unifying model of non-critical concentration fluctuations is briefly presented and some implications for our understanding of aqueous solutions are discussed which follow from the description of relevant ultrasonic spectra in terms of this model. Also considered in the light of non-critical fluctuations are surfactant solutions close to the critical micelle concentration. An alternative model for the theoretical representation of solutions of pre-micellar and improper micellar aggregates is indicated.

\* \* \*

### **On the correlation between thermal and electric properities of C-SiC ceramic composites**

KAŹMIERCZAK-BALATA A., MAZUR J.,  
DREWNIAK Ł., BODZENTA J.

The carbon silicon carbide (C-SiC) ceramic composites have an outstanding thermomechanical and thermochemical properties, required for lightweight constructions and future engine components. The carbon/carbon silicon carbide (C/C-SiC) composites are high ceramic friction materials used for high speed and high energy braking. The nanostructures of SiC are applied as a basic material for high temperature electronic devices. It was found that pure material is stable up to 2000°C and electronic devices based on SiC perform in temperatures over 200°C.

In this work seven samples of C-SiC composites were examined. A few of them were covered with Au layer and others contained epitaxially growth graphene layer. The samples were produced by sintering method. The thermal diffusivity was determined by continuous wave photothermal technique. The front sample surface was illuminated by the intensity modulated light and the temperature disturbance caused by absorption of this light was detected by infrared (IR) radiometry from the rear surface of the sample.

\* \* \*

### **Volumetric properties of electrolytes in water – 2-methoxyethanol mixed solvent**

KLIMASZEWSKI K., BALD A.

Density of KCl, KBr, NaCl and NaBr solutions in mixtures of water with 2-methoxyethanol has been measured at 298.15 K. The values of the apparent molar volumes of tested electrolytes were calculated on the basis of density values. Data from the measurements of the electrical conductivity indicated incomplete dissociation of electrolytes in mixtures with a high alkoxyethanol content which is related to the low permittivity of such mixtures. Therefore, in order to determine the value of the partial molar volume, was used an appropriately modified Redlich-Meyer equation. This equation takes into account the presence of in solution: free ions and undissociated forms of electrolyte (ion pairs) and contains three parameters which should have been determined. Two of these parameters, i.e. the partial molar volume of the electrolyte in of a completely dissociated form and the partial volume of the ion-pairs, are very important for description of the electrolyte-solvent interactions.

\* \* \*

### **Sensitivity of phase transition of ferronematics in combined electric and magnetic fields**

KOPCANSKY P., TOMASOVICOVA N., TIMKO M.,  
MAJOROSOVA J., ZAVISOVA V., KONERACKA M.,  
TOMCO L., MITROVA Z., JADZYN J., CHAUD X.

Ferronematics are colloidal suspensions of nematic liquid crystals and magnetic nanoparticles. We have studied the combined influence of electric and magnetic fields on the orientational structure of ferronematics based on a thermotropic nematic 4-trans-4'-n-hexyl-cyclohexyl-isothiocyanato-benzene (6CHBT). The 6CHBT liquid crystal have been dissolved in in phenyl isothiocyanate and doped with the rod-like or chain-like magnetic particles. In such mixture, the phase transition from isotropic to nematic phase is via droplet state, i.e. coexistence of nematic and isotropic phase. The obtained results showed that combination of the electric and magnetic fields can change the character of phase transition from isotropic to nematic phase via droplet state in such systems. Moreover, the magneto-dielectric measurements of structural transitions showed the magnetic field induced shift of phase transition temperature from isotropic to droplet state. All results have direct influence on the construction of liquid crystalline sensors of magnetic field.

\* \* \*

### **Characterization of pure and modified TiO<sub>2</sub> layer on glass and aluminum support by beam deflection spectrometry**

KORTE D., PAVLICA E., BRATINA G., FRANKO M.

TiO<sub>2</sub> thin films used as photocatalysts in environmental application were studied by beam deflection spectroscopy (BDS) coupled to multiparameter fitting of a novel theoretical model to the experimental data, as well as by AFM measurements. Two groups of films were prepared: pure and N-doped TiO<sub>2</sub> deposited on glass support, as well as

pure, N- and C-doped TiO<sub>2</sub> deposited on aluminum support. The results show a correlation between the thermal parameters of the examined films and their transport properties such as value of energy band gap, carrier life time or concentration and kind of introduced dopants. Furthermore, the material's thermal conductivity and thermal diffusivity depend on the porosity and the surface roughness of the material, which usually change when the material is used as photocatalysts in water purification processes. Thermal and mechanical properties of the thin film photocatalysts also depend on the support to which the TiO<sub>2</sub> layer is deposited. All of these features determine the performance of the TiO<sub>2</sub> films in their photocatalytic application and can be deduced from BDS measurements. This work confirms the suitability of BDS as a method for non-contact and nondestructive evaluation of thin film TiO<sub>2</sub> photocatalysts, which is also the basis for improving their efficiency.

\* \* \*

#### **Diffusion beyond the Stokes-Einstein theory: effects of inhomogeneity and nonlinear friction**

LISÝ V., TÓTHOVÁ J., GLOD L.

The Langevin equation describing the Brownian motion of particles is studied in the case when the friction is modeled not by the usually used Stokes force but nonlinearly depends on the particle velocity. The intensity  $D(v)$  of the thermal random force driving the particles then also depends on the velocity, and the equation of motion should contain an additional “spurious” force proportional to the derivative of  $D(v)$ . This force is chosen in the kinetic representation, for which the stationary probability density for velocities is the Maxwell distribution and simultaneously the generalized Einstein relation is obeyed. A general formula for the diffusion coefficient of the particle is obtained and then specified for various models of dissipative forces studied in the literature, e.g., those for active Brownian motion. We also discuss the effects of inhomogeneity in the diffusion and a choice of the appropriate stochastic calculus. The correspondence between the kinetic representation for  $D(v)$  and the Stratonovich concept in the description of spatially dependent diffusion has been established.

\* \* \*

#### **Optimization of a thermal lens microscope for detection in a microfluidic chip**

LIU M., NOVAK U., PLAZL I., FRANKO M.

The optical configuration of a thermal lens microscope (TLM) was optimized for the detection in a microfluidic chip. Influence of the flow velocity on the TL signal was analyzed theoretically and experimentally. Optimizations regarding the pump and probe beam parameters (beam waists, offsets, mode mismatching) were made. It was found that an appropriate pump-probe beam offset for certain flow velocity will provide not only a higher sensitivity but also a better response linearity of TLM over three orders of magnitude of concentration. Diffraction-limited pump beam excitation is advantageous for space-resolved measurement, while a larger pump beam with ten times lower power density is favorable for higher sensitivity. As an

application, this optimized TLM was used to study the diffusion of ferroin between two aqueous streams in a microreactor chip. The diffusion profiles at different distances from the mixing point were recorded by scanning the TL signal along the cross-section of the microchannel. High sensitivity of TLM enabled accurate detection of small changes in absorbance ( $7 \times 10^{-6}$  AU in a 50  $\mu\text{m}$  channel) which provides a powerful tool for depicting the governing transport characteristics in a microfluidic chip by incorporating convection and diffusion terms along with the parabolic velocity profile.

\* \* \*

#### **Photoacoustic Transformation at Opposite-Interaction of Electromagnetic Waves in Magnetoactive Two-Layer**

MITYURICH G., SVIRIDOVA V., SERDYUKOV A.

The paper considers thermo-optical generation of sound in magnetoactive two-layered medium in the conditions of the tunnel electromagnetic interference of electromagnetic waves. It studies the conditions of complete suppression of the amplitude of photoacoustic signal depending on the type of polarization of opposite-interacting waves, the difference of their initial phases, the values of intensity of either opposite-interacting beam. The given paper provides an effective method of management the thermo-optical generation of sound in magnetoactive structural element in the conditions of the tunnel electromagnetic interference of electromagnetic waves.

\* \* \*

#### **Study of molecular packing properties of dense simple liquids based on acoustic and thermophysical data**

POSTNIKOV E., GONCHAROV A., MELENT'EV V.

We analyze the inverse reduced fluctuations: a form of ratio for speeds of sound for the substance and the model ideal gas of the same molecules at the same density and temperature. We have shown that there exists a relatively wide (up to 25% of reduced density range for each phase) region, where the dependence of relative fluctuation on density has strong exponential character along a liquid-vapour coexistence curve of simple liquids within the region corresponding to dense liquids. These results support the assumption that the structure of a uniform cell packing for substances' molecules and the the model of hole theory of liquids. This conclusion is confirmed by direct numerical calculation using lattice fluid model with mean-field quadratic potential energy of attractive interactions.

\* \* \*

#### **The use of multifrequency induction heating for temperature distribution control**

SMALCERZ A., GOLAK S., PRZYLUCKI R.

The paper presents possibilities of controlling temperature field distribution in induction heated charge. The change of the distribution was obtained with use of sequential one-, two-, and three-frequency heating. The study

was conducted as a multi-variant computer simulations of strongly coupled to each other fields: the electromagnetic and the temperature fields. For the analysis, the professional calculation package using the finite element method Flux 3D was used.

The problem of obtaining an appropriate temperature distribution in a heated charge of, a complex shape is very important in many practical applications. A typical example is a quenching of gears. For such applications, it is required to obtain, a surface and in desired depth, uniform temperature distribution on the tooth face and top land and on the bottom land of the gear. The obtained temperature should have proper distribution and value. Achieving such a defined distribution is very difficult. During the study more than 50 different calculation variants were examined.

\* \* \*

### Crack depth evaluation by infrared lock-in thermography

STREZA M., DADARLAT D., LONGUEMART S.

This paper describes an approach to evaluate linear opened surface cracks by means of lock-in thermography. The presence of a defect or a discontinuity close to a heated region causes anomalies of heat flow distribution. Highlighting these disturbances provides information on position and depth of discontinuities.

For describing and interpreting the interactions between a modulated laser induced heating spot and a crack, a 3D heat flow simulation model has been developed by using a FEM computer package. The model considers an aluminium metallic block with a linear open crack of length  $L$ , width  $l$  and depth  $h$ , the circular heat source of power  $P$  being located at a distance  $d$  from the crack. Amplitude and phase of the temperature modulation of the irradiated surface, at the excitation frequency  $f$ , have been extracted by computing a temporal discrete Fourier for each pixel of the temperature image. In order to enhance the thermal contrast the proposed procedure exploits the second spatial derivative of the amplitude image. The minimum of this quantity provides information concerning the crack geometry. Information regarding the depth of the crack can be revealed by changing thermal diffusion length of thermal wave in the vicinity of the crack.

Aluminium test specimens with linear open cracks have been investigated and the results are compared with FEM simulation. This approach allows a classification of the cracks according to their depths.

\* \* \*

### Heating Characteristics of Magnetic Fluids based on various carrier liquid. Concentrations

TIMKO M., MOLCAN M., RAJNAK M., KOVAC J., KOPCANSKY P., SKUMIEL A., JOZEFCAK A., MAJOROSOVA J., ZAVISOVA V., KONERACKA M., TOMCO L., MITROVA Z., JADZYN J., CHAUD X.

The heating ability of magnetic fluids based on various carrier liquid with different magnetic particle concentrations was studied. The calorimetric measurements were carried out in an alternating magnetic field up to  $2500 \text{ A}\cdot\text{m}^{-1}$

amplitude and of 500 kHz frequency. The obtained law-type dependence of the temperature increase rate,  $(dT/dt)t = 0$ , on the amplitude of the magnetic field indicates the presence of superparamagnetic particles in the tested samples since  $n = 2$ . The specific absorption rate (SAR) defined as the rate of energy absorption per unit mass increases with an increase of the volume fraction of the dispersed nanoparticles and with increasing of density of carrier liquid.

\* \* \*

### Theoretical analysis of the influence of dopants on the temperature dependence of thermal diffusivity of single crystals

TREFON-RADZIEJEWSKA D., BODZENTA J.

The thermal diffusivity dependence on temperature of YAG, YVO<sub>4</sub> and GdCOB single crystals was determined. Samples were doped with different concentrations of rare earth ions such as ytterbium, neodymium, thulium, and transition metal ions such vanadium. Determination of the thermal diffusivity was based on an analysis of propagation of thermal wave in the sample.

Obtained results confirmed, that the influence of doping on the thermal diffusivity of investigated materials strongly depends on temperature.

The thermal diffusivity of all investigated samples decreases with increasing of sample temperature from 30°C to 300°C, but the drop in the thermal diffusivity is the highest for pure single crystals. The introduction of dopant ions into a crystal lattice leads to a significant reduction in the thermal diffusivity at lower temperatures in comparison with pure crystals. However, the influence of dopants becomes less pronounced with increasing temperature, and in the case of weakly doped crystals it becomes negligible at higher temperatures.

The interpretation of obtained thermal diffusivity dependence on temperature for pure and doped single crystals was based on the Debye model of lattice thermal conductivity of solids. It was assumed that the decrease in thermal diffusivity with temperature and increasing concentration of impurities is caused by shortening of the phonons mean free path due to phonon-phonon and phonon-point defect scatterings.

\* \* \*

### Peculiarities of hydrogen bond formation in solutions of proton acceptors in ethylene glycol: calorimetric study

VARFOLOMEEV M., ZAITSEVA K., SOLOMONOV B.

Hydrogen bonds significantly influence on physical-chemical properties of organic compounds in the liquid state. They can be formed between different molecules or between the fragments of one molecule. Thousands of works were devoted to the determination of the thermodynamic functions of hydrogen bonding. Based on experimental data different empirical relationships were obtained. They allow us to calculate the free energy and enthalpy of hydrogen bonding between molecules. However, it should be noted that all these empirical relationships can be applied only to the calculation of the thermodynamic functions of for-

mation of bimolecular H-complexes in an inert solvent (usually carbon tetrachloride), or in an environment of “pure base”. Apparently they couldn’t be applicable to the analysis of solute-solvent hydrogen bonds in self-associated liquids, such as water, aliphatic alcohols, formamide, glycols, etc., the use of which is of great practical and theoretical importance. This is due to the fact that the formation of solute-solvent hydrogen bonds in self-associated liquids differs from the process of the formation of H-complexes of 1 : 1 in an inert solvent. The presence of a very small number of experimental data on the thermodynamic functions of hydrogen bonding in self-associated solvents due to the limitations of traditional method of H-bond study IR and NMR spectroscopy.

In present work we have studied hydrogen bonding of proton acceptors in solutions of ethylene glycol by the solution calorimetry method. Enthalpies of solution at infinite dilution of nitriles, ketones, amines, ethers and esters in ethylene glycol were measured ( $T = 298.15$  K). Based on the experimental data enthalpies of solvation of proton acceptors were determined. Obtained values were compared with enthalpies of solvation of studied molecules in methanol and water. Enthalpies of hydrogen bonding of proton acceptors in ethylene glycol solutions were first time determined. It was shown that for weak proton acceptors enthalpies of hydrogen bonding in ethylene glycol can be positive (endothermic process). Reasons of such behavior were discussed.

\* \* \*

#### **Investigations of subsurface damaged layer by piezoelectric photothermal spectroscopy**

ZAKRZEWSKI J., MALIŃSKI M., STRZAŁKOWSKI K., FIRSZT F.

Photothermal spectroscopy has found the wide range of application as a method of monitoring nonradiative recombination of excited carriers generated by optical absorption in semiconductors. One of them is surface quality monitoring after different procedure of sample preparation. Mechanical, chemical and thermal treatment of surface can create the damaged subsurface layer which introduces additional defects states in the material. Lately, it was shown that piezoelectric spectroscopy can be a useful tool to monitor the state of surface. In this method, the stress and

strain of a sample due to the absorption of electromagnetic radiation is detected by a piezoelectric transducer.

The amplitude and phase of piezoelectric spectra are necessary for the proper interpretations, however in some cases, they do not give the clear results. To minimize uncertainty, the new procedure of measurements is proposed. It involves the detection of the signal in front and rear configurations and measurements with illumination at different surfaces of the samples. Four pairs of amplitude and phase spectra are obtained which are interpreted to receive the consistent optical and thermal parameters. The model based on Blonskij’s one was chosen to calculate the thermal and optical parameters of investigated material.

\* \* \*

#### **High pressure physicochemical properties of biodiesel components**

ŻARSKA M., DZIDA M.

The main components of biodiesel derived from coconut oil or babassu oil such as ethyl caprylate, ethyl caprate, ethyl laurate, and ethyl myristate are studied. The speeds of sound were measured within the temperatures from 293 to 318 K and at pressures from 0.1 to 101 MPa. The densities and heat capacities were measured under atmospheric pressure in the temperature range from 283 to 353 K and 286 to 341 K, respectively. The densities, heat capacities, isentropic and isothermal compressibilities, thermal expansions, and internal pressures as functions of temperature and pressure were calculated using experimental results. The densities of ethyl esters decrease with increasing their molecular weight below intersection temperature while at higher temperatures densities increase with increasing molecular weight. The intersection temperature probably moves toward higher temperatures with increasing pressure. For a given temperature, differences between densities of the ethyl esters increases with increasing pressure. With increasing molecular weight the effect of temperature on thermal expansivity decreases with pressure. It is approximately independent of temperature for ethyl laurate and ethyl myristate at pressures higher than 80 and 60 MPa, respectively. The pressure and temperature dependence of internal pressure of the esters is the most similar to that of alkanes.

## 7 Integrated Optics – Sensors, Sensing Structures and Methods

### Szczyrk, Poland, February 25 – March 1, 2013

#### Abstracts

##### **Application of Photonic Crystal Fiber in optical fiber current sensors**

BARCZAK K.

Optical fibers may be applied in the technique of measuring electric current., particularly as so-called optical current transducers (OCT). They are small in size, cheap, light and safe. Their sensitivity, however, connected with the Faraday effect, is rather poor, and they are also susceptible to considerable disturbances (deformations of the fibers). Their sensitivity can be increased by lengthening the path of effect, but this involves also an increase of noise and greater fluctuations of the measured signal due to changes of internal stresses induced by external factors. These negative phenomena can be reduced by applying fibers less sensitive to deformations [1, 2]. The author supposes that photonic crystal fiber (PCF) with a glass core may display such properties. Orifices in the cladding can to some extent prevent the propagation of deformations from outside the core, and thus decrease the birefringence of the fiber induced by elastooptic effects. The papers presents preliminary measurements of typical photonic crystal fiber exposed to mechanical vibrations.

\* \* \*

##### **Optical and electronic properties of thin boron-doped diamond films grown by pacvd: boron level studies**

BOGDANOWICZ R., FICEK M., ŚMIETANA M., JASIŃSKI J., GNYBA M., RYL J., SOBASZEK M., GOLUŃSKI Ł.

Significant attention of researchers in the world is focused on so-called “non-active” anodes such as boron doped diamond (BDD) electrodes. They could be used effectively in electrochemical sensors designed for environment monitoring and biomedical applications. Due to these advantages authors are going to use the BDD as a sensing film in optical sensors, including fibre optic sensors. The preliminary studies of electronic and optical properties has been performed to determine the best parameters of CVD synthesis and film for optical purposes. The main parameter of the BDD films, boron level, was particularly investigated. The thin film microcrystalline diamond have been deposited in a MW PE CVD process on silicon wafers and glass substrates. Surface resistivity of BDD electrode was studied using four-point probe measurements. The molecular structure of BDD films as well as the sp<sup>3</sup>/sp<sup>2</sup> phase ratio were determined by Raman spectroscopy. The scanning electron microscope was utilized to investigate the film morphology. The surface analysis made by SEM showed a small variation of surface morphology for boron-doped films. Studies of optical constants, thickness and optical

energy band gap were performed using spectroscopic ellipsometry in the wavelength range from 260 to 820 nm. The  $\Delta$  and  $\Psi$  were fitted with a double-layer Tauc-Lorentz (TL) dispersion model.

\* \* \*

##### **Silicon carbide on silicon photodiode stacks for sensing applications in the ultraviolet to visible range**

BORECKI M., KOCIUBINSKI A.,  
KORWIN-PAWLOWSKI M. L., DUK M., SZMIDT J.

The capabilities of optoelectronic sensors can be extended by use of spectral data processing. A popular range of wavelengths that are used is from UV at 260 nm, where DNA absorbs radiation, to the visible, where the results can be directly observed by the human eye. This wide range can be examined with a spectrophotometer, but in many situations the use of that kind of instrument is not desirable because of its cost and complexity. On the other hand, sensing of a sample using one beam with different wavelengths at the same time is required when the sample is in movement or in a dynamical state, as happens in analyzing the sedimentation of a mixture. Using LEDs as light sources is advantageous because of their low cost and ease of electronic modulation which makes the sensor more immune to background radiation. The LEDs radiation can be coupled to beam with a fiber optic bundle. There are a few ways of constructing the device that allows in one optical axis simultaneous detection of a UV and VIS radiation beam. First is the use of commercially available components, such as Si and Si UV enhanced photodiodes with dichroic mirrors. Second is the use of hybrid technology to build photodiode matrices that use SiC and Si components or Si UV enhanced components and matrices of optical filters. The third possibility is the use of stack of detectors using one transparent-to-VIS radiation element that absorbs and detects UV radiation [1]. For such a construction we made SiC photodiodes which were transparent-to-VIS radiation and had an active area of 10 mm<sup>2</sup>. As the back Si photodiode we used a large area commercial photodiode – BPYP 44A that is sensitive from 400 to 1100 nm. Both photodiodes were mounted in the center of 1" printed board rings fitting into a SM1 tube from THORLABS. This enables of use of standard THORLABS micromechanical components for the construction of detectors and for positioning and coupling of optical fibers. We characterized the fabricated SiC photodiodes, which showed a sensitivity of 0.2 A/W in the range from 275 nm to 375 nm and a transparency for VIS radiation almost flat and of 20%. These structures can be connected to dual channel optical power meters or to dedicated optoelectronic interfaces.

##### Acknowledgments

This work was made with the support of NCBIR grant No. N R02 0008 06/2010 2 “New optoelectronic devices for

intelligent classification of organic and biological liquids” [1] M. Borecki, M. Bełłowska, K. Kopczyński, Z. Mierczyk, J. Szmidt, “Multilayer semiconductor photodetector”, Polish Patent application # Z-397020 of November 17, 2011.

\* \* \*

### **Influence of the glass matrix of tellurite glasses on the spectroscopic properties**

BURTAN B., REBEN M.

The Pr<sup>3+</sup> doped oxyfluoride glass-ceramic containing PbF<sub>2</sub> nanocrystals has been presented. Transparent glass ceramic was obtained by heat treating the glass from the SiO<sub>2</sub>-PbO-Na<sub>2</sub>O-PbF<sub>2</sub> system at the first crystallization temperature. The thermal conditions of PbF<sub>2</sub> nanocrystallization in the oxyfluoride glass have been presented. Ceramization of glass was studied by DTA/DSC, XRD and TEM methods. X-ray diffraction analysis of the transparent glass-ceramic revealed that the PbF<sub>2</sub> nanocrystals are precipitated in the glass matrix. It has been found that nanocrystallization of PbF<sub>2</sub> strongly depends on the ratio between the components and amount of PbF<sub>2</sub>. Formation of the PbF<sub>2</sub> nanocrystals have been obtained as a low phonon host for rare earth active ions in oxide glassy matrix. The optical properties of glass were determined by UV-VIS spectroscopy in transmission which was carried out in order to assess the absorption spectra of the rare earth doped glasses. The spectral dependence of ellipsometric angles of the oxyfluoride glass samples, have been studied. The influence of nanocrystallization of PbF<sub>2</sub> onto changes of refractive index of oxyfluoride glass were examined. The optical measurements were conducted on Woolam M2000 spectroscopic ellipsometer, in spectral range of 190–1700 nm.

\* \* \*

### **Polarimetric optical fiber sensors for dynamic strain measurements in composite materials**

DOMAŃSKI A. W., BIEDA M., LESIAK P., SIERAKOWSKI M., MAKOWSKI P., POCZEŚNY T., PROKOPCZUK K., SOBOTKA P., WOLIŃSKI T. R.

Optical fiber Bragg grating sensors are the most popular sensing systems for strain monitoring including also composite materials. However, there are some disadvantages in their applications due to limited points of measurements, troubles with temperature desensitization and relatively high cost of detecting systems. In comparison, polarimetric optical fiber sensors possess possibility of temperature compensation, dynamic and integral system of strain monitoring as well as low cost of photo-detecting elements. In the paper we present results of an experimental analysis with different kinds of birefringent optical fibers leading to an optimal set-up for dynamic strain monitoring in composite materials.

#### **Acknowledgment**

This work is supported by the NCBiR grant no. PBS1/B5/20/2012

\* \* \*

### **Temperature cross-sensitivity for highly refractive index sensitive nanocoated Long-Period Gratings**

GROCHOWSKI J., MYŚLIWIEC M., MIKULIC P., BOCK W. J., ŚMIETANA M.

There is a demand for highly sensitive refractive index (RI) devices simultaneously insensitive to temperature changes, especially when applied for biosensing purposes. We investigated here temperature sensitivity of Long-Period Gratings (LPGs) coated with 100 nm-thin silicon nitride (SiN<sub>x</sub>) film. The LPGs with period of 500 μm were induced in standard Corning SMF28 fibre using electric arc and then nanocoated with SiN<sub>x</sub> by Radio Frequency Plasma Enhanced Chemical Vapour Deposition (RF PECVD) method. As a sensorial effect we investigated resonance shift with temperature variations (range from 20°C to 70°C) for selected external media, i.e. water (n<sub>D</sub> = 1.3281) and glycerine (n<sub>D</sub> = 1.4552). For peak observed at 1510 nm, the RI sensitivity of 3147 nm/RIU and temperature sensitivity of 0.22 nm/°C (H<sub>2</sub>O) were obtained which give temperature-RI sensitivity ratio (T/RI) of 6.99e-5 RIU/°C. For peak observed at 1560 nm, when the LPG was immersed in glycerine, the sensitivities of 857 nm/RIU and 0.19 nm/°C (glycerine) were observed which leads to T/RI ratio of 2.24e-4 RIU/°C. The influence of the nanocoating is discussed and compared to the results obtained for bare LPGs.

\* \* \*

### **Modal birefringence measurements of the planar optical waveguide**

GUT K.

The modal birefringence can be determined basing on the interference of scattered light may be utilized. By adjusting the polarization input beam appropriately, it was possible to excite the modes TE<sub>i</sub> and TM<sub>i</sub> simultaneously. A definite pair of modes was selected by changing the angle of the beam input to the prism. Placing an additional polarizer in front of the camera we can observe oscillations of the intensity of light (in the direction of its propagation) connected with the interference of scattered light from the modes TE and TM. Determining the distance between the first and the last distinct maximum (minimum) and the number of oscillations between them we determine the beat length and can calculate the modal birefringence of the investigate waveguide.

\* \* \*

### **Multispectral systems of imaging scenery in critical infrastructure protection, threat identification and recognition**

KAROL M., SZUSTAKOWSKI M., ZYCKOWSKI M., MARKOWSKI P., CIURAPINSKI W.

The surveillance system more often use devices working on different bands that are used not only for detection but also to recognize and identify threats. The paper presents the possibility of protection, detection and identification of risks, achieved through the use of multispectral detection systems in critical infrastructure facilities. The authors consider the benefits of active sensors based on other wavebands, such as millimeter-wave radar, terahertz cameras,

infrared camera and shows the ability to detect and identify the target using data collected by these sensors. Also discussed the issue of fusion data from different sensors and opportunities that the whole system obtained by application of data fusion.

\* \* \*

### **Hyperspectral Imaging Infrared Sensor Used For environmental monitoring**

KASTEK M., ŻYCKOWSKI M., PIĄTKOWSKI T.,  
DULSKI R.

The paper presents the detection of gases using an infrared imaging Fourier-transform spectrometer (IFTS). The Telops company has developed the IFTS instrument HyperCam, which is offered as short or long wave infrared device. The principle of HyperCam operation and methodology of gases detection has been shown in the paper, as well as theoretical evaluation of gases detection possibility. The calculations of optical path between IFTS device, cloud of gases and background have been also discussed. The variation of a signal reaching the IFTS caused by the presence of a gas has been calculated and compared with the reference signal obtained without the presence of a gas in IFTS's field of view. Verification of theoretical result has been made by laboratory measurements. Some result of the detection of various types of gases has been also included in the paper.

Keywords: gas detection, hyperspectral detection, imaging Fourier-transform spectrometer, stand-off detection.

\* \* \*

### **Effect of temperature on upconversion luminescence in Yb3+/Tb3+ co-doped germanate glass**

KOCHANOWICZ M., DOROSZ D., ŻMOJDA J.,  
MILUSKI P., DOROSZ J.

In the article effect of temperature on the cooperative energy transfer in germanate glass doped with Yb3+/Tb3+ under 976 nm laser diode pumping was investigated. The optimization of Tb3+ concentration on the upconversion luminescence was determined. Measured strong luminescence at 489, 543, 586, 621 corresponding to 5D4→7FJ ( $J = 6, 4, 3$ ) transitions and luminescence at 381, 415, 435 nm resulting from 5D3, 5G6→7FJ ( $J = 6, 5, 4$ ) transitions. The highest upconversion emission intensity was obtained in the germanate glass doped with 0.7 Yb2O3/0.7 Tb2O3. The effect of temperature on the luminescent properties of germanate glass in the range of 5–250°C indicates the presence of competing phenomena: an increase in the effective absorption cross-section of Yb3+ ions donor as a function of temperature and migration of energy between pairs of ions Yb3+ – Yb3+ and of multiphonon excitation levels 7FJ.

\* \* \*

### **Fabrication and Characterization of 4H-SiC Light Emitting Diode**

KOCIUBIŃSKI A., DUK M., KWIETNIEWSKI N.,  
BORECKI M.

Silicon carbide is a suitable semiconductor to manufacture LEDs which emit in the blue-violet part of spectrum.

In this paper, we discuss 4H-SiC LED fabrication and characterization. We used implantation technique to obtain p-n junction in the epitaxy layer. Ohmic contacts were deposited on the epilayer structure using evaporation, etching and liftoff. Various characteristics of the devices were measured including current vs. voltage, contact resistance, and output spectra.

\* \* \*

### **The concept of an optical sensor system for the temporary blood chamber volume measurements in the POLVAD-EXT**

KONIECZNY G., PUSTELNY T.

The presentation includes the preliminary research of the newly developed temporary blood volume measurements optical system for using in the in the extracorporeal Polish Ventricular Assist Device (POLVAD-EXT). The proposed solution is basing on the measurement of the light reflected from the membrane in different configurations of light emitters and light detectors analyzed using the principal component analysis (PCA) method for estimating the blood volume. The measurement technique used in the experiment is described, measurement circuit and preliminary static measurements results of the developed measurement system are included and preliminary analyses of the measurements are shown. Future development plans are stated.

\* \* \*

### **Long period gratings usage possibility on planar gradient waveguides**

KOTYCZKA T., ROGOZIŃSKI R.

In this analysis the possibility of long period gratings usage as visible range spectrum filters on planar gradient waveguides (LPWG) was presented. The influence of technological processes as well as gratings parameters on received resonance profiles and spectrum characteristics was shown. These calculations was based on the real material properties (BK7 glass doped with potassium ions) as well as specific technological processes.

\* \* \*

### **Hidden object detection system based on fusion of THz and VIS images**

KOWALSKI M., PALKA N., PISZCZEK M.,  
SZUSTAKOWSKI M.

Image fusion can be used in wide range of security applications like detection and identification of hidden objects. This is a very urgent and demanding problem at the present time. Finding objects hidden under clothes is one of key issues of public places security. The aim of image fusion is to combine information from multiple images of the same scene. The result of the fusion is a new image which gives more information and is more suitable for human perception. Image fusion can be used for further image-processing tasks such as segmentation, feature extraction and object recognition. Our goal is to build a system for harmless humans screening and detection of hidden objects. Visual detection of hidden objects can be achieved using various imaging methods but some of these methods are harmful

for humans. The THz radiation band can be applied for screening humans because this radiation does not point any harmful ionizing effects so it is safe for human beings. In the paper we demonstrate results of various fusion methods applied for THz and VIS images. The research is focused on dangerous objects detection – guns, knives and bombs hidden under some popular types of clothing.

\* \* \*

### Measurement of longitudinal strain sensitivity in highly birefringent polymer microstructured optical fibers

KUCZKOWSKI M., LESIAK P., SZELAĞ M., BIEDA M., ERTMAN S., BUDASZEWSKI D., DOMAŃSKI A., WOLIŃSKI T.

In this work, we present results of our research on highly birefringent polymer fibers based on photonic crystal fiber. These new fibers are characterized by a very low Young modulus in comparison to standard silica fibers. Our research was concentrated to measure the value of the longitudinal strain sensitivity in four polymer fibers made by Kiriama Pty Ltd.

By applying strain we can observe periodic changes (sine like) in light polarization described by the Stokes vector. The amplitude of the signal depends on the azimuth of the input polarization introduced to the fiber and the period of the function depends on the fiber birefringence. In this paper, a comparison between strain sensitivity in different highly birefringent polymer fibers has been presented.

\* \* \*

### All-optical steering of self-traped beams in liquid crystalline planar waveguide

LAUDYN U. A., KWAŚNY M., SALA F. A., KARPIERZ M. A.

Optical reorientational nonlinearity in nematic liquid crystals causes that spatial solitary waves (nematicons) are created for a light power of a few tenths of milliwatts. The low absorption allows to observe nematicons propagation distance over lengths in a few millimeter range. Additionally, it was demonstrated that a weak signal beam can be guided in a channel formed by a nematicon. We report on the experimental studies the possibility of steering of such nematicons by another beams in chiral structure being the array of planar waveguides. Because of a low power necessary for steering, such a configuration can be applied in all-optical switching and routing elements for integrated optical systems, where electric fields cannot be used.

\* \* \*

### Measurement of temperature sensitivity in highly birefringent polymer microstructured optical fibers

LESIAK P., SZELAĞ M., KUCZKOWSKI M., ERTMAN S., BUDASZEWSKI D., DOMAŃSKI A., WOLIŃSKI T.

In this work, we present results of our research on highly birefringent polymer fibers based on photonic crystal fiber. Our research was concentrated to measure the value of the temperature sensitivity in four polymer fibers made by Kiriama Pty Ltd.

Typical highly birefringent photonics crystal fibers are very low sensitive to temperature changes. Fibers based on poly(methyl methacrylate) are characterized by high thermal sensitivity. These value is somewhat larger than those for silica fiber and are consistent with the values expected on the basis of the bulk polymer properties. In this paper, a comparison between sensitivity in different highly birefringent polymer and standard fibers has been presented.

\* \* \*

### Balance of polarization in a hybrid fiber optic sensor

MARKOWSKI P., SZUSTAKOWSKI M., ŻYCKOWSKI M., KAROL M.

This paper presents the effect of light polarization on a hybrid fiber optic sensor. Hybrid sensor is defined as a combination of interferometer sensor and modalmetric sensor. Hybrid sensor system based on effect outputting interferometric sensor of interference as a result of changes made by the modalmetric system. Balance polarization in the arms of classical interferometer sensor leads to improved contrast interference pattern at the output of the interferometer. Aim of this study is to present a balance light polarization effects in a hybrid system.

\* \* \*

### Photonic crystal fibers selectively infiltrated with polymers

MILEŃKO K., WOLIŃSKI T. R.

Photonic crystal fibers are a special type of optical fibers, where cladding is composed of periodically arranged micron sized air channels while the fiber core is defined as a defect in the structure. The structure of the photonic crystal fiber allows for the air channels to be infiltrated with liquid materials such as liquid crystal (LC), alcohol and polymers. In this way we obtain new class of fibers that play important role in sensing applications because the infiltration enhances the sensitivity of the fiber to external physical parameters.

In this paper we will demonstrate the air channels of photonic crystal fibers selectively infiltrated with polymer materials. Two polymer materials used for infiltration are: Polydimethylsiloxane (PDMS) that is a silicone elastomer and silicon oil DC-704. The influence of the infiltration on the light propagation properties of the fibers and its applications in sensing will be also presented.

\* \* \*

### Pattern Recognition Algorithm for Eye Tracker Sensor Video Data Analysis

MURAWSKI K., RÓŻANOWSKI K.

The aim of this paper is to discuss research conducted at the Military Institute of Aviation Medicine as well as the Military University of Technology. The essence of the research is to develop patterns recognition algorithm for the analysis of data received from Eye Tracker sensor. Data analysis can be done using computer vision algorithms. In the article the state of the art pattern recognition algorithms was presented. Particular attention was paid to the

possibilities and limitations of their use. The main part of the paper is the presentation of the author's pattern recognition algorithm. Received results of their operation were compared with the results obtained by the commercial system to track the activity of the eye. An operation of developed method is also illustrated by examples.

\* \* \*

### **An Infrared Sensor for Monitoring of Meibomian Gland Dysfunction**

MURAWSKI K., RÓŻYCKI R., MURAWSKI P.,  
MATYJA A., REKAS M.

Meibomian Gland Dysfunction (MGD) is the one of the most common eye disorders observed in clinical practice. It applies to almost 50% of the population, especially people using contact lenses. It is believed that MGD is the most common cause of abnormal stability and integrity of the tear film. Despite this, there is no commercially available equipment for the diagnosis. The article proposes the construction of an optical sensor and a computer system for the rapid, non-invasive diagnosis of MGD. The designed hardware and software as well as preliminary results of clinical research are also described.

\* \* \*

### **Effect of wet etching of arc-induced Long-period Gratings on their refractive index sensitivity**

MYŚLIWIEC M., GROCHOWSKI J., KROGULSKI K.,  
MIKULIC P., BOCK W. J., ŚMIETANA M.

High performance optical refractive index (RI) sensors are important for advanced chemical and biological sensing applications. In this paper we present highly RI sensitive operation of Long Period Gratings (LPG). LPGs were fabricated with electric arc technique using Corning SMF28 optical fibre. As was previously reported, reduction of LPG cladding diameter results in increase of RI sensitivity and the red shift of resonant wavelengths. We used wet chemical etching in hydrofluoric (HF) acid solution to improve operation of the sensor. Optigrating software by Optiwave was used to simulate mentioned sensitivity improvement effect. As a result of simulations and experiments we obtained significant increase in RI sensitivity by applying appropriate etching time resulting in desired fibre thinning. What is more, we discuss further improvement of the sensitivity by deposition of nanocoating on the surface of the LPGs.

\* \* \*

### **Partial spectrum method for detection of covered materials in the TDS-THz setup**

PALKA N.

A signal processing method which can identify covered materials in a reflection configuration in the Time Domain Spectroscopy (TDS) setup is reported. THz radiation can transmit through most of covering materials like plastic foils, paper, clothes, etc. and can be detected after reflection from a covered substance (e.g. explosive material). In case of TDS setup, spectrum of the impulse is heavily deformed and the spectral features of substances are hard or impossible to identify. The proposed method based on the

fact that the TDS signal reflected from a covered sample consists of two peaks and some "waves" after the second peak, which carry the spectral information about the sample. FFT analysis of this part of the signal reveals spectral features of the sample. The presented method is reference-free and bases only on analysis of the signal reflected from the sample. The method is restricted to frequencies in the range 0.4–1.8 THz and, therefore only some materials with characteristic features in this range, like RDX-based explosives, lactose, paraaminobenzoic acid and tartaric acid can be analyzed. We covered the materials with foils, paper and cotton and obtained good results for solid and powder samples. The method is sensitive to atmospheric water vapour.

\* \* \*

### **Terahertz spectra of materials measured by an Hot Electron Bolometer-based system**

PALKA N., WALCZAKOWSKI M., ZAGRAJEK P.,  
TRZCINSKI T., SZUSTAKOWSKI M., SYPEK M.

Development and implementation of terahertz (0.1–10 THz) technology in security area is connected with unique features in this part of electromagnetic spectrum. Many explosives (e.g., Hexogen, Penthrite, Trinitrotoluene, Octogen) and drugs have characteristic transmission/reflection features in the THz range, what is of great importance for security and defense related applications. We report on measurements of transmission spectra of chosen materials (Hexogen, Sucrose, Tartaric Acid) in the range 0.7–2.5 THz. The measurements were carried out by means of a setup, which bases on an Optical Parametric Oscillator (OPO) combined with a Hot Electron Beam (HEB) Bolometer. The setup consists of commercially available tunable OPO from MSquared working in the range 0.7–2.5 THz with repetition rate 53 Hz, duration of the impulse of about 10 ns and energy 10 nJ. The beam was detected by a HEB from Scontel ( $NEP \sim 10^{-13} \text{ W/Hz}^{-1/2}$ ) in a Pulse Tube cryocooler. The spectra was compared to results obtained from a standard Time Domain Spectroscopy setup (Teraview TPS 3000). Only small discrepancies between spectra measured by both methods are observed. For the range 0.7–2 THz typical features can be identified using both methods.

\* \* \*

### **Laser photography – examples of processing of image information**

PISZCZEK M.

Modern cameras working in various spectral bands can be successfully applied in many areas of our lives. Observational capabilities of modern cameras are widely utilized in the area of public security systems while measurement functionalities of cameras are the domain of Machine Vision Systems used mainly in industry applications. The development of vision techniques and data processing methods changed the perception of vision systems as a multifunction, observation and measurement systems. Laser photography device (LPD), developed in the Institute of Optoelectronics, Military University of Technology is an example of such imaging device. One of properties of time-spatial framing method is the ability to make complex spatial analysis

of observed scene thanks to proper control of acquisition parameters and metadata processing. This property can be very useful especially in real-time spatial and geographic information systems. The paper presents theoretical applications of time-spatial framing method and first experimental results.

\* \* \*

### **Laser photography system – selected functionalities**

PISZCZEK M., KAROL M., KOWALSKI M., RUTYNA K., ZARZYCKI M., SZUSTAKOWSKI M.

The laser photograph is an advanced experimental system acquisition and distribution of spatial information. Its key element is a laser photograph device (UFL) use active illumination to observe the scene. The system also includes other information units such as VIS, IR, NIR cameras, and millimeter wave Radar. The study characterizes the individual components of the hardware and the system as a whole. Particular attention was given to the firmware so that it is possible to control system components as well as, a preliminary analysis of the results.. The presented system draws attention not only on the properties offered by the UFL but also the possibility of synergistic effects of the system as a whole through the fusion of data from different sensor units.

\* \* \*

### **Laser photography device – spatial parameters of imaging**

PISZCZEK M., KOWALSKI M., KAROL M., RUTYNA K., ZARZYCKI M., SZUSTAKOWSKI M.

Laser photography device (LPD) is an imaging device developed in the Institute of Optoelectronics, Military University of Technology and it is an example of modern image acquisition device. The LPD allows to define a 3D observation scene thanks to short-time scene illumination and image acquisition method. This device works according to time-spatial framing method. In the paper, basics of time-spatial framing method are explained. Special attention is given to time parameters of device and their influence to spatial parameters of registered images. In this paper the laser photography device and results of chosen experiments are presented and described. Experimental results presented in the paper show the potential and possibilities of using the laser photography device as a camera for observation and measurement applications.

\* \* \*

### **Imaging with laser photography camera during limited visibility**

PISZCZEK M., ZARZYCKI M., RUTYNA K., KOWALSKI M., KAROL M., SZUSTAKOWSKI M.

Present systems for vision monitoring very often employ many different imaging devices working in various spectral ranges e.g. visible (VIS), near infrared (NIR), infrared (IR). Every single spectral range deliver slightly different relevant information for process of threats detection, recognition and identification. However, technology

progress still does not satisfy every need. Finding new vision solutions, capable for imaging in very difficult conditions (adverse weather conditions, partial occultation of observed scene) is still one of the most urgent and demanding task for researchers. One of the possible solutions is using a camera working with time-spatial framing method. This method is able to minimize the impact of adverse factors on image acquisition process. Laser photography device (LPD) is a camera developed in the Institute of Optoelectronics, Military University of Technology working according to the time-spatial framing method. The paper presents theoretical basics and initial results of the laser photography device tests.

\* \* \*

### **Optical sensor to monitor pupillary light reflex (PLR)**

RÓŻANOWSKI K., MURAWSKI K.

This paper describes the design of an optical sensor intended to analyze pupillary light reflex (PLR). It also presents the results of physiological adaptation mechanisms in human eye, i.e. response of the iris to changes in the intensity of light that falls on the retina of the eye under conditions of sensory deprivation. PLR is a closed loop nerve reflex. It controls the amount of light that reaches the retina. Based on the test results, an optical sensor was designed, fabricated and correctly calibrated. In comparative tests with the use of F<sup>2</sup>D Fit-For-Duty, a commercial system by AMTech, selected pupillographic parameters were primarily evaluated (baseline pupil diameter, oscillations, reflex latency, maximum reaction time, pupil constriction time, pupil dilation time, and constriction amplitude) under conditions of diminished alertness, reduced ability to concentrate, increasing fatigue, and drowsiness.

The solution comes as part of a mobile pupillography device intended to be assembled in cars and airplanes to identify conditions of lower alertness, reduced ability to concentrate, increasing fatigue and drowsiness in drivers, pilots and traffic controllers, and to trigger alarm and preventive measures, if necessary.

\* \* \*

### **Design of a Multisensor System for a High Integrated Driver Fatigue Monitoring Technology**

RÓŻANOWSKI K., SONDEJ T., MURAWSKI K.

This paper describes the architecture and use of the multisensor system for estimating the driver's fatigue. These sensors are an element of the integrated system of monitoring driver behaviour, also including the measurement of environmental conditions and driving technique. Sensors described in the article are designed to work under simulated and real conditions (during driving a vehicle). To estimate driver's fatigue we used the subsystem for measurement of eye closing frequency and oculomotor activity (Eye Activity Research Subsystem - EYEARS), personal medical packet (PMP) and non-contact electrocardiography signal measurement module. In addition, for comparison purposes, we used commercial solutions called faceLAB (system for tracking and testing the activity of the eyes)

and FlexComp (system for physiological activities measurement). The article presents a description of the sensors, the way they communicate with the master device Logger Box and their location in the car. It also presents a typical test procedure and sample signals recorded during driving.

\* \* \*

### **Tunability of discrete diffraction in photonic liquid crystal fibers**

RUTKOWSKA K., LAUDYN U., JUNG P.

Recently, photonic liquid crystal fibers, PLCFs [i.e. photonic crystal fibers (PCFs) infiltrated with liquid crystals (LCs)] have gained significant amount of scientific attention. Their uniqueness results from the specificity of the PCF-host structures, as well as from the special optical properties of LCs. The latter can be dynamically adjusted by: (i) external fields and factors (including electric and magnetic fields, temperature, strain and pressure) and (ii) optical beam itself (i.e. when nonlinear effects take place in LCs).

In this work, the results of theoretical analyses and experimental tests on the light propagation in PLCFs are presented. While refractive index of the inclusion is higher than that of silica glass, analyzed photonic structure can be considered as a matrix of mutually parallel waveguide channels. This connotes discrete light propagation to be observed in PLCF, with the output beam profile strongly dependent on geometrical and optical properties of both the beam (e.g. wavelength and beam size) and the fiber (e.g. periodicity and index contrast). Importantly, changes in discrete propagation can be obtained dynamically (by varying optical power and/or by applying external fields and factors). In nonlinear case, under specific conditions, discrete spatial soliton can be obtained, paving thus the way for all-optical switching to be developed in PLCFs.

\* \* \*

### **Efficiency of diamond thin-film sensor on an Si-based integrated Mach-Zehnder interferometer**

SOBASZEK M., JĘDRZEJSKA-SZCZERSKA M., BOGDANOWICZ R., PŁOTKA P.

Mach-Zehnder is one of the commonly used two-beam interferometer in fiber-optic technology. This is because this interferometer can be easily elaborated with the use of conventional couplers or as an integrated optic device. Mach-Zehnder interferometer has measurement and reference arms and the output signal is the sum of optical signals from those two arms. By the use of special designed construction of measurement arm it is possible to determine many physical or chemical quantities. In this article, authors will present the result of theoretical investigation upon the new design of measurement arm in Mach-Zehnder interferometer. Authors will compare and contrast parameters of two devices. The first, which consists of SiO<sub>2</sub> and diamond layer and the second made from Si and diamond. The technology of devices made from SiO<sub>2</sub> and diamond is known but still very promising, on the other hand the technology of silicon optics is quite new. Therefore the

comparing and contrasting parameters of Mach-Zehnder interferometer made with the use of those two technology is very interesting. In this article, authors will present preliminary research which will be the base for constructing fiber optic sensor with the use of Mach-Zehnder interferometer.

\* \* \*

### **The possibility of using functional optical coherence tomography for scattering materials with nanoparticles evaluation and testing**

STRAKOWSKI M., KRASZEWSKI M., SOBASZEK M., ANTONIUK P.

The optical coherence tomography is a non-invasive and non-destructive method for scattering materials evaluation and testing. It produces tomography images of the devices inner structure with micrometer resolution. The functional OCT covers a number of different additional techniques of measurement signal analysis, which expand the range of OCT applications. The main OCT extensions are the Doppler analysis, polarization sensitive analysis and spectroscopic analysis. All these methods enable a better characterization of the tested materials and investigation the phenomena occurs inside the sample. Our research interests are focused on polarization sensitive (PS-OCT) and spectroscopic OCT (Sc-OCT). The PS-OCT delivers the information about the local changes of the devices optical anisotropy. However, the Sc-OCT gives the local spectral scattering characteristic of the tested sample. The combination of those two methods can be very useful for examination of scattering materials with nanoparticles. The size, the shape and the concentration of the nanoparticle dopants change the scattering features, which can be investigated by the use of functional OCT. In this paper we present our preliminary studies on the OCT application for nano-composite materials examination. Also the experimental results have been shown.

\* \* \*

### **Hybrid input-output systems of light for applications in integrated optics devices**

STRUK P., PUSTELNY T., GOŁASZEWSKA K., BORYSIEWICZ M. A., EKIELSKI M., PIOTROWSKA A.

The paper presents the results of investigations concerning input-output systems of light for applications in structures of integrated optics. The input-output system used in described planar waveguides are in the form of prism couplers and photonics structures with grating couplers. The first part of the paper contains numerical analyses of photonics structures with grating couplers aiming at an optimization of their geometrical parameters. The second part presents the practical realization as well as experimental tests of the planar optical waveguide with the hybrid input-output system. As the input system of the light a prism coupler was applied, and in the case of the output system – a photonic structure with grating coupler was used. The investigated planar wave guides with the input-output structures were made of a wide energy band gap semiconductor – zinc oxide (ZnO).

\* \* \*

### Comparison of standard and polymer polarimetric fiber sensors for stress monitoring in composite materials

SZELAĞ M., LESIAK P., AWIETJAN S., BOCZKOWSKA A., KUCZKOWSKI M., ERTMAN S., BUDASZEWSKI D., DOMAŃSKI A., WOLIŃSKI T.

In this work, we present results of our research on optical fiber sensors based on standard and plastic polarimetric fibers embedded into composite material samples. We manufactured two composite samples. In each sample six fibers were embedded: four plastic fibers made by Kiriama Pty Ltd, one Blaze-photonic fiber and one standard fiber HB1500 made by Nufern. Both ends of the composite plate were support articulated, so we could observe only bending effect. The composite plate was deformed up to 10 mm. In this paper, influence of bending effect on light polarization parameters such as: the Stokes vector, azimuth, ellipticity, degree of polarization, phase difference, power on standard and polymer polarimetric fiber sensors placed in composite material has been investigated.

\* \* \*

### Analyzing of transmission properties in birefringent gradient index MMI structures with absorbing layer

SZEWCZUK A., BŁAHUT M.

The aim of this study is to examine the possibility of the use in optical sensors technology polarization effects occurring in ion exchanged waveguides. The investigated gradient index MMI structures are produced in the  $K^+Na^+$  ion exchange process. Investigation was carried out for both polarizations TE and TM. This distinction into two polarizations is necessary because for the gradient index MMI structures produced in the  $K^+Na^+$  ion exchange process, propagation condition of light are different for each of the polarization. We have stated that birefringence of these waveguides manifests itself in different modal fields distributions and in particular in the different input field image position. Furthermore, the wave of TE and TM polarization will react differently with the sensing layer. This results in different operating characteristics for each of polarization.

\* \* \*

### Investigations of MMI structures covered by bromocresol purple for ammonia detection

SZEWCZUK A., BŁAHUT M.

The basis of the presented research is a gradient index MMI structure (Fig. 1) with multimode section length relating to the length of the input field image formation. Multimode section is covered with material which changes its optical parameters (refractive index and extinction coefficient) when is exposed to appropriate gaseous environments. The coating parameters changes have an effect on light propagation condition in structure and can be seen as the output signal variation. In this configuration, there are several factors responsible for changes in the registered

output signal. Variation of coating layer refractive index influence the mode properties of the multimode waveguide, which result in the input field image position and the output signal value. In the case of the absorption layer the refractive index changes have also influence on the modal attenuation value by changing the shape of the modal field distribution, which leads to a reduction or increase of the amount of light penetrating the absorbing sensing layer. This effect is different for different modes. Depending on the properties of a sensor layer – its refractive index and extinction coefficient – there is observed the predominance of the first (phase) or second (absorbing) effect. In this paper, we focus on the absorbing effect occurring in bromocresol purple sensing layers used to measure the gas concentrations.

\* \* \*

### Photonic liquid crystal fibers with polymers

TEFELSKA M., MILEŃKO K., BUDASZEWSKI D., DOMAŃSKI A. W., ERTMAN S., ORZECZOWSKI K., RUTKOWSKA K., SIARKOWSKA A., SIERAKOWSKI M., NOWINOWSKI-KRUSZELNICKI E., DĄBROWSKI R., MERGO P.

Photonic liquid crystal fibers with polymers constitute a new solution based on liquid crystals and microstructured polymer optical fibers (mPOFs) opening up new areas in innovative sensing and photonic devices applications. Compared with their silica-based microstructured fibers, it is easier to fabricate exotic mPOFs by extrusion or drilling at low temperature; their nonlinearity is potentially stronger, the range of available polymers that may be drawn is more diverse and the biocompatibility of polymers is often better. Liquid crystals due to their attractive properties i.e., the high birefringence, high electro-optic and thermo-optic effects are a very good candidate for mPOF infiltration to obtain tunable all-in-fiber innovative photonic devices. The paper will discuss basic properties and possible applications of the polymer photonic liquid crystal fibers that will arise from their high optical tunability with external and internal factors. Current research effort is directed towards two main solutions: photonic crystal fibers and mPOF-based structures, both infiltrated with liquid crystals of tailored optical properties.

\* \* \*

### Processing of measured data from integrated optics low-coherent sensors

WRÓBEL M., KARPIENKO K., JĘDRZEJEWSKA-SZCZERSKA M.

In this article authors present data processing techniques of measured signal from integrated optical sensors. We propose the use of dedicated software to support data analysis. It is shown that the use of such data processing can improve metrological parameters of the low-coherent sensor. The results of theoretical investigation and experimental work confirm that described techniques can be an effective method for improving the signal processing in low-coherent measurement sensors.

\* \* \*

### White upconversion in Yb<sup>3+</sup>/Tm<sup>3+</sup>/Ho<sup>3+</sup> – doped antimony-germanate glasses

ŻMOJDA J., DOROSZ D., KOCHANOWICZ M., MILUSKI P., DOROSZ J.

Glasses from the system SiO<sub>2</sub> – Sb<sub>2</sub>O<sub>3</sub> – GeO<sub>2</sub> – Na<sub>2</sub>O triply doped with Yb<sup>3+</sup>/Tm<sup>3+</sup>/Ho<sup>3+</sup> ions were melted and characterized. Fabricated glasses was exhibiting simultaneous multicolour upconversion luminescence. Strong blue 1G<sub>4</sub>→3H<sub>6</sub> (Tm<sup>3+</sup>), green 5F<sub>4</sub>→5I<sub>8</sub> (Ho<sup>3+</sup>) and red 5F<sub>5</sub>→5I<sub>8</sub> (Ho<sup>3+</sup>) upconversion emission bands have been measured under 980 nm excitation at room temperature. The dipole-dipole interaction between Yb<sup>3+</sup>/Tm<sup>3+</sup> and Yb<sup>3+</sup>/Ho<sup>3+</sup> ions have a non-resonant character hence, the phonon-assisted energy transfer has been investigated. Ytterbium ions plays important role in enhancing of upconversion emission efficiency. Colour coordinates (CIE-1931) and correlated colour temperature CCT of different molar ratio of active ions were calculated.

\* \* \*

### Modalmetric fiber optic sensor for security of collections

ŻYCZKOWSKI M.

The main aim of the work is the integration of modalmetric sensor in one arm. The proposed sensor is a classical fiber optic cable which measuring arm is upgraded with multimode sensor. The theme of research is to assess the impact of modal changes to detection of mechanical disturbances and comparative assessment of detection capabilities interferometer sensor, modalmetric sensor and hybrid

sensor for different types of excitations. The described tests, concerned area of basic research. The results can be used in the construction sensors with corresponding characteristics and properties to the appropriate applications as specially to protect of collection.

\* \* \*

### Quantum key as fiber optic security sensor

ŻYCZKOWSKI M.

he paper describes the methodology of identification interference in optical fiber, it also presents the technology known as "QKD" Quantum Key Distribution. It is based on the technology of constant comparison the input quantum characteristics of light source with its characteristics at the end of fiber optic. Methodology of presented work includes the evaluation of the functional objectives through assumptions to develop laboratory models. The paper presents the model of a system based on the comparison of polarization states of light quanta using two asymmetric Mach-Zender interferometers as transmitting and receiving systems to enable compensation of polarization state changes at the input and output of fiber optic sensing cable. On the base of continuous monitoring the state of the reference signal, the specific changes natural or ambient effects on the fiber, will attempt to identify interference in the optical waveguide as a change in the polarization of the quantum states of the light. The authors indicate the possibility of using such a fiber optic sensor as a security sensor to protect of wide critical infrastructure facilities.

## Calendar of events

### Conferences

#### Conferences in Poland

XLI Winter School on Vibroacoustical Hazards Suppressions, February 25–March 1, 2013, Szczyrk

XLII Winter School on Wave and Quantum Acoustics, February 25–March 1, 2013, Szczyrk

XX Conference on Acoustical and Biomedical Engineering, April 15–19, 2013, Zakopane  
<http://www.iab2013.agh.edu.pl/>

XXX Symposium on Hydroacoustics, May 13–16, 2013, Jastrzębia Góra  
<http://www.sha2013.pl/>

XVI International Conference on Noise Control, May, 26–29, 2013  
<http://www.ciop.pl/28242.html>

XV International Symposium on Sound Engineering and Tonmeistering, 27–29 June 2013, Kraków  
<http://www.isset2013.agh.edu.pl/>

International Conference on Auditory Display 2013, July 6–10, 2013, Lodz University of Technology  
<http://icad2013.com/>

LX Open Seminar on Acoustics, 9–13 August 2013, Rzeszów – Polańczyk  
<http://www.pta.univ.rzeszow.pl/osa2013/>

#### International Conferences

16–18 January, Frejus, France 12th Anglo-French Physical Acoustics Conference (AFPAC 2013)  
<https://www.sfa.asso.fr/afpac2013.html>

15–17 March, Merano, Italy EAA Winter School  
<http://www.aia-daga.eu.html>

18–21 March, Merano, Italy AIA-DAGA 2013 / EAA EUROREGIO  
<http://www.aia-daga.eu>

17–20 March, Les Arcs, France Electroceramics for End-users VII (PIEZO 2013)  
<http://www.piezoinstitute.com>

23–25 April, Marrakech, Morocco 1st Euro-Mediterranean Conference on Structural Dynamic and Vibroacoustics (ME-

DYNA 2013

<http://www.medyna2013.com>

1–4 May, Singapore 3rd International Congress on Ultrasonics (ICU 2013) concurrently organized with the 32nd International Symposium on Acoustical Imaging (AI 2013)  
<http://www.epc.com.sg/ICU 2010.pdf>

20–23 May, Hong Kong 2nd Symposium on Fluid-Structure-Sound Interactions and Control  
<http://www.fssc2013.com/>

26–29 May, Krakow-Rytko, Poland 11th International Conference on Active Noise and Vibration Control Methods (MARDiH-2013)  
<http://www.vibrationcontrol.pl>

26–31 May, Vancouver, Canada 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)  
<http://www.icassp2013.com>

02–07 June, Montreal, Canada 21st International Congress on Acoustics (ICA 2013), 165th Meeting of the Acoustical Society of America, and 52th Annual Meeting of the Canadian Acoustical Association  
<http://www.ica2013montreal.org>

09–11 June, Toronto, Canada International Symposium on Room Acoustics (ISRA 2013)  
<http://www.isra2013.com>

23–28 June, Corfu, Greece 1st International Conference and Exhibition on Underwater Acoustics  
<http://www.uam-conferences.org>

07–11 July, Bangkok, Thailand 20th International Congress on Sound and Vibration (ICSV20)  
<http://www.icsv20.org>

09–09 July, Glasgow, UK Invertebrate Sound and Vibration (ISV 2013)  
<http://www.isv2013.org>

30 July–03 August, Stockholm, Sweden Stockholm Music Acoustics Conference (SMAC 2013)  
<http://www.european-acoustics.org/smac-2013>

26–28 August, Denver, USA NOISE-CON 13  
<http://www.inceusa.org>

27–30 August, Denver, USA Wind Turbine Noise 2013  
<http://www.inceusa.org>

04–06 September, Portoroz, Slovenia 12th International Conference “Application of Contemporary Non-destructive testing in Engineering” (ICNDT 2013)  
<http://lab.fs.uni-lj.si>

08–13 September, Pirenopolis, Brazil International Bioacoustics Congress (IBAC 2013)  
<http://www.ibacbrasil.com/Home.html>

15–18 September, Innsbruck, Austria Internoise 2013  
<http://www.internoise2013.com>

9–11 October, Hangzhou, China 4th Pacific Rim Underwater Acoustics Conference (PURAC 2013)  
<http://pruac.zju.edu.cn/index.htm>

10–15 November, New Delhi, India Acoustics 2013 New Delhi  
<http://www.acoustics2013newdelhi.org>

2–6 December, San Francisco, USA 166th Meeting of the Acoustical Society of America  
<http://www.acousticalsociety.org>

## 2014

5–9 May, Providence, USA 167th Meeting of the Acoustical Society of America  
<http://www.acousticalsociety.org>

06–10 July, Beijing, China 21th International Congress on Sound and Vibration (ICSV21)

07–12 September, Krakow, Poland Forum Acusticum 2014  
<http://www.fa2014.pl/>

06–10 October, Prague, Czech Republic 11th European Conference on Non Destructive Testing  
<http://www.ecndt2014.com/>

27–31 October, Indianapolis, USA 168th Meeting of the Acoustical Society of America  
<http://www.acousticalsociety.org>

16–19 November, Melbourne, Australia Internoise 2014  
<http://www.internoise2014.org>



XX Conference on Acoustical and Biomedical Engineering  
15 -19 April 2013

Organizers:

Polish Acoustical Society, Krakow Division

AGH University of Science and Technology  
Faculty of Mechanical Engineering and Robotics  
Department of Mechanics and Vibroacoustics

<http://www.iab2013.agh.edu.pl>

---

## XXX SYMPOSIUM ON HYDROACOUSTICS

Jastrzębia Góra, May 13-16, 2013



Hotel Astor  
ul. Rozewska 38  
84-104 Jastrzębia Góra

Organizers:

Polish Acoustical Society, Gdansk Division  
Gdansk University of Technology -  
Faculty of Electronics, Telecommunications and Informatics  
Polish Naval Academy in Gdynia - Institute of Hydroacoustics

[www.sha2013.pl](http://www.sha2013.pl)



**OSA 2013**  
**60 OPEN SEMINAR**  
**on ACOUSTICS**

Rzeszów – Polańczyk, 9-13.09.2013

[www.pta.univ.rzeszow.pl/osa2013](http://www.pta.univ.rzeszow.pl/osa2013)

## AIMS AND SCOPE

Archives of Acoustics, the peer-reviewed quarterly journal publishes original research papers from all areas of acoustics like:

- acoustical measurements and instrumentation,
- acoustics of musics,
- acousto-optics,
- architectural, building and environmental acoustics,
- bioacoustics,
- electroacoustics,
- linear acoustics,
- noise and vibration,
- nonlinear acoustics,
- psychoacoustics,
- physical and chemical effects of sound,
- physiological acoustics,
- speech processing and communication systems,
- speech production and perception,
- transducers,
- ultrasonics,
- underwater acoustics,
- quantum acoustics.

## INFORMATION FOR AUTHORS

*ARCHIVES OF ACOUSTICS* the journal of Polish Academy of Sciences (a quarterly founded in 1976), publishes original papers concerning all fields of acoustics. Basically, the papers should not exceed 40 000 typographic signs (about 20 typewritten pages).

Submitted papers must be written in good English.

Follow this order when typing manuscripts: Title, Authors, their affiliations and e-mails, Abstract, Keywords, Main text, Acknowledgments, Appendix, References, Figure Captions and then Tables.

Special care must be attached to the following rules:

1. Text should be submitted with double-lined spacing and preferably in a 12 point font size. The text should be in single-column format.
2. The title of the paper should be as short as possible.
3. Full names and surnames should be given.
4. The full postal address of each affiliation, including the country name should be provided. Affiliations should contain the full postal address, as well as an e-mail address of one author designated as corresponding author.
5. The text should be preceded by a concise abstract (less than 200 words).
6. Keywords should be given. It is also desirable to include a list of notations.
7. The formulae to be numbered are those referred to in the paper, as well as the final formulae. The formula number should be written on the right-hand side of the formula in round brackets.
8. All notations should be written very distinctly.
9. Vectors should be denoted by semi-bold type. Trigonometric functions are denoted by sin, cos, tan and cot, inverse trigonometric functions – by arcsin, arccos, arctan and arccot; hyperbolic functions are denoted by sinh, cosh, tanh and coth.
10. References in the text (author(s) and year of publication) are to be cited between parentheses.
11. All cited works within the text must be included in the References Section, and arranged in alphabetical order by authors' last name. Number them from 1 to n.
12. Acknowledgments should be included in a section before the references.
13. Items appearing in the reference list should include the surname and the initials of the first name of the author, the full title of the paper/book in English followed by the information on the original paper language.  
In case of a book, the publisher's name, the place and year of publication should be given.  
In case of a periodical, the full title of the periodical, consecutive volume number, current issue number, pages, and year of publication should be given; the annual volume number must be semi-bold so as to distinguish it from the current issue number.
14. For more information on references see <http://acoustics.ippt.pan.pl/Instructions.pdf>

Each manuscript is initially evaluated by the Editor and then allocated to an Associate Editor according to specific subject area. Upon receiving the Associate Editor's list of potential reviewers, the Editorial Office forwards it to reviewers. The Editor and the Associate Editors make the final recommendation. Their opinions are the basis for the Editorial Committee to determine whether the paper can be accepted for publication or not.

The Word or L<sup>A</sup>T<sub>E</sub>X source and pdf files of the final version of articles will be requested after acceptance for publication. High-quality illustrations (format bitmap (gif, jpg) or PostScript, with resolution no lower than 300 dpi) should be sent along with the manuscript.

After publication CD-ROM containing PDF version of article will be sent to authors.

*EDITORIAL BOARD  
ARCHIVES OF ACOUSTICS*

## SUBSCRIPTION

Subscription orders for "Archives of Acoustics" may be sent to:

Editorial Office "Archives of Acoustics"  
Pawińskiego 5B, 02-106 Warszawa, Poland  
phone: (48-22) 826 12 81 ext. 206  
fax (48-22) 826 98 15  
e-mail: akustyka@ippt.pan.pl

Information about subscription are available in <http://acoustics.ippt.pan.pl>

# Instructions for citation and bibliographic references

References in the text (author(s) and year of publication) are to be cited between parentheses.

**1.** Direct quotations should be between quotation marks.

**2.** Indirect quotations (i.e. when the author's idea is explained but not directly quoted) do not use quotation marks or italics.

Examples:

- a. The research of mufflers was started by Davis (1954).
- b. To increase the muffler's acoustical performance, the assessment of a new acoustical element – a reverse-flow mechanism with double internal perforated tubes – was proposed and investigated by this author (Munjaj, 1987),

**3.** When an author or group of authors has more than one publication in the same year a lower case letter is added to the date.

Example:

In two recent studies (Rakowski, 1993a; 1993b) it was suggested that...

**4.** Multiple author citation:

a. Two authors: both names should be listed in each citation.

Example: (Rakowski, Miyazaki, 2007)

b. Three or more authors: use the first author's name and "*et al.*"

Example: (Dobrucki *et al.*, 2005)

**5.** To cite a document produced by an organization, the first time write out the name of the organization in full and give the acronym or abbreviation in square brackets. For subsequent references you may use the acronym or abbreviation.

Example:

First time: Sanitary politics (United Nations High Commissioner for Refugees [UNHCR], 2006, p. 35)...

Then: Established politics (UNHCR, 2006, p. 45)...

In the reference list write the full name of the organization.

**6.** When a source has no identified author, cite the first two or three words of the title (in italics if it is a book, between quotation marks if it is an article) followed by the year and the page.

Examples:

In the recent book (*Encyclopedia of Physics*, 1993, p. 164)...

In the article ("*The relation of pitch ...*", 1979, p. 67) it is held that...

The in-text reference should match the start of the reference in the bibliography/reference list.

**7.** Using quotations in the text may also be as it is shown in the following examples:

Examples:

a. "Field measurements of bistatic scattering strength (BSSS) are difficult and expensive to acquire at sea, in real conditions." (Blondel, Pace, 2009, p. 101)

b. Blondel and Pace (2009) held that "field measurements of bistatic scattering strength (BSSS) are difficult and expensive to acquire at sea, in real conditions" (pp. 101)

**8.** To cite a work that was discovered in another work, observe the following examples:

a. Brown (1967), cited by Smith (1970, p. 27), found...

b. It was found (Brown, 1967, cited by Smith, 1970, p. 27) that...

**9.** To cite e-mail messages:

If necessary, cite them within the body of the text as a personal communication.

Example:

a. ...and this point was conceded (S. Weyna, personal communication, August 22, 2001), ...

## Presenting bibliography/reference list

### General matters

a. Reference list should be arranged alphabetically by author surname. Put all the entries in one long alphabetical list. Do not list books, journal articles, websites, etc. in separate sections. Number them from 1 to n.

Example:

- 1.
- 2.
- 3.

b. When an author or group of authors has more than one publication in the same year a lower case letter is added to the date. This letters will be used when citing this source in the text.

Example:

- Zou D., Crocker M.J (2009a), *Response of a plate to PZT actuators*, Archives of Acoustics, **34**, 1, 13–23.  
Zou D., Crocker M.J (2009b), *Sound power radiated from rectangular plates*, Archives of Acoustics, **34**, 1, 25–39.

c. Book, encyclopedia and article titles should be italicized.

### 1. Books

**Author, initials (year), Title of the book [language of publication – if other than English], Edition, if later than the first, Publisher, Place of publication.**

Examples:

- American Psychological Association (2001), *Publication manual of the American Psychological Association*, 5th Ed., Washington D.C. [Author].
- Encyclopedia of Physics* (1993), 2nd edition, McGraw-Hill, New York.
- Varshneya A. (1944), *Fundamentals of inorganic glasses*, p. 111, Academic Press, New York.
- Moore M.H., Estrich S., McGillis D., Spelman W. (1984), *Dangerous offenders: the elusive target of justice*, Harvard University Press, Cambridge.
- Strunk W., White E.B. (1979), *The elements of style*, 3rd Ed., Macmillan, New York.
- Engel Z., Piechowicz J., Stryczniewicz L. (2003), *The fundaments of industrial vibroacoustics* [in Polish], WIMIR AGH, Kraków.
- Nowicki A. (1995), *Basics of Doppler Ultrasonography*, [in Polish:] *Podstawy ultrasonografii dopplerowskiej*, PWN, Warszawa.

List up to six (6) authors. If there are seven (7) or more, list the surname and initials of the first one and then “et al.”

### 2. Edited book

**Editor, initials (Ed.), (year), Title of the book, Edition if later than the first one, Volumes if there are more than one, Publisher, Place of publication.**

Examples:

- Crocker M.J. [Ed.], (2007), *Handbook of Noise and Vibration Control*, John Wiley & Sons, Inc., New York.
- Klinke R., Hartman R. [Eds.], (1983), *Hearing – physiological bases and psychophysics*, Springer, Berlin, Heidelberg, New York, Tokyo.

### 3. Chapter in edited book

**Author of chapter, initials (year), Title of chapter, [in:] Title of the book, Name of the Editor/s Initials [Ed.], pp. start and end of chapter page numbers, Publisher, Place of publication.**

Example:

- Rakowski A. (1991), *Context-dependent intonation variants*, [in:] *Music, language, speech and brain*, Sundberg J., Nord L., Carlson R. [Eds.], pp. 203–211, MacMillan Press, London.
- Berlincourt D.A., Curran D.R., Jaffe H. (1964), *Piezoelectric and Piezomagnetic Materials and Their Function in Transducers*, [in:] *Physical Acoustic*, Mason W.P. [Ed.], vol. 1, part. A, pp. 169–270, Academic Press, New York.

### 4. Journal article

**Author, initials (year), Title of the article [language of publication – if other than English], Title of the journal, Volume number (bold face), issue number, start and end page numbers of article.**

Write directly page numbers (not preceded by p. or pp.),

Examples:

- Houtsma A. (2007), *Experiments on pitch perception: Implications for music and other processes*, Archives of Acoustics, **32**, 475–490.
- Sek A., Moore B. C. J. (1995), *Frequency discrimination as a function of frequency, measured in several ways*, Journal of the Acoustical Society of America, **97**, 2479–2486.
- Ranachowski P., Rejmund F., Pawelek A., Piątkowski A. (2005), *Studies of cordierite material under compressive load at different temperatures and after thermal shock* [in Polish], Ceramics, **89**, 101–115.

## 5. Conference paper in published proceedings

Example:

- a. Rakowski A., Miśkiewicz A. (2002), *Pitch discrimination of low-frequency tones*, Proceedings of 7th International Conference on Music Perception and Cognition , pp. 538–540, Sydney.

## 6. Dissertations in congresses and meetings (unpublished)

Example:

- a. Tucker S. (2003) , *An ecological approach to the classification of transient underwater acoustic events: perceptual experiments and auditory models*, Ph.D. Thesis, Department of Computer Science, University of Sheffield.
- b. Salamon R. (2009), *Contemporary military sonar system*, Dissertation presented during the 56th Open Seminar on Acoustics, OSA 2009, Goniądz, Poland.

## 7. Electronic sources

**Author, initials (year), Title, Retrieved month, day, year, from Internet address.**

If no date is shown on the document, authors should use n.d. (no date)

If the author is not given, authors should begin the reference with the title of the document.

If a document is part of a large site such as a university or a government department's website, give the name of the parent organization and the relevant department before the web address.

Examples:

- a. Cox T. (n.d.), *Sound quality*, from [www.acoustics.salford.ac.uk/res/cox/sound\\_quality/](http://www.acoustics.salford.ac.uk/res/cox/sound_quality/)
- b. Deciding your future (2000), Retrieved September 5th, 2001, from University of Portsmouth, Careers Service: [www.port.ac.uk/departments/careers/plancareer/deciding-your-future.htm](http://www.port.ac.uk/departments/careers/plancareer/deciding-your-future.htm)
- c. Alexander J., Tate M.A. (2001), *Evaluating web resources*, Retrieved August 21st, 2001, from Widener University, Wolfgram Memorial Library: [www2.widener.edu/Wolfgram-Memorial-Library/webevaluation/webeval.htm](http://www2.widener.edu/Wolfgram-Memorial-Library/webevaluation/webeval.htm)

Authors should not write the web address (URL) within the text of the paper, it should appear only in the reference list/bibliography. To cite this source within the text they should use the author's name (if the reference has one) or the first few words of the website title.

## 8. Audiovisual sources: music

**Author, initials (Date of copyright), Title of the song. On Title of the album [medium of recording], Location: Label (Recording date if different from copyright date)**

Example:

- a. Puccini G. (1990), *Nessun dorma*, On Carreras Domingo Pavarotti in concert [CD], Decca, London.