

# **VIBRATIONS**

## **IN PHYSICAL SYSTEMS**

**VOLUME 30, No. 1**

**POZNAŃ 2019**





# VIBRATIONS IN PHYSICAL SYSTEMS

## **EDITORS**

Czesław CEMPEL (honorary editor)

Tomasz STRĘK (editor in-chief)

Maciej TABASZEWSKI (editor)

## **GUEST EDITOR**

Wojciech ŁAPKA

## **TECHNICAL EDITORS**

Paweł FRITZKOWSKI

Tomasz HERMANN

Hubert JOPEK

## **PUBLISHER**

Institute of Applied Mechanics,  
Poznan University of Technology

## **EDITORIAL OFFICE**

Poznan University of Technology

Institute of Applied Mechanics

ul. Jana Pawła II 24, 61-139 Poznań, Poland

tel. +48 61 665 23 01

[vibsys@put.poznan.pl](mailto:vibsys@put.poznan.pl)

[vibsys.put.poznan.pl](http://vibsys.put.poznan.pl)

**ISSN 0860-6897**

Poznan, Poland 2019



# VIBRATIONS

IN PHYSICAL SYSTEMS

## EDITORIAL BOARD

Vladimir I. ALSHITS  
Jan AWREJCEWICZ  
Wojciech BATKO  
Romuald BĘDZIŃSKI  
Tadeusz BURCZYŃSKI  
Czesław CEMPEL  
Enzo CIANCIO  
Evgen CZAPLA  
Zbigniew DĄBROWSKI  
Marian W. DOBRY  
Antoni GAJEWSKI  
Joseph GRIMA  
Jan HOLNICKI-SZULC  
Erik V. JANSSON  
Jarosław JĘDRYŚIAK  
Stefan JONIAK  
David JOU  
Jerzy KALETA  
Jan KOŁODZIEJ  
Tomasz ŁODYGOWSKI

Krzysztof MAGNUCKI  
Krzysztof MARCHELEK  
Stanisław MATYSIAK  
Wolfgang MUSCHIK  
Józef NIZIOŁ  
Andrzej RADOWICZ  
Stanisław RADKOWSKI  
Liliana RESTUCCIA  
Hamid M. SEDIGHI  
Roman STAROSTA  
Tomasz STRĘK  
Tomasz SZOLC  
Maciej TABASZEWSKI  
Franciszek TOMASZEWSKI  
Andrzej TYLIKOWSKI  
Tadeusz UHL  
Jerzy WARMIŃSKI  
Józef WOJNAROWSKI  
Alexandr YAVLENSKY

## COPYRIGHT / OPEN ACCESS

Articles published in Vibrations in Physical Systems will be Open-Access articles distributed under the terms and conditions of the Creative Commons Attribution License (CC BY).



## CONTENTS

<b>1. Agnieszka Paula PIETRZAK .....</b>	<b>2019101</b>
Musicians' Daily Sound Exposure Assessed by Full-Day Dosimetry	
<b>2. Anna SNAKOWSKA, Łukasz GORAZD, Jerzy JURKIEWICZ.....</b>	<b>2019102</b>
Evaluation of the Transmission and the Scattering Matrix Applicability to the Mufflers Analysis	
<b>3. Kamil WÓJCIAK, Patryk GAJ, Joanna KOPANIA, Grzegorz BOGUSŁAWSKI .....</b>	<b>2019103</b>
The Impact of the Damper Blade Position on the Generated Noise and Pressure Losses in Plenum Box with Swirl Diffusers	
<b>4. Andrzej DOBRUCKI, Stefan BRACHMAŃSKI, Maurycy KIN .....</b>	<b>2019104</b>
Objective and Subjective Evaluation of Musical and Speech Recordings Transmitted by DAB+ System	
<b>5. Krzysztof J. OPIELIŃSKI, Marcin WRZOSEK, Józef NICPOŃ, Przemysław PODGÓRSKI, Tomasz ŚWIETLIK .....</b>	<b>2019105</b>
Ultrasound Tomography Imaging: Results of Breast Phantom Study and Resolution Estimation	
<b>6. Wiktor STASZEWSKI, Tadeusz GUDRA .....</b>	<b>2019106</b>
The Effect of Dynamic Focusing of the Beam on the Acoustic Field Distribution Inside the Ultrasonic Ring Array	
<b>7. Dorota CZOPEK, Jerzy WICIAK.....</b>	<b>2019107</b>
Vibrotactile Amplitude Discrimination on the Wrist of Visually Impaired People	
<b>8. Iwona KOCHAŃSKA, Jan H. SCHMIDT.....</b>	<b>2019108</b>
Simulation of Direct-Sequence Spread Spectrum Data Transmission System for Reliable Underwater Acoustic Communications	
<b>9. Patryk GAJ, Joanna KOPANIA, Kamil WÓJCIAK, Grzegorz BOGUSŁAWSKI .....</b>	<b>2019109</b>
Assessment of Sound Absorbing Properties of Composite Made of Recycling Materials	
<b>10. Milena KACZMAREK-KLINOWSKA, Tomasz HORNOWSKI .....</b>	<b>2019110</b>
Cyclodextrine vs D-Glucose in the Solutions of the Derivative of 1,4-DHP	

<b>11. Katarzyna A. URBANIEC .....</b>	<b>2019111</b>
A Prototype of Chinese Aspirated Consonants Pronunciation Training System Based on Multi-Resolution Cochleagram	
<b>12. Dariusz BISMOR .....</b>	<b>2019112</b>
Analysis and Comparison of Vibration Signals from Internal Combustion Engine Acquired Using Piezoelectric and MEMS Accelerometers	
<b>13. Andrzej MIŚKIEWICZ.....</b>	<b>2019113</b>
Scientific Legacy of Professor Andrzej Rakowski in Current Studies of Pitch Discrimination in Music	
<b>14. Katarzyna BARUCH, Tadeusz KAMISIŃSKI, Aleksandra MAJCHRZAK.....</b>	<b>2019114</b>
Measurement of Perforated Panels at a Scaled Measurement Setup	
<b>15. Tadeusz KAMISIŃSKI, Krzysztof BRAWATA .....</b>	<b>2019115</b>
Sound Strength G Prediction in Orchestra Pit Based on Barron and Lee "Revised Theory"	
<b>16. Joanna Maria KOPANIA, Grzegorz BOGUSŁAWSKI, Patryk GAJ, Kamil WÓJCIAK .....</b>	<b>2019116</b>
Noise Radiation from Circular Rods at Low-Moderate Reynolds Number	
<b>17. Krzysztof RUDNO-RUDZIŃSKI .....</b>	<b>2019117</b>
Required Attenuation of Aircraft Noise in Buildings in the Light of Data from the Chopin Airport Monitoring	
<b>18. Łukasz GORAŻD, Anna SNAKOWSKA .....</b>	<b>2019118</b>
Simulation of a Single Mode Wave Generation in Cylindrical Systems Applying Numerical Methods	
<b>19. Krzysztof Władysław ZUBER, Krzysztof J. OPIELIŃSKI .....</b>	<b>2019119</b>
Animal Mimicry for Covert Communication with Arbitrary Output Distribution: Beyond the Assumption of Ignorance	
<b>20. Anna PREIS, Jan FELCYN, Piotr KOKOWSKI, Michał GAŁUSZKA, Paweł LIBISZEWSKI .....</b>	<b>2019120</b>
Estimating the Population Exposed to Transportation Noise: a Case Study on Poznań City	
<b>21. Damian WĘGRZYN, Piotr WRZECIONO .....</b>	<b>2019121</b>
Problem of Placing the Organ Pipes on the Windchest	
<b>22. Jan ŻERA .....</b>	<b>2019122</b>
Timbre Solfege and Auditory Profile Analysis	

<b>23. Karol LISTEWNIK .....</b>	<b>2019123</b>
A Design of an Acoustic Coupler for Calibration of Hydrophones at Low Frequencies	
<b>24. Adam KACZOROWSKI, Maria KOMASA .....</b>	<b>2019124</b>
Influence of Noise on Children's Sense of Hearing	
<b>25. Krzysztof KACHNIARZ, Marcin LEWANDOWSKI .....</b>	<b>2019125</b>
Automatic Recognition of Artificial Reverberation Settings in Speech Recordings	
<b>26. Michał LUCZYŃSKI .....</b>	<b>2019126</b>
Classes of Tonality of Signals in the Aspect of Active Elimination of Tonal Components	
<b>27. Mirosław MEISSNER .....</b>	<b>2019127</b>
Prediction of Low-Frequency Sound Field in Rooms with Complex-Valued Boundary Conditions on Walls	
<b>28. Bartłomiej CHOJNACKI, Tadeusz KAMISIŃSKI, Klara JUROS Daniel KACZOR .....</b>	<b>2019128</b>
Coupled Speakers Directivity Measurements for Small Acoustic Omnidirectional Source Development	
<b>29. Paweł DZIECHCIŃSKI .....</b>	<b>2019129</b>
A Computer Model for Calculating the Speech Transmission Index Using the Direct STIPA Method	
<b>30. Jędrzej BOROWSKI, Krzysztof J. OPIELIŃSKI .....</b>	<b>2019130</b>
Model of Multipath Propagation of Ultrasonic Pulses in Soft Tissue Using Divergent Beam Tomography Method	
<b>31. Mateusz KUCHARSKI, Stefan BRACHMAŃSKI .....</b>	<b>2019131</b>
Coding Effects on Changes in Formant Frequencies in Japanese Speech Signals	
<b>32. Konrad MIERZEJEWSKI, Lucyna LENIOWSKA .....</b>	<b>2019132</b>
Design of Control System for Active Vibration Suppression of Trapezoidal Plate	
<b>33. Piotr STARONIEWICZ .....</b>	<b>2019133</b>
Tests of Basic Voice Stress Detection Techniques	
<b>34. Maciej SABINIOK, Krzysztof OPIELIŃSKI .....</b>	<b>2019134</b>
Comparison of the Image Compounding Methods for the Multi-Angle 3-D Ultrasound Imaging	

<b>35. Jan H. SCHMIDT, Aleksander M. SCHMIDT, Iwona KOCHAŃSKA.....</b>	<b>2019135</b>
Performance of Coherent Modulation Scheme Used in Acoustic Underwater Communication System	
<b>36. Janusz PIECHOWICZ, Jakub PIECHOWICZ.....</b>	<b>2019136</b>
Development of Predictive Model for Vibro-Acoustic Protections in Industrial Hall	
<b>37. Tsuyoshi USAGAWA, IRWANSYAH .....</b>	<b>2019137</b>
Binaural Speech Segregation System on Single Board Computer	
<b>38. Mateusz GAWLIK, Wiesław WSZOLEK .....</b>	<b>2019138</b>
Intonation Accuracy and Pitch Stability During Crescendo as the Voice Quality and Singer's Experience Indicator Among Choral Singers	
<b>39. Grzegorz BOGUSŁAWSKI, Joanna Maria KOPANIA, Patryk GAJ, Kamil WÓJCIAK .....</b>	<b>2019139</b>
Determination of Sound Power Level by Using a Microphone Array and Conventional Methods	
<b>40. Wojciech ŁAPKA.....</b>	<b>2019140</b>
Tuning the Selected Acoustic Helicoidal Resonator with a Short Flat Bar – Numerical Analysis	
<b>41. Piotr Z. KOZŁOWSKI.....</b>	<b>2019141</b>
Effectiveness of Acoustic Banners Depending on their Arrangement in the Concert Hall – Case Study	
<b>42. Roman TROJANOWSKI, Jerzy WICIAK.....</b>	<b>2019142</b>
Numerical Analyses of the Effectiveness of an Integrated Disc Based Piezoelectric Sensor-Actuator	
<b>43. Przemysław PLASKOTA.....</b>	<b>2019143</b>
Research of Acoustical Impedance of Human Skin	
<b>44. Tomasz ŚWIETLIK, Krzysztof J. OPIELIŃSKI.....</b>	<b>2019144</b>
Analysis of Doppler Tomography in Circular Geometry as a Novel Method of Imaging Tissue Cross-Sections <i>in vivo</i>	
<b>45. Aleksandra KLIMEK.....</b>	<b>2019145</b>
Fluid-Fluid Phononic Crystal with Elastic Coat Working in Audible Frequencies	

## Musicians' Daily Sound Exposure Assessed by Full-Day Dosimetry

Agnieszka Paula PIETRZAK  
Warsaw University of Technology,  
Institute of Radioelectronics and Multimedia Technology,  
ul. Nowowiejska 15/19, 00-665 Warsaw, a.pietrzak@ire.pw.edu.pl

### Abstract

In day-to-day life student musicians are often exposed to high-level sounds, that may be damaging for their hearing. At music universities, there are various obligatory activities, such as playing in the symphony or wind orchestra, playing in chamber ensembles or a Big Band. Many hours of individual practice is also mandatory. In every one of these activities, sound pressure levels are often high and the daily sound exposure levels often exceed the permissible limit of 85 dB. In this study, the exemplary results of the daily sound exposure are shown for a group of music students. In contrast to other studies, the measurements were carried out throughout the entire duration of the typical workday, so the calculated daily sound exposure level ( $L_{EX,8h}$ ) is based on factual data, not estimations. Data concerning the A-weighted equivalent continuous sound pressure level in the function of time are presented for musicians playing the flute, clarinet, double-bass, percussion, trombone, trumpet and French horn, for every activity (e.g. lesson, rehearsal, concert). The daily sound exposure level values ( $L_{EX,8h}$ ) are determined for each instrument.

**Keywords:** sound exposure, musicians, noise dosimetry, full-day dosimetry

### 1. Introduction

In their everyday work, musicians devote a lot of time to practicing the instrument - both solo and in chamber ensembles or orchestras. The sound pressure levels during these practice, as well as during performances, are often high and can lead to hearing loss. Over the past decades, musicians' exposure to sound has been the subject of many studies and measurements.

Studies available in the literature indicate that sounds occurring during playing in an orchestra and often exceed 90 dB [1-6]. Musicians playing wind instruments, especially brass, are exposed to the highest sound pressure levels [1, 7-9].

The research conducted so far has initially focused on sound exposure during performances with orchestras only. Exposure to sound during solo exercises was considered for the first time by Royster [7]. Measurement during individual exercises were carried out also by Laitinen [1] and Schmidt [3]. However, there is little work considering overall sound exposure during all activities involving musicians during a typical working day. Previously mentioned work involves measuring the sound pressure levels during a single event during the day, such as a rehearsal or concert. Exposure to sound (daily, weekly or annual) is either not determined, or based only on single event data, and does not take into account the fact that musicians are daily involved in a wide variety of music activities, rehearsals and hours of solo practice. Smith, Nielsen and Grimshaw conducted a study that takes into account all of the musicians' activities during the day [10]. They performed a two-day continuous

dosimeter measurement for a group of musicians, to indicate that the musicians were exposed to doses of sound exceeding the current standards.

The aim of this study was to estimate the daily noise exposure level ( $L_{EX,8h}$ ), according to ISO 9612 standard [11] for a group of Fryderyk Chopin University of Music students, based on the measurement of the whole day of their musical activity, and to analyze how  $L_{EX,8h}$  values depend on the type of activity and the type of instrument.

Due to the musical nature of the sounds being the subject of the research, in the following part of this study author uses the term “sound exposure” instead “noise exposure” which occurs in the ISO 9612 standard.

## 2. Measurement method

The research was conducted for a group of students from the Fryderyk Chopin University of Music in Warsaw playing flute, clarinet, double bass, drums, trombone, trumpet and horn. The musicians were asked to indicate the most typical day in their opinion, in terms of the type of classes and the number of hours spent at the university. On the day indicated by the musicians, after their arrival at the university, they were equipped with a portable dosimeter, microphones were mounted on their arms, and they received instructions on how to use the device properly. The survey participants were asked to fill in a table during the measurement, in which they specified the type of instrument, the type of activity (solo exercise, playing in a chamber ensemble, an orchestra) and the time of starting and ending of a given activity. The research was conducted for 7 musicians playing the flute, clarinet, trumpet, French horn, trombone, percussion and double bass. Due to the small size of the studied group, the results are exemplary.

The measurements were carried out using two two-channel personal noise dosimeters (Svantek SV 102+), equipped with SV 25D microphones. The dosimeters were calibrated using the BK 4231 calibrator. The microphones were placed on both sides of the head, on the musicians' shoulders, at a distance of about 10 cm from the ears. The devices recorded values of A-weighted equivalent continuous sound pressure level  $L_{Aeq,T}$  [11] for every one second of measurement time ( $L_{Aeq,1s}$ ). The resulting material was checked for artifacts such as accidental hitting of the microphone and samples with artifacts were not considered for further analysis. Based on the obtained  $L_{Aeq,1s}$  values and the duration of all activities, daily sound exposure level, i.e. A-weighted noise exposure level normalized to an 8-hour working day ( $L_{EX,8h}$ ) [11], was determined for each instrument.

## 3. Results

The recorded  $L_{Aeq,1s}$  waveforms as a function of time for individual instruments are shown in Fig. 1a-g. The musical activities indicated by the musicians in the questionnaires were divided into 5 categories: solo practice - marked in green on the figures, playing in chamber ensembles - yellow, symphony orchestra - red, wind orchestra - blue and Big Band - violet. The range of time spent on each of the above mentioned activities is marked in Fig. 1a-g with rectangles in aforementioned colours.

The graphs show the  $L_{Aeq,1s}$  values separately for the right (red) and left ear (blue), but the level difference between the right and left ear is not analyzed in this work (see [12]).

In the case of the flutists (Fig. 1a), it can be seen that the highest  $L_{Aeq,1s}$  values occur during individual practice, sometimes exceeding even 120 dB. Most of the day is spent playing in chamber ensembles, where  $L_{Aeq,1s}$  often exceeds 100 dB. For a musician playing the clarinet (Fig. 1b) during solo exercises,  $L_{Aeq,1s}$  values are lower than when playing in a wind orchestra (blue rectangle) or in a chamber ensemble. The biggest exposure to sound during the measured day was the rehearsal of a chamber ensemble lasting more than 2 hours, during which the  $L_{Aeq,1s}$  values ranged from 80–100 dB. In the case of the trumpet (Fig. 1c), as in the case of the flute, the highest  $L_{Aeq,1s}$  values occur during the solo practice. In the case of the trumpet, the  $L_{Aeq,1s}$  values during this activity are in the range of 100–110 dB. During rehearsals of a symphony orchestra (red rectangle),  $L_{Aeq,1s}$  sometimes also exceed 100 dB, but there are also many times when the values of  $L_{Aeq,1s}$  are low – these are moments when the musicians do not play because they listen to the conductor's speaking. For a trombone player (Fig. 1d), the highest  $L_{Aeq,1s}$  values occur during an hour and a half long Big Band rehearsal (violet rectangle), exceeding 110 dB. One can also pay attention to the blue rectangle – a rehearsal of a wind orchestra, during which in the first and second half of the time a different repertoire is practiced and the course of  $L_{Aeq,1s}$  has a different character for these parts.

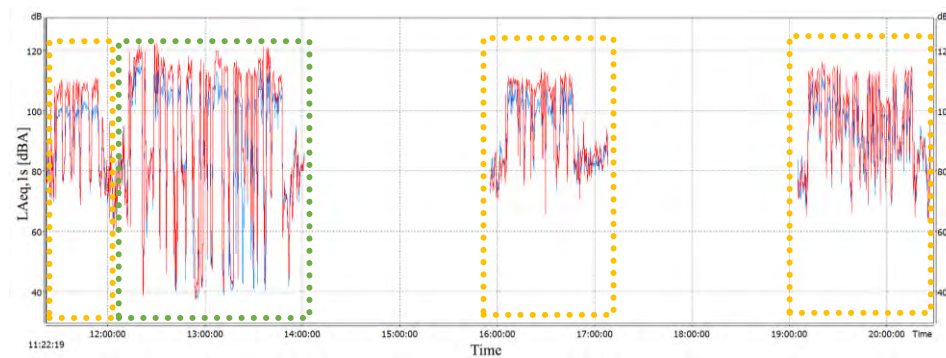


Figure 1a. Flute

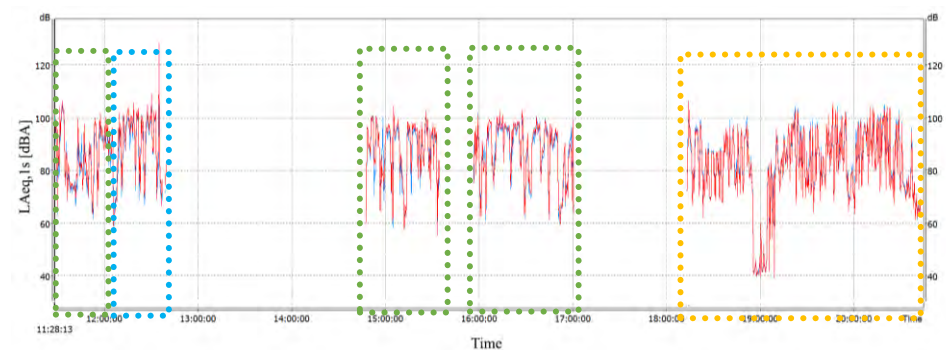


Figure 1b. Clarinet

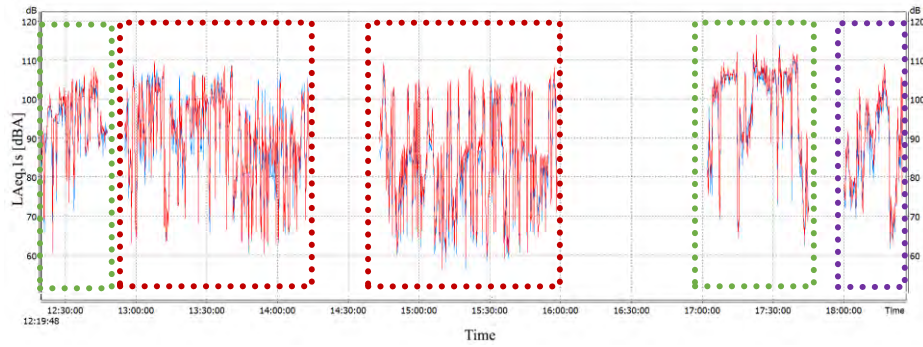


Figure 1c. Trumpet

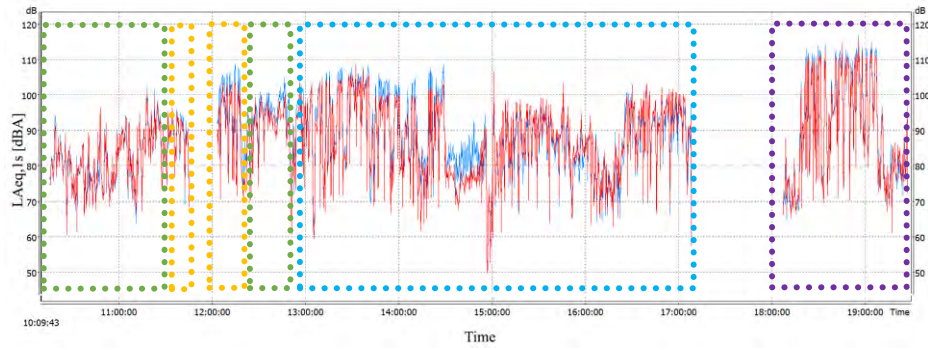


Figure 1d. Trombone

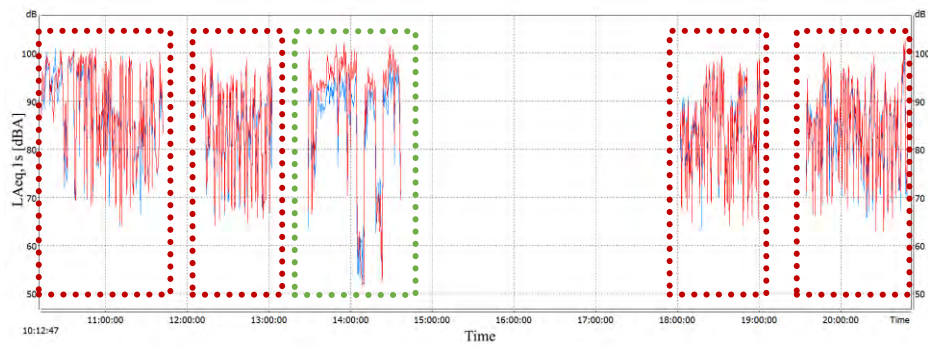


Figure 1e. French horn



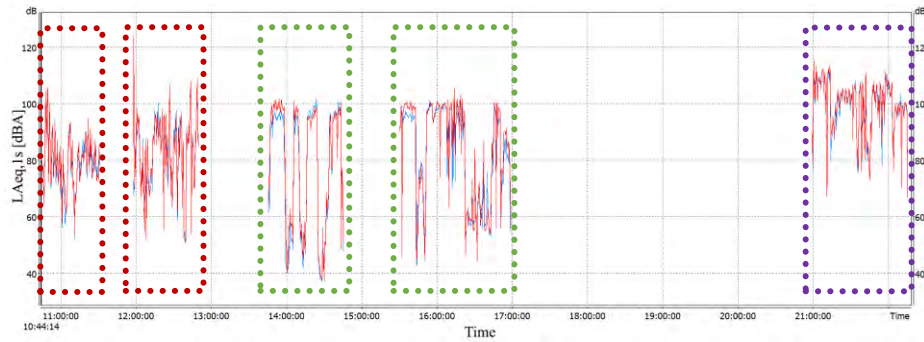


Figure 1f. Percussion

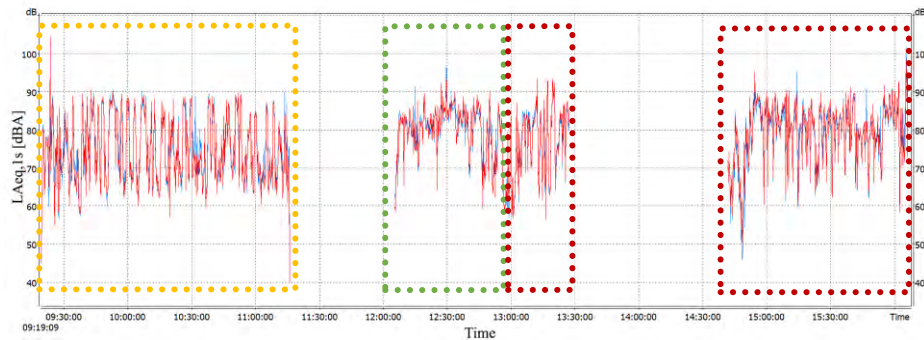


Figure 1g. Double-bass

Figure 1a-g. A-weighted equivalent continuous sound pressure level  $L_{Aeq,1s}$  in the function of time (left ear – blue line, right ear – red line). Different types of activities are marked with rectangles in different colours: individual practice – green, chamber ensembles – yellow, symphony orchestra – red, wind orchestra – blue, Big Band – violet

On the figure concerning the French horn (Fig. 1e), it can be noticed that the course of  $L_{Aeq,1s}$  during solo practice is of a different character than during rehearsals of a symphony orchestra. When playing solo, high sound pressure level is maintained for longer periods of time compared to a symphony orchestra. For both of these activities,  $L_{Aeq,1s}$  values often exceed 90 dB. For the musician playing the percussion (Fig. 1f), the largest exposure took place during a Big Bang rehearsal, where the level often exceeded 100 dB. As in the case of French horn, during solo practice of percussionist,  $L_{Aeq,1s}$  has high values for longer periods of time, than during orchestra rehearsals. The graph concerning the double-bass player (Fig. 1g) shows that  $L_{Aeq,1s}$  values to which a bassist is exposed are lower than those of other instruments and only occasionally exceed 90 dB. The lowest values of  $L_{Aeq,1s}$  for double bass occur during a chamber ensemble rehearsal, and the highest during symphony orchestra rehearsal.

Based on the recorded values of  $L_{Aeq,1s}$  and on duration of sound exposure for each instrument, a daily sound exposure level,  $L_{EX,8h}$ , was determined in accordance with ISO 9612 standard [11]:

$$L_{EX,8h} = L_{Aeq,T} + 10 \lg(T_e / T_0) \quad [\text{dB}] \quad (1)$$

where  $T_e$  is the effective duration, in hours, of the working day,  $L_{Aeq,T}$  [dB] is the A-weighted equivalent continuous sound pressure level for  $T_e$ , and  $T_0$  is reference duration of 8 hours.

Obtained data are presented in Tab. 1. Effective duration  $T_e$  for each instrument is the total time of recording, i.e. the sum of time of every measured event (not the continuous time from start to end of the day). The A-weighted equivalent continuous sound pressure level  $L_{Aeq,T_e}$  is calculated over the entire measurement duration  $T_e$ . In Tab. 1 there is also information about activities that took place during the day in question for each instrument. “Individual” – meaning solo practice, “chamber” – playing in chamber ensembles, “SO” – playing in Symphony Orchestra, “WO” – Wind Orchestra, “Big Band” – playing in Big Band ensemble. The determined daily sound exposure levels ( $L_{EX,8h}$ ) are given in the last column.

Table 1. Type of music activities during the day of measurement, A-weighted equivalent continuous sound pressure level  $L_{eq,T_e}$ , effective duration  $T_e$ , and daily sound exposure level  $L_{EX,8h}$  determined for each instrument

Instrument	Activities	$L_{eq,T_e}$ [dB]	$T_e$ [h, min]	$L_{EX,8h}$ [dB]
Flute	individual, chamber	102.9	5 h 13 min	101.0
Clarinet	individual, chamber, WO	91.4	5 h 29 min	89.8
Trumpet	individual, SO, Big Band	96.1	4 h 2 min	93.1
Trombone	individual, WO, Big Band	93.9	7 h 54 min	93.8
French horn	individual, SO	86.8	5 h 43 min	85.3
Percussion	individual, SO, Big Band	94.1	5 h 23 min	92.4
Double-bass	individual, chamber, SO	78.7	4 h 43 min	76.4

Conducted measurements show that on a typical university day, students playing flute, clarinet, trumpet, trombone, horn, and percussion are exposed to sounds that exceed the permissible limit of 85 dB during the day. Although musicians usually spend about 5 hours at university, the sound pressure level during this time is so high that the  $L_{EX,8h}$  is exceeded. Musicians playing the flute are exposed to the highest A-weighted sound pressure levels, which during their playing often significantly exceeds 100 dB and the level of  $L_{EX,8h}$  is 101 dB. Assuming a 3-dB exchange rate in accordance with the safety guidelines, the safe working time at this level is less than 15 minutes. For musicians playing the trumpet, trombone and percussion,  $L_{EX,8h}$  exceeds 90 dB. The results are consistent with the literature, which indicates that the greatest exposure to sounds is for the wind instruments players. For musicians playing the clarinet and French horn,  $L_{EX,8h}$  is also above the 85 dB limit. The only instrument for which the daily sound

exposure level does not exceed the permissible limit is the double bass. Although the exposure time is comparable to other instruments, the A-weighted sound pressure level generated by the double bass is low and there is no risk of over-exposure to sound.

#### 4. Conclusions

Students of the Fryderyk Chopin University of Music playing the flute, clarinet, trumpet, trombone, French horn, and percussion in the course of daily activities are exposed to sounds of a level exceeding the 85 dB limit specified in the standards. Such high levels of sound can affect musicians' hearing and can lead to temporary and permanent noise induced hearing loss. The greatest exposure to sound occurs when playing the flute, in which case the daily sound exposure level ( $L_{EX,8h}$ ) is 101 dB. The permissible limit is also exceeded for wind instruments: clarinet 89.8 dB, trumpet 93.1 dB, trombone 93.8 dB, French horn 85.3 dB, and percussion 92.4 dB. Among the studied instruments, only musicians playing on double-bass are not at the risk of excessive exposure to sound. This results indicate that daily sound exposure levels are highest for musicians playing wind instruments, which is consistent with the results of the studies cited in the introduction. The value of the daily sound exposure level depends on the type of instrument as well as the type of musical activity. Some instrumentalists are exposed to the highest A-weighted sound pressure level during solo practicing (flute, trumpet, horn), others (clarinet, trombone, drums, double bass) have the highest levels when playing in chamber ensembles and orchestras (symphony orchestra and Big Band). For the musicians playing the clarinet, the highest levels are in chamber ensembles, for trombonists and percussionists when playing in the Big Band, and in the case of double bass – when playing in a symphony orchestra. The presented analyses are based on a measurement carried out for a small group of people, so their character is rather exemplary. However, it can be seen that full-day noise dosimetry provides information about musicians' actual exposure to sounds during a variety of daily activities, unlike studies that consider only one type of noise. It is planned to continue the study for a larger group of musicians.

#### Acknowledgments

This work was supported by the statutory grant 504/04064/1034/40.00 from the Warsaw University of Technology.

#### References

1. H. M. Laitinen, E. M. Toppila, P. S. Olkinuora, K. Kuisma, *Sound Exposure Among the Finnish National Opera Personnel*, *Applied Occupational and Environmental Hygiene*, **18** (2003) 177 – 182.
2. J. Lee, A. Behar, H. Kunov, W. Wong, *Musicians' noise exposure in orchestra pit*, *Applied Acoustics*, **66** (2005) 919 – 931.
3. J. H. Schmidt, E. R. Pedersen, P. M. Juhl, J. Christensen-Dalsgaard, T. D. Andersen, T. Poulsen, J. Bælum, *Sound Exposure of Symphony Orchestra Musicians*, *Ann. Occup. Hyg.*, **55** (2011) 893 – 905.

4. E. Kozłowski, J. Zera, R. Mlynski, *Sound levels on stage during performances of music school symphony and wind symphony orchestras*, The 20th International Congress on Sound and Vibration, Bangkok, Thailand 2013.
5. M. Pawlaczyk-Łuszczynska, A. Dudarewicz, M. Zamojska, M. Śliwinska-Kowalska, *Evaluation of Sound Exposure and Risk of Hearing Impairment in Orchestral Musicians*, International Journal of Occupational Safety and Ergonomics, **17** (2011) 255 – 269.
6. M. Jasinski, A. Pietrzak, J. H. Shin, J. Zera, *Exposure of Music Students to Sound in Large Music Ensembles*, Audio Engineering Society, 2015.
7. J. D. Royster, L. H. Royster, M. C. Killion, *Sound exposures and hearing thresholds of symphony orchestra musicians*, The Journal of the Acoustical Society of America, **89** (1991) 2793 – 2803.
8. D. McBride, F. Gill, D. Proops, M. Harrington, K. Gardiner, C. Attwell, *Noise and the classical musician.*, BMJ, **305** (1992) 1561 – 1563.
9. J. Lee, A. Behar, H. Kunov, W. Wong, *Noise exposure of opera orchestra players*, Canadian Acoustics, **31** (2003) 78 – 79.
10. K. H. Smith, T. B. Neilsen, J. Grimshaw, *Full-day noise exposure for student musicians at Brigham Young University*, Boston, Massachusetts, 2017.
11. EN ISO 9612:2009. *Acoustics - Determination of Occupational Noise Exposure - Engineering Method*, Comite Europeen de Normalisation.
12. A. P. Pietrzak, J. Zera, G. Makarewicz, *The Risk of Asymmetrical Noise Exposure Among Music Students*, Joint Conference - Acoustics, IEEE, Ustka, 2018.

## Evaluation of the Transmission and the Scattering Matrix Applicability to the Mufflers Analysis

Anna SNAKOWSKA

*AGH University of Science and Technology, Faculty of Mechanical Engineering  
and Robotics, Krakow, Poland, anna.snakowska@agh.edu.pl*

Łukasz GORAŁD

*AGH University of Science and Technology, Faculty of Mechanical Engineering  
and Robotics, Krakow, Poland, lukasz.gorald@agh.edu.pl*

Jerzy JURKIEWICZ

*AGH University of Science and Technology, Krakow, Poland,  
jerzy.jurkiewicz@agh.edu.pl*

### Abstract

The analysis of the acoustical systems can be carried out based on a number of different formalisms, of which applied frequently are the transfer matrix formalism, in which the chosen state variables are the sound pressure  $p$  and the sound velocity  $v$ , and the scattering matrix formalism adopting the sound pressures  $p^+$  and  $p^-$  and the sound velocities  $v^+$  and  $v^-$  of waves propagating through an element in both directions. Even though, they are mathematically equivalent, i.e. there exists the unequivocal transformation from one to another there are some advantages and disadvantages in applying one or the other to analyse mufflers or other acoustic system, especially when the propagation of a multimode wave is assumed and numerical calculations are indispensable. In the paper the transformation between the formalisms has been derived and applied to analyse the phenomena at a junction between the tail pipe and the chamber and also in mufflers composed of one or two chambers. The more flexible for numerical calculations seems the scattering matrix formalism, especially when the number of propagating modes differs on both sides of a junction. On the other hand the transmission matrix formalism is suitable for analysing systems constituting a cascade. The sources of the advantages and disadvantages of both formalisms are explained. The results obtained can be helpful in the effective design of silencers with specific properties.

**Keywords:** transmission matrix, scattering matrix, multimode wave, silencers and mufflers

### 1. Introduction

The examination of the properties of acoustic systems, in particular silencers, can be carried out using the transmission, impedance, admittance or the scattering matrix. The use of one of these formalisms depends on the choice of the state variables. In each of these formalisms, a relationship is sought between the state variables. In the study of muffler attenuation properties, these are the values of state variables on selected cross-sections of inlet and outlet pipes. In the case of analysing the wave propagation through all the elements of the muffler in the low frequencies approximation, i.e. assuming only the plane wave propagation, these matrices have the dimension of  $2 \times 2$ . However, the analysis carried out in this way is correct only if the Helmholtz number for each of the muffler elements does not exceed the cut-off frequency of the first waveguide mode.

In the case of silencers with cylindrical symmetry, it is the value of  $ka = 1.84$  for axial asymmetric excitation and  $ka = 3.83$  for symmetric excitation. However, even if these conditions are fulfilled, the analysis will be subject to some errors resulting from neglecting the effects of the near field on the discontinuities of the boundary condition, *i.e.* for example on the junction of the inlet/outlet pipe with the chamber.

Formalities previously mentioned are mathematically equivalent, that is there is unequivocal transformation of one matrix to another. However, they are not equivalently convenient to adopt, especially if the muffler testing goes beyond the plane wave approximation, accounting the propagation of the higher waveguide modes. For cylindrical symmetry, this will be the so-called Bessel modes.

The article analyses two formalisms in terms of their advantages and disadvantages -- the formalism of the  $\mathbf{T}$  transmission matrix and the  $\mathbf{S}$  scattering matrix. The analysis was carried out for silencers consisting of one or more chambers with different radii. For a given frequency, the number of propagating modes (cut-on modes) depends on the radius of an element, so it can change from one element to another. This fact, as will be shown, significantly impedes the analysis by means of the transmission matrix  $\mathbf{T}$ . On the other hand, the  $\mathbf{T}$  matrix is particularly convenient to describe the cascade system of which it is a product of individual elements' matrices [1]. There scattering matrix  $\mathbf{S}$  do not have this convenient property, but in turn it can be adopted without problems for a variable number of modes propagating in individual elements of the silencer [2, 3].

The silencers have been subject of numerous research and publications. At first they have been examined within the plane wave approximation [1, 4-6], however of increasingly complex geometry [5, 6] and with application of additional sound attenuating materials such as linings or perforates surfaces [7, 8]. Constructing mufflers of large radii or adapted to suppress high frequencies caused the necessity of analyzing the propagation of modes other than the plane wave [2, 3, 9-11]. Frequently, the multimode wave has been assumed in chambers and the incident plane wave in the inlet pipe [9, 11]. To adjust the acoustic pressure and the velocity at junctions between subsequent elements the mode matching method (MMM) has been applied [9-12].

## 2. Transformation of the transmission matrix into the scattering matrix

Let us consider a simple muffler composed of two tail pipes (inlet and outlet) of the same radius  $a$  connected to the expansion chamber of the radius  $b$  in which a multimode wave propagates (Fig. 1).

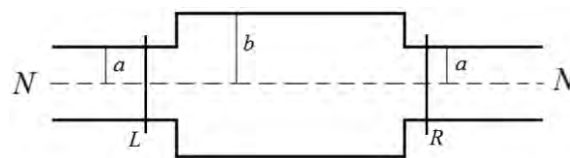


Figure 1. Schematic representation of a simple muffler

The relation between the modal pressures and modal velocities on cross sections depicted as L (left) and R (right) are given by the transfer matrix.

$$\begin{bmatrix} \mathbf{P}^R \\ \mathbf{V}^R \end{bmatrix} = \mathbf{T} \times \begin{bmatrix} \mathbf{P}^L \\ \mathbf{V}^L \end{bmatrix} \quad (1)$$

where

$$\begin{bmatrix} \mathbf{P}^R \\ \mathbf{V}^R \end{bmatrix} = \begin{bmatrix} \mathbf{P}_1^R \\ \vdots \\ \mathbf{P}_N^R \\ \mathbf{V}_1^R \\ \vdots \\ \mathbf{V}_N^R \end{bmatrix} \quad \begin{bmatrix} \mathbf{P}^L \\ \mathbf{V}^L \end{bmatrix} = \begin{bmatrix} \mathbf{P}_1^L \\ \vdots \\ \mathbf{P}_N^L \\ \mathbf{V}_1^L \\ \vdots \\ \mathbf{V}_N^L \end{bmatrix} \quad (2)$$

The elements of the muffler form a cascade, and so the transfer matrix of the system is a product of transfer matrices representing subsequent elements, which are: a straight duct of a given length, a junction from a tail pipe to an expansion chamber, then through the expansion chamber to the junction on the right and finally along the exhaust pipe. The mathematical formula of this cascade transfer matrix is

$$\mathbf{T} = \mathbf{T}_1 \times \mathbf{T}_{M1} \times \mathbf{T}_2 \times \mathbf{T}_{M2} \times \mathbf{T}_3 \quad (3)$$

where  $\mathbf{T}_1 = \mathbf{T}_3$  and  $\mathbf{T}_{M1} = \mathbf{T}_{M3}^{-1}$ . The transfer matrices  $\mathbf{T}_1$ ,  $\mathbf{T}_2$  and  $\mathbf{T}_3$  express the change of phase of the wave travelling across a straight duct that is

$$\mathbf{T}_i = \begin{bmatrix} [\text{diag}(\cos(k_{z,n}^i l^i))] & [\text{diag}(i/Y_n^a \cdot \sin(k_{z,n}^i l^i))] \\ [\text{diag}(iY_n^i \cdot \sin(k_{z,n}^i l^i))] & [\text{diag}(\cos(k_{z,n}^i l^i))] \end{bmatrix} \quad (4)$$

where  $k_{z,n}^i$  is the axial wave number of the  $n$ -th mode in  $i$ -th straight element of length  $l^i$  and the transmission matrix of a junction  $\mathbf{T}_M$  is

$$\mathbf{T}_M = \begin{bmatrix} \mathbf{F} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} \end{bmatrix} \quad (5)$$

where the elements of the  $\mathbf{F}$  and  $\mathbf{G}$  matrices have been calculated applying the mode matching method [3] at the junction and are equal

$$F_{p,m} = \frac{1}{S} \iint_{S_B} \Psi_m^A(q, \varphi) \Psi_p^{B*}(q, \varphi) q dq d\varphi, \quad m = 1, 2, 3 \dots \quad (6)$$

$$G_{m,p} = \frac{1}{S} \iint_{S_B} \Psi_p^B(q, \varphi) \Psi_m^{A*}(q, \varphi) q dq d\varphi, \quad m = 1, 2, 3 \dots \quad (7)$$

where  $\Psi_m^A(q, \varphi)$  and  $\Psi_p^B(q, \varphi)$  are the mode shape functions of the tail pipes and the chamber, respectively.

The scattering matrix describes the relation between the ingoing and outgoing modal pressures of the element under consideration („black box”).

$$\begin{bmatrix} \mathbf{P}^{L-} \\ \mathbf{P}^{R+} \end{bmatrix} = \begin{bmatrix} \mathbf{S}^{11} & \mathbf{S}^{12} \\ \mathbf{S}^{21} & \mathbf{S}^{22} \end{bmatrix} \times \begin{bmatrix} \mathbf{P}^{L+} \\ \mathbf{P}^{R-} \end{bmatrix} \quad (8)$$

where  $\mathbf{P}^{L+}$  and  $\mathbf{P}^{R-}$  represent the one column matrices of the in-going waves modal pressures on the selected cross sections of the muffler (*cf.* Fig. 1), that is

$$\begin{bmatrix} \mathbf{P}^{L-} \\ \mathbf{P}^{R+} \end{bmatrix} = \begin{bmatrix} P_1^{L-} \\ \vdots \\ P_N^{L-} \\ P_1^{R+} \\ \vdots \\ P_N^{R+} \end{bmatrix} \quad \begin{bmatrix} \mathbf{P}^{L+} \\ \mathbf{P}^{R-} \end{bmatrix} = \begin{bmatrix} P_1^{L+} \\ \vdots \\ P_N^{L+} \\ P_1^{R-} \\ \vdots \\ P_N^{R-} \end{bmatrix} \quad (9)$$

On the left side the ingoing waves propagate in the positive direction of the muffler axis, while on the right side cross section in the opposite direction and that determines the signs „+” and „-”.

Derivation of the transfer matrix allows calculating the scattering matrix by means of the linear transformation. The procedure of transforming the transfer matrix  $\mathbf{T}$  into the scattering matrix  $\mathbf{S}$  is composed of the following mathematical operations: at first, the one column matrices of the total acoustic pressure  $\mathbf{P}^R$  and  $\mathbf{P}^L$  are decomposed into the matrices of the ingoing and outgoing pressures  $\mathbf{P}^{R+}$  and  $\mathbf{P}^{R-}$ , and  $\mathbf{P}^{L+}$  and  $\mathbf{P}^{L-}$ . Similarly with the acoustic velocity, introducing the matrix of the modal admittances  $\mathbf{Y}^a$ , where  $y_m = k_{z,m}/k\rho_0 c$

$$\begin{bmatrix} \mathbf{P}^R \\ \mathbf{V}^R \end{bmatrix} = \underbrace{\begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}}_{\mathbf{T}} \times \begin{bmatrix} \mathbf{P}^L \\ \mathbf{V}^L \end{bmatrix} \quad (10)$$

$$(\mathbf{P}^{R+} + \mathbf{P}^{R-}) = T_{11} \times (\mathbf{P}^{L+} + \mathbf{P}^{L-}) + T_{12} \times \mathbf{Y} \times (\mathbf{P}^{L+} - \mathbf{P}^{L-}) \quad (11)$$

$$\mathbf{Y} \times (\mathbf{P}^{R+} - \mathbf{P}^{R-}) = T_{21} \times (\mathbf{P}^{L+} + \mathbf{P}^{L-}) + T_{22} \times \mathbf{Y} \times (\mathbf{P}^{L+} - \mathbf{P}^{L-}) \quad (12)$$

$$\mathbf{Y}^a = \begin{bmatrix} y_1 & & 0 \\ & \ddots & \\ 0 & & y_N \end{bmatrix} \quad (13)$$

This, in turn, allows separating the expressions for the in-going and out-going waves



$$\begin{aligned} \mathbf{P}^{R+} + \mathbf{P}^{R-} &= \underbrace{(\mathbf{T}_{11} + \mathbf{T}_{12} \times \mathbf{Y})}_{\mathbf{X}^+} \times \mathbf{P}^{L+} + \underbrace{(\mathbf{T}_{11} - \mathbf{T}_{12} \times \mathbf{Y})}_{\mathbf{X}^-} \times \mathbf{P}^{L-} \\ \mathbf{P}^{R+} - \mathbf{P}^{R-} &= \underbrace{\mathbf{Y}^{-1} \times (\mathbf{T}_{21} + \mathbf{T}_{22} \times \mathbf{Y})}_{\mathbf{W}^+} \times \mathbf{P}^{L+} + \underbrace{\mathbf{Y}^{-1} \times (\mathbf{T}_{21} - \mathbf{T}_{22} \times \mathbf{Y})}_{\mathbf{W}^-} \times \mathbf{P}^{L-} \end{aligned} \quad (14)$$

To simplify the notation the matrices  $\mathbf{X}$  and  $\mathbf{W}$  were introduced

$$\begin{aligned} \mathbf{P}^{R+} + \mathbf{P}^{R-} &= \mathbf{X}^+ \times \mathbf{P}^{L+} + \mathbf{X}^- \times \mathbf{P}^{L-} \\ \mathbf{P}^{R+} - \mathbf{P}^{R-} &= \mathbf{W}^+ \times \mathbf{P}^{L+} + \mathbf{W}^- \times \mathbf{P}^{L-} \end{aligned} \quad (15)$$

what leads to the final derivation of the scattering matrix with the following sub-matrices as a result of some simple mathematical transformation

$$\begin{aligned} \mathbf{S}^{11} &= -(\mathbf{X}^- - \mathbf{W}^-)^{-1} \times (\mathbf{X}^+ - \mathbf{W}^+) \\ \mathbf{S}^{12} &= 2 \cdot (\mathbf{X}^- - \mathbf{W}^-)^{-1} \\ \mathbf{S}^{21} &= \frac{1}{2} \left( (\mathbf{X}^+ + \mathbf{W}^+) - (\mathbf{X}^- + \mathbf{W}^-) \times (\mathbf{X}^- - \mathbf{W}^-)^{-1} \times (\mathbf{X}^+ - \mathbf{W}^+) \right) \\ \mathbf{S}^{22} &= (\mathbf{X}^- + \mathbf{W}^-) \times (\mathbf{X}^- - \mathbf{W}^-)^{-1} \end{aligned} \quad (16)$$

where

$$\mathbf{X}^+ = \mathbf{T}_{11} + \mathbf{T}_{12} \times \mathbf{Y} \quad \mathbf{X}^- = \mathbf{T}_{11} - \mathbf{T}_{12} \times \mathbf{Y} \quad (17)$$

$$\mathbf{W}^+ = \mathbf{Y}^{-1} \times (\mathbf{T}_{21} + \mathbf{T}_{22} \times \mathbf{Y}) \quad \mathbf{W}^- = \mathbf{Y}^{-1} \times (\mathbf{T}_{21} - \mathbf{T}_{22} \times \mathbf{Y}) \quad (18)$$

Experimental eduction of the scattering matrix is especially easy applying the single mode synthetiser [13] as the necessity of calculating the invers matrix is avoided.

### 3. Application to the muffler description

As has been mentioned in the Introduction, for a given frequency of excitation, the number of cut-on modes depends on the radius of an element and grows with the radius. Thus, the number of propagating modes  $N$  can change from one element to another. This fact significantly impedes the analysis by means of the transfer matrix  $\mathbf{T}$  because the matrix is then of a rectangular, not of a square shape and the procedure demands derivation of the reverse matrix  $\mathbf{T}^{-1}$ . On the other hand, the  $\mathbf{T}$  matrix is particularly convenient to describe the cascade system. In such a case the matrix of a whole system is a product of individual elements' matrices [1]. The scattering matrix  $\mathbf{S}$  do not have this convenient property, but in turn it can be adopted without problems for a variable number of modes propagating in individual elements of the silencer [2, 3], as has been illustrated in Fig. 2.

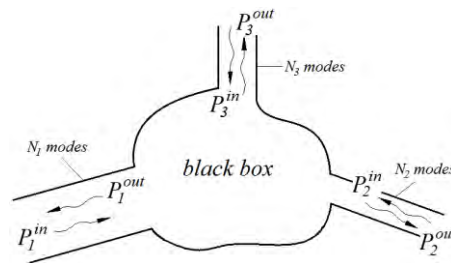


Figure 2. Scheme of a not-cascade multi-port with three connected straight ducts of different radii

The results of numerical calculations carried out according to the analytical solutions are presented below - in Fig. 3 for a muffler with one expansion chamber and in Fig. 4 – for a muffler composed of two expansion chambers. To minimise errors resulting from neglecting the effects of the near field on the junctions of the inlet/outlet pipe with the chambers, when applying the mode matching method it has been taken into account from a few up to several dozen of the excited modes, depending on the difference in radii of the joint elements. Finally, the transmission loss TL has been calculated adopting the transmission matrix (dotted and interrupted lines) and the scattering matrix method (continuous line). The numbers below the scheme of the muffler depict the relative radiuses (for example R: 1, 2, 1 and R: 1, 3, 1 in Fig. 3) and length of the expansion chamber and the connecting pipe (for example L: 5, 4, 7 and L: 6, 4, 2.5 in Fig. 4).

The Helmholtz number is chosen for the inlet/outlet pipe, so up to  $ka = 1.92$  in Fig. 3a and  $ka = 1.3$  in Fig. 3b when the first radial Bessel mode becomes cut-on the plane wave is the only mode propagating through the muffler.

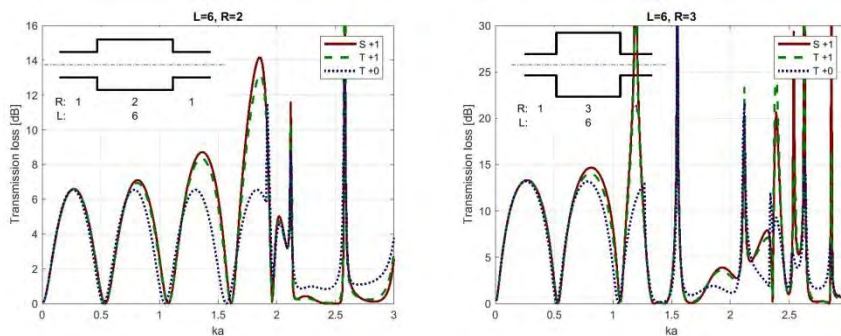


Figure 3. Transmission loss as a function of the Helmholtz number  $ka$  in the inlet pipe calculated by means of the transmission matrix of a size equal to the number of the cut-on modes (T+0 - dotted line) and exceeding it by one (T+1 - interrupted line) and the scattering matrix of a size exceeding by one the number of the cut-on modes (S+1 - continuous line)

However, the discrepancies between the results accounting only for the plane wave and neglecting the propagation of the attenuated first Bessel mode are visible for the Helmholtz number much smaller than these based on the cut-on frequencies. It means that having regard to the propagation of at least one cut-off mode the calculations better reflect the real muffler sound attenuation properties. According to some additional results, not presented in the paper, taking into consideration a larger number of the cut-off modes do not change the transmission loss significantly. It comes from the fact, that each of the subsequent mode is attenuated more efficiently and is extinguished practically in the vicinity of the junction.

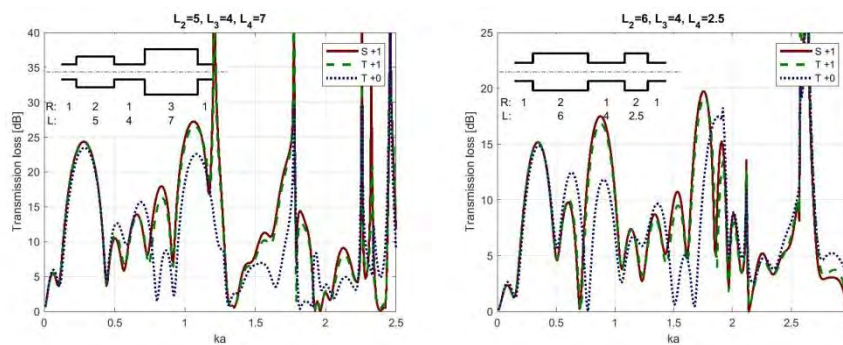


Figure 4. Transmission loss of a muffler composed of two expansion chambers as a function of the Helmholtz number  $ka$  in the inlet pipe calculated by means of the transmission matrix of a size equal to the number of the cut-on modes ( $T+0$  - dotted line) and exceeding it by one ( $T+1$  - interrupted line) and the scattering matrix of a size exceeding by one the number of the cut-on modes ( $S+1$  - continuous line)

#### 4. Conclusions

Research on the mufflers sound attenuation properties is still a subject of intensive scientific investigations [14] carried out by means of analytical, numerical and experimental methods. The appearance of computers with high computational power as well as the effective numerical methods has significantly intensified the numerical examination of acoustic systems with complex geometry, including silencers. However, the analytical methods are still the subject of research because of their high level of generality.

An example of a new analytical approach to the description of acoustic systems is the adaptation of the electric network method, known in acoustics as the multiport method, and within it the methods of the dispersion matrix. As shown in the article, this method is more general than the transmission matrix method commonly used in the silencer theory and can also be used for the variable number of cut-on modes in individual elements of the damper. On the other hand, the transmission matrix is particularly convenient for describing cascade systems. To sum up, the choice between the scattering and transmission matrix should depend on the property and structure of the investigated acoustic system.

## References

1. M. L. Munjal, *Acoustics of ducts and mufflers*, Willey 2 ed., New York 2014.
2. A. Snakowska., K. Kolber., Ł. Gorazd, J. Jurkiewicz, Derivation of an acoustic two-port scattering matrix for a multimode wave applying the single-mode generator, *IEEE Digital Xplore Library*, (2018) 294 – 298.
3. A. Sitel, J-M Ville, F. Felix, *Multiload procedure to measure the acoustic scattering matrix of a duct discontinuity for higher order mode propagation conditions*, J. Acoust. Soc. Am., **120** (2006) 2478.
4. M. L. Munjal, *Plane wave analysis of side inlet/outlet chamber mufflers with mean flow*, Applied Acoustics, **52** (1997) 165 – 175.
5. T. Kar, M. L. Munjal. *Generalized analysis of a muffler with any number of interacting ducts*, Journal of Sound and Vibration, **285** (2005) 585 – 596.
6. C. D. Gaonkar, M. L. Munjal, *Theory of the double-tuned side-inlet side-outlet muffler*, Noise Control Engineering Journal, **66** (2018) 489 – 495.
7. T. Elnady, M. Abom, S. Allam, *Modeling perforates in mufflers using two-ports*, Journal of Vibration and Acoustics, **132** (2010).
8. M. L. Munjal, K. Narayana Rao, A. D. Sahasrabudhe, *Aeroacoustic analysis of perforated muffler components*, Journal of Sound and Vibration, **114** (1987) 173 – 188.
9. R. Glav, P. Regaudb, M. Abom, *Study of a folded resonator including the effects of higher order modes*, J. Sound. Vib., **273** (2004) 777 – 792.
10. R. Kirby, P. T. Williams, J. Hill, *A three dimensional investigation into the acoustic performance of dissipative splitter silencers*, J. Acoust. Soc. Am., **135** (2014) 2727 – 2737.
11. F. D. Denia, A. Selamet, M. J. Martínez, F. J. Fuenmayor, *Sound attenuation of a circular multi-chamber hybrid muffler*, Noise Control Engineering Journal, **56**(5) (2008) 356 – 364.
12. A. Snakowska, J. Jurkiewicz, Ł. Gorazd, *A hybrid method for determination of the acoustic impedance of an unflanged cylindrical duct for multimode wave*, J. Sound. Vib., **396** (2017) 325 – 339.
13. A. Snakowska, Ł. Gorazd, J. Jurkiewicz, K. Kolber, *Generation of a single cylindrical duct mode using a mode synthesiser*, Applied Acoustics, **114** (2016) 56 – 70.
14. M. L. Munjal, *Recent advances in muffler acoustics*, International Journal of Acoustics and Vibration, **18** (2013) 71 – 85.

## The Impact of the Damper Blade Position on the Generated Noise and Pressure Losses in Plenum Box with Swirl Diffusers

Kamil WÓJCIAK

*Institute of Power Engineering - Thermal Technology Branch "ITC" in Lodz,  
113 Dabrowskiego Street, 93-208 Lodz, Poland, [kamil.wojciak@itc.edu.pl](mailto:kamil.wojciak@itc.edu.pl)*

Patryk GAJ

*Institute of Power Engineering - Thermal Technology Branch "ITC" in Lodz,  
113 Dabrowskiego Street, 93-208 Lodz, Poland, [patryk.gaj@itc.edu.pl](mailto:patryk.gaj@itc.edu.pl)*

Joanna KOPANIA

*Lodz University of Technology,  
266 Piotrkowska Street, 90-924 Lodz, Poland, [joanna.kopania@p.lodz.pl](mailto:joanna.kopania@p.lodz.pl)*

Grzegorz BOGUSŁAWSKI

*Lodz University of Technology,  
266 Piotrkowska Street, 90-924 Lodz, Poland, [grzegorz.boguslawski@p.lodz.pl](mailto:grzegorz.boguslawski@p.lodz.pl)*

### Abstract

The article presents the effect of setting the throttle in the plenum box with swirl diffuser on the generated noise and the pressure losses. The sound power levels of the source was determined using the precision method in the reverberation room in accordance with the PN-EN ISO 3741:2011 standard. The examination was carried out for damper blade position set vertically and horizontally for full opening and at a 45 degree angle in both directions. In order to confirm the universality of the conducted experiment, the tests were carried out for two different face swirl diffuser with different shapes of air control blades and holes in the plate and for different flow rates. The multi-criteria optimization methods were used to select the best throttle position.

**Keywords:** HVAC, noise emission, sound power level, pressure loss, optimization

### 1. Introduction

The Polish standard PN-N-01307:1994 specifies the limit values for noise at workplaces due to its nuisance determined. It states that the equivalent sound level A for administrative rooms, design offices for theoretical work, data processing and other similar purposes should not exceed 55 dB [1]. One of the sources of noise in this type of rooms is ventilation [2, 3], mainly from grilles and diffusers [4, 5].

Smaller flow resistance in ventilation systems allows you to use less powerful fans. It allows to save money on purchase of a fan and reduce electricity expenses. In the case of a company with an extensive ventilation system, this can result in a large saving of money.

Manufacturers of ventilation products use various solutions to mount the damper blade in the plenum box, often without wondering if it is the best, because no such

studies have been carried out. The purpose of the tests is to check which damper blade position in the box is better, vertically or horizontally, and which setting direction it is better for a given setting. The noise and flow resistance were taken into account as the criterion of the best location.

## 2. Test object

The test object is insulated plenum box with dimensions of 600 x 600 mm and 380 mm height with two different swirl diffusers. The diffusers differed in the construction of the air control blades and the tilt angle in the plate. The diffuser A has a smaller tilt angle, fin under the blades and air guides. The blades from the B diffuser are smooth. A visual comparison of diffusers and blades is shown in Figure 1. All air control blades were set in one direction - the swirl was set to the right. The measurements were carried out for three volume flows: 600, 1000, 1400 m<sup>3</sup>/h. During measurements, the damper position and direction in the plenum box were changed, as shown in Figure 2. Set damper blade in vertical and horizontal directions for full opening and 45 degrees in both directions.

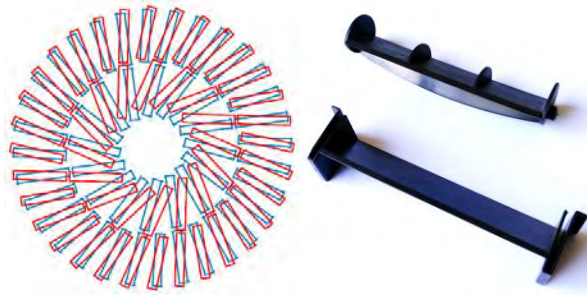


Figure 1. Differences in the shape of swirl diffusers holes and air control blades (blue holes and up blade form diffuser A, red holes and down blade from diffuser B)

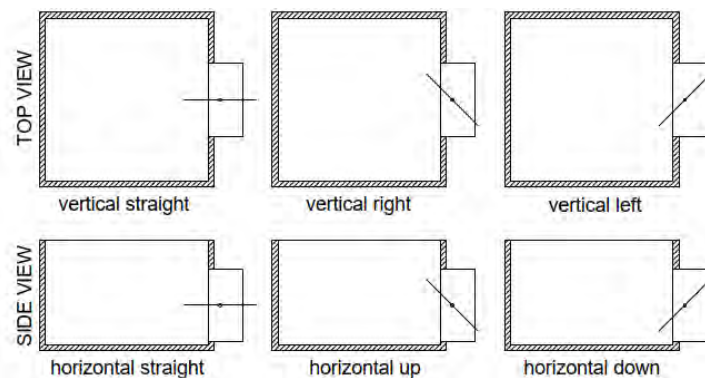


Figure 2. Tested damper blade position in plenum box was tested during measurements

### 3. Experiment

The measurements were made in a reverberation room (Figure 3) with a volume of  $237.0 \text{ m}^3$  and an area of  $231.5 \text{ m}^2$  which non-parallel walls. The diffusers connected to the centrifugal fan through three absorption silencers with a total length of 3.3 meters outside the chamber.

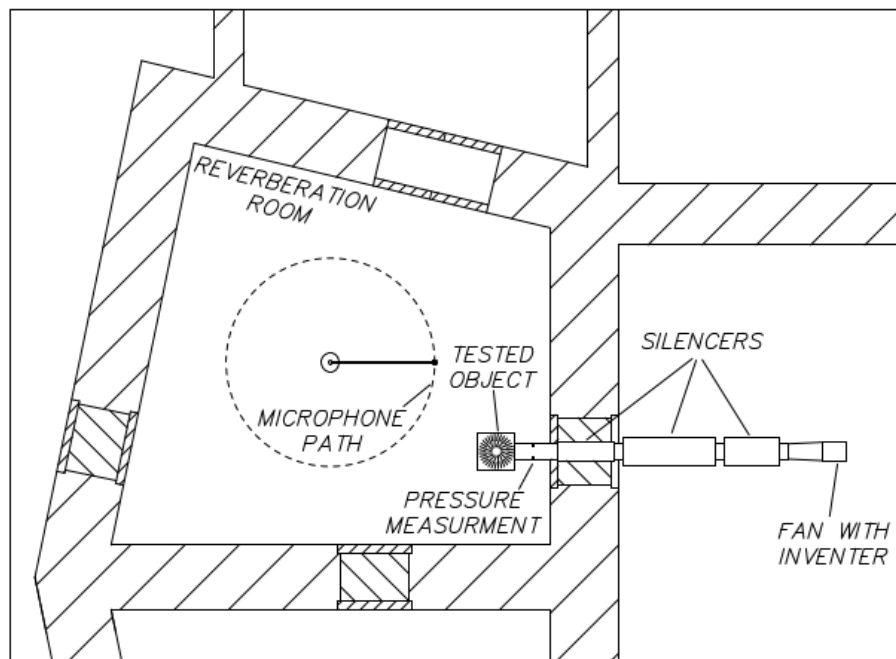


Figure 3. Reverberation room scheme

The volumetric flow was set by changing the rotation frequency of the fan motor. For this purpose, a three-phase inverter was connected to the motor. Flow velocity were measured using the Testo 420 balometer.

The static pressure drop on the expansion box together with the front plate was measured 300 mm on the channel in front of the damper in four evenly located points around the circumference and the environment. A differential electronic pressure transducer was used for this. The tested object in the reverberation room is shown in Figure 4.



Figure 4. Plenum box with swirl diffusers under test

The noise generated is determined by sound power, measured and calculated in accordance with PN-EN ISO 3741:2011 "Acoustics - Determination of sound power levels and acoustic energy levels of noise sources based on sound pressure measurements - Precision methods in reverberation chambers" [6]. The Nor 140 measuring set with the Nor850 software and the Nor265 rotary table was used for the measurements. The sound pressure was measured at twelve uniformly spaced points in a circle with a radius of 1.7 meters (circumference 10.7 m). Measured in 1/3 octaves in the range from 50 Hz to 10,000 Hz. The measurement time is set to 15 seconds. Shorter measurement time had no effect on standard deviation, and more points measurements reduced them. And in the case of background measurement time set to 30 seconds to determine the background correction  $K_1$ . Background noise measured for a stand without flow. Reverberation was measured for four omnidirectional loudspeaker settings with three microphone settings every 120 degrees. All calculations of the sound power level were made using a previously made calculation sheet.

Before and after all measurements were made, the background level was measured and calibrated using the Bruel & Kjaer 4231 calibrator. After measuring each setting, the temperature, relative humidity and atmospheric pressure necessary for calculating the sound power were recorded.

#### 4. Results

The results of measurements and calculations of the sound power level correlated with curve A and pressure loss are presented in the form of graphs for three volume flows in the Figure 5.



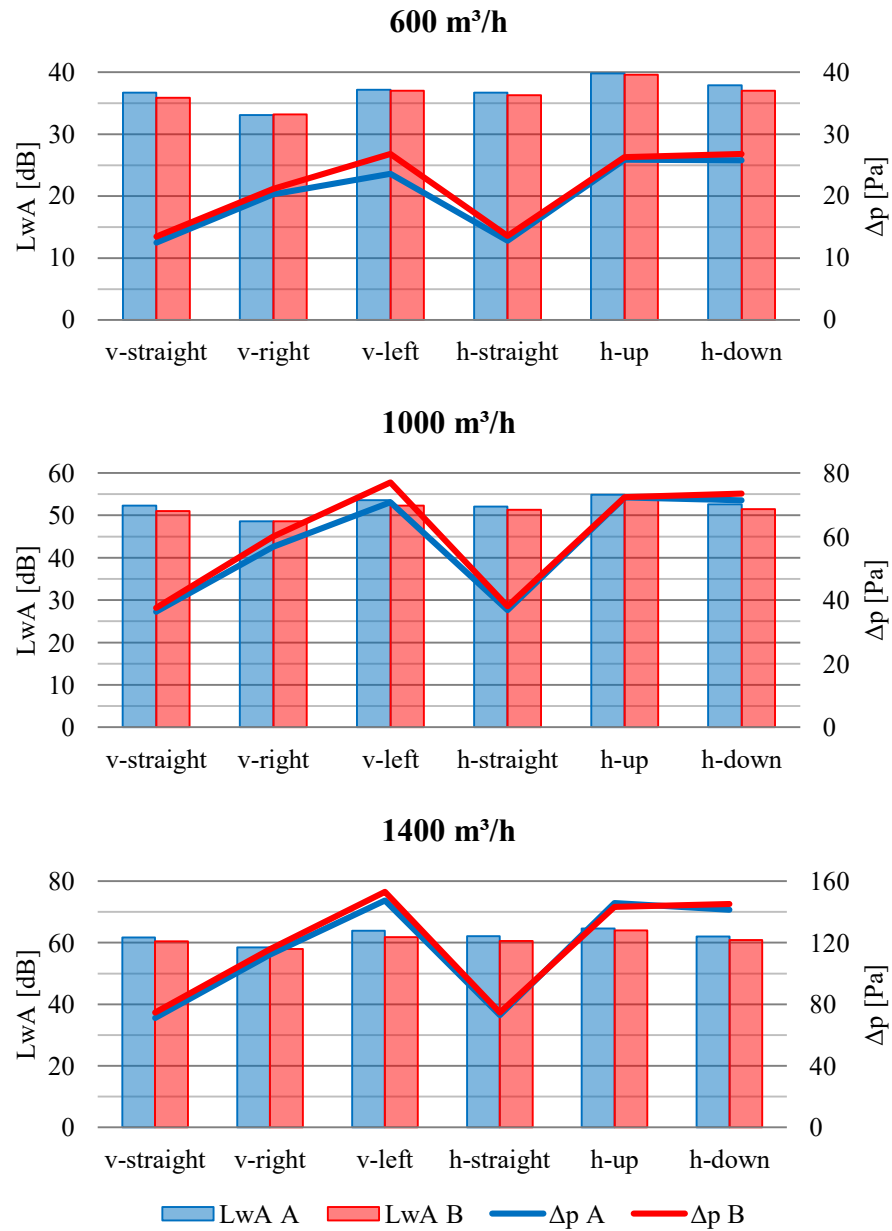


Figure 5. Results of measurements of sound power and pressure losses for various damper blade position in plenum box, volume flow and swirl diffusers (v means vertical and h horizontal)

The spectrum of the sound power level for the volume flow 1000 m<sup>3</sup>/h air diffuser B for various damper blade settings is shown in Figure 6. The right throttle setting reduces the noise level for medium and high frequencies. Changing to the left direction increases the high frequency levels. The down adjustment is similar to that for full opening, and turning to the up increase the noise levels in the whole range.

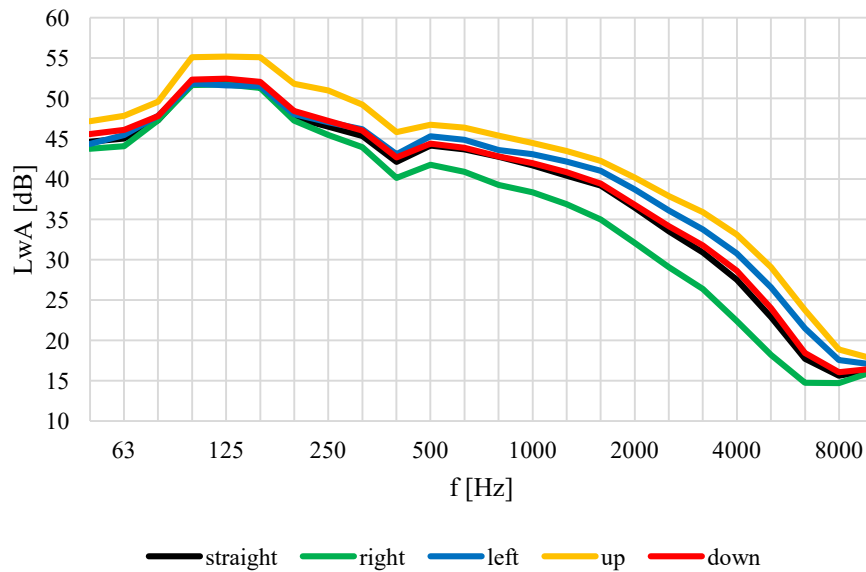


Figure 6. Acoustic power level spectrum for different blade settings with diffuser B and volume flow 1000 m<sup>3</sup>/h

Because values have different units and are of different order and scope, normalization is necessary [7]. The following formula was used for this purpose, which normalizes the value from 0 to 1 in a dimensionless form.

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

Where  $x_{\min}$  and  $x_{\max}$  are the largest and smallest value of  $x$ .

To determine the best point considering both parameters, we used the formula

$$F_{\text{glob}} = \sqrt{L_{\text{wA-norm}}^2 + \Delta p_{\text{norm}}^2} \quad (2)$$

This pattern determines the distance from the ideal point, which consists of the minimization of the component parameters.

The results of normalization and the global objective function calculations are presented in Table 1.

Table 1. Results after normalization and calculation of the global function

swirl diffuser	A			B		
600 m³/h						
	$L_{wA-norm}$	$\Delta p_{norm}$	$F_{glob}$	$L_{wA-norm}$	$\Delta p_{norm}$	$F_{glob}$
vertical-straight	0.54	0.00	0.54	0.42	0.00	0.42
vertical-right	0.00	0.58	0.58	0.00	0.58	0.58
vertical-left	0.61	0.83	1.03	0.59	1.00	1.16
horizontal-straight	0.54	0.02	0.54	0.48	0.01	0.48
horizontal-up	1.00	1.00	1.41	1.00	0.96	1.39
horizontal-down	0.72	0.99	1.22	0.59	1.00	1.16
1000 m³/h						
	$L_{wA-norm}$	$\Delta p_{norm}$	$F_{glob}$	$L_{wA-norm}$	$\Delta p_{norm}$	$F_{glob}$
vertical-straight	0.59	0.00	0.59	0.41	0.00	0.41
vertical-right	0.00	0.57	0.57	0.00	0.57	0.57
vertical-left	0.79	0.96	1.24	0.64	1.00	1.19
horizontal-straight	0.56	0.01	0.56	0.47	0.01	0.47
horizontal-up	1.00	1.00	1.41	1.00	0.88	1.33
horizontal-down	0.63	0.98	1.16	0.50	0.91	1.04
1400 m³/h						
	$L_{wA-norm}$	$\Delta p_{norm}$	$F_{glob}$	$L_{wA-norm}$	$\Delta p_{norm}$	$F_{glob}$
vertical-straight	0.52	0.00	0.52	0.42	0.00	0.42
vertical-right	0.00	0.54	0.54	0.00	0.53	0.53
vertical-left	0.89	1.00	1.34	0.63	1.00	1.18
horizontal-straight	0.59	0.02	0.59	0.43	0.00	0.43
horizontal-up	1.00	0.98	1.40	1.00	0.88	1.33
horizontal-down	0.57	0.92	1.09	0.48	0.90	1.02

## 5. Conclusions

Setting the damper blade in the plenum box with swirl diffusers not only affects the pressure loss, but also the noise emitted. This allows to increase the comfort of work and reduce costs by reducing the flow resistance.

Due to the noise, the best position of the damper blade is to close it in accordance with the swirling of the air control blades - in this case to the right. Appropriate setting

can give over 6 dB less noise. To achieve the lowest possible pressure loss, open the damper blade completely.

According to the method of multi-criteria optimization, it is best to fully open the damper, very close to the minimum there is a damper blade position to the right (less noise, but greater pressure losses).

Vertical throttle mounting is better than horizontal. In terms of noise emissions and pressure losses. Regulation in accordance with the swirl (in this case the right) is better. If the damper is set horizontally, it is better to regulate it by directing it down - pressure losses are similar, but it is quieter.

All the above-mentioned applications are independent of the volume flow and the front of the diffuser.

Appropriate mounting of the damper in the expansion box by the manufacturers and its proper control by the users allows to improve flow properties.

## References

1. PN-N-01307:1994, *Noise - Limit values for noise in the work environment - Requirements for performing measurements*. (PN-N-01307:1994, Hałas - Dopuszczalne wartości hałasu w środowisku pracy - Wymagania dotyczące wykonywania pomiarów).
2. *ASHRAE Handbook, HVAC Applications (SI), Chapter 48, Sound and Vibration Control*, American Society of Heating, USA, 2011.
3. Siemens Building Technologies, *Minimizing Excessive Sound in Ventilation System Design*, USA, 2009.
4. Ramalingam, Srinivasan, *Simulation and experimental validation of airborne and structure-borne noise transmission in hvac plenums*, Theses and Dissertations - Mechanical Engineering, 2012.
5. F. Qin, J. Chen, Z. Chen, *Acoustic characterization and prediction for fan-duct-plenum-room integrations*, *Applied Acoustics*, **67**(3) (2006) 201 – 229.
6. PN-EN ISO 3741:2011, *Acoustics - Determination of sound power levels and sound energy levels of noise sources using sound pressure - Precision methods for reverberation test rooms*.
7. M. Ostwald, *Basics of design optimization*, Poznań University of Technology Publishing House, Poznań, 2005. (M. Ostwald, *Podstawy optymalizacji konstrukcji*, Wydawnictwo Politechniki Poznańskiej, Poznań, 2005).

## Objective and Subjective Evaluation of Musical and Speech Recordings Transmitted by DAB+ System

Andrzej DOBRUCKI, Stefan BRACHMAŃSKI,  
Maurycy KIN

*Wrocław University of Science and Technology,  
Department of Acoustics and Multimedia, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław,  
andrzej.dobrucki@pwr.edu.pl, stefan.brachmanski@pwr.edu.pl,  
maurycy.kin@pwr.edu.pl*

### Abstract

The results of research on the sound quality of various kinds of music as well as speech signals transmitted via Digital Audio Broadcasting Plus system are presented. The results showed that bitrate values significantly influence the results of quality assessment, i.e. the overall audio quality as well as a timbre are dependent on the bitrate. The additional conclusion is that the CCR method is more accurate for sound assessment for higher bitrate values and this fact has been verified by standard deviation values of obtained results. The speech signals were additionally examined with PESQ method. The results have shown that the assumed quality of 4 MOS for speech could be achieved at 48 kbps. This fact was confirmed by both: subjective and objective research.

**Keywords:** sound quality, broadcasting, sound attributes

### 1. Introduction

The development of digital media technology finds the wide area of application. One of this application is Digital Audio Broadcasting (DAB) which allows not only a simple listening to the radio programs, but also some redundant-like information such as traffic, news, weather reports, city maps etc. It simply causes a growth of signal volume so it is necessary to apply some lossy compression of the transmitted signals to make all of the DAB advantages possible to work [1].

The final quality of audio signals is mainly determined by the parameters of the coding algorithms, codec quality and it is mostly independent from the parameters of transmission path in “traditional” analog meaning. It is important to evaluate how the coding as well as the transmission process influence the Quality of Service (QoS) and Quality of Experience (QoE). The first one is usually assumed as a synonymous for network performance which reflects the characteristics of devices and their parameters. However, the QoS does not reflect the real quality of signals perceived by the users, except for breaks in the signal, tracking errors etc. It was necessary to introduce a new means reflecting the impressions of listeners so the Quality of Experience can really describe the telecommunication channel from the user’s point of view.

The objective quality evaluation is based on a comparison between psychoacoustic representations of tested signals and these representations are created by the mathematical

model taking into account the various aspects of hearing process including the masking phenomenon. Such psychoacoustic model defines which components or information of sound could be perceived, or not, by human ear in very particular situations. It can be found that there is a good correlation between the sound quality evaluation measured with the both: objective and subjective methods [1 - 5] although these results refer to the overall quality of sound, only, and some details of auditory image important from the aesthetical point of view have not been examined.

The main aim of the paper is to find the minimal value of bitrate of DAB+ transmission which guarantees similar hearing impressions that obtained by listening to an original signal. The results may indicate the preferences of the average listener and can give directions in the field of sound production, especially prepared for radio programs.

## 2. Experiment

The research on the sound quality of music and speech signals transmitted via DAB+ was performed during the emission in Wroclaw, Poland. In the experiment, the subjective quality assessment has been provided for speech and musical signals. In addition, the objective PESQ method was used for speech examination. This musical signals' assessment has been provided with standard listening tests using a Comparison Category Rating (CCR) method [6-8]. The listening tests have been provided with the staff of experts containing twelve people, aged from 26 to 31 years old. The listeners had their hearing loss no more than 5dB with the reference to normal hearing, what had been confirmed previously with audiometric tests. The listeners had participated various listening tests previously and all of them have been working as recording engineers, recording and radio producers. Tests were performed in accordance with EBU and ITU recommendations [6, 7]. The sound material contained various kinds of musical styles, from classical (chamber and symphonic) to heavy rock, and all of 13 samples were typical for the profile of National Polish Radio. The subjects of research were an overall sound quality and sound color impression. The paradigm of stimuli for musical signals was based on CCR procedure [6]: the sound samples were presented in pairs, in randomized order. The subjects were asked to rate the impairment of the second sample, in relation to the first sample, on a seven-level scale, extending from -3 (much worse) to 3 (much better). The length of sample was about 15 seconds and the musical structures of the samples were taken into account during a sample preparation. For speech signals the ACR test methodology defined in ITU [6] was used: the speech samples were rated using a scale from 1 to 5 in descriptive intervals: from excellent (5) to bad (1) according to the listeners' judgments, without comparison to the pattern. The test lists of Polish language used in the subjective speech quality measurements were created in Department of Acoustics and Multimedia, Wroclaw University of Science and Technology [5]. Each list was divided into 10 groups each with 5 tasks.

Sound samples were recorded at the analog output of the consumer set DAB Sangean DPR-26 receiver so the signal is passed by transmission channel: multiplex – DAB transmitter – receiver. The musical samples were coded with HE AAC system [1], and

transmitted with six bitrates (136 kbps, 128 kbps, 96 kbps, 80 kbps, 64 kbps and 48 kbps) with sampling frequency of 48 kHz while the speech signals were coded with five bitrates (128 kbps, 96 kbps, 64 kbps, 48 kbps and 24 kbps). As reference signals, CD recordings with the identical samples to the broadcasted ones, have been used. In addition, the same samples have been tested after passed by a simulation channel of typical FM broadcasting. The ZOOM H4n PRO was used as the recording machine.

The signals have been stored on CDR and presented with the use of CD-player (Pioneer PD-201) and a pair of active loudspeakers-boxes (TLC Pro-AMS 1). The listening sessions were provided in the recording studio of Department of Acoustics and Multimedia at Wrocław University of Science and Technology.

### 3. Results and analysis

#### 3.1. *The overall sound quality of music*

The statistical treatment by the means of ANOVA testing with the statistical power set on the critical  $p$  value of 0.05 has shown that the results could be averaged over the signals for all listeners in a group referring to particular musical style. Moreover, the variances of results obtained for the chosen groups of sound material as well as for all bitrates have been found as homogeneous and confirmed by Bartlett test ( $\chi^2 < \chi_{\alpha}^2 = 5.99$  at  $\alpha = 0.05$ ). Thus, it allowed to average the results over the twelve listeners in testing group and their five repetitions for four types of the musical program. The results of the overall quality evaluation and their standard deviations for 95% confidence obtained with the use of CCR method for the distinguished kinds of music are presented in Fig. 1. The results of statistical treatment have indicated a significance effect of bit-rate values ( $p < 0.007$ ) and a musical style ( $p < 0.005$ ). It can be seen that the quality of broadcasted musical material gets worse when the bitrate value decreases. It should also be noted that this dependence is not linear, and the signals with the lowest bitrate used in the experiment featured the worst quality, with the CMOS equal to -2. Comparing those results with the standard FM broadcasting one can assume that the similar sound quality as for FM can be achieved at 80 kbps or 96 kbps bitrates depending on the musical style.

#### 3.2. *Sound color assessment of musical samples*

Fig. 2 presents the dependence of sound color impression obtained for different bit-rate values and the ranges of 95% confidence. Listeners' task was to evaluate the naturalness of sound color, or the change of this attribute. It can be seen that the changes of sound color get worse significantly at 48 kbps bitrate. Another interesting fact is that the sound color impression is almost the same for various kinds of musical signals for different bit-stream values, and again, this is unlike the spatial attributes of sound which featured the quality strongly dependent on types of signals [2, 3]. It means that the use of spectral band replication processing (with the HE AAC v. 1 system) makes the sound quality acceptable by the listeners even at the comparison paradigm.

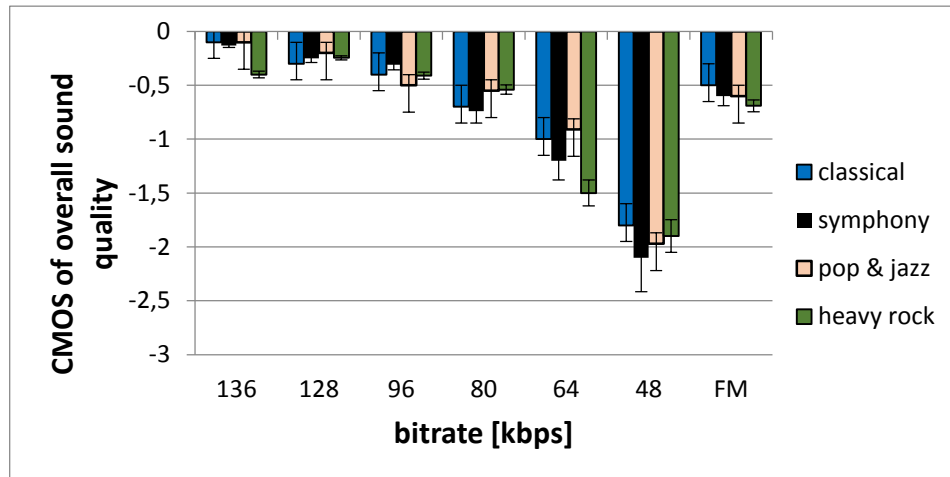


Figure 1. Results of overall sound quality obtained by the CCR method for different bitrates and various musical programs

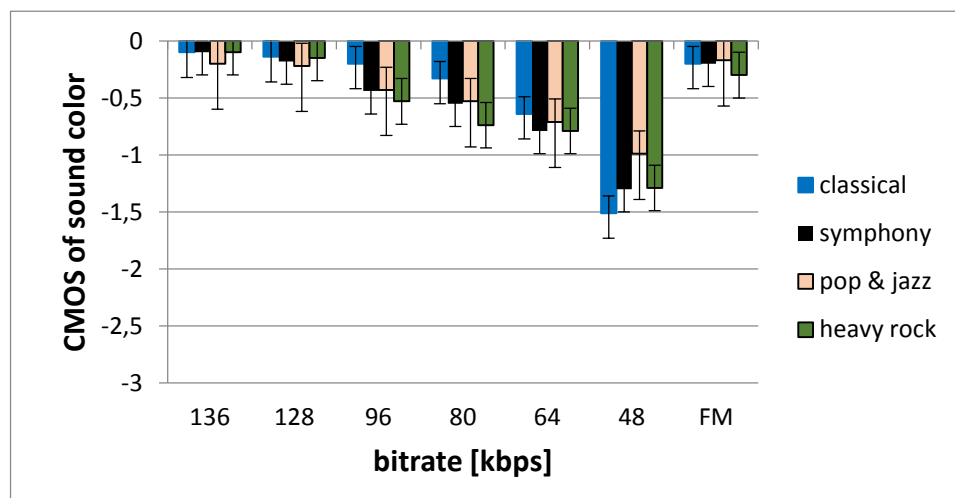


Figure 2. The dependence of sound color on the bitrate value

The results of statistical treatment indicated a significant effect of bitrate values ( $p < 0.007$ ) only, so the significant degradation of sound quality has been observed when the bitrate values decrease. It simply suggests that for some specific groups of listeners and for specific radio program profiles, DAB+ system should be set with special parameters. For popular and jazz music when the signals are not so complex, the satisfactory sound quality is guaranteed by the bitrate of 64 kbps with SBR processing. For



more complex sound structures and aesthetical requirements (as a symphonic music, for example), the good quality may be obtained when the bit-stream is of 96 kbps, or higher.

The results obtained for FM transmission have indicated that the degradation of timbre impression is comparable to those obtained at the highest bitrates. It can also be noted that the observed changes of sound color in dependence of the bitrate have their deviation being smaller in comparison to the evaluation of the presented spatial attributes of transmitted musical signals. This fact means that the changes of spectra are more univocally interpreted by listeners from the aesthetical point of view in comparison to the spatial attributes of sound [2-4].

### 3.3. Speech signal evaluation

The testing speech material contains the lists of sentences spoken by male (M) and female (F) voices. The listeners assessed coded speech samples and hidden reference signals. The PESQ and MOS (signed as “Brachmanski”) [5, 8], averaged over male and female voices, are shown in Fig. 3. This picture shows also results obtained in previous research for the MUX 1 on Gdańsk area (signed as “Gilski”) [9, 10].

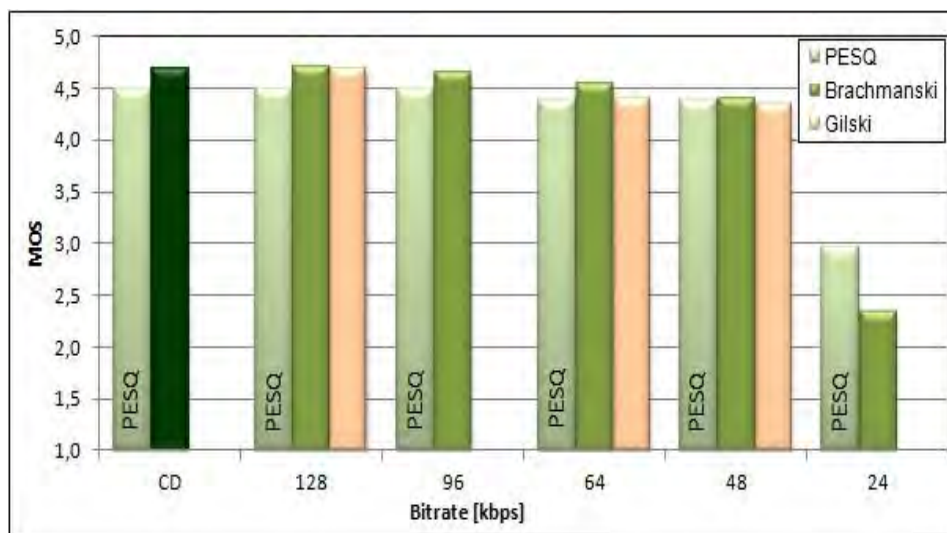


Figure 3. The MOS and PESQ results of speech signals transmitted for various bitrates

It can be shown, that for the speech signals, the minimal bit rate value which guarantees the subjective acceptable quality (4.0 MOS) has been obtained at 48 kbps bitrate. The results for this part of experiment show the good correlation between the results obtained by the objective and subjective methods of quality assessment, as this fact had been confirmed previously in the literature [3, 5]. The small shift in scores may be explained by the listeners' sensitivity for the changes in timbre of speaking voices,

but these changes are less important for the objective measurement than aspects of speech intelligibility and clarity, in a meaning of a cognitive comparison.

#### 4. Discussion

The results of overall quality assessment of musical signals have shown a similar tendency among the bit-rate values as the notes of timbre evaluation, so it may be concluded that the good sound quality can be achieved by the use of HE AAC v.1 system at 80 kbps bitrate, if an audio material does not include a very special stereophonic tricks. It may suggest that the return of creating so called “radio version” of musical piece is recommended.

As it has been reported in literature [11] the most important factor influencing the over-all audio quality comes from the sound color impression (timbre). The other sound attributes, as spatial impression, panoramic spread, a listener envelopment etc. are less significant [2, 12, 13], although the information of all of these factors may be distorted when some processing is not transparent. However, the universally preferred values for these attributes are unlikely to exist but if the reverberation level is too low it will be easy indicated and influencing the overall sound quality [12].

The results of overall quality assessment have shown also that the greatest differences for all kinds of music have been observed for the higher bitrates. On the base of these results it could be noted that the most degradation of sound quality might occur for signals with complex time/frequency structure, i.e. symphony as well as a heavy rock pieces. This fact means that the changes of spectra are more univocally interpreted by listeners from the aesthetical point of view.

For speech signals, the results have indicated that at bitrate of 48 kbps, or higher, the MOS index takes the value of 4.5. This index is only a little bit higher than PESQ values. This fact suggest that speech quality of broadcasted voices is almost the same as the quality of original signals.

It also should be noted that the dependence of the bitrate on the subjective evaluation is not linear. It clearly means that signals with complex spectra and the time structures can be taken into account and treated as a specific material, so some tools useful for so called simple signals (as a speech, for example) cannot be applied to more complex musical structures. The reflection of this complexity in time and spectral domains could be reduced when compression of data is relatively high. During the editing process of musical recording, some effects introduced for creating certain aesthetic impressions, for example dynamic compression, are accompanied by additional processing such as equalization, stereo enhancement or reverberations. For this reason, it is still unclear whether average listeners experience the claimed negative effects of signal processing in actual recorded music productions [14]. On the other hand, signal processing with multi-layered sound effects are widely applied in a process of mastering of popular music in an attempt to boost perceived loudness. The higher degrees of compression of dynamics, for example, have been used to make musical pieces to be perceived in many different situations and environments (noisy places, stereo cars, computer loudspeakers or

mp3 players) [15] and many daily-musical-consumers do not pay attention for the lower quality of music instead of focusing on the groove and emotional aspects of the text, without reference to the pattern.

The quality of the classical FM broadcasting is almost equal to the quality of DAB+ at the bitrate value of 80 kbps, or higher. For the lower values of bitrate (48 and 64 kbps) the FM transmission offers a higher sound quality, so the additional information offered by the digital broadcasting system is the advantage of DAB+ what may suggest that these bitrates may be used for speech-based programs when particular aesthetic aspects are not important.

## 5. Conclusions

If one assumed that the sound quality of tested musical programs has been assessed as a little bit degraded in comparison to the original quality (with CCR value of - 0.5) the sufficient sound quality can be unconditionally guaranteed with the bitrate values of 80 kbps, or more. It can be seen that these bit-rates assure the perception of the music at the satisfactory level.

The speech signals can be transmitted with lower bitrates (the very good quality has been obtained at 48 kbps in both: objective and subjective evaluations). Moreover, the increasing the bitrate value does not make the speech quality higher significantly in a DAB+ system.

The obtained results can suggest that it is necessary to prepare the musical recording in a particular way when they have to be reproduced in proper media taking into account the technical parameters of telecommunication channel. The similar preprocessing was widely applied in classical analogue radio broadcasting (dynamic compression as well as equalization) in order to achieve almost the same sensations when listening to the radio program as to recording.

## Acknowledgment

This work was supported by Wroclaw University of Science and Technology, grant no. 0401/0045/18, and National Center of Research and Development, grant no. PBS3/A3/19/2015.

## References

1. W. Hoeg, T. Lauterbach, *Digital Audio Broadcasting, Principles and Applications of Digital Audio*, 2<sup>nd</sup>ed.; Wiley, England, 2003.
2. M. Kin, *Subjective evaluation of sound quality of musical recordings transmitted via DAB+ system*. Proc. of the 134<sup>th</sup> AES Convention, Rome, Italy, 2013, preprint 8874.
3. S. Brachmanski, M. J. Kin, *Quality evaluation of sound broadcasted via DAB+ system based on a single frequency network*. Proc. of the 144<sup>th</sup> AES Convention, Milano, Italy, 2018, preprint 1004.

4. A. Dobrucki, M. Kin, *Subjective and objective evaluation of sound quality of radio programs transmitted via Digital Audio Broadcast (DAB+) system*. Proc. of the ICA, Montreal, Canada, **19** (2013), doi.org/10.1121/14799170.
5. S. Brachmanski, *Quality evaluation of speech AAC and HE-AAC coding*. Joint Conference - Acoustics Ustka 2018, IEEE, 2018, 30 – 33.
6. ITU-T: Recom, P.800, *Method for subjective determination of transmission quality*. Geneva, Switzerland.
7. EBU Technical Recommendation R22-1999, *Listening Conditions for the Assessment of Sound Programme Material*. EBU Geneva, Switzerland.
8. S. Brachmanski, *Selected problems of speech transmission quality assessment* (in Polish), Oficyna Wyd. Politechniki Wrocławskiej, Wrocław 2015.
9. Gilski P., *DAB vs DAB+ Radio Broadcasting: a Subjective Comparative Study*, Archives of Acoustics, **42**(4) (2017) 715 – 723.
10. P. Gilski, J. Stefański, *Subjective and Objective Comparative Study of DAB+ Broadcast System*, Archives of Acoustics, **42**(1) (2017) 3 – 11.
11. D. Ko, W. Woszczyk, *Virtual acoustics for musicians: subjective evaluation of a virtual acoustic system in performance of string quartets*. J. Audio Eng. Soc., **66**(9) (2018) 712 – 723.
12. B. De Man, K. McNally, J. D. Reiss, *Perceptual evaluation and analysis of reverberation in multitrack music production*. J. Audio Eng. Soc., **65**(1/2) (2017) 108 – 116.
13. F. Zotter, M. Frank, *Efficient phantom source widening*. Arch. of Acoustics, **38** (2013) 27 – 37.
14. J. Hjortkjaer, M. Walter-Hansen, *Perceptual effects of dynamic range compression in popular music recordings*. J. Audio Eng. Soc., **62**(1/2) (2014) 37 – 41.
15. F. Rumsey, *Recording In The Light Of New Technology*. J. Audio Eng. Soc., **63**(12) (2015) 1053 – 1057.

## Ultrasound Tomography Imaging: Results of Breast Phantom Study and Resolution Estimation

Krzysztof J. OPIELIŃSKI

*Department of Acoustics and Multimedia, Faculty of Electronics, Wrocław University of Science and Technology, Wyb. Wyspiańskiego 27, 50-370 Wrocław, Poland, krzysztof.opielinski@pwr.edu.pl*

Marcin WRZOSEK

*Department of Internal Diseases with a Clinic for Horses, Dogs and Cats, Faculty of Veterinary Medicine, Wrocław University of Environmental and Life Sciences, Plac Grunwaldzki 47, 50-366, Wrocław, Poland, marcin.wrzosek@upwr.edu.pl*

Józef NICPOŃ

*Center of Experimental Diagnostics and Innovative Biomedical Technologies, The Faculty of Veterinary Medicine, Wrocław University of Environmental and Life Sciences, Plac Grunwaldzki 47A, 50-366, Wrocław, Poland, jozef.nicpon@upwr.edu.pl*

Przemysław PODGÓRSKI

*Department of General Radiology, Interventional Radiology and Neuroradiology, Wrocław Medical University, Borowska 213, 50-556 Wrocław, Poland, przemyslaw.podgorski@umed.wroc.pl*

Tomasz ŚWIETLIK

*Department of Acoustics and Multimedia, Faculty of Electronics, Wrocław University of Science and Technology, Wyb. Wyspiańskiego 27, 50-370 Wrocław, Poland, tomasz.swietlik@pwr.edu.pl*

### Abstract

In order to improve breast cancer detection rates, new and better imaging methods are required. Currently, the ultrasound tomography (UT) as non-invasive and safe hybrid method may contribute to achieving a new standard for breast cancer diagnostics. The aim of the paper was to analyse the imaging ability of tissue-like media structure found in female breast using the developed novel ultrasound computer-assisted tomographic scanner. Measurements were performed on commercial breast biopsy phantoms due to their well-defined structure with inclusions mimicking glandular tissue with lesions, as well as on the simple agar phantom. Obtained magnetic resonance images (MRI), conventional ultrasound images (US) or X-ray computed tomography (CT) images of the measured media sections were used for comparison. The analysis of the obtained results and carried out theoretical considerations have allowed to estimate the resolution of soft tissue UT imaging.

**Keywords:** ultrasound tomography, breast phantoms, estimation of imaging resolution

### 1. Introduction

It is estimated that the detection of the breast cancer with a size less than one inch gives a 98% probability of survival. Referring physician has at his/her disposal, apart from

palpation, imaging tests: mammography (MMG), conventional ultrasound B-mode scanning (US) and magnetic resonance imaging (MRI). Each of these methods has its advantages but also disadvantages and limitations [1].

In order to improve breast cancer detection rates, new and better imaging methods are required. Currently, the ultrasound tomography (UT) as non-invasive and safe hybrid method may contribute to achieving a new standard for breast cancer diagnostics [2-5]. The hybrid method means that 3 complementary ultrasound images, representing a distribution of transmission, reflection and scattering acoustic parameters are simultaneously reconstructed. The fusion of these images allows a qualitative and quantitative characterization of breast tissue in the entire breast volume [5]. There is no need any contrast agent in the UT method. The aim of the paper was to analyse the imaging ability of tissue-like media structure found in female breast using the novel ultrasound computer-assisted tomographic scanner developed by DRAMINSKI S.A. company in cooperation with Wroclaw University of Science and Technology team [5, 6]. Measurements was performed on commercial breast biopsy phantoms due to their well-defined structure with inclusions mimicking glandular tissue with lesions, as well as on the simple self-made agar phantom. Optical, magnetic resonance, conventional US or X-ray computed tomography (CT) images of the measured media sections were used for comparison. The analysis of obtained results and carried out theoretical considerations have allowed to estimate the resolution of soft tissue UT imaging.

## 2. Estimation of spatial resolution

The estimation of the smallest malignant tumour which can be detected in the breast of woman using an ultrasound tomography scanner is very important due to effective breast cancer screening. First, we have to consider resolution limitations resulting from a given scan resolution and properties of ultrasound wave.

Ultrasound tomography transmission quantitative images (which are the basis for the automatic recognition of tumours and estimating their malignancy) are reconstructed from projection measurements of the ultrasound pulse runtime and amplitude transmitted through the breast in water [5]. The horizontal plane resolution in the ultrasound transmission tomography (UTT) can be define as the pixel size. The minimal pixel size is determined by the number of ultrasonic transducers in the tomography ring array and by the distance between its centres (so called pitch), as well as by the ultrasonic wavelength in the breast tissue. The constructed tomography array [6] consists of 1024 piezoelectric transducers evenly distributed on the inside of the 260 mm diameter ring with the pitch about 0.8 mm and 2 MHz resonant frequency. The ultrasound wavelength in the breast for 2 MHz is  $\lambda \approx 0.75$  mm. The half of the wavelength can be estimated as 0.4 mm. This is the basis for determining the horizontal plane resolution 0.4 x 0.4 mm in our ultrasound tomography transmission quantitative images. However, it does not mean that such small breast tissue lesions are always detected. This horizontal plane resolution (in the XY plane) should be understood only as a limitation of the device. It means that the minimum size of the heterogeneity visualised in the ultrasound transmission tomography image cannot be smaller than the pixel size 0.4 x 0.4 mm. The same applies

to the vertical resolution (slice thickness) resulting from the height of piezoelectric transducer and the vertical scan step (array movement) of the measurement of coronal breast sections in the Z-axis direction. The vertical scan step is 1 or 2 mm (depending on the breast size) but the ultrasound beam in the vertical plane is wider, due to the large height of the elementary transducer ( $b = 18$  mm). The effective angle of the vertical beam divergence resulting from the 3 dB decrease of the ultrasound wave pressure relative to the pressure on the beam axis can be calculated as about 2 using the formula  $2\varphi_{3\text{dB}} = 2\arcsin(0.442 \cdot \lambda/b)$ . The vertical ultrasound far field of ultrasonic rectangular transducer with height of 18 mm calculated from the equation  $l_{0z} = 0.35b^2/\lambda$  is about 150 mm. Thus, the real vertical resolution can be geometrically estimated as about 5 mm. Finally, we obtain the voxels of breast tissue limited to size  $0.4 \times 0.4 \times 5$  mm.

### 3. Estimation of contrast resolution

The foregoing considerations, while important and necessary, but do not give the answer to the question: What the smallest malignant tumours can be detected in breasts of women using the ultrasound tomography scanner? Answering this question, we have to estimate the contrast resolution of ultrasound tomography images and verify it by medical testing on a large group of patients healthy and with lesions of different sizes and malignancy degrees. The problem is that the contrast resolution is not constant and depends not only on the scanning and image reconstruction methods but also on the structure of the imaged tissue. This applies of course not only to the UT method but also to all of the previously mentioned.

In the paper, we have estimated the contrast resolution of ultrasound tomography images by means of calculations of projection values of the ultrasound speed and attenuation after passing the ultrasound pulse through the path from the emitter to receiver along the diameter of the array ring of transducers (Fig. 1):

$$c_{\text{projection}} = \frac{L_{\text{path}}}{t_{\text{water}} + t_{\text{breast}} + t_{\text{lesion}}} = \frac{L_{\text{water}} + L_{\text{breast}} + L_{\text{lesion}}}{L_{\text{water}}/c_{\text{water}} + L_{\text{breast}}/c_{\text{breast}} + L_{\text{lesion}}/c_{\text{lesion}}} \quad (1)$$

$$\alpha_{\text{projection}} = \frac{\alpha_{\text{water}} \cdot L_{\text{water}} + \alpha_{\text{breast}} \cdot L_{\text{breast}} + \alpha_{\text{lesion}} \cdot L_{\text{lesion}}}{L_{\text{water}} + L_{\text{breast}} + L_{\text{lesion}}} \quad (2)$$

The model of the simulation of projection measurements consists of the circular section of a homogenous breast glandular tissue located in the plane of the ring array filled with water. Then, the small circular section of a homogeneous lesion can be inserted in the breast centre (Fig. 1). The acoustic parameter values of the *in vivo* breast tissue were adopted for calculations from the literature [7]. Unfavorable measuring conditions for a large coronal section of the breast at its base were assumed. In this way, the share of the lesion in the projection speed and attenuation values is small. Acoustic parameters of water correspond to the temperature of 25°C. Figure 2 shows the calculation results of the accuracy of ultrasound speed and attenuation measurements of projection values. The Y-axis represents the values of the difference  $\Delta p = (p_{\text{BreastWithLesion}} - p_{\text{BreastWithoutLesion}})$ , where  $p$  is the ultrasound speed or attenuation projection value. This difference defines the accuracy of the projection value measurement that is necessary to

detect the lesion. The  $X$ -axis represents the lesion size from 0.1 – 10 mm. The calculations were carried out for 6 different values of the ultrasound speed and attenuation representing breast tissue of different densities, while the ultrasound speed and attenuation of lesion is constant (Fig. 1). These values are shown in Fig. 2 in the form of  $(c_{\text{lesion}} - c_{\text{breast}})$  or  $(\alpha_{\text{lesion}} - \alpha_{\text{breast}})$  on the right of graphs. The interpretation of graphs from Fig. 2 is now very clear. For example, if we want to detect the lesion with the size  $\geq 5$  mm in ultrasound tomography images and the local value of the ultrasound speed in the lesion differs from the value in the surrounded glandular breast tissue by about 5 m/s, we need to ensure the ultrasound speed measurement accuracy at least  $\Delta c \approx 0.1$  m/s. We can refer possible to detect lesion sizes to these accuracy values using graphs in Fig. 2.

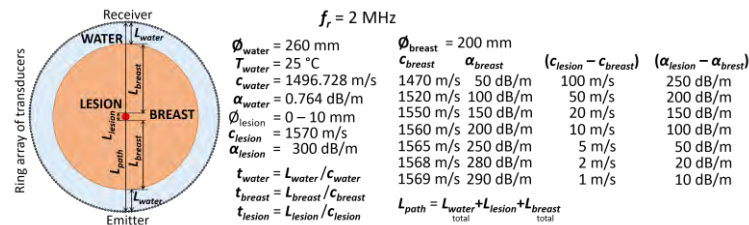


Figure 1. The model of the simulation of projection measurements used to estimate the contrast resolution of UT images

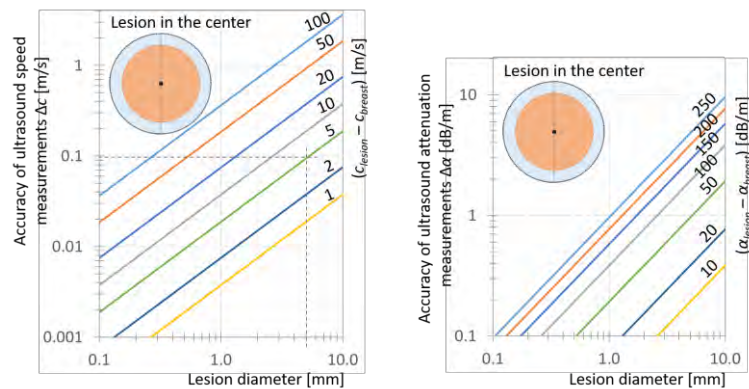


Figure 2. Calculation results of the accuracy of ultrasound speed and attenuation measurements of projection values performed during the UT breast scanning

#### 4. Measurement results

The designed phantom of the agarose gel with 3 holes filled with water ( $\phi$  7 mm) and one filled with water and pieces of the agar gel ( $\phi$  4 mm) has been measured (Fig. 3) in order to assess the accuracy of the projection measurements of the ultrasound speed and attenuation using the ultrasound tomography scanner developed by DRAMINSKI S.A. The agar gel characterizes by the ultrasound speed similar to the soft tissue and by very low attenuation so that the impact of noise and signal distortions as well as adverse



phenomena accompanying the ultrasonic wave propagation on measured values are small. The measured projection values are presented after the transformation from the divergent scanning geometry to the parallel one. Projection values of ultrasound speed were calculated from measurements of the ultrasonic pulse runtime, and attenuation projection values – from measurements of the pulse amplitude after transition.

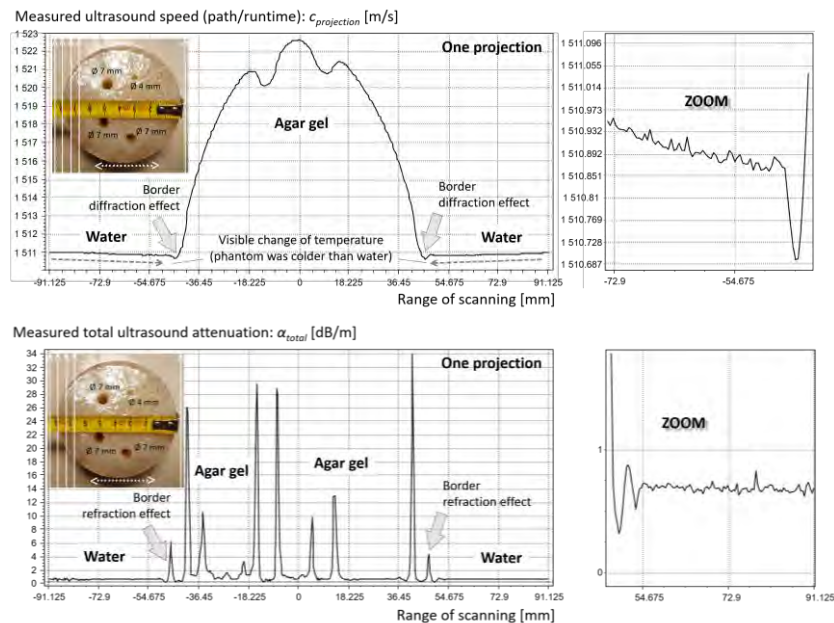


Figure 3. Projection measurements of the ultrasound speed and attenuation using the ultrasound tomography scanner developed by DRAMIŃSKI S.A.

The changes of the pulse amplitude reflects not only the attenuation of ultrasound but also its weakness after passing through the boundaries of different acoustic impedances ( $\alpha_{total} = \alpha_{projection} + \alpha_{weakness}$ ) and are impossible to separate. Therefore, we can expect distortions of reconstructed local values of the ultrasound attenuation distribution in the tomography images. Based on Fig. 3 (see ZOOMs) we can estimate the accuracy of measurements of the ultrasound speed projection as  $\Delta c \approx 0.02$  m/s and the accuracy of measurements of the ultrasound attenuation projection as  $\Delta \alpha \approx 0.1$  dB/m. Figure 4 shows ultrasound tomography images of the tested agar gel phantom section reconstructed from projection measurements as presented in Fig. 3 together with the distributions of local values of the ultrasound speed and attenuation along the dotted  $y$ -line marked in the images. Based on Fig. 4 (see ZOOMs) we can estimate the accuracy of the reconstructed local ultrasound speed as  $\Delta c \approx 0.3$  m/s and local ultrasound attenuation as  $\Delta \alpha \approx 8$  dB/m in this case.

Three commercial breast biopsy phantoms with various acoustic parameters were selected to verify possibilities of the ultrasound tomography imaging in the context of the spatial and contrast resolution (Fig. 5 – Fig. 8).

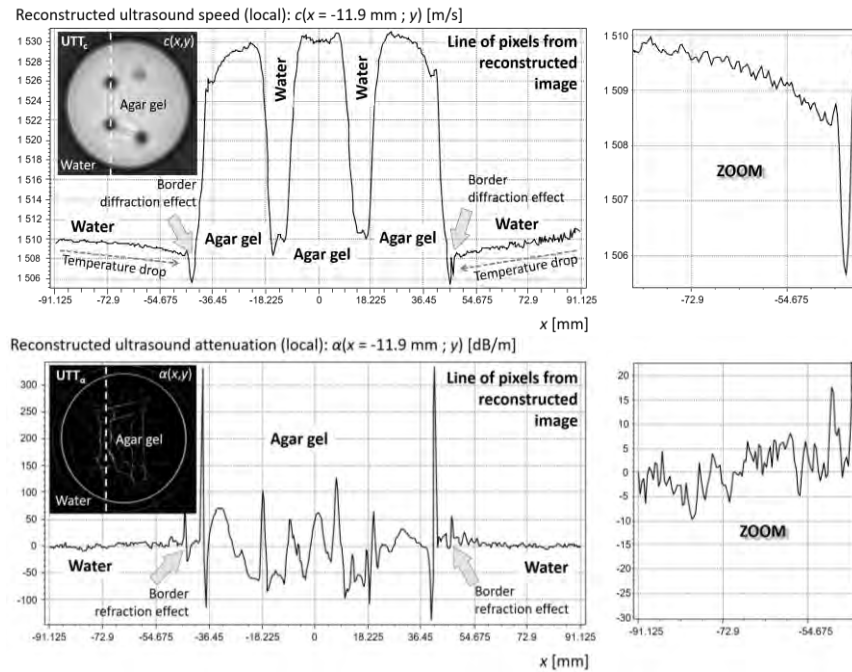


Figure 4. Ultrasound tomography images of the tested agar gel phantom section reconstructed from projection measurements together with distributions of local values of the ultrasound speed and attenuation along the dotted  $y$ -line marked in the images

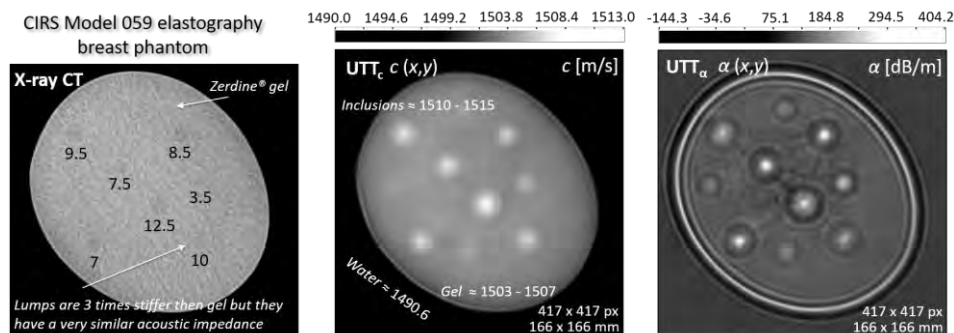


Figure 5. Comparison of X-ray CT image with UTT images of the same section of CIRS Model 059 elastography breast phantom with inclusions of the acoustic impedance very similar as in the phantom gel (low attenuation of ultrasound)

The reconstructed ultrasound images of the ultrasound speed and attenuation distribution in the selected section of these objects were compared with reference X-ray CT image, magnetic resonance images and the conventional US image.

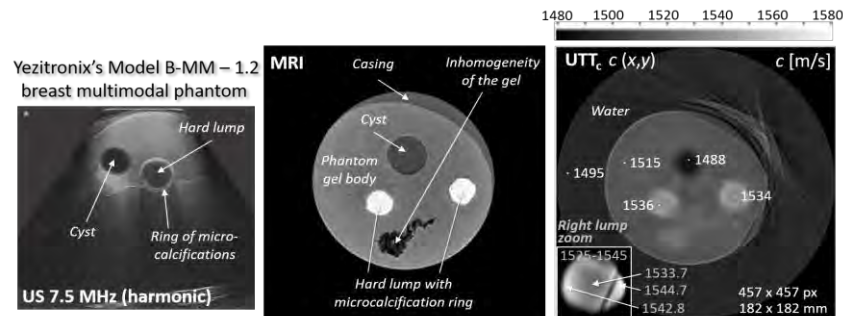


Figure 6. Comparison of harmonic conventional US and magnetic resonance image with sound speed UTTc image of the same section of Yezitronix's Model B-MM – 1.2 breast multimodal phantom with inclusions mimicking cyst and hard lumps surrounded by micro-calcifications (mid attenuation of ultrasound)

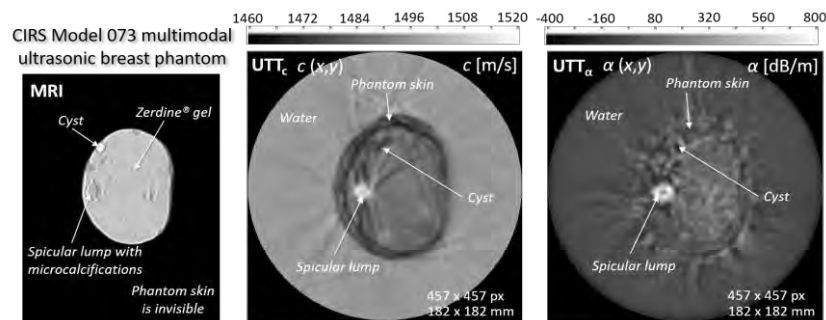


Figure 7. Comparison of magnetic resonance image with UTT images of the same section of CIRS Model 073 breast multimodal phantom with a skin and inclusions mimicking cyst and spicular cancer lump (high attenuation of ultrasound)

## 5. Discussion and Conclusions

Calculations show that the detection of heterogeneities in coronal breast sections of size about  $d_{lesion} \geq 5$  mm differing by the ultrasound speed barely 1 m/s from the surrounding tissue is possible using the ultrasound tomography scanner developed by DRAMIŃSKI S.A. in cases of not noisy projection measurements of the ultrasound speed (Fig. 2). It was confirmed by measurements of the elastography breast phantom (Fig. 5). The size of a detected heterogeneity decreases linearly with an increase in the difference of speeds and in general is  $d_{lesion} \geq 0.5$  mm for  $(c_{lesion} - c_{breast}) \geq 10$  m/s due to the limitation of the horizontal plane resolution. Edges of small heterogeneities with a similar ultrasound speed in relation to their environment are blurry in UTTc images (Fig. 5, Fig. 6) because ultrasound beam rays pass then by short chords of quasi-circular shapes. This reduces the contrast resolution in such places, in accordance with calculations (Fig. 2). Despite this, it is possible to detect micro-calcifications due to their large value of the ultrasound speed (Fig. 6). The relatively low vertical resolution (5 mm) is not

a disadvantage, because it allows us to increase the vertical step of the ultrasound ring array movement to 2 mm and thus reducing the time needed to scan the whole breast. A few millimeter lesions above and below the scanned coronal breast section will be still detected (see Fig. 5). In the case of extremely dense breasts, highly absorbing ultrasound, we should expect a deterioration in contrast resolution due to the decrease of the signal-to-noise ratio and the impact of phenomena accompanying the ultrasonic wave propagation (refraction, scattering, reflection, diffraction) to precise measurements of the pulse runtime and amplitude as well as affecting adversely the image reconstruction process (see distortions of UTT images in Fig. 7).

According to calculations, the detection of heterogeneities in coronal breast sections of size about  $d_{\text{lesion}} \geq 2.5$  mm differing by the ultrasound attenuation about 10 dB/m from the surrounding tissue is possible in cases of not noisy projection measurements of the ultrasound attenuation (Fig. 2). The size of this way detected heterogeneity decreases linearly with an increase in the difference of attenuations. However, we should expect a falsification of reconstructed local values of the ultrasound attenuation because the projection measurement of the total weakness of the transmitted signal. This is shown in Fig. 4, where large attenuation values are exposed in projection measurements despite a very low attenuation of ultrasound in the agar gel. It is the result of weakening the signal passing through the boundaries of the water and gel. For this reason, very low ultrasound attenuation values (e.g. in water) can be reconstructed as negative (Fig. 5, Fig. 7).

Quantitative UTTc and UTT $\alpha$  images are currently used to develop the algorithm of the automatic recognition of breast tumors and estimation their degree of malignancy in screening test, taking into account the age of the woman and her breast density.

## References

1. R. M. Kaplan, F. Porzsolt, *The natural history of breast cancer*, Arch. Intern. Med., **168**(21) (2008) 2302 – 2303.
2. N. Duric, *Ultrasound Tomography: A Breast Imaging Modality Whose Time Has Come*, Med Phys, **42**(6) (2015) 3699 – 3699.
3. B. Malik, R. Terry et al., *Quantitative transmission ultrasound tomography: Imaging and performance characteristics*, Med Phys, **45**(7) (2018) 3063 – 3075.
4. H. Gemmeke, T. Hopp et al., *3D Ultrasound Computer Tomography: Hardware Setup, Reconstruction Methods and First Clinical Results*, Nucl Instrum Methods Phys Res A, **873** (2017) 59 – 65.
5. T. Milewski, M. Michalak et al., *Hybrid Ultrasound Tomography Scanner - a Novel Instrument Designed to Examine Breast as a Breast Cancer Screening Method*, Biomed J Sci & Tech Res, **14**(4) (2019) 1 – 5.
6. K. J. Opieliński, P. Pruchnicki et al., *Multimodal ultrasound computer-assisted tomography: An approach to the recognition of breast lesions*, Comput Med Imaging Graph, **65** (2018) 102 – 114.
7. F. A. Duck, *Physical properties of tissue – a comprehensive reference book*, Academic Press, London 1990.

## **The Effect of Dynamic Focusing of the Beam on the Acoustic Field Distribution Inside the Ultrasonic Ring Array**

Wiktor STASZEWSKI\*

*Wrocław University of Science and Technology, Wybrzeże Wyspiańskiego 27,  
50-370 Wrocław, wiktor.staszewski@pwr.edu.pl*

*\*T. Marciniak Lower Silesian Specialist Hospital – Emergency Medicine Centre,  
Gen. Augusta Emila Fieldorfa 2, 54-049 Wrocław*

Tadeusz GUDRA

*Wrocław University of Science and Technology, Wybrzeże Wyspiańskiego 27,  
50-370 Wrocław, tadeusz.gudra@pwr.edu.pl*

### **Abstract**

This paper presents the analysis of readings acquired from the ultrasonic ring array used in tomography for the diagnosis of female breast tissue. In addition, this paper also presents the results for the acoustic field distribution simulation, acquired through a method of summing up all acoustic fields generated by each of the elementary transducers of the ring array. The change in acoustic field pressure level when changing activation frequency (2 MHz, 3MHz, 4MHz) of the elementary ultrasonic transducers for the sector consisting of 32 and 64 ultrasonic transducers was studied. By changing the time of activation of individual transducers, a change in the natural position of the focus inside the ultrasonic ring array was observed. For the sector consisting of 32 ultrasonic transducers the relation between the echo coming from the wires of the wire pattern and the level of noise and distortion on the ultrasonographic image for different locations of the focus of the central transducers was studied. The results were compared with the simulations of the acoustic field, which were conducted using MATLAB software. This research is the continuation of studies [13, 14] aimed at choosing the optimal focus and number of transducers in ultrasonic ring array with the goal of receiving the best possible quality of images of cross-sections of the female breast.

**Keywords:** ultrasonic ring array, acoustic field distribution, dynamic focusing of ultrasonic beam

### **1. Introduction**

Ultrasonic tomography makes early diagnosis of pathological changes in biological tissues possible. There are currently a couple of research centres across the world that analyse the construction and results of ultrasonic tomograph prototypes [1 - 8]. Researchers from the Wrocław University of Science and Technology [6, 7, 9 - 14] are conducting research connected to the use of the ultrasonic transmission tomograph to diagnose female breast tissue. With the aim of receiving the best possible image of the cross-section of the female breast, the acoustic field generated by each of the elementary transducers of the array during dynamic focusing of the ultrasonic beam was analysed. The position of the focus inside the ultrasonic ring array was changed by changing the time of activation of the ultrasonic elementary transducers. The simulation presented in this paper attempted to use the ring array, adapted for ultrasonic transmission tomography, for ultrasonographic imaging. This paper analyses the focusing of the ultrasonic beam generated by either 32 or 64 ultrasonic transducers for their different activation frequencies. Moreover,

an analysis of acquired ultrasonographic images when measuring the wire pattern during the activation of only one section of the ultrasonic ring array consisting of 32 elementary transducers has been conducted [11, 13, 14].

## 2. Construction of the ultrasonic ring array

The ultrasonic ring array made of 32 sections. A section consists of 32 ultrasonic elementary transducers, which are adapted to work in water. The transducers can serve two purposes: they can work as transmitters and as receivers. Each ultrasonic elementary transducer measures 0.5 [mm] x 18 [mm] and is 1 [mm] thick [11]. The beam of the ultrasonic ring array is equal to 130 [mm], and the pitch between the transducers measures 0.3 [mm]. During tests the water had a temperature of 25°C.

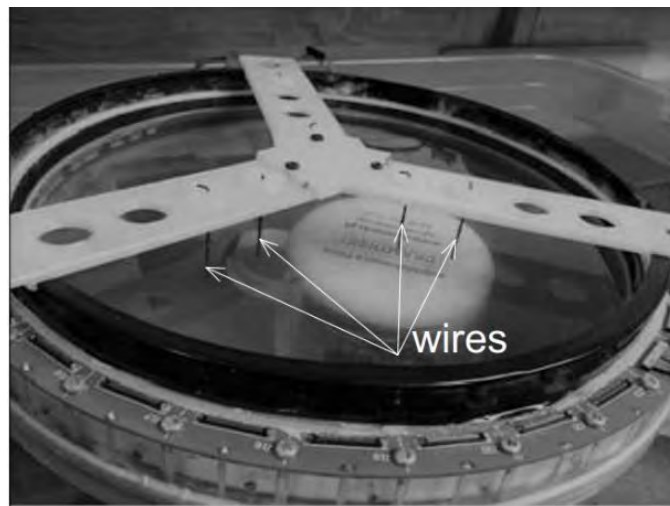
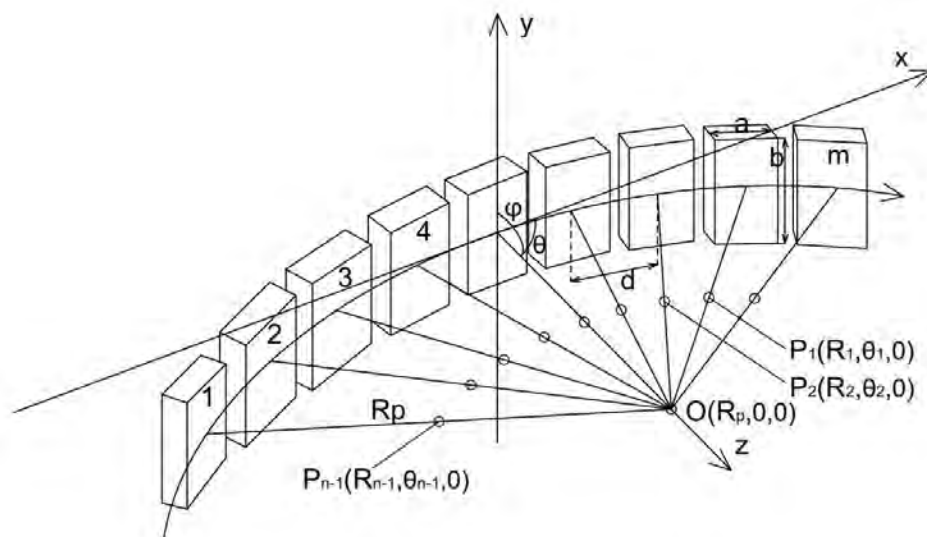


Figure 1. Ultrasonic ring array together with the wire pattern submerged in water that were used in calculations, as viewed from top

## 3. Calculation method

The range of the near field can be calculated based on the following dependency:  $l_o = 0.35a^2/\lambda$ . In the case of resonance frequency  $f_r = 2$  MHz and the width of the ultrasonic elementary transducer  $a = 0.5$  mm the near field range is equal to approximately 12 mm. Beyond this distance a precise analysis of the acoustic field in the central transducer axis in a horizontal plane is possible.

With the goal of determining the acoustic field distribution generated by the curvilinear array of the ultrasonic elementary transducers, a method of calculating the sum of all acoustic fields was applied, as a sum of geometric transformations of fields calculated for all the elementary transducers of the sector (Fig. 2). Simulations for the distribution of acoustic field were conducted using an algorithm obtained according to formula (1).



In this case, the following equation was applied[11, 13, 14]:

$$L_p = 20 \cdot \log \left( \left| \sum_{i=0}^{m-1} -\frac{j\rho ckV_a}{2\pi R_i} ab \cdot e^{j(\omega t - kR_i)} \left( \frac{\sin\left(\frac{u_i a}{2}\right)}{\frac{u_i a}{2}} \right) \left( \frac{\sin\left(\frac{wb}{2}\right)}{\frac{wb}{2}} \right) \right| \cdot \frac{1}{p_o} \right) \quad (1)$$

where:

- $\rho$  – density of the medium,
- $c$  – ultrasonic wave propagation velocity in the medium,
- $k = 2\pi/\lambda$  – wavenumber,
- $\lambda = c/f$  – wavelength,
- $f$  – resonance frequency of the ring array transducers,
- $\omega = 2\pi f$  – cycle resonance frequency of pulsations,
- $t$  – time,
- $V_a$  – acoustic velocity,
- $p_o$  – reference pressure ( $p_o = 1 \mu\text{Pa}$ ),
- $a$  – width of the elementary rectangular transducer,
- $b$  – length of the elementary rectangular transducer,
- $m$  – number of transducers in the ring array sector,
- $u_i = 2\pi \cdot \sin(\theta_i)/\lambda$ ,
- $w = 2\pi \cdot \sin(\varphi)/\lambda$ ,
- $R_i, \theta_i, \varphi$  – polar coordinates of the point  $P(R, \theta, \varphi)$ , corrected in respect of the location  $(i+1)$ th transducer in sector.

The geometrical transformation of the location of every point of the field  $P(R, \theta, \varphi)$  in equation (1) is carried out symmetrically for every  $i + 1$ th transducer in the ring array sector (Fig. 2) by turning the point around the symmetry axis of the array with coordinates  $O(r = R_p, \theta = 0)$  parallel to the  $Y$  axis (where  $R_p$  denotes the array's internal radius) by an appropriate multiplicity of the angle  $\beta = i \cdot 2\pi/N$  (where  $N$  is the number of all the transducers in the ring array). Revised coordinates for  $P(R, \theta, \varphi)$  for following revolutions in the polar system (occurring in formula (1)) can be determined using formulae [11, 13, 14]:

$$\begin{cases} r_i = \sqrt{r^2 + 4R_p^2 \sin^2 \frac{\beta_i}{2} + 2rR_p \cos(\theta - \beta_i) - 2R_p \cos \theta} \\ \theta_i = \arctg \left( \frac{r \sin(\theta - \beta_i) + R_p \sin(\beta_i)}{r \cos(\theta - \beta_i) + 2R_p \sin^2 \frac{\beta_i}{2}} \right) \end{cases} \quad (2)$$

$$\beta_i = \left( \frac{m-1}{2} - i \right) \cdot \beta \quad (3)$$

where  $i = 0, \dots, m-1$ .

#### 4. Results

When taking the readings with the use of linear phased (steered) scanning, the ultrasonic ring array was submerged in distilled water with a temperature of 25°C. The method of dynamic focusing was applied; by changing the time of activation of the ultrasonic elementary transducers the position of the focus was changed. This caused the focus to move closer or farther apart along the axis between two central transducers and a point where natural focus occurs due to the curvature of the ring array. During readings, the section of the ring array consisting of 32 ultrasonic elementary transducers was used. The readings were taken for nine different distances between the focus and central transducers. Each reading was taken in identical conditions. The wires of the wire pattern had the same thickness and were made of the same material. The results of the readings are presented in (Fig. 3). Next, the acquired ultrasonographic images of the wire pattern were analysed.



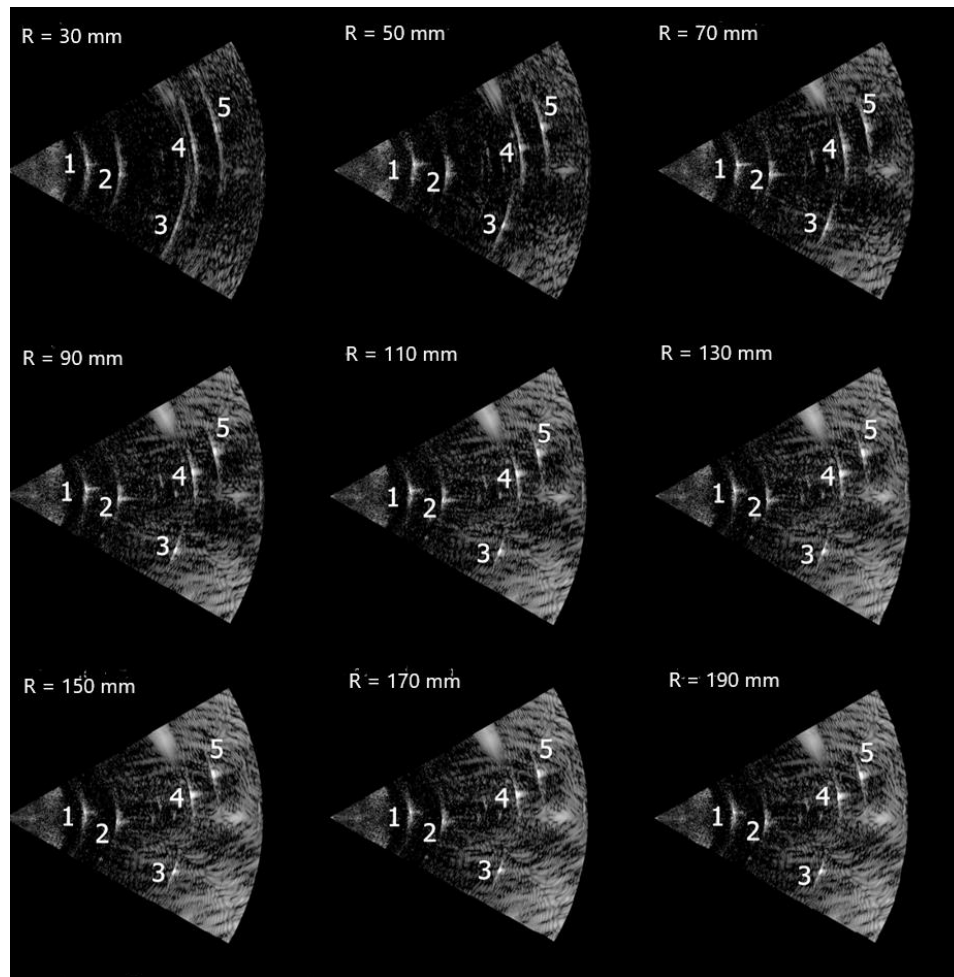


Figure 3. Results of readings of the wire pattern using the ultrasonic ring array with  $n = 32$  activated ultrasonic transducers with the focus at the following intervals  $R = 30$  mm,  $R = 50$  mm,  $R = 70$  mm,  $R = 90$  mm,  $R = 110$  mm,  $R = 130$  mm,  $R = 150$  mm,  $R = 170$  mm and  $R = 190$  mm with numbered echoes coming from the rods [see also 14]

Table (1) shows the analysed data from the ultrasonographic images (Fig. 3) during dynamic focusing. The relationship between an average level of noise and distortion occurring around the analysed inclusion and the brightness of the inclusion coming from the wires of the pattern was recorded using a greyscale. The results for each of the identified inclusion are presented in percentages in the table below.

Table 1. The relationship between an average level of noise and distortion to the brightness of the inclusion coming from the wire of the wire pattern

R [mm]	point 1	point 2	point 3	point 4	point 5	average
30	32.74	18.95	60.44	56.76	59.96	45.77
50	26.46	17.10	48.16	38.52	52.79	36.61
70	28.69	38.96	58.47	51.59	57.37	47.01
90	30.45	33.21	66.70	55.87	60.46	49.34
110	32.91	36.81	58.36	61.53	64.08	50.74
130	38.61	37.74	55.87	62.84	62.99	51.61
150	37.78	35.91	55.82	60.91	61.18	50.32
170	40.28	46.92	60.08	60.08	59.61	53.40
190	42.63	45.72	61.11	59.60	61.18	54.05

Next, it was analysed how a change in activation frequency of the ultrasonic elementary transducers affected the acoustic pressure in the focus. A simulation of acoustic field pressure for activation frequencies 2 MHz, 3 MHz and 4 MHz was conducted using formula (1) for 32 and 64 activated ultrasonic transducers. Moreover, it was checked how a  $10^\circ$  inclination of the beam affected the acoustic pressure in the focus. The results are presented in graphs 4 and 5.

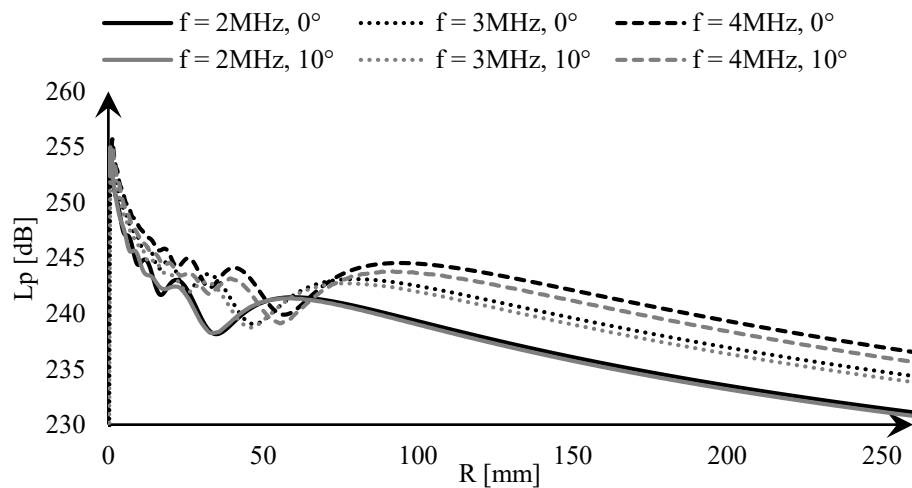


Figure 4. Results of the acoustic field distribution calculations in the form of  $L_p(z)$  with a number of activated transducers  $m = 32$ , for  $f = 2$  MHz, 3 MHz and 4 MHz without and with a  $10^\circ$  inclination of the beam

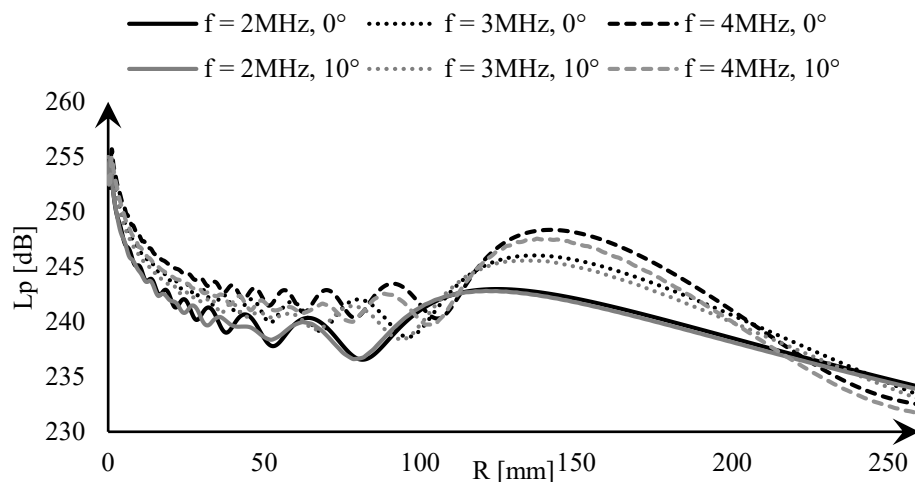


Figure 5. Results of the acoustic field distribution calculations in the form of  $L_p(z)$  with a number of activated transducers  $m = 64$ , for  $f = 2$  MHz, 3 MHz and 4 MHz without and with a  $10^\circ$  inclination of the beam

## 5. Conclusions

After analysing the results of the wire pattern readings, it can be stated that the smallest ratio of noise and distortion level to the echo coming from the inclusion occurs for the focus positioned 50 [mm] away from central transducers, in which average worth of this relationship for all echoes equalled 36.61[ %]. Based on analysis of these readings, it can be concluded that the most beneficial focus for analysing the cross-sections of the female breast occurs 50 [mm] away from central transducers, because the acquired ultrasonographic images had the biggest contrast.

Based on the analysis of sections consisting of 32 and 64 ultrasonic elementary transducers it can be concluded, that the level of acoustic pressure in the focus occurring due to the natural curvature of the ring array increases together with activation frequency of the ultrasonic elementary transducers despite using the same ultrasonic transducers. In the case of a  $10^\circ$  inclination of the beam, the level of acoustic pressure decreased on average by 1 [dB] for three of the analysed frequencies. This means that the beam can be inclined in this range without deterioration of its parameters. That is why, when acquiring an image consisting of 32 sections made up of 32 transducers, a significant reduction in noise and distortion can be achieved by averaging out the pixels.

## References

1. M. Birk, E. Kretzek, P. Figuli et al., *High-speed medical imaging in 3D ultrasound computer tomography*, IEEE Trans. Parallel Distrib. Syst., **27**(2) (2016) 455 – 467.
2. F. A. Duck, *Physical Properties of Tissue - A Comprehensive Reference Book*, Academic Press London 1990.

3. N. Duric, P. Littrup, L. Poulo et al., *Detection of breast cancer with ultrasound tomography: first results with the Computed Ultrasound Risk Evaluation (CURE) prototype*, Med. Phys., **34**(2) (2007) 773 – 785.
4. N. Duric, P. Littrup, S. Schmidt et al., *Breast imaging with the SoftVue imaging system: first results*, In: J. G. Bosch, M. M. Doyley (Eds.), Medical Imaging: Ultrasonic Imaging, Tomography, and Therapy, Proc of SPIE.SPIE, p. 8675 (2013) 86750K-1-8.
5. R. Entekin, P. Jackson, J. R. Jago, B. A. Porter, *Real Time Spatial Compound Imaging in Breast Ultrasound: Technology and Early Clinical Experience*, Medica-Mundi, **43**(3) (1999) 35 – 43.
6. T. Gudra, K. Opieliński, *The ultrasonic probe for investigating of internal object structure by ultrasound transmission tomography*, Ultrasonics, **44** (2006) e679 – e683.
7. T. Gudra, K. Opieliński, *The multi-element probes for ultrasound transmission tomography*, Journal de Physique IV, **137** (2006) 79 – 86.
8. T. Gudra, K. J. Opieliński, *Sposób wizualizacji struktury wewnętrznej ośrodka i urządzenie do realizacji tego sposobu*, Patent, Polska, Nr 210202.
9. R. Jirik, I. Peterlík, N. Ruiter et al., *Sound-speed image reconstruction in sparse-aperture 3-D ultrasound transmission tomography*, IEEE Trans. Ultrason. Ferroelectr. Freq. Control, **59**(2) (2012) 254 – 264.
10. V. Z. Marmarelis, J. Jeong et al., *High-resolution 3-D imaging and tissue differentiation with transmission tomography*, Acoust. Image, **28** (2007) 195 – 206.
11. K. J. Opieliński, *Application of Transmission of Ultrasonic Waves for Characterization and Imaging of Biological Media Structures*, Printing House of Wrocław University of Science and Technology, Wrocław 2011 [in Polish].
12. K. J. Opielinski, P. Pruchnicki, T. Gudra, J. Majewski, *Full angle ultrasound spatial compound imaging*. In: Proceedings of 7<sup>th</sup> Forum Acusticum 2014 Joined with 61<sup>st</sup> Open Seminar on Acoustics and Polish Acoustical Society –Acoustical Society of Japan Special Session Stream [CD-ROM], Krakow: European Acoustics Association (ISSN 2221-3767), 2014.
13. W. Staszewski, T. Gudra, K. J. Opieliński, *The acoustic field distribution inside the ultrasonic ring array*, Archives of Acoustics, **43**(3) (2018) 455 – 463.
14. W. Staszewski, T. Gudra, K. J. Opieliński, *The effect of dynamic beam deflection and focus shift on the acoustic field distribution inside the ultrasonic ring array*, Archives of Acoustics, 2019 (article submitted to the editorial office).

## **Vibrotactile Amplitude Discrimination on the Wrist of Visually Impaired People**

Dorota CZOPEK

*AGH University of Science and Technology,  
al. Mickiewicza 30, 30-059 Kraków, dorota.czopek@agh.edu.pl*

Jerzy WICIAK

*AGH University of Science and Technology,  
al. Mickiewicza 30, 30-059 Kraków, wiciak@agh.edu.pl*

### **Abstract**

Support for spatial orientation and mobility of the blind and partially sighted people is still a hot and unsolved topic. It is well known that in supporting devices non-acoustic communication is recommended due to the role of the sense of hearing in the spatial orientation of the visually impaired people. Knowledge about discriminative ability of touch is crucial when designing devices which use tactile interfaces. The paper presents results of psychophysical studies connected with vibration perception on the wrist of blind and partially sighted people. The presented research examined the ability to identify the changes of amplitude of vibration on the wrist. Research was carried out on students of Special Schools for the Blind and Partially Sighted Children in Krakow. Thirty blind, partially and normal sighted people were examined. Transformed up/down method was used to determine vibrotactile amplitude discrimination threshold. Thresholds were obtained on the ventral wrist at 5 frequencies: 25, 32, 63, 125 and 250 Hz. Results were examined to find factors which could influence the threshold value.

**Keywords:** vibration perception, vibrotactile amplitude discrimination, blind people

### **1. Introduction**

Visual stimuli constitute approximately 80% of all the information reaching the human brain [1]. People that could not perceive visual stimuli use the active cooperation of all the other senses to compensate the lack of sight. This mechanism is called the compensation. Blind people use other senses: sense of touch, sense of hearing, sense of smell and additional skills called algorithms and obstacle sense. Obstacle sense is describe as feeling a tingling, or „shadow” of the obstacle when that person is near the obstacle. Since there are no additional receptors on the body of blind people it is considered a synthesis of sharpened hearing and touch (like hearing the reflected sound and feeling the changes in the air flow on skin). Despite the popular opinion, the most important sense of a blind person when moving is not touch but hearing. Because hearing is the long-range sense. Therefore, devices supporting spatial orientation should use vibration interfaces. These devices should not occupy user's hands because of the role of touch in movement, but also for practical reasons [2]. Therefore, the researchers decided to create a vibrating bracelet supporting spatial orientation.

The hereby paper presents the part of the psychophysical research carried out to specify vibration perception on the wrist. The starting point of the described tests were

the results of the measurements of absolute vibration thresholds on the proximal phalange and ventral wrist [3, 4]. The presented research examined the ability to identify the changes of the vibration amplitude on the wrist for selected frequency. The measurements were carried out using the adaptive method 1 up/2 down on a measuring stand specially designed for the needs of the described test.

## 2. Methodology

The research method including the measuring stand description, psychophysical methods used in experiment measurement procedure is described in detail below.

### 2.1. Measuring stand

The schematic diagram of the measuring stand for testing the differential vibration thresholds on the wrist using the adaptive method 1 up/2 down is shown in Figure 1. The measuring procedure was implemented and controlled by computer with LabVIEW software. Vibration stimuli were sent through 4-channel voltage output module NI 9263 (National Instrument) to Apart MB-150 amplifier and generated by the mini modal shaker TMS 2004E (The Modal Shop). The measurement conditions were controlled by the accelerometer PCB M354C03 and pressure sensor and connected with the computer through the 4-channel input module NI 9234. The masking signal (pink noise) was sent through the computer sound card to the Beyerdynamic DT 770 pro headphones.

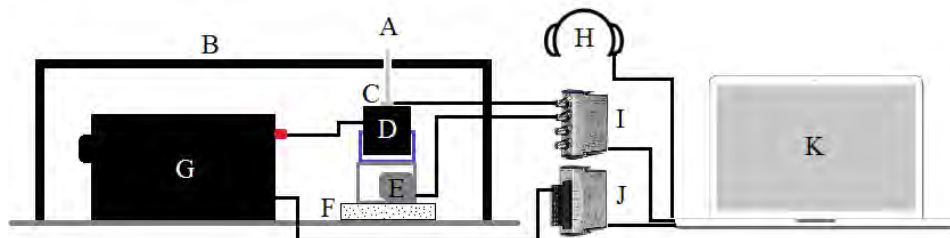


Figure 1. Schematic diagram of the measuring stand for testing the differential vibration thresholds on the wrist [5].

- A – probe, B – support of the subject's forearm,
- C – accelerometer PCB M354C03, D – modal shaker TMS 2004E,
- E – pressure sensor, F – vibration insulation,
- G – amplifier Apart MB-150,
- H – closed headphones Beyerdynamic DT 770 pro,
- I – 4-channel input module NI 9234,
- J – 4-channel voltage output module NI 9263,
- K – computer with LabVIEW software

## 2.2. Psychophysical adaptive procedures

Discrimination of changes in the vibration amplitude is the ability of the sense of feeling to perceive changes in vibration amplitude over time. In other words, this ability allows a person to notice the difference in amplitude of two vibration signals [6].

Adaptive up/down method was chosen to study the vibrotactile amplitude discrimination threshold on the wrist [7, 8]. The method was selected based on pilot studies carried out using the method of constant stimuli. Five frequencies were selected for the study: 25; 31.5; 63; 125; 250 Hz. The frequencies were selected based on results and analyzes of absolute vibration thresholds on the ventral wrist described by Wiciak et al. [4]. Vibrations of selected frequencies stimulate FAI mechanoreceptors (25; 31.5 and 63 Hz) and FAII mechanoreceptors (125 and 250 Hz).

The algorithm of up/down method with two interval forced-choice (2IFC) was implemented in the LabVIEW environment [9]. The difference between the standard (reference) and variable (compared) stimulus in each first attempt was such that the subjects clearly felt it and provided the correct answer. In this phase of the algorithm, after each correct answer there was a reduction of the difference between the standard and variable stimulus by a double step value. After the first incorrect answer (first reversal), the main phase of procedure 1 up/2 down followed (Figure 2).

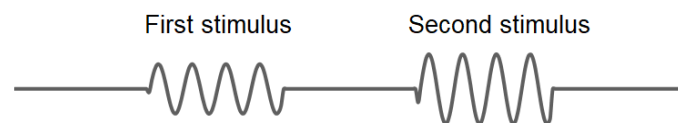


Figure 2. Vibration amplitude discrimination – stimulus sequences given by the modal shaker to the wrist

In the next attempt, the difference between the stimuli was greater by the value of the step. From now on, the reduction of the difference between stimuli in the trial could only take place after two successive correct answers and was the value of the step. However, after each wrong answer, or a sequence of correct and wrong answer, the difference between the stimuli in the sample was increased by the step value. The end of the study occurred after 8 retreats. The test conditions are summarized in Table 1.

Table 1. Vibrotactile amplitude discrimination threshold – test conditions

		Vibrotactile amplitude discrimination threshold
Stimulus	Standard stimulus	Sinusoidal signal with frequencies 25; 31.5; 63; 125; 250 Hz and intensities equal to the vibrotactile perception threshold for a given frequency
	Variable stimulus	Stimuli level varying in the range of 100 – 150 dB ref. $10^{-6}$ m/s <sup>2</sup>
	Probe	Flat cylinder with a diameter of 5 mm, without surroundings, with controlled skin contact force in the range of 0.1 – 0.2 N
Method		Adaptive method 1 up/2 down
Task		Two interval forced-choice (2IFC)
Analysis		Vibrotactile amplitude discrimination threshold for each tested frequency was calculated as the arithmetic mean of the stimulus reversal values without the first
Measure		dB ref. $10^{-6}$ m/s <sup>2</sup>

### 3. Results analysis and discussion

The research group consisted of 30 people: 10 blind people, 10 partially sighted people and 10 normally sighted people. There were 5 women and 5 men aged 18 to 28 years in each group. Median value of the amplitude discrimination threshold received for 30 people is shown in Figure 3.

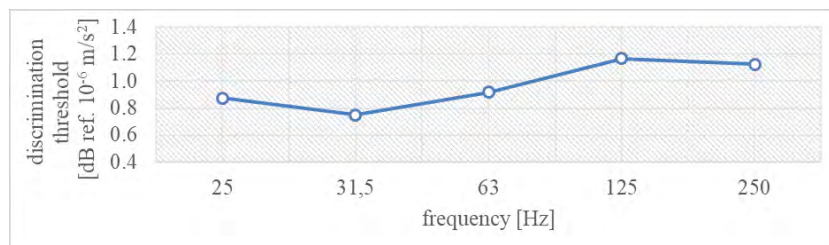


Figure 3. Median value of vibrotactile amplitude discrimination thresholds measured on the wrist

The main purpose of the results analysis was to check if the values of vibrotactile amplitude discrimination thresholds for blind people differ from those for people with normal sight. Additional analysis was conducted to check whether the frequency and the gender of the subject affects the value of the vibrotactile amplitude discrimination threshold. Nonparametric tests were used for the analysis because the obtained threshold values do not follow the normal distribution. The influence of individual factors such as the degree of visual disability of the subjects as well as



stimulus frequency (more than two groups) were tested using the Kruskal-Wallis test (1) and gender (two groups) were tested using the Kolmogorov-Smirnov test (2) [10].

$$H = \frac{12}{N(N+1)} \sum_{i=1}^p \frac{R_i^2}{n_i} - 3(N+1), \quad (1)$$

where:  $H$  – Kruskal-Wallis test;  $N$  – number of all observations;  $p$  – number of compared groups;  $R_i$  – the sum of ranks in a given group;  $n_i$  – number of observations in a given group.

$$D_n = \max_x [F_n(x) - F(x)], \quad (2)$$

where:  $F_n$  – empirical distribution function.

The results were also analyzed in terms of creating design guidelines for devices supporting spatial orientation of the blind and partially sighted people with vibration interfaces.

Figure 3 shows the median value of vibrotactile amplitude discrimination thresholds depending on the frequency. There were differences in the threshold values in the range of 0.4 dB. To check if there are statistically significant differences between the results for the five frequencies the Kruskal-Wallis test was performed. The analyzed results follow the assumptions required to perform the Kruskal-Wallis test. The null hypothesis ( $H_0$ ) and the alternative hypothesis ( $H_a$ ) were stated as follows:

$H_0$  – values of vibrotactile amplitude discrimination thresholds do not depend on the frequency;

$H_a$  – values of vibrotactile amplitude discrimination thresholds depend on the frequency in at least two of the observed groups.

Value  $H = 12.96364$  ( $p = 0.0115$  is less than 0.05) is in the critical area of statistics. As a result, the null hypothesis can be rejected, and it can be assumed that there is a statistical difference between vibrotactile amplitude discrimination thresholds for different frequencies, for at least two measured frequencies. Further analyzes (Table 2) showed statistically significant lower thresholds for 31.5 Hz versus 125 Hz. No statistically significant differences were found between the other frequency pairs [5].

Table 2. P-value for multiple (two-sided) comparisons  
Kruskal-Wallis Test:  $H = 12.96364$   $p = 0.0115$

	25 Hz	31.5 Hz	63 Hz	125 Hz	250 Hz
25 Hz		1.000000	1.000000	0.228343	1.000000
31.5 Hz	1.000000		1.000000	<b>0.006752*</b>	0.234639
63 Hz	1.000000	1.000000		0.577850	1.000000
125 Hz	0.228343	<b>0.006752*</b>	0.577850		1.000000
250 Hz	1.000000	0.234639	1.000000	1.000000	

Figure 4 compares the median values of vibrotactile amplitude discrimination thresholds for men and women. It was noticed that men achieved lower or the same as women thresholds for low frequencies (received by FAI mechanoreceptors). The opposite situation occurs at high frequencies (received by FAII mechanoreceptors).

The Kolmogorov-Smirnov test was performed to check whether the observed differences of thresholds in individual groups were statistically significant.

The null hypothesis ( $H_0$ ) and the alternative hypothesis ( $H_a$ ) were stated as follows:

$H_0$  – values of vibrotactile amplitude discrimination thresholds do not differ between women and men;

$H_a$  – values of vibrotactile amplitude discrimination thresholds differ between women and men.

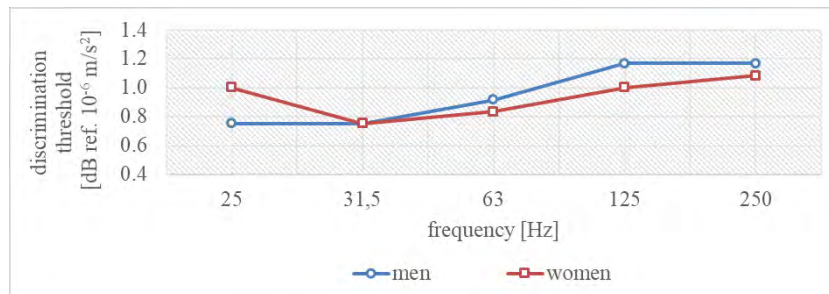


Figure 4. Median value of vibrotactile amplitude discrimination thresholds measured on the wrist depending on the gender of the subjects

According to the Kolmogorov-Smirnov test the p value for comparison between different frequencies is greater than 0.1 ( $\alpha = 0.05$ ). Therefore, the null hypothesis could not be rejected. For all tested frequencies, it was found that the observed differences in the results of men and women are not statistically significant. Further analyzes were conducted for all subjects without considering the division by gender.

Figure 5 compares the median values of vibrotactile amplitude discrimination thresholds for the blind, partially sighted and normally sighted people.

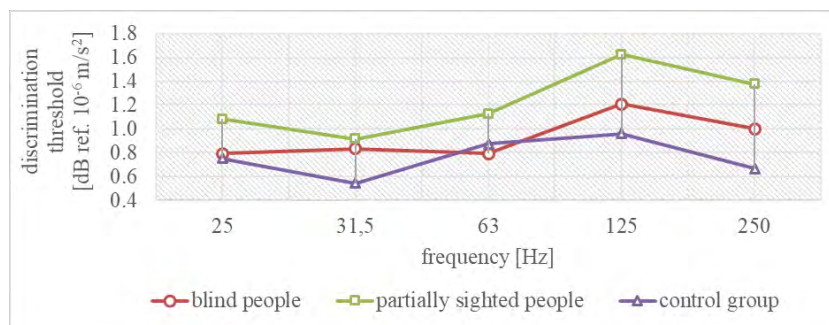


Figure 5. Median value of vibrotactile amplitude discrimination thresholds measured on wrist depending on the degree of visual disability of the subjects

To check if there were statistically significant differences between the results for the blind, partially sighted and normally sighted people the Kruskal-Wallis test was performed. The null hypothesis ( $H_0$ ) and the alternative hypothesis ( $H_a$ ) were stated as follows:

$H_0$  – values of vibrotactile amplitude discrimination thresholds do not depend on the degree of visual disability of the subjects;

$H_a$  – values of vibrotactile amplitude discrimination thresholds depend on the degree of visual disability of the subjects in at least two of the observed groups.

For all tested frequencies the value of  $H$  is outside the critical range (Table 3). As a result, null hypothesis cannot be rejected, and it can be assumed that values of vibrotactile amplitude discrimination thresholds do not depend on the degree of visual disability of the subjects for following frequencies: 25; 31.5; 63; 125 and 250 Hz.

Table 3. The Kruskal–Wallis test by ranks

Frequency	Kruskal-Wallis Test – grouping variable – degree of visual disability
25 Hz	$H = 2.724989$ $p = 0.2560$
31.5 Hz	$H = 5.339776$ $p = 0.0693$
63 Hz	$H = 3.410659$ $p = 0.1817$
125 Hz	$H = 1.323420$ $p = 0.5160$
250 Hz	$H = 8.145166$ $p = 0.0170$

### 3. Conclusions

The paper presents the results of the research on vibration perception of blind, partially sighted and normally sighted people. Based on the research presented in the paper vibrotactile amplitude discrimination thresholds for frequencies 25; 31.5; 63; 125 and 250 Hz were determined. Additional analysis was conducted to check whether the frequency and the gender of the subject affect the value of vibrotactile amplitude discrimination threshold. Based on the analyzes and statistical tests it was found that:

- just-noticeable difference of vibration amplitude on the wrist is equal around 1 dB;
- there are no statistically significant differences between the most tested frequency pairs. Vibrotactile amplitude discrimination threshold is statistically significant lower for 31.5 Hz versus 125 Hz;
- there are no statistically significant differences between values of vibrotactile amplitude discrimination thresholds for men and women;
- there are no statistically significant differences between values of vibrotactile amplitude discrimination thresholds for blind, partially sighted and normally sighted people.

In addition, the analysis allowed the development of design guidelines for vibration interfaces considering the human factor, with emphasis on the needs of visually impaired people. When designing vibration interfaces worn on the wrist using 25 frequencies; 31.5; 63; 125 and 250 Hz or similar, observe the following:

- it is recommended to use signals with different levels of amplitude, because changes in this vibration signal parameter are well recognized by the human sensory system;
- it is recommended to use amplitude differences greater than 1 dB for signals with frequencies up to 63 Hz and amplitude differences greater than 2 dB for signals with frequencies above 63 Hz.

### Acknowledgments

The project described in this paper has been executed within the project No. 16.16.130.942 in Department of Mechanics and Vibroacoustics of AGH University of Science and Technology in Cracow, Poland.

### References

1. T. Majewski, *Methodological guide for teachers working with children with vision impairment in the integration system*, Warsaw, PWZN, 1997.
2. J. Wiciak, B. Borkowski, D. Czopek, *A System for Determination of Areas Hazardous for Blind People Using Wave-vibration Markers Final Conclusions and Technical*, Application Acta Physica Polonica A, **123.6** (2013) 1101 – 1105.
3. E. Bogusz, G. Mrozik, E. Skrodzka, *Investigation of Vibratory Perceptron Thresholds in Blind and Visually Handicapped People in Chosen Areas of the Palm and the Wrist*, Acta Physica Polonica A, **121(1A)** (2012) A19 – A23.
4. J. Wiciak, B. Borkowski, D. Czopek, R. Trojanowski, A. Okarmus, L. Zalewska, E. Bogusz, A. Furman, E. Skrodzka, *Blind in a big city: system for determination of hazardous and important areas for blind and visually impaired people using wave-vibration markers*, Kraków: AGH publishing houses, 2012.
5. D. Czopek, T. Korbziel, B. Kukulski, P. Małecki, P. Pawlik, B. Stępień, T. Wszolek, W. Wszolek, *New methods of signal processing in the selected vibroacoustics problems*, Monographs of the Department of Process Automation of the AGH in Krakow 2016, ISBN: 978-83-64755-20-0.
6. E. Ozimek, *Sound and its perception. Physical and psychoacoustic aspects*, Warsaw-Poznan: PWN 2002.
7. H. Levitt, *Transformed Up-Down Methods in Psychoacoustics*, The Journal of the Acoustical Society of America, **49(2 (2))** (1970) 467 – 477.
8. F. A. Kingdom, N. Prins, *Psychophysics. Practical Introduction*, Academic Press, 2010.
9. M. Chruściel, *LabVIEW in practice*, Legionowo: BTC Publishing House, 2008.
10. S. McClave, *Probability and Statistics for Engineers*, Boston: PWS-KENT Publishing Company 1986.

## Simulation of Direct-Sequence Spread Spectrum Data Transmission System for Reliable Underwater Acoustic Communications

Iwona KOCHAŃSKA

*Gdansk University of Technology, Faculty of Electronics, Telecommunication and Informatics, Department of Marine Electronics Systems,  
G. Narutowicza 11/12, 80-233 Gdansk, Poland, iwona.kochanska@pg.edu.pl*

Jan H. SCHMIDT

*Gdansk University of Technology, Faculty of Electronics, Telecommunication and Informatics, Department of Marine Electronics Systems,  
G. Narutowicza 11/12, 80-233 Gdansk, Poland, iwona.kochanska@pg.edu.pl*

### Abstract

Underwater acoustic communication (UAC) system designers tend to transmit as much information as possible, per unit of time, at as low as possible error rate. It is a particularly difficult task in a shallow underwater channel in which the signal suffers from strong time dispersion due to multipath propagation and refraction phenomena. The direct-sequence spread spectrum technique (DSSS) applied successfully in the latest standards of wireless communications, gives the chance of reliable data transmission with an acceptable error rate in a shallow underwater channel. It utilizes pseudo-random sequences to modulate data signals, and thus increases the transmitted signal resilience against the inter symbol interference (ISI) caused by multipath propagation. This paper presents the results of simulation tests of DSSS data transmission with the use of different UAC channel models using binary spreading sequences.

**Keywords:** underwater acoustics communication, UAC, direct-sequence spread spectrum, DSSS, PRBS

### 1. Introduction

Underwater acoustics communication (UAC) systems for both civil and military applications operate in tough propagation conditions, especially when the transmission is performed in shallow waters. The signal transmitted in shallow UAC channel suffers from time dispersion, caused by multipath propagation due to signal reflections from sea surface, sea bottom and other objects present in water, and by the refractions phenomenon due to the significant changes of sound velocity as a function of depth. Moreover, the signal suffers from frequency dispersion caused by the Doppler effect – much more pronounced in ultrasonic signals than in the case of electromagnetic ones [1, 2].

To ensure reliable communication in a shallow UAC channel, it is necessary to implement techniques used in modern wireless communication systems. Among them, spread spectrum (SS) techniques deserve special attention. Originally, the spread spectrum technique was used in military applications including guidance and communication systems. With the introduction and the evolution of cellular mobile radio systems, spread spectrum communications have been applied as an efficient technique in many practical systems since the late 1980s [3]. The many features of pseudo random signal processing techniques that are important for spread spectrum communications include the ability to

cope with multipath propagation, the resistance to interference, and the potential of sharing spectrum resources with other users. Spread-spectrum systems have low probability of intercept (LPI). Transmitter-receiver pairs using independent random carriers can operate in the same bandwidth with minimal co-channel interference. Moreover, a SS systems have cryptographic capabilities when the data modulation cannot be distinguished from the carrier modulation, and the carrier modulation is effectively random to an unwanted observer [4]. Among the SS systems can be distinguished: direct-sequence spread spectrum (DSSS) systems, frequency hopping spread spectrum (FHSS) systems and time hopping spread spectrum (THSS) systems. This paper focuses on DSSS systems only.

There are several reported UAC systems using DSSS technique. In [5] a DSSS signaling has been used to increase the SNR per data symbol and resolve multipath components for single-user, short range application. The system utilized 10 kHz of bandwidth and short spreading codes in order to combat multipath effects in shallow water. The research described in [6] includes discussions of multiuser access and report favorable results in the 1–2 kHz frequency band at long ranges for a single user. In [7] the authors proposed the UAC modem. Two Gold codes of length 2047 are used for spreading the differentially coded data bits. The bit rate of the system is 100 bps and the carrier frequency is set to 12 kHz. The proposed modem was tested in the Baltic sea at approximate range about 3000 m. The authors obtained error-less transmission of blocks of 200 data bits.

The DSSS-based UAC systems described in literature use specific PN sequences and a fixed transmission bandwidth. To the best of the authors' knowledge, there are no publications presenting the analysis of underwater DSSS systems performance depending on its bandwidth or PN sequence used. This paper presents a comparison of DSSS-based UAC system performance due to its bandwidth and spreading sequence. In the simulation tests performed, the bit error rate (BER) of data transmission was estimated. The channel was modeled as an additive white Gaussian noise (AWGN) channel, the Rician fading channel, and as replay channel simulated by impulse responses of UAC channel measured during the inland water experiment. The results allow for rough estimation of the transmission bandwidth and the type and the rank of the spreading sequence necessary to obtain the assumed BER.

## 2. Spread Spectrum technique

The foundations of spread spectrum systems are provided by the Shannon-Hartley theorem in terms of the channel capacity:

$$C = B \log_2 \left( 1 + \frac{S}{N} \right) \quad (1)$$

Where:  $C$  denotes the capacity in bits per second of a channel perturbed by additive white Gaussian noise,  $S$  is the average received signal power,  $N$  is the average noise power, and  $B$  is the bandwidth available to the band-limited system. It can be simplified to equation:

$$C \sim 1.44B \frac{S}{N} \quad (2)$$

A reduction in the signal-to-noise ratio (SNR) can be compensated by proportionally increasing the bandwidth  $B$ . This is a basic rule for spread spectrum systems, which spread data signals over a much wider frequency band compared to the minimum bandwidth required to transmit the information [4].

The main idea of DSSS is to modulate data signals with a pseudo random signal called the spreading sequence. A transmit signal using BPSK modulation can be written as [4]:

$$s(t) = A_s d(t) c(t) \cos(2\pi f_0 t + \theta) \quad (3)$$

Where  $A_s$  denotes the amplitude of the signal,  $d(t)$  is the data signal,  $c(t)$  is the spreading sequence,  $f_0$  is the carrier frequency, and  $\theta$  is the carrier phase.

### 3. Spreading sequences

For simulation tests of the DSSS system, three types of spreading sequences were chosen, namely m-sequences, Gold codes and Kasami codes.

M-sequences or maximal length sequences are one of the most important classes of pseudo random sequences. They have good pseudo randomness properties and can be generated by a linear feedback shift register. Its name corresponds to the fact, that they are the maximal length sequence that can be generated by a shift register of a given length [4].

Gold codes and Kasami codes are both derived from m-sequences. These codes have good cross-correlation properties and thus are useful for multi-user ranging systems, such as GPS system. Gold codes are formed by combining two specific m-sequences of the same length  $2^n - 1$  such that their absolute cross-correlation is less than or equal to  $2^{(n+2)/2}$ , where  $n$  is the size of the linear feedback shift register used to generate the m-sequence. They are combined this way, that one m-sequence is added with all shifts of other sequence using modulo two addition.

Kasami codes are formed by decimating an m-sequence by taking every  $2^{n/2} + 1$  bit from the periodic sequence. Combining the cyclic decimated sequence with all shifts of the original one, a set of  $n$  sequences with good cross-correlation properties is created.

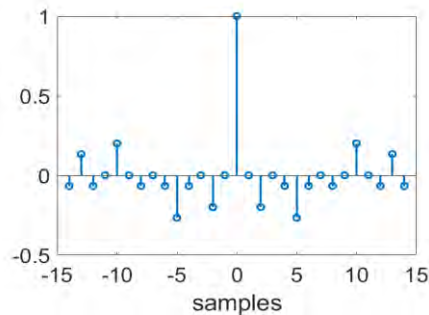


Figure 1. Autocorrelation function of m-sequence of rank  $L = 4$

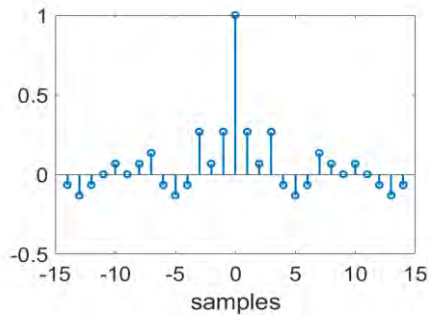


Figure 2. Autocorrelation function of Gold code of rank  $L = 4$

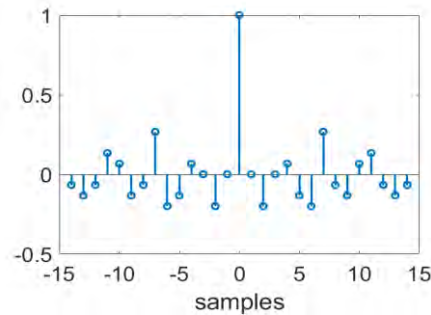


Figure 3. Autocorrelation function of Kasami code of rank  $L = 4$

Fig. 1-3 show autocorrelation functions of m-sequences, Gold codes, and Kasami codes. It is clearly seen that the autocorrelation properties of Kasami codes and Gold codes are worse than the autocorrelation property of m-sequence of the same length. On the other hand, there are very few m-sequences of any given length with good cross-correlation properties. Gold codes and Kasami codes extend the set of binary PN sequences that have controlled cross-correlation properties, and thus are suitable for wireless communication systems with code-division multiple access (CDMA) [4].

### 3. Simulation of DSSS data transmission

DSSS data transmission simulations were performed in the Matlab environment to estimate the bit error rate (BER) in the propagation conditions similar to those in stationary UAC channel. Fig. 4 shows a block diagram of DSSS modulator. Each of the data bits is multiplied with PN sequence. The product is upsampled by a factor of  $R = f_s/B$ , where  $f_s$  is the sampling frequency and  $B$  denotes the system bandwidth. Next, the signal is passed through the binary phase shift keying (BPSK) block and the pulse shaping filter. The output is complex-value digital signal that is used for simulation of baseband transmission in UAC channel. At the receiver side (Fig. 5), the input signal is again filtered by the pulse shaping filter and passed through the BPSK demodulator. Next, the real-value digital signal is downsampled and the matched filtration is performed by a filter with coefficients corresponding to the PN code used in the transmitter. Information detection is performed by summing up the amplitudes of samples from the range corresponding to the length of a single PN sequence and checking if the result is a positive (1) or negative (0) value.



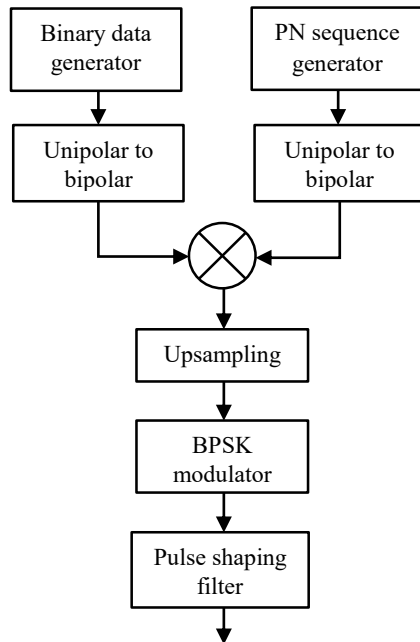


Figure 4. DSSS transmitter block diagram

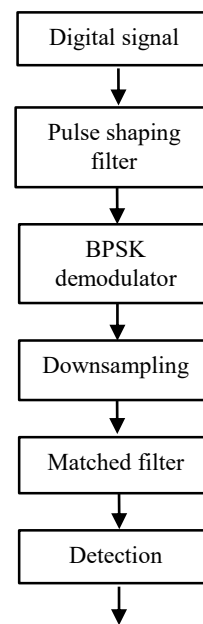


Figure 5. DSSS receiver block diagram

Three kinds of channel models were used, namely the AWGN channel model with varying signal-to-noise ratio (SNR), the Rician fading model, and UAC channel impulse responses measured during the inland water experiment. In each case, the sampling rate was  $f_s = 200$  kHz and the signal bandwidth was in the range from 2 kHz to 10 kHz. M-sequences, Gold codes and Kasami codes with lengths in the range of  $2^3 - 1$  to  $2^{10} - 1$  were used as spreading sequences. In each communication test 20 kbits were transmitted, which allowed to estimate the bit error rate at the level of  $10^{-3}$ . For each spreading sequence length a maximum bandwidth was found for which  $\text{BER} < 10^{-3}$  was obtained.

During the tests using AWGN channel the SNR was varying from -10 dB to -30 dB. Fig. 6 shows the results obtained for the values of SNR: -12 dB, -18 dB, -24 dB and -30 dB. In case of SNR = -12 dB all spreading sequences of rank equal or higher than 5 allow for reliable transmission in the maximum tested band (10 kHz). Tests performed for SNR = -18 dB and SNR = -24 dB have shown, that using m-sequences and Kasami codes give better results than Gold codes. In case of SNR = -30 dB the reliable transmission is possible only for Kasami codes and m-sequences of rank 9 or 10.

Real impulse responses (IR) gathered during the inland water experiment were used as another channel model in simulation tests. The impulse response were measured in May, 2017 in Wdzydze Lake. The detailed description of the experiment can be found in [8, 9]. The impulse responses were measured at the distances of 300 m and 500 m. The exemplary IRs are presented in Fig. 7. Fig. 8 shows the rank of spreading sequences

and signal bandwidth for which  $\text{BER} < 10^{-3}$  was obtained during the simulation tests. It is clearly seen that increasing the signal bandwidth means that a higher-order spreading sequence must be used to obtain the same BER. In the case of an impulse response with a larger number of multi-path reflections (Fig. 8b), the reliable transmission ( $\text{BER} < 10^{-3}$ ) is possible using at least the spreading sequence of rank 5 (if the signal bandwidth is equal to 3 kHz). In case of impulse response with less multipath reflections, the same BER can be obtained in 2 kHz bandwidth with the use of spreading sequence of length  $2^3 - 1$ .

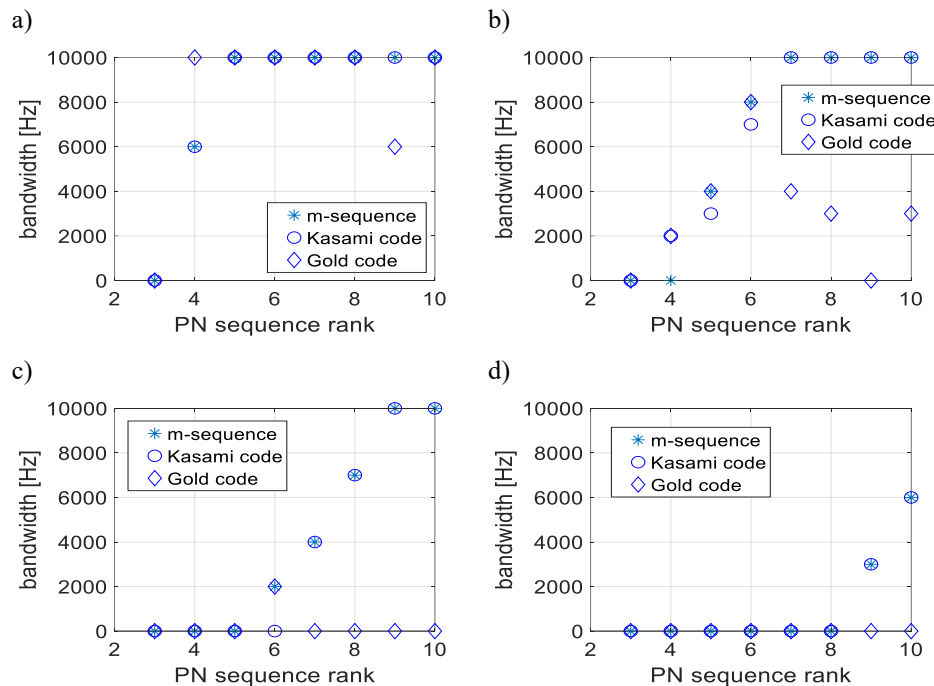


Figure 6. Maximum transmission bandwidth for which  $\text{BER} < 10^{-3}$ , as a function of PN sequence rank; AWGN channel model with  $\text{SNR} = -12$  dB (a),  $-18$  dB (b),  $-24$  dB (c), and  $-30$  dB (d).

The third channel model was Rician fading channel in two configurations: with 3 and 5 signal paths. The calculations were performed for rank of spreading sequences up to 9 due to the limited computational capabilities of the computer hardware used for simulations. The results tests are shown in Fig. 9. In case of 3 signal paths the data transmission with  $\text{BER} < 10^{-3}$  was possible for spreading sequence of rank 5 or higher, except for rank 8, in the case of which the transmission with  $\text{BER} < 10^{-3}$  could not be obtained. Surprisingly for the PN sequence of rank 5 and 6, the Gold's codes and the Kasami codes proved to be much more effective than m-sequences. In case of 5 signal paths no reliable data transmission was possible, except for m-sequence of rank 10 and signal bandwidth equal to 2 kHz.

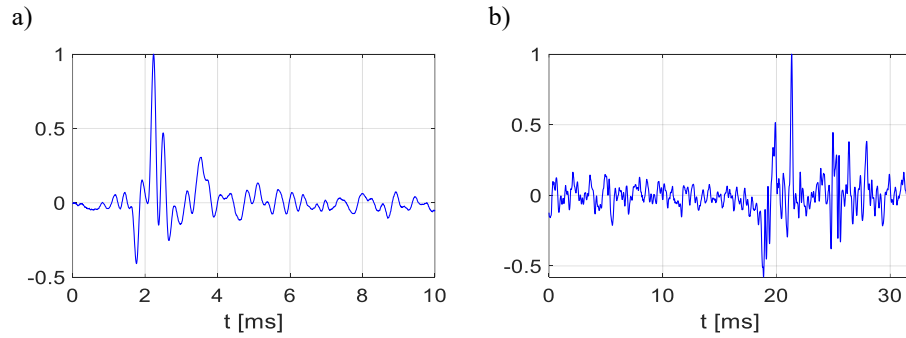


Figure 7. Impulse response of UAC channel at a distance  $d = 300$  m (a),  $d = 500$  m (b)

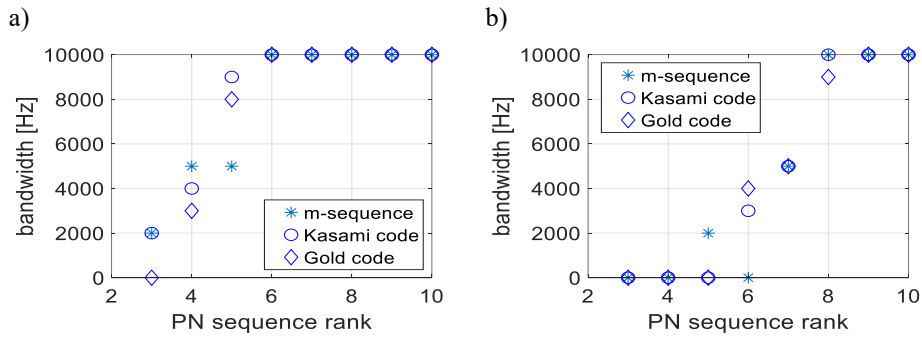


Figure 8. Maximum transmission bandwidth for which  $\text{BER} < 10^{-3}$ , as a function of PN sequence rank; IR of UAC channel at a distance  $d = 300$  m (a) and  $d = 500$  m (b)

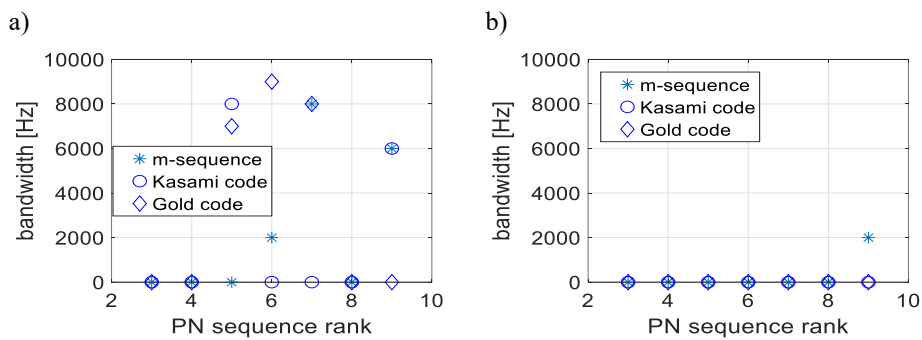


Figure 9. Maximum transmission bandwidth for which  $\text{BER} < 10^{-3}$ , as a function of PN sequence rank; Rician fading channel model with 3 (a) and 5 (b) reflections

#### 4. Conclusions

The simulation tests of DSSS data transmission with the use of three UAC channel models were performed. M-sequences, Kasami codes, and Gold codes of rank varying from 3 to 10 were used as spreading sequences. For all channel models increasing the signal bandwidth means that a higher-order spreading sequence must be used to obtain the same data transmission quality. Although Kasami codes and Gold codes have worse autocorrelation properties than m-sequences, their application allowed to achieve reliability similar to that obtained in the system using m-sequences. In tests with the AWGN channel m-sequences and Kasami codes ensured better performance than Gold codes. Reliable transmission was possible even for  $\text{SNR} = -30$  dB in the 2 kHz band and using a spreading sequence of rank 9. In tests with the Rician channel, the transmission was successful if 3 signal paths were simulated. Increasing the number of fading paths makes it impossible to obtain the  $\text{BER} < 10^{-3}$ .

The results of simulation tests performed allow for rough estimation of the transmission bandwidth and the spreading sequence necessary to obtain the desired BER in DSSS-based UAC system.

#### References

1. I. Kochanska, J. Schmidt, M. Rudnicki, *Underwater Acoustic Communications in Time-Varying Dispersive Channels*, Federated Conference on Computer Science and Information Systems, **8** (2016) 467 – 474.
2. I. Kochanska, I. Nissen, J. Marszał, *A method for testing the wide-sense stationary uncorrelated scattering assumption fulfillment for an underwater acoustic channel*, The Journal of the Acoustical Society of America, **143**(2), EL116 – EL120.
3. R. A. Scholtz, *The Origins of Spread-Spectrum Communications*, IEEE Transactions on Communications, **COM-30**, No. 5, May 1982.
4. H. J. Zepernick, A. Finger, *Pseudo Random Signal Processing: Theory and Application*, John Wiley & Sons Ltd, 2005.
5. J. H. Fischer, R. Kendrick et al., *A high data rate, underwater acoustic data-communication transceiver*, in Proc. Oceans '92, Newport, RI.
6. G. Loubet, V. Capellano, R. Filipiak, *Underwater spread-spectrum communications*, in Proc. Oceans '97, Halifax, Canada.
7. E. M. Sozer, et. al., *Direct sequence spread spectrum based modem for under water acoustic communication and channel measurements*, Proceedings of Oceans '99 MTS/IEEE, 1999.
8. I. Kochanska, J. Schmidt, A. Schmidt, *Measurement of impulse response of shallow water communication channel by correlation method*, Hydroacoustics, **20** (2017) 149 – 158.
9. I. Kochanska, J. Schmidt, *Estimation of Coherence Bandwidth for Underwater Acoustic Communication Channel*, Proceedings of 2018 Joint Conference – Acoustics, 2018.

## **Assessment of Sound Absorbing Properties of Composite Made of Recycling Materials**

Patryk GAJ

*Institute of Power Engineering*

*OTC – Thermal Technology Branch „ITC” in Lodz, 93-208 Łódź, ul Dąbrowskiego 113,  
patryk.gaj@itc.edu.pl*

Joanna KOPANIA

*Lodz University of Technology, 90-424 Lodz, ul Piotrkowska 266,  
joanna.kopania@p.lodz.pl*

Kamil WÓJCIAK

*Institute of Power Engineering*

*OTC – Thermal Technology Branch „ITC” in Lodz, 93-208 Łódź, ul Dąbrowskiego 113,  
kamil.wojciak@itc.edu.pl*

Grzegorz BOGUSŁAWSKI

*Lodz University of Technology, 90-424 Lodz, ul Piotrkowska 266,  
grzegorz.boguslawski@p.lodz.pl*

### **Abstract**

In modern world we are searching methods to reuse most of industrial disposals produced during manufacturing. Some kinds of materials, like scraps from acoustic foam, however are not so easily utilized using recycling methods for its primary usage. Disposals produced during the manufacturing process can be compressed and reused as sound absorbing material.

The purpose of this article is to examine sound absorbing properties of material made of acoustic foam disposals and compare it with sound protection materials, which are commonly used. Sound absorbing damping were tested using Kundt's tube and reverberation room examination method. Tests were carried out according to standard PN-EN ISO 354:2005, for reverberation room examination, and according to Bruel&Kjaer electronic instruments technical manual for Kundt's tube examination.

**Keywords:** sound absorbing materials, recycling, sound absorption coefficient, composite materials

### **1. Introduction**

Everlasting increase of consumptions of goods by the humans leads to production rise, however it is also connected to escalation of waste production leading to growing industrial pollution.

In XXI century we can observe growing trend to utilize as much waste as possible by recycling. However even if some materials are widely reused, others, like acoustic foam, are omitted in recycling utilization process.

Moreover because of everlasting technological improvements in various areas of services people are surrounded by noise generation. Cars, industrial machinery, communal transport etc. generates high amount of infrasound, which even if not

perceived by our ears, has negative impact on our health conditions [1]. Our internal organs have specific density, which results in certain frequencies have impact on them. People, which are exposed on constant low frequency noise are reporting more often problems with headaches, fatigue, ringing in the ears and heart problems. It was reported that low-frequency sound of value equal to 75 dB is assumed as critical for human health, omitting specific personal sound vulnerability [2].

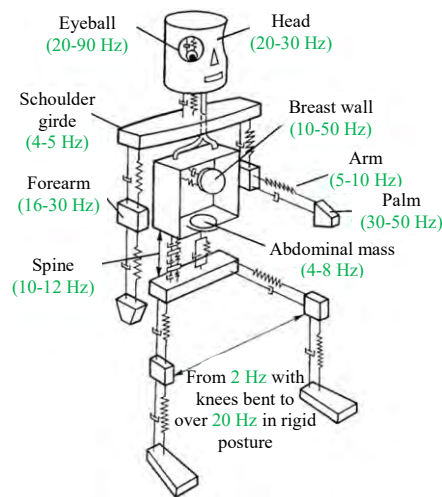


Figure 1. Characteristic frequencies of human organs vibrations [3]

What is more, infrasound, due to its long wave nature, they propagates well and are not dampened in the air. For example wave of frequency equal to 10 Hz decays 0.1 dB per kilometer. Because of wavelength of infrasound typical acoustic isolation are not so effective and other damping materials has to be used [4, 5].

## 2. Methodology research

The research was carried out in Institute of Power, Thermal Technology Branch “ITC” laboratory. Two examination methods were used. Test were performed using four types of material: three commonly used and recycling composite made of mix of scraps from polyurethane and polyethylene foams. Components ratio of scraps is classified by the producer. Materials were prepared for the tests accordingly. Thickness of tested materials are presented in Table 1.

Table 1. Thickness of examined materials

Material Type	Thickness [mm]	Density [kg/m <sup>3</sup> ]
Acoustic Foam (Plain)	56	35
Acoustic Foam (Pyramids)	55	28
Glasswool	49	35
Recycling Composite	53	40

First phase was checking physical absorption coefficient with impedance tube method (Figure 1). SVAN 912 analyzer was used to record test results. Circular samples with 100 mm diameter were prepared for examination.

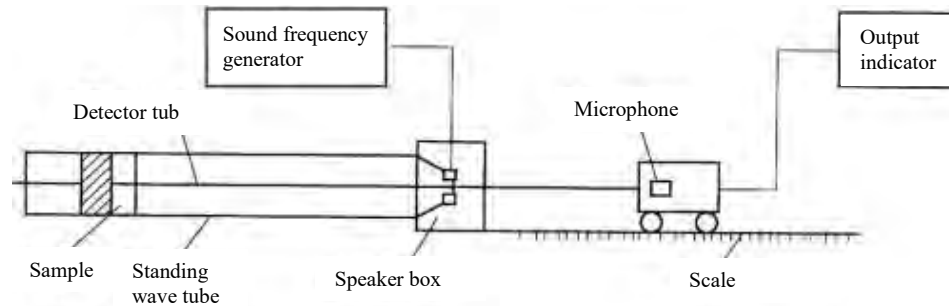


Figure 2. Impedance tube scheme [6]

Measuring equipment consists of tube with 100 mm diameter, which is connected to the speaker box. Microphone probe is inserted via the hole in the speaker and connected to the trolley, which is moving on track with scale.

Physical sound absorption coefficient  $\alpha_f$  is calculated using following formula (1).

$$\alpha_f = 1 - \left( \frac{n-1}{n+1} \right)^2 \quad (1)$$

Where:

$\alpha_f$  – physical sound absorption coefficient,

$p_{\max}/p_{\min}$  – maximal/minimal value of acoustic pressure,

$n$  – value of maximal acoustic pressure divided by minimal acoustic pressure.

To obtain meaningful results examinations were carried out for three samples of every kind of material

Second phase of test was examination of materials using reverberation room method. At start samples were inserted into reverberation room for 24 hours to acclimatize to the atmospheric conditions in chamber.

In order to measure reverberation sound absorption coefficient for every sample two measurements were conducted: first for reverberation of chamber and second – reverberation of chamber with sample placed on the floor of the chamber.

For every test 10 m<sup>2</sup> of material was placed on the floor of the reverberation chamber according to standard PN-EN ISO 354:2005 [7]. For examination nor140 acoustic analyzer connected to the rotating boom has been used.

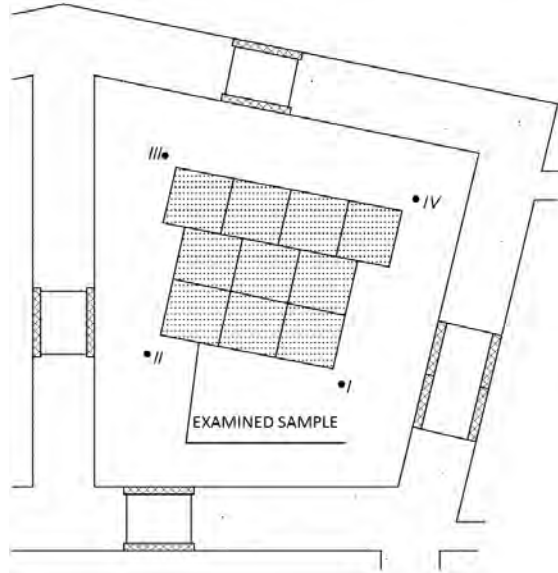


Figure 3. Setup of sample in reverberation room

To calculate reverberation sound absorption coefficient of material, firstly equivalent area of sound absorbing area within reverberation room, without tested material must be calculated (formula 2):

$$A_1 = \frac{55.3 \times V}{c \times T_1} - 4 \times V \times m_1 \quad (2)$$

where:

$V$  – volume of the reverberation room [ $\text{m}^3$ ],

$c$  – speed of sound in specific temperature [ $\text{m/s}^2$ ],

$T_1$  – reverberation time of the empty reverberation room [s],

$m_1$  – acoustic power damping coefficient [ $1/\text{m}$ ].

After that, equivalent area of sound absorbing area within reverberation room, with tested material must be calculated (formula 3):

$$A_2 = \frac{55.3 \times V}{c \times T_2} - 4 \times V \times m_2 \quad (3)$$

Which leads to calculating equivalent area of sound absorbing area of the tested material (formula 4):

$$A_T = A_2 - A_1 \quad (4)$$



Reverberation absorption coefficient is calculated according to formula 5:

$$\alpha_s = \frac{A_T}{S} \quad (5)$$

where:

$S$  – area of examined material

What is important to remember that unlike to physical sound absorption coefficient, reverberation sound absorption coefficient can be greater than one due to e.g. diffraction effects on the edges of examined material and additional sound absorbing surface of sidewalls. This is why it can't be presented as percentage value [8].

### 3. Research results

On the chart (Figure 4) are presented results of sound absorption examination conducted via Kundt's tube method. As it is presented it can be noticed that all of tested materials doesn't show sound absorbing properties in range up to 200 Hz ( $\alpha < 0,3$ ). However, after that range, we can notice increase of absorbing properties. The best overall results are presented by sample made of recycling acoustic materials. Only glasswool was more effective than composite in range of 400-500 Hz.

It is worth to acknowledge, that results for sample made of acoustic foam with pyramids were non-uniform and were highly dependent on from which part of sheet sample was cut off. Placement of pyramids within examined sample had high influence on value of absorption coefficient.

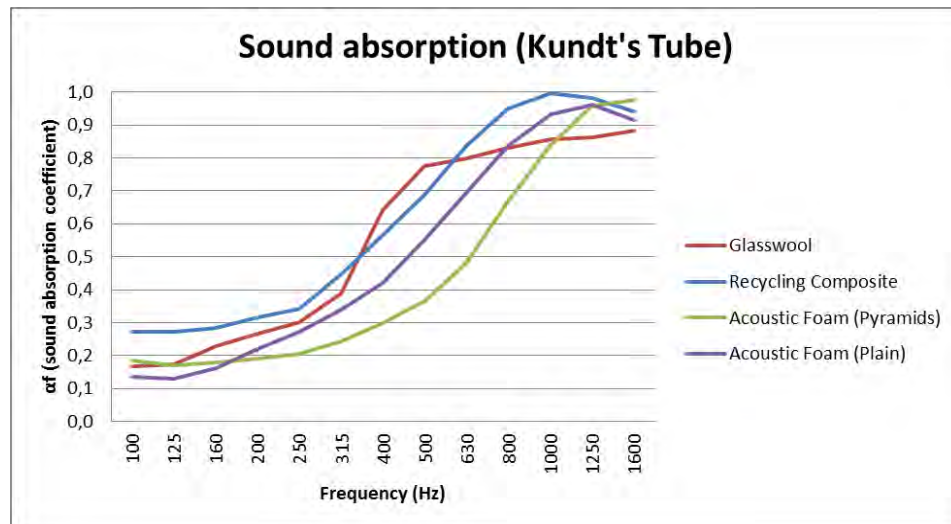


Figure 4. Physical sound absorption coefficient of examined materials

On another chart (figure 5) are presented results of sound absorption examination conducted by reverberation room method. As it is presented it can be noticed that all of examined materials doesn't show sound absorbing properties up to 125 Hz. After this frequency we can observe increasing sound absorbing properties of materials, mainly for recycle composite, glasswool and plain acoustic foam. Recycling composite presents the best overall absorbing properties. Again only glasswool presented better results than composite in limited frequency range (200-315 Hz).

Tests of acoustic foam with pyramids exposed lower than other materials ability to absorb sounds.

It is worth to remember that results of this examination cannot be presented in percentage form, because coefficient  $\alpha$  is greater than 1 and it is used mostly to check trend of sound absorbing properties of the materials, where bigger samples could be used.

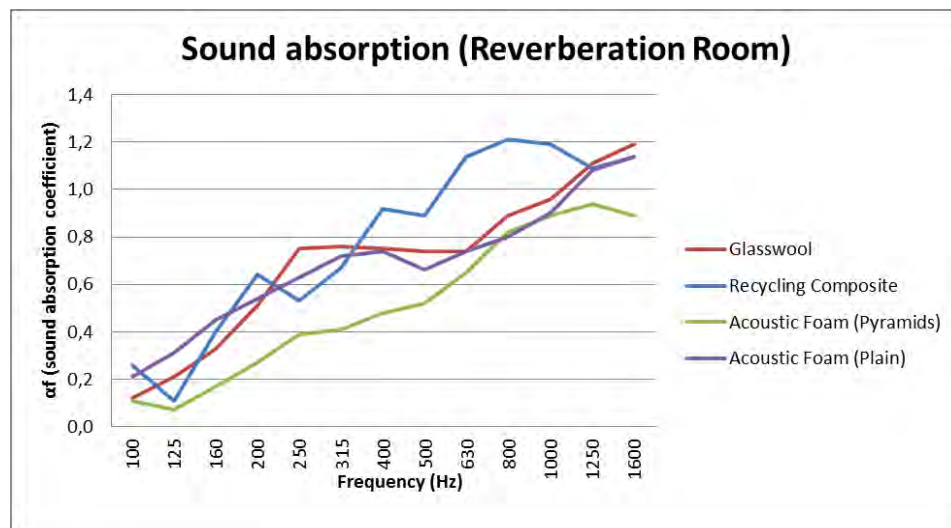


Figure 5. Reverberation sound absorption coefficient of examined materials

#### 4. Conclusions

Conducted examinations has shown that recycling acoustic material composite has comparable, or even superior sound absorbing properties than more commonly used materials in tested frequency range.

However it is worth remembering, that this kind of material has to be used in conditions with no high temperature and humidity, because this parameters could damage this kind of material.

What is more, production process of recycling composite is based on using scraps of previously used in company preparing acoustic materials, so acoustic properties of this type of composite can vastly differ according to materials used in manufacturing process of composite and its ration.

Tests also exposed lower absorbing properties, in comparison to the other materials, of acoustic foam with pyramids, which is one of the most well –known and widely used type of sound absorbing material in music industry and soundproofing the flats.

## References

1. J. Kopania, P. Gaj, *Badania nad tłumieniem hałasu w zakresie niskich częstotliwości słyszalnych (wytyczne konstrukcyjno-materiałowe). Etap I. Badania akustyczne materiałów kompozytowych na potrzeby tłumienia dźwięków w układach wentylacyjnych*, Łódź, Polska, (2018) 3 – 5.
2. N. Rajishore, P. Rajiv, *Acoustic Textiles, Textile Science and Clothing Technology*, doi 10.1007/978-981-10-1476-5\_1.
3. H. Gierke, A. Brammer, *Effects of shock and vibrations on humans*, McGraw-Hill Education, 2002.
4. C. Zwikker, C. W. Kosten, *Sound absorbing materials*, Elsevier; 1949.
5. H. V. Fuchs, M. Möser, *Sound absorber*, G. Muller, M. Möser, (eds.), *Handbook of Engineering Acoustics*, Springer-Verlag Berlin Heidelberg, (2013) 165 – 213, doi:10.1007/978-3-540-69460-1\_8.
6. *Brüel&Kjær electronic instruments*, Brüel&Kjær, Denmark 1997, 469 – 470.
7. PN-EN ISO 354:2005, *Acoustics – Measurement of sound absorption in a reverberation room*.
8. Y. Na, J. Lancaster, J. Casali et al., *Sound absorption coefficients of micro-fiber fabrics by reverberation room method*, Text Res J., **77**(5) (2007) 330 – 335.



## Cyclodextrine vs D-glucose in the Solutions of the Derivative of 1,4-DHP

Milena KACZMAREK-KLINOWSKA

*Institute of Acoustics, Faculty of Physics, Adam Mickiewicz University,  
ul. Uniwersytetu Poznańskiego 2, 61-614 Poznań, mkacz@amu.edu.pl*

Tomasz HORNOWSKI

*Institute of Acoustics, Faculty of Physics, Adam Mickiewicz University,  
ul. Uniwersytetu Poznańskiego 2, 61-614 Poznań, hornaku@amu.edu.pl*

### Abstract

Hydrophilic and hydrophobic cyclodextrins (CDs) have found a lot of applications in medicine pharmacology, food processing and cosmetology. They can function as a drug carrier material and parent host molecules, increasing drug biocompatibility, optimizing the efficiency of drug activity, and controlling drug release at the desired level. The effectiveness of CD complexation depends on many factors such as the type and the size of both the CD molecule itself and the guest molecule, pH of the solution, and temperature. In aqueous solution of glucose the aggregation of molecules can occur leading to the formation of CD-like structures. In the paper the possibility of the formation of the inclusion complexes of CD and glucose with Nimodipine was investigated with the aid of ultrasonic spectroscopy. By comparing the efficiency of industrial saccharides and glucose in the formation of the inclusion complexes the cost effectiveness of the use of glucose as a substitute for CD can be determined.

**Keywords:** Cyclodextrin,  $\alpha$ -D-glucose, Nimodipine, ultrasounds, host-guest complexes

### 1. Introduction

The interest in the multicomponents systems has possibly influenced the multidisciplinary approach to the study of supramolecular compounds as well as stimulated their practical applications. To optimise the efficacy of drug activity and control its release, some drug carrier materials, such as cyclodextrins (CDs) have been designed. CDs are small cyclic maltooligosaccharides with six to nine glucose residues. They have the shape of torus-like macro ring. It is generally accepted that in aqueous solutions CDs form “inclusion complexes” where water molecules located in the central cavity are replaced by a lipophilic guest molecule. CDs can form complexes with chemical compounds of appropriate size. If the guest is the wrong size, it will not fit properly into the cyclodextrin cavity and some of its functional groups will not penetrate the cavity. The CDs complexes are relatively stable, however, they show poor water solubility (in comparison with simple sugars) and precipitate easily from the solution as crystal form.

CDs are often modified by exchanging the outer hydrogen atoms in hydroxyl groups with the ones that bond with oxygen. In this way the cyclodextrine derivatives such as  $\beta$ -cyclodextrin and methyl- $\beta$ -cyclodextrin (MBCD) can be produced which have better

solubility in aqueous solution (several dozen times). The MB CD was used in the present study [1-3].

$\alpha$ -D-glucose also known as dextrose belongs to the class of organic compounds known as hexoses and is the end product of photosynthesis. The cyclic form which is present in highest concentrations in aqueous solutions arises from the open-chain form by an addition reaction between the aldehyde group and hydroxyl group. Moreover, glucose molecules in aqueous solutions undergo the process called mutarotation. The ring then unfolds back to the open chain form with inverse location of groups at the first carbon atom. The process takes place as long as necessary to reach equilibrium in the solution.

Nimodipine (NM) is a 1,4-dihydropyridine derivative calcium ion channel blocker, applying to improve the blood circulation in the acute cerebrovascular disease. Administered by oral tablets Nimodipine has the low bioavailability, about 13%. Furthermore, the drug is water-insoluble but is soluble in alcohol and decomposes easily under the light. The bioavailability, water-solubility and light stability of Nimodipine could be enhanced by the inclusion of cyclodextrin [4, 5].

In this paper the possibility of the formation of inclusion complexes of MB CD or  $\alpha$ -D-glucose with Nimodipine was investigated with the aid of ultrasonic spectroscopy. The efficiency of complex formation in both solutions is also determined.

## 2. Materials and Methods

The following chemicals were used in this work: Nimodipine (98% purity) was obtained from Sigma Aldrich, ethanol (purity of 99.9%) from POCH, metylo- $\beta$ -CD (purity > 98%) from Wacker Chemie GMBH, and  $\alpha$ -D-glucose (purity > 98%) from Sigma Aldrich. All chemicals were used without further purification. The solutions were prepared with double-distilled and deionized water. First, the standard solutions were prepared but due to the poor solubility of Nimodipine in water it was decided to use aqueous-alcoholic solvent as an environment for the formation of inclusion complexes. Nimodipine is well soluble in ethanol. If the water were to be the only solvent, the preparation of the solutions of appropriate sugar and Nimodipine would require very low concentrations of the drug. But then because of the small volume of the ultrasonic measuring cell (0.7 mL) the measurement error would be very high. For this reason the initial solutions were prepared as binary and, after the solubility analysis, it was decided that the ratio of their components ( $\alpha$ -D-glucose or MB CD to water and NM to ethanol) of 1:750 for each sample is sufficient for carrying out the reliable acoustic measurements and does not lead to the saturation of the solution.

The volume of the sample was 5 mL and the volumes of the solvents were calculated from the expression:

$$V_t = \frac{1 - x_2}{c_1} + \frac{x_2}{c_2} \quad (1)$$

where  $x_2$  is the molar concentration of Nimodipine;  $c_1$  and  $c_2$  are the molarities of the solutions: water with sugars and ethanol with the drug, respectively. It should be noted that for each solution the volumes of the components within the concentrations range of

0–1 mole fraction of Nimodipine are normalized and they sum to 5 mL. The 29 samples for acoustic measurements were prepared according to the procedure described above. Because of the difference in molar mass between MBCD and  $\alpha$ -D-glucose, the amount of  $\alpha$ -D-glucose in solution was seven times bigger than the amount of MBCD thereby it can be assumed that the sugar contents in both solutions were comparable.

The ultrasonic method was used which makes possible to study elastic properties of liquids and thereby their change in rigidity in order to check whether the cavitants are formed in the studied system of sugar (MBCD or  $\alpha$ -D-glucose) + NM. From the speed of sound,  $c$ , and the density,  $\rho$ , the isentropic compressibility can be calculated from the Laplace equation  $k_S = 1/\rho c^2$ . Isentropic compressibility can be transformed into its molar counterpart via the relation  $K_{S,m} = V_m k_S$ , where  $V_m$  is the molar volume. The efficiency of the formation of supramolecular complexes can be determined from the excess molar isentropic compressibility,  $K_{S,m}^E$ , that is the difference between molar isentropic compressibility of the studied mixture calculated from Laplace equation and the molar compressibility for the ideal mixture obeying the Raoult law:

$$K_{S,m}^E = K_{S,m} - K_{S,m}^{id} \quad (2)$$

Since the ideal isentropic compressibility does not obey so-called ideal mixing law, it must be evaluated from the well known thermodynamic identity [6]:

$$K_{S,m}^{id} = \sum_i x_i \left\{ K_{S,i}^* - T \cdot A_{p,i}^* \left[ \left( \sum_i x_i A_{p,i}^* / \sum_i x_i C_{p,i}^* \right) - (A_{p,i}^* / C_{p,i}^*) \right] \right\} \quad (3)$$

where  $A_{p,i}^*$  is the product of the molar volume  $V_{m,i}$  and the isobaric expansivity  $\alpha_{p,i}^*$ ,  $C_{p,i}^*$  is the isobaric molar heat capacity,  $K_{S,i}^*$  is the product of the molar volume  $V_i$  and the isentropic compressibility  $k_{S,i}^*$  referred to “pure”  $i$ -th liquid component. In our case the “pure” components are the initial solutions, e. g. for  $x_2 = 0$  and  $x_2 = 1$ . This principle was also applied to the evaluation of excess molar volume,  $V_m^E$ , using molar masses  $M_{x_1}$ ,  $M_{x_2}$  and densities,  $\rho_{x_1}$ ,  $\rho_{x_2}$ , of the components

$$V_m^E = \frac{M}{\rho} - \left[ (1 - x_2) \frac{M_{x_1}}{\rho_{x_1}} + x_2 \frac{M_{x_2}}{\rho_{x_2}} \right] \quad (4)$$

The measurements of ultrasound velocity and density were carried out in the temperature range of 288.15–313.15 K. Molar heat capacity was also measured since its value is necessary to calculate ideal isentropic compressibility from Eq. (3).

Speed of ultrasonic wave was measured by a resonance method using the ResoScan<sup>TM</sup>System (Germany) apparatus. The ultrasonic speed is determined from the change in resonance frequencies which arise when the ultrasonic signal is transmitted through the path length of the measurement cell. The change in resonance frequencies occur when ultrasonic wave frequency is tuned within a range of 7.3–8.4 MHz. The system permits the measurements of the ultrasonic velocity with the accuracy of  $\pm 0.01$  m/s with temperature precision of  $\pm 0.01^\circ\text{C}$ . The density was measured using a microprocessor gauge of the DMA 38 type by Anton Paar. This instrument permits

density measurements up to  $3 \times 10^3 \text{ kg/m}^3$  with an accuracy of  $\pm 0.2 \text{ kg/m}^3$ . The isobaric molar heat capacity was measured by using DSC Q2000 (TA Instruments) with temperature accuracy of  $\pm 0.01 \text{ K}$ . The heating rate was  $2 \text{ K/min}$ . Each sample was heated from  $270 \text{ K}$  to  $330 \text{ K}$ .

### 3. Results and discussion

The results of the velocity measurements for MBCD + NM and  $\alpha$ -D-glucose + NM systems as a function of the concentration of Nimodipine (in mole fraction) are shown in Figures 1a and 1b.

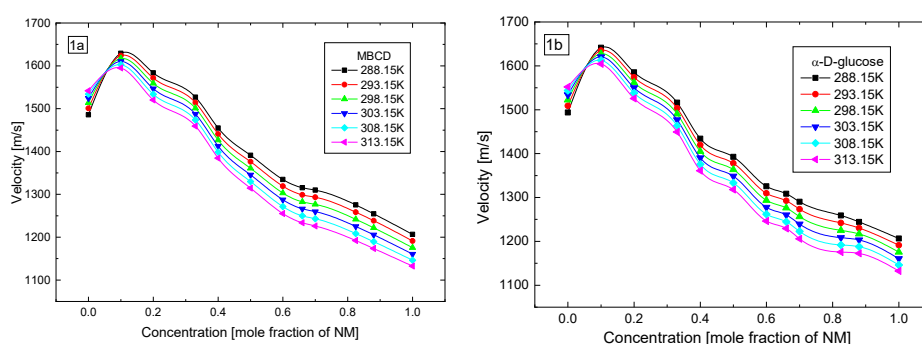


Figure 1. Ultrasound velocity as function of Nimodipine concentration: a) in the system of MBCD + Nimodipine, b)  $\alpha$ -D-glucose + Nimodipine

For both systems the velocity isotherms shown in Figures 1a and 1b intersect and then reach their maxima at the concentration of 0.1 mole fraction of Nimodipine which is due to the introduction of the ethanol to the system (clathrates formation in ethanol-water mixture). The value of the ultrasound velocity at maximum depends on temperature. For higher concentrations of Nimodipine the ultrasound velocity decreases monotonically.

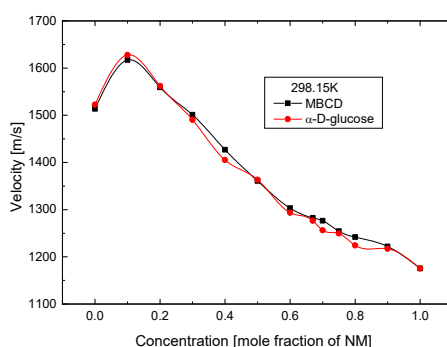


Figure 2. Ultrasound velocity vs Nimodipine concentration at the temperature of  $298.15 \text{ K}$  in the systems of MBCD + NM and  $\alpha$ -D-glucose + NM



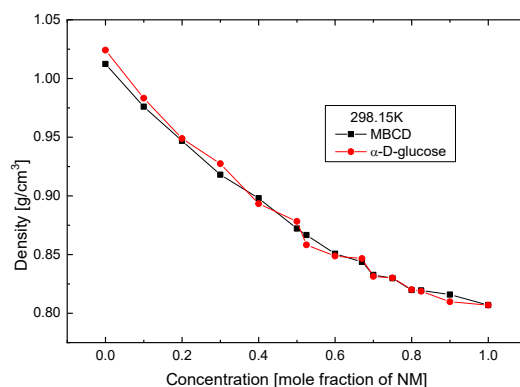


Figure 3. Density vs Nimodipine concentration at the temperature of 298.15 K in the systems of MBCD + NM and  $\alpha$ -D-glucose + NM

Figure 2 shows the dependence of ultrasound velocity and Figure 3 the dependence of density on the concentration of Nimodipine in both systems, with MBCD and  $\alpha$ -D-glucose, at the same temperature of 298.15 K. The greatest difference in ultrasound velocities in both systems amount to only 21.8 m/s, so it can be said that it is not velocity alone that can affect the efficiency of complexation process of one sugar or the other with Nimodipine.

The results of the measurements of molar heat capacity necessary for the evaluation of the excess molar isentropic compressibility are shown in Figure 4.

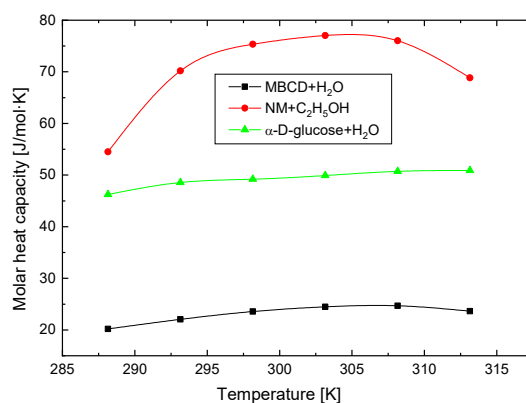


Figure 4. Molar heat capacities for the initial solutions

It should be noted that while density and velocity results were close to each other for both systems, their molar heat capacities differ significantly. It is related to the difference in the structure between MBCD and  $\alpha$ -D-glucose. Although glucose imitate cyclodextrin, mutaroration does not lead to the formation of equally strong glycosides bonds as in CDs.

The excess molar isentropic compressibilities plotted against mole fraction of Nimodipine (presented in Figure 5) show the influence of the reactions taking place in the studied solutions on the final structure of the complexes that are formed in sugar/drug solutions.

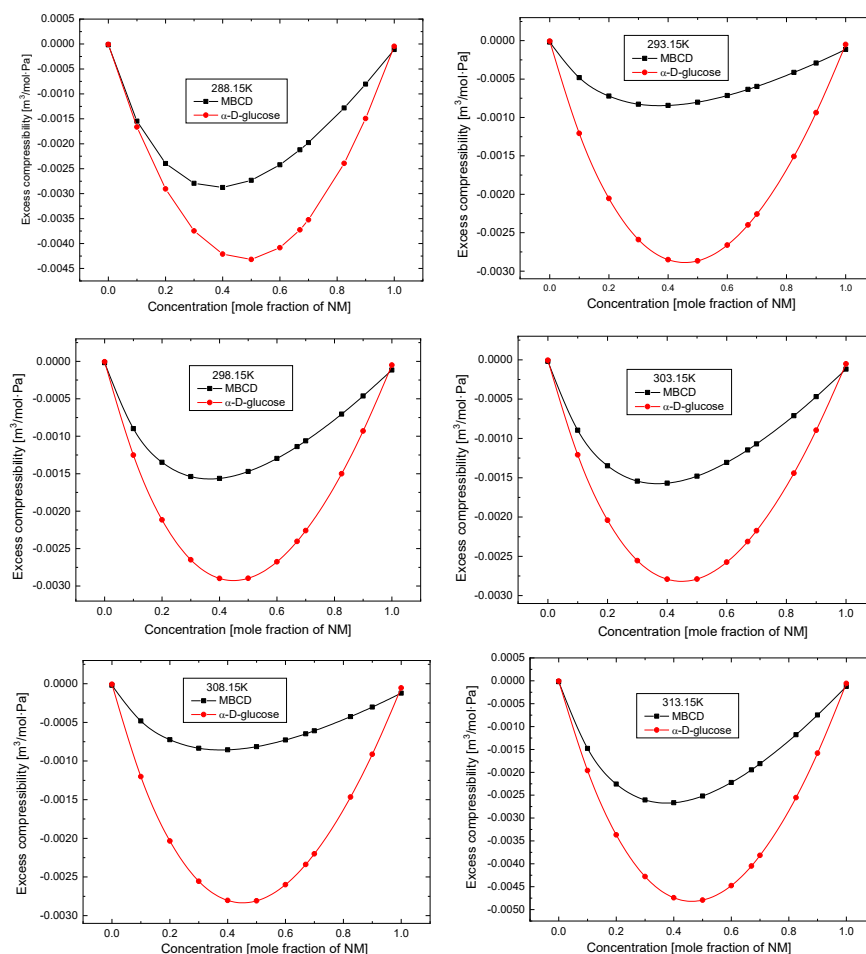


Figure 5. Excess molar compressibility,  $K_{S,m}^E$ , for MBCD + Nimodipine and D-glucose + Nimodipine plotted against mole fraction of Nimodipine at various temperatures: 288.15, 293.15, 298.15, 303.15, 308.15, and 313.15 K to 313.15 K

For initial solutions the molar isentropic compressibility of the real mixture is of course equal to that of ideal mixture so the excess compressibility has the value of zero. With the addition of the drug the excess molar compressibility for both solutions takes the negative values for the entire composition ranges and all temperatures. The shift

from ideality is in the direction of the enhanced rigidity. The minimum of excess molar compressibility occur at 0.5 mole fraction of Nimodipine indicating that the compound formed is composed of one molecule of the host (sugar) and one molecule of the guest (drug).

The analysis of the temperature dependence of excess molar compressibility shown in Figure 5 reveals that both solutions (cyclodextrin + Nimodipine and  $\alpha$ -D-glucose + Nimodipine) reach their minimum values at the lowest (288.15 K) and the highest (313.15 K) temperatures. Thus regardless of host molecule these outermost temperatures make the best conditions for the complexation process inside both polysaccharides but in favor of the 288.15 K because too high temperature may destabilize the properties of the drug. It also worth noting that the values of the excess molar isentropic compressibility are lower for the solution with  $\alpha$ -D-glucose than that with MBCD for all the studied temperatures. It may point out to the more frequent complex formation in the systems with glucose, independently on external conditions.

To determine the efficiency of complexation process through the amount/dose of drug encapsulated in the complex, in possibly better absorbed form, the masses of the sugar (MBCD or  $\alpha$ -D-glucose) and the drug (Nimodipine) in the solutions have to be calculated. The results of the calculation are presented in Table 1.

Table 1. Masses of the components of the solutions for the samples studied;  $m_1$  is the mass of sugar (MBCD or  $\alpha$ -D-glucose) and  $m_2$  the mass of NM

Solution studied	Temperature [K]	$m_1$ [mg]	$m_2$ [mg]
MBCD + NM	288.15	143.77	31.60
	293.15	149.96	30.90
	298.15	151.23	30.76
	303.15	150.39	30.85
	308.15	148.28	31.09
	313.15	147.04	31.23
$\alpha$ -D-Glucose + NM	288.15	116.71	35.82
	293.15	125.72	34.78
	298.15	128.34	34.62
	303.15	128.71	34.58
	308.15	127.21	34.74
	313.15	122.79	35.19

To achieve comparable efficiency in the same temperature the complex of MBCD + Nimodipine require more sugar and the complex of  $\alpha$ -D-glucose + Nimodipine – more drug. However, it should be remembered that MBCD has the structure designed for complexation while glucose is very reactive as chemical. So the lower values of the excess molar compressibility for the solution of Nimodipine with  $\alpha$ -D-glucose can be attributed to other processes taking place in this solution apart from complexation. But even that does not indicate that MBCD is a better candidate for the host molecule than  $\alpha$ -D-glucose and both sugars are comparable in terms of formation inclusion compounds.

#### 4. Conclusions

The ultrasonic and volumetric investigations of the systems MBCD + Nimodipine and  $\alpha$ -D-glucose + Nimodipine lead to the following conclusions on the complexation processes in the solutions: (1) The properties of the components of the solutions have direct influence for the complexation process; (2) It was found that both sugars (MBCD and  $\alpha$ -D-glucose) are able to form complexes with Nimodipine; (3) The values of the excess molar isentropic compressibility are negative for the entire composition ranges and all temperatures regardless the type of sugar used in the solution what suggests that formed structure is packed, relatively rigid and reaches the energetic balance for the stoichiometric ratio of 1:1.

The better knowledge about the inclusion complexation processes and their dynamics in the such specific systems as sugars and a 1,4-dihydropyridine derivative will make possible for them to find more applications in medicine.

#### References

1. J. Szejtli, *Utilization of cyclodextrins in industrial products and processes*. J. Mater Chem., **36** (1997) 575 – 587.
2. K. Kawakami, A. Fujita, T. Mikami, H. Yoshii, V. Paramita, T. L. Neoh, T. Furuta, *Formation of rice flavor powder with  $\alpha$ -cyclodextrin by spray drying*. Eur. Food Res. Technol., **229** (2009) 239 – 245.
3. R. Bhardwaj, R. T. Dorr, J. Blanchard, *Approaches to reducing toxicity of parenteral anticancer drug formulations using cyclodextrins*. PDA J. Pharm. Sci. Technol., **54** (2000) 233 – 239.
4. K. Sygnecka, C. Heine, N. Scherf, M. Fasold, H. Binder, C. Scheller, H. Frank, *Nimodipine enhances neurite outgrowth in dopaminergic brain slice co-culture*, Int. J. Dev. Neurosci., **40** (2015) 1 – 11.
5. Yong-Zhong Du, Jia-Guo Xu, Ling Wang, Hong Yuan, Fu-Qiang Hu, *Preparation and characteristics of hydroxypropyl- $\beta$ -cyclodextrin polymericnanocapsules loading nimodipine*, Eur. Polym. J., **45** (2009) 1397 – 1402.
6. G. C. Benson, O. Kiyohara, *Evaluation of excess isentropic compressibilities and isochoric heat capacities*, J. Chem. Thermodynamics, **11** (1979) 1061 – 1064.
7. F. Eggers, U. Kaatze, *Broadband ultrasonic measurement techniques for liquid*, Meas. Sci. Technol., **7** (1996) 1 – 9.

## A prototype of Chinese Aspirated Consonants Pronunciation Training System Based on Multi-Resolution Cochleagram

Katarzyna A. URBANIEC

*AGH University of Science and Technology,*

*Department of Mechanics and Vibroacoustics,*

*al. Mickiewicza 30, 30-059 Kraków, urbaniec@agh.edu.pl*

### Abstract

Many Mandarin Chinese learners, especially those whose mother tongue's phonological system differs significantly from Chinese phonological system, find it challenging to learn pronunciation of Chinese phonemes. Yet pronunciation training in language class settings is limited. It is therefore essential to develop computer-aided training system to help learners practice Chinese pronunciation without teacher's assistance. In this article I introduce a prototype of Chinese pronunciation training system that specifically focuses on phoneme substitution errors related to aspiration of consonants. I describe feature extraction process based on multi-resolution cochleagram (MRCG), a psychoacoustic model of basilar membrane excitation pattern, and architecture of recurrent neural network (RNN) used for mispronounced phonemes detection. The system achieves 96.12% and 98.58% accuracy rate in detecting phoneme substitution errors and determining aspiration length respectively. Proposed system may be particularly useful for learners of Slavic and Romance origin, since in their mother tongues aspiration is not a distinctive feature.

**Keywords:** computer-aided pronunciation training (CAPT), mispronunciation detection, phoneme substitution

### 1. Introduction

Since China became a world's largest economy and opened up to the world, encouraging economic cooperation and foreign investment, more and more people have been required or willing to learn Mandarin Chinese as a second language. Because of its phonological and tonal systems, Chinese is found to be difficult to learn, especially by beginner-level learners, whose mother tongue is not a tonal language [1, 2]. Shortage of qualified teachers and limited pronunciation training in formal class settings make it even more challenging. A solution to these problems are computer-aided language learning (CALL) systems incorporating computer-assisted pronunciation training (CAPT), which identify a specific pronunciation error in an utterance and provide a corrective feedback without any human assistance [3].

In most of approaches to CAPT, mispronunciation detection is performed by extending automatic speech recognition (ASR) technologies such as Hidden Markov Models (HMMs), Artificial Neural Networks (ANNs) or hybrid HMM-ANN systems [4]. Many of them employ Log-Likelihood Ratio (LLR) between non-native like and native-like models to detect pronunciation errors. The most representative system using LLR is Goodness Of Pronunciation (GOP) introduced by Witt and Young in 2000. The GOP measure incorporates a set of HMMs trained using mel-frequency cepstral coefficients (MFCC) and provides score for each phone in an utterance. A phone-specific threshold is set based on global GOP statistics and applied to each of the scores to decide, which of the

phones are mispronounced [5]. GOP methodology was used in numerous works, e. g. [6-12]. Above-mentioned LLR-based scoring methods achieve good pronunciation errors detection. Nevertheless, they don't provide user with diagnostic feedback about specific errors made by them such as phoneme insertion, deletion and substitution [13]. This feedback is obtained by using so called extended recognition networks (ERN). ERN is a representation of the canonical pronunciations and possible mispronunciations of a word. Using this representation reduces computational cost of ASR algorithms by avoiding searches in superfluous phone paths [14, 15]. A corrective feedback is also obtained by using ANNs, which map phonemic context information and acoustic features into phonemic posterior probabilities [1, 11, 16, 17].

Difficulties in learning proper Chinese pronunciation encountered by non-English native speakers are rarely discussed. There is only one work mentioning pronunciation problems specific for Mandarin Chinese learners of Slavic origin [18]. These problems are caused predominantly by lack of aspirated consonants in phonological systems of Slavic languages. Since there is no system designed to meet Slavic learners needs available, it is particularly important to develop one.

The aim of this work is to design and implement a prototype pronunciation training system that specifically focuses on pronunciation errors related to aspiration: deaspiration of aspirated consonants and voicing of voiceless not aspirated consonants. The system incorporates multi-resolution cochleagram (MRCG), a psychoacoustic model of basilar membrane excitation pattern [19]. Features extracted from MRCG are fed to recurrent neural network (RNN) detecting mispronounced phonemes and providing feedback about length of aspiration. With this knowledge user can alter the aspiration length adequately and improve their pronunciation skills. Proposed system can also help learners of Romance origin, since in their mother tongues aspiration is not a distinctive feature and aspiration-related errors are common among them [18].

## 2. System design

Proposed CAPT system is meant to help learners, who do not have access to qualified teachers, practice and obtain native-like pronunciation of Chinese aspirated consonants. Multi-resolution cochleagram (MRCG), composed of four 64-channel cochleagrams at different resolutions (hereafter referred to as CG1, CG2, CG3 and CG4), was used to model frequency selectivity properties of human cochlea and ensure, that extracted features and differences between them may be in fact perceived by human.

### 2.1. Speech corpus

The speech corpus used in this study is a subset of of native speech corpus AISHELL-1 [21]. Based on analysis of Mandarin Chinese phonological system, 12 phonemes were chosen to be included in designed CAPT system: 6 aspirated consonants ( $k^h$ ,  $t^h$ ,  $ts^h$ ,  $tʂ^h$ ,  $tɕ^h$ ,  $p^h$ ) and their not aspirated equivalents ( $k$ ,  $t$ ,  $ts$ ,  $tʂ$ ,  $tɕ$ ,  $p$ ) [1]. All above-mentioned consonants may occur only as syllable-initial phonemes and there are no consonant clusters in Mandarin language [20]. Thus, setting the length of analyzed signal to 180 ms ensures, that only one consonant will be included in this signal and the remaining frames not corresponding to any of these consonants will correspond to vowels

or glides. 360 utterances spoken by 15 male speakers and 15 female speakers were used: 300 of them constitute a training and validation data set and remaining 60 constitute a testing data set.

## 2.2. MRCG-based classifier

The mispronunciation detection system was implemented in Python. Speech signals were divided into 20 ms frames overlapping by 50%. For each signal 256-dimensional MRCG features and their deltas and double deltas were calculated using algorithm described in [19]. Kurtosis (Kurt) and spectral moments of 2<sup>nd</sup>, 3<sup>rd</sup> and 4<sup>th</sup> order ( $M_2$ ,  $M_3$ ,  $M_4$ ) were also calculated for each channel of CG1, CG2, CG3, CG4, and their deltas and double deltas. Since MRCG is known to be noise robust [19, 22], signal filtering was omitted. To reduce computational complexity and feature redundancy, dimensionality reduction using recursive feature elimination (RFE) method was performed. 35 most relevant features shown in Table 1 were chosen and used in further analysis.

Table 1. 35 most relevant MRCG-features used in further analysis

Most relevant MRCG-features according to RFE results			
25 <sup>th</sup> CG1 channel	63 <sup>rd</sup> CG3 channel	33 <sup>rd</sup> $\Delta$ CG3 channel	32 <sup>nd</sup> $\Delta\Delta$ CG2 channel
26 <sup>th</sup> CG1 channel	64 <sup>th</sup> CG3 channel	34 <sup>th</sup> $\Delta$ CG3 channel	10 <sup>th</sup> $\Delta\Delta$ CG3 channel
29 <sup>th</sup> CG1 channel	20 <sup>th</sup> $\Delta$ CG1 channel	24 <sup>th</sup> $\Delta$ CG4 channel	15 <sup>th</sup> $\Delta\Delta$ CG3 channel
36 <sup>th</sup> CG1 channel	25 <sup>th</sup> $\Delta$ CG1 channel	22 <sup>nd</sup> $\Delta\Delta$ CG1 channel	22 <sup>nd</sup> $\Delta\Delta$ CG3 channel
58 <sup>th</sup> CG1 channel	52 <sup>nd</sup> $\Delta$ CG1 channel	49 <sup>th</sup> $\Delta\Delta$ CG1 channel	$M_2(\Delta\Delta$ CG4)
26 <sup>th</sup> CG2 channel	24 <sup>th</sup> $\Delta$ CG2 channel	59 <sup>th</sup> $\Delta\Delta$ CG1 channel	$M_4(\Delta\Delta$ CG3)
42 <sup>nd</sup> CG2 channel	31 <sup>st</sup> $\Delta$ CG2 channel	64 <sup>th</sup> $\Delta\Delta$ CG1 channel	Kurt(CG2)
58 <sup>th</sup> CG2 channel	60 <sup>th</sup> $\Delta$ CG2 channel	4 <sup>th</sup> $\Delta\Delta$ CG2 channel	Kurt( $\Delta\Delta$ CG1)
62 <sup>nd</sup> CG3 channel	32 <sup>nd</sup> $\Delta$ CG3 channel	8 <sup>th</sup> $\Delta\Delta$ CG2 channel	

Above-mentioned features were normalized and fed to recurrent neural network (RNN) implemented in Keras and trained for 225 epochs. The RNN's architecture is shown in Figure 1. After each recurrent layer (gated recurrent unit, GRU, or long-short term memory, LSTM), dropout layers were added to avoid model's overfitting. Output layer is a dense layer with softmax activation function, trained under cross-entropy regime. Based on obtained probability distributions, each frame of analyzed signal was classified into one of 13 classes:  $k$ ,  $t$ ,  $ts$ ,  $t_s$ ,  $t_e$ ,  $p$ ,  $k^h$ ,  $t^h$ ,  $ts^h$ ,  $t_s^h$ ,  $t_e^h$ ,  $p^h$  and vowel.

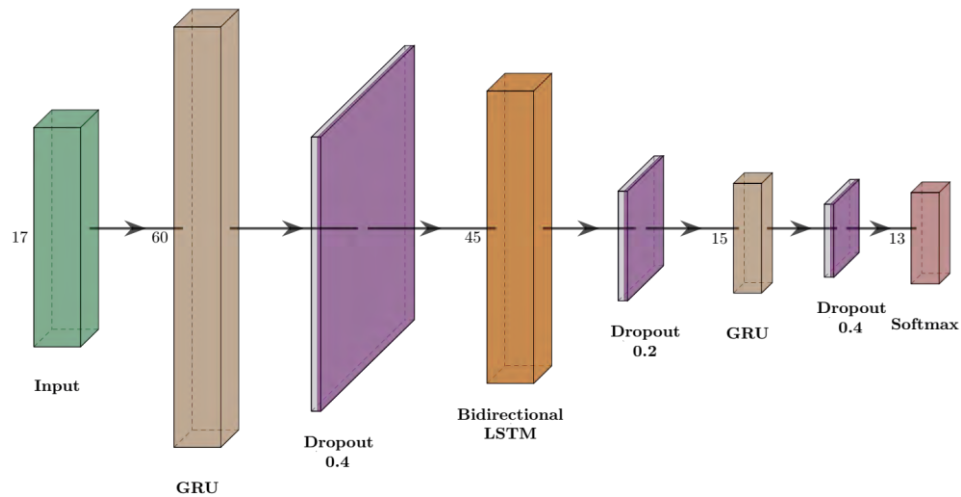


Figure 1. Diagram of RNN architecture

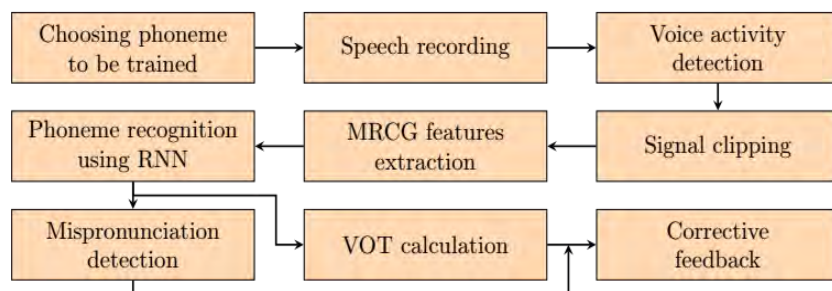


Figure 2. Block diagram of proposed CAPT system

Based on the frame length and the number of frames classified as not a vowel, voice onset time (VOT) was estimated. VOT value combined with information about possible mispronunciations detected by RNN is used to give user a corrective feedback about the reason of incorrect pronunciation (phoneme substitution, wrong aspiration length or both). The block diagram of designed CAPT system is shown in Figure 2.



### 2.3. Evaluation metrics

The performance of proposed MRCG-RNN-based CAPT system is evaluated using three metrics, namely recall, precision and accuracy:

$$Recall = \frac{N_C}{N_{PH}} \cdot 100\% \quad (1)$$

$$Precision = \frac{N_C}{N_D} \cdot 100\% \quad (2)$$

$$Accuracy = \frac{N_{FC}}{N} \cdot 100\% \quad (3)$$

where  $N_{PH}$  is number of utterances containing analyzed phoneme,  $N_C$  is number of correctly detected phonemes,  $N_D$  is number of detected phonemes (both correctly and incorrectly),  $N_{FC}$  is number of frames classified correctly, and  $N$  is number of frames analyzed.

### 3. Results and discussion

Table 2 shows the performance of implemented RNN classifier. Precision and recall were calculated for each of 12 consonants incorporated in proposed CAPT system. International Phonetic Alphabet (IPA) is one of the most popular phonetic transcription system, but for Mandarin there is no one accepted IPA system [23]. Thus, in Table 2 consonants are displayed using both IPA and pinyin (romanization system for Mandarin Chinese).

The system achieves high recall and precision rates for recognition of all analyzed phonemes. Its performance is significantly better than performance of state-of-the-art CAPT system described in [2]. RNN classifier used to detect mispronunciations achieves high accuracy rate of 96.12%. The worst results were obtained for  $ts^h$  (recall 86.67%) and  $k^h$  (precision 85.30%). The reason for this may be that observed VOT differences for  $ts^h$ - $ts$  and  $k^h$ - $k$  pairs were relatively small for signals contained in used data sets. Classifier's performance could be improved by increasing number of samples in training data set.

To evaluate system's performance in determining aspiration length, VOT value estimated using classification results of single frames was compared with reference value obtained manually. The results show that proposed system achieves 98.58% accuracy rate in aspiration length determination. It should be noted, however, that the accuracy of approach to VOT estimation used in this work is limited by frame length and using other approach may be worth considering [24].

Table 2. Performance evaluation results of RNN classifier used in proposed CAPT system

Consonant		Precision	Recall
Pinyin	IPA		
<i>b</i>	<i>p</i>	100%	96.67%
<i>p</i>	<i>p<sup>h</sup></i>	87.88%	96.67%
<i>d</i>	<i>t</i>	100%	96.67%
<i>t</i>	<i>t<sup>h</sup></i>	100%	93.33%
<i>c</i>	<i>tʂ<sup>h</sup></i>	100%	86.67%
<i>z</i>	<i>ts</i>	93.55%	96.67%
<i>j</i>	<i>tɕ</i>	93.55%	96.67%
<i>q</i>	<i>tɕ<sup>h</sup></i>	100%	96.67%
<i>ch</i>	<i>tʂ<sup>h</sup></i>	96.77%	100%
<i>zh</i>	<i>tɕ</i>	87.88%	96.67%
<i>g</i>	<i>k</i>	100%	100%
<i>k</i>	<i>k<sup>h</sup></i>	85.30%	96.67%

#### 4. Conclusions

I have proposed a prototype CAPT system focusing on pronunciation errors related to aspiration made by Mandarin Chinese learners. It incorporates MRCG, a psychoacoustic model of basilar membrane excitation pattern. Using this model in feature extraction process ensures that the system uses the same features as humans do to differentiate phonemes. MRCG is also known to be noise robust, therefore proposed system is suitable to be used in an environment with relatively low signal to noise ratio.

The system achieves high recall, precision and accuracy rates in phoneme recognition and pronunciation errors detection (86.67-100%, 85.30-100% and 96.12% respectively). Obtained results are significantly better than results described in other works, e. g. [2]. The system also achieves high accuracy rate in aspiration length determination (98.58%). It makes the system particularly useful for Mandarin Chinese learners of Slavic and Romance origin, since proper aspiration is difficult for them.

Although the system's performance is satisfactory, an additional evaluation using non-native speech corpus should be performed. Moreover, increasing the number of signals included in data set used for RNN's training could also improve its performance. Decreasing frame length to 10 ms or using other approach to VOT estimation should also be considered in future work.

## References

1. W. Li et al., *Improving non-native mispronunciation detection and enriching diagnostic feedback with DNN-based speech attribute modeling*, IEEE International Conference on Acoustics, Speech and Signal Processing, (2016) 6135 – 6139.
2. H.-C. Liao et al., *A prototype of an adaptive Chinese pronunciation training system*, System, **45** (2014) 52 – 66.
3. C. Molina et al., *ASR based pronunciation evaluation with automatically generated competing vocabulary and classifier fusion*, Speech Communication, **51** (2009) 485 – 498.
4. X. Qian, H. Meng, F. Soong, *A Two-Pass Framework of Mispronunciation Detection and Diagnosis for Computer-Aided Pronunciation Training*, IEEE/ACM Transactions on Audio, Speech, and Language Processing, **24.6** (2016) 1020 – 1028.
5. S. M. Witt, S. J. Young, *Phone-level pronunciation scoring and assessment for interactive language learning*, Speech Communication, **30** (2000) 95 – 108.
6. B. Mak et al., *PLASER: pronunciation learning via automatic speech recognition*, Proceedings of Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics, (2003) 23 – 29.
7. A. Neri, C. Cucchiaroni, H. Strik, *ASR corrective feedback on pronunciation: does it really work?*, Proceedings of Interspeech, (2006) 1982 – 1985.
8. J. Zheng et al., *Generalized Segment Posterior Probability for Automatic Mandarin Pronunciation Evaluation*, IEEE International Conference on Acoustics, Speech, and Signal Processing, (2007) 201 – 204.
9. C. Cucchiaroni, A. Neri, H. Strik, *Oral proficiency training in Dutch L2: The contribution of ASR-based corrective feedback*, Speech Communication, **51** (2009) 853 – 863.
10. H. Strik et al., *Comparing different approaches for automatic pronunciation error detection*, Speech Communication, **51** (2009) 845 – 852.
11. W. Hu et al., *Improved Mispronunciation Detection with Deep Neural Network Trained Acoustic Models and Transfer Learning based Logistic Regression Classifiers*, Speech Communication, **67** (2015) 154 – 166.
12. G. Huang et al., *English Mispronunciation Detection Based on Improved GOP Methods for Chinese Students*, Proceedings of International Conference on Progress in Informatics and Computing, (2017) 425 – 429.
13. Sh. Mao et al., *Applying Multitask Learning to Acoustic-Phonemic Model for Mispronunciation Detection and Diagnosis in L2 English Speech*, IEEE International Conference on Acoustics, Speech and Signal Processing, (2018) 6254 – 6258.
14. A. M. Harrison et al., *Implementation of an Extended Recognition Network for Mispronunciation Detection and Diagnosis in Computer-Assisted Pronunciation Training*, Proceedings of ISCA International Workshop on Speech and Language Technology in Education, (2009) 45 – 48.

15. W.-K. Lo, Sh. Zhang, H. Meng, *Automatic Derivation of Phonological Rules for Mispronunciation Detection in a Computer-Assisted Pronunciation Training System*, Proceedings of Interspeech, (2010) 765 – 768.
16. Sh. Mao et al., *Unsupervised Discovery of an Extended Phoneme Set in L2 English Speech for Mispronunciation Detection and Diagnosis*, IEEE International Conference on Acoustics, Speech and Signal Processing, (2018) 6244 – 6248.
17. K. Li, X. Qian, H. Meng, *Mispronunciation detection and diagnosis in L2 English speech using multidistribution deep neural networks*, IEEE/ACM Transactions on Audio, Speech, and Language Processing, **25.1** (2017) 193 – 207.
18. N. F. Chen et al., *Large-scale characterization of non-native Mandarin Chinese spoken by speakers of European origin: Analysis on iCALL*, Speech Communication, **84** (2016) 46 – 56.
19. J. Chen, Y. Wang, D. Wang, *A Feature Study for Classification-Based Speech Separation at Low Signal-to-Noise Ratios*, IEEE International Conference on Acoustics, Speech and Signal Processing, (2014) 1993 – 2002.
20. L.-H. Wee, M. Li, *Modern Chinese Phonology* In W. S.-Y. Wang, Ch. Sun, *The Oxford handbook of Chinese linguistics*, Oxford University Press, (2015) 474 – 489.
21. H. Bu et al., *AISHELL-1: An open-source Mandarin speech corpus and a speech recognition baseline*, 20th Conference of the Oriental Chapter of the International Coordinating Committee on Speech Databases and Speech I/O Systems and Assessment, (2017) 1 – 5.
22. X.-L. Zhang, D. Wang, *Boosted deep neural networks and multi-resolution cochleagram features for voice activity detection*, Proceedings of Interspeech, (2014) 1534 – 1538.
23. K. K. Y. Lam, C. K. S. To, *Speech sound disorders or differences: Insights from bilingual children speaking two Chinese languages*, Journal of Communication Disorders, **70** (2017) 35 – 48.
24. Ch.-Y. Lin, H.-Ch. Wang, *Automatic estimation of voice onset time for word-initial stops by applying random forest to onset detection*, Journal of the Acoustical Society of America, **130.1** (2011) 514 – 525.

## Analysis and Comparison of Vibration Signals from Internal Combustion Engine Acquired Using Piezoelectric and MEMS Accelerometers

Dariusz BISMOR

*Institute of Automatic Control, Silesian University of Technology,  
ul. Akademicka 16, 44-100 Gliwice, Poland; [dariusz.bismor@polsl.pl](mailto:dariusz.bismor@polsl.pl)*

### Abstract

Condition monitoring of vehicles with internal combustion engine is of immense importance due to high number of vehicles with such engines and their importance to transport and economy. As many persons use a vehicle which is old and inexpensive, a condition monitoring system designed for such vehicles cannot be expensive. Unfortunately, condition monitoring of engines is usually based on the use of vibration signals, which are acquired by accelerometers. Piezoelectric accelerometers are the most commonly used for this purpose, and such accelerometers are not cheap. However, an alternative exists in the form of microelectromechanical systems (MEMS) accelerometers, which are much cheaper, but have narrower frequency characteristics. This paper describes preliminary results of a research on feasibility of use of MEMS accelerometers for condition monitoring and failure detection in internal combustion engines.

**Keywords:** MEMS accelerometers, vibration signal processing, spectrogram, Wigner-Ville spectrum

### 1. Introduction

Condition monitoring is a widely-used technique that allows for early detection of machinery faults or failures in order to reduce maintenance costs and downtime. Application of condition monitoring allows to improve reliability and productivity of the equipment and safety of its operators [1]. A comprehensive condition monitoring program includes detection, diagnostics and prognosis.

Condition monitoring usually applies to a number of machines that operate under heavy environmental conditions, like high temperatures, high humidity, high load, etc. Among these machinery, condition monitoring of internal combustion engines (ICE) is a particularly difficult task due to complexity of ICEs [2]. However, it is also very important due to importance of transport for the economy [3].

The age of motor vehicles in use still increases, and the older a vehicle, the higher the probability of its defect or failure; therefore, condition monitoring of vehicles is of vital importance. This, however, requires a special monitoring hardware to be installed in the vehicle permanently. Such hardware cannot be expensive, if it is intended to be installed in old, cheap vehicles.

Although different techniques also exist, diagnostics and condition monitoring of ICEs are mainly based on vibration and acoustic signals [1]. Vibration signals are acquired by means of accelerometers, which need to be installed on a flat surface using screw joints or a special wax. Commonly used piezoelectric accelerometers allow to acquire signals in the frequency range up to 15 kHz, or more. These transducers are not cheap, with prices

starting at around \$300, and reaching several thousands dollars. This makes application of a system based on such accelerometers in older cars economically unreasonable. Fortunately, less expensive alternative exists in the form of MEMS accelerometers.

Nowadays, microelectromechanical systems (MEMS) is a technology that allows for manufacturing of microscopic devices, particularly those with moving parts. This technology allowed for production of low-cost accelerometers, and recently also microphones [4]. The application of MEMS accelerometers and microphones in a vehicle condition monitoring system seems attractive, even if the accelerometers usually have substantially narrower frequency ranges than piezoelectric devices. However, it is necessary to elaborate on minimum requirements concerning the MEMS accelerometers first, and to compare their performance with piezoelectric accelerometers. Indeed, this paper describes some preliminary results of comparison of cheap MEMS accelerometers and good quality piezoelectric accelerometers in application to ICE monitoring.

## 2. MEMS accelerometers and controller board

The first stage of the research was to select sensors suitable for data acquisition in the designed application. Regarding the accelerometer, the basic requirements were three-axial measurement and wide acceleration range. However, the greatest concern was the frequency range of the measured acceleration. Unfortunately, the price of MEMS accelerometers is proportional to the frequency range.

Based on analysis of frequencies corresponding to ICE operation in the rotational speed between 0 and 6000 rpm, vehicle's wheel rotational speed, and considering that higher frequencies are generated by rolling bearings, the ADXL345 accelerometer by *Analogue Devices* was selected [5]. This accelerometer is a 3-axial MEMS device, with sampling frequency up to 3.2 kHz, configurable acceleration ranges from 2 to 16 g, and supply voltage from 3 to 5 V. The device is available on a small PCB, simplifying its installation, and comes at a price less than \$3.

Two ADXL345 accelerometers were connected to the STM32F407G-Discovery board, which was selected as an optimal choice for rapid development regime. The STM Discovery board is equipped with the STM32F407VGT6 microcontroller, clocked at 168 MHz, what is sufficient to receive the accelerometer's data with 3.2 kHz sampling rate. The project assumed that the device will be installed in a vehicle; therefore, the controller board needed to be protected against harsh environmental conditions, including vibrations, moisture condensation and dust. The simplest solution was to use a plastic case, inside which the controller board was fixed together with additional hardware components, like power supply and flash memory. The design was carefully tested in a number of experiments, using a digital signal analyzer. Figure 1 presents the exploded view of the device, and the MEMS accelerometer fixed to a small piece of rectangular pipe, which simplifies its installation.



Figure 1. The controller board in its casing (left) and the MEMS accelerometer mounted on a piece of rectangular pipe (right)

### 3. Measurement setup

To evaluate quality of the signals from MEMS accelerometers, good-quality piezoelectric devices needed to be used as reference accelerometers. For this purpose, the PCB 622B01 accelerometers were selected, which are high-quality ceramic shear devices, with frequency range 0.2 Hz to 15 kHz. The PCB 622B01 accelerometers will be further referred to as the “reference accelerometers”. The reference accelerometers were connected to a National Instruments NI-9234 sound and vibration input module. The acquired signal was downsampled (using a proper decimation technique, i.e. with anti-aliasing filtration) to 3.2 kHz.

The accelerometers were installed on a passenger vehicle's internal combustion engine chassis, using a wax. One accelerometer was placed on the screw securing the cylinder head cover. The other was placed on the lower part of the engine's block. The measurements were acquired with the first set of accelerometers, and the engine was cooled down, so the wax could work again. Then, the second set of accelerometers was waxed in the same positions, and the second set of measurements was acquired. The results presented and discussed below are based on the measurements from the accelerometers placed on the cylinder head cover only.

Each experiment was began with starting the engine and allowing it to run idle for several seconds. Next, the engine rotational speed was slowly increased, up to 5000 rpms, and than slowly decreased to idle. Finally, a small engine defect was introduced, in a form of one cylinder misfire, by unplugging one of the engine's sparks. Figure 2 presents the time plot of the signal acquired from the MEMS accelerometer during the experiment.

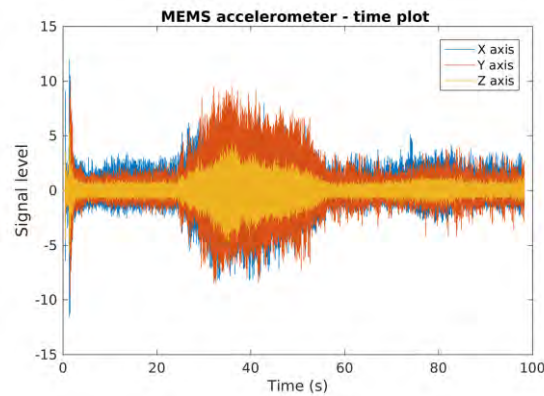


Figure 2. Time plot of the signals acquired from the MEMS accelerometer

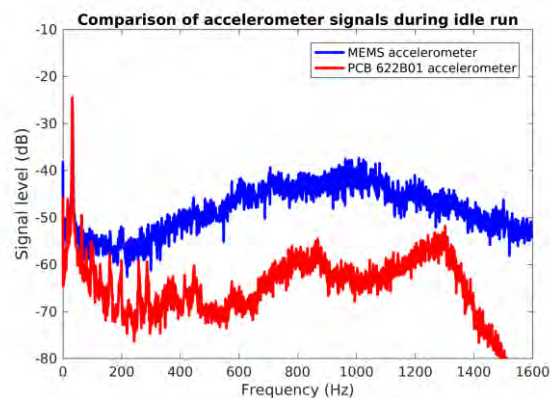


Figure 3. Spectra of the accelerometer signals during ICE idle run

#### 4. Results

The first analysis considered the part of the signal when the engine was running idle. The engine's rotational speed was approximately constant; therefore the analysis could have been performed by plotting power spectra of the signals from the accelerometers, as presented in Fig. 3. Each spectrum was obtained by dividing the acquired signal into six segments and averaging the spectra of each segment.

The fundamental engine frequency, equal to double the engine rotational speed, is clearly visible in the figure. As the engine idle run was at around 950 rpm, the frequency is equal approximately 32 Hz. A number of harmonics can also be observed in both the reference and the MEMS accelerometer signal. It can also be noticed that the MEMS accelerometer produced a substantially higher level of signal in frequency range between 200 and 1200 Hz. (The decreasing level of the reference accelerometer spectrum is due to anti-aliasing filtration performed during the signal decimation).



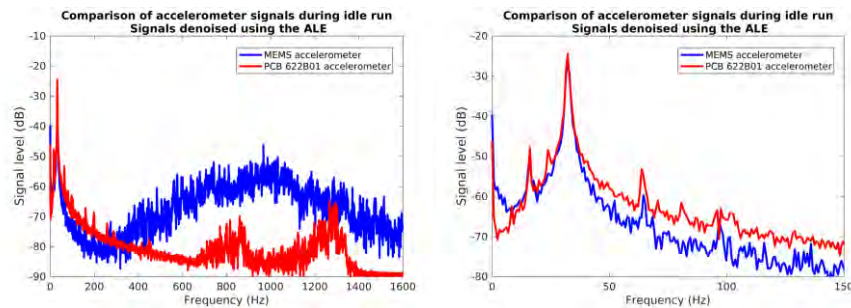


Figure 4. Spectra of the accelerometer signals during ICE idle run after de-noising using the ALE (left). Zoomed part of the spectrum (right)

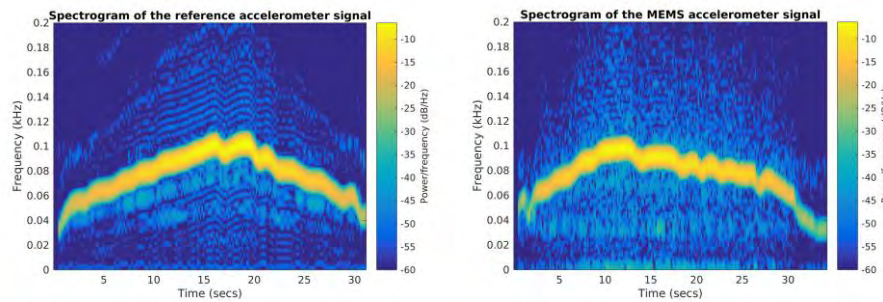


Figure 5. Spectrograms of the accelerometer signals during up-down engine run after denoising using the ALE. Reference sensor (left), and MEMS sensor (right)

A simple technique to enhance signal quality is called the *adaptive line enhancer* (ALE) [6]. This technique proved very effective in extracting sinusoidal signals embedded in a large amount of noise. Therefore, both the signals were filtered using the ALE with 1000 parameters. Figure 4 shows the spectra of the signals thus processed. It can be noticed that the enhancement worked very well, and the spectra show a very good level of similarity in the frequency range between 0 and 150 Hz. However, the levels of the signals differ significantly in the frequency range between 400 and 1300 Hz. This is probably due to poorer quality of the MEMS accelerometers, and it needs to be further researched.

As the ALE technique worked well, it was also applied for the part of the signal, when the engine's rotational speed was increased up to 5000 rpm, and decreased down to idle. However, the signal obtained during this phase cannot be treated as stationary, and therefore time-frequency analysis should be applied. One of the simplest and well-known techniques of time-frequency analysis is the spectrogram [7]. Spectrograms of the signals acquired with the reference and MEMS accelerometers are presented in Fig. 5. For clarity of the presentation, only the frequency range from 0 to 200 Hz has been shown. In this range, the fundamental frequency of the engine is clearly visible. Moreover, the spectrogram of the reference sensor signal shows also a number of other curves, which may depict harmonics of the fundamental frequency. These curves are not visible in the MEMS sensor signal, which appears to contain more noise.

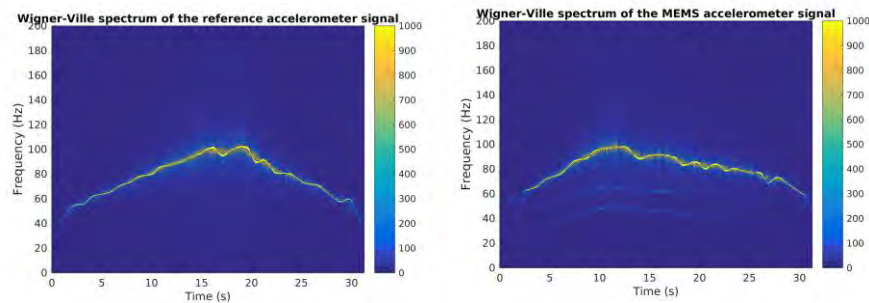


Figure 6. Wigner-Ville spectrum of the accelerometer signals during up-down engine run after de-noising using the ALE. Reference sensor (left), and MEMS sensor (right)

The main disadvantage of the spectrogram is its poor resolution, what is clearly visible in Fig. 5, where the power of the signal is spread over a range of frequencies. Therefore, a better distribution was used in the form of the Wigner-Ville spectrum [7] – see Fig. 6. In the Wigner-Ville spectrum, the engine's fundamental frequency is easily visible and can be precisely determined. Moreover, the spectrum of the MEMS sensor signal does not appear so noisy, and both the spectra show a couple of subharmonics. On the other hand, the harmonics, which were clearly visible in the spectrogram of the reference sensor signal, disappeared. Nevertheless, the overall quality of the result is much better in case of the Wigner-Ville spectrum, compared to the spectrograms, and both the signals present comparable amount of details.

Finally, the last part of the signal was analyzed. It was the part where the engine defect, in the form of one cylinder misfire, was introduced. As the engine was running at approximately the same rotational speed, this parts of the signals can be considered to represent stationary processes, and can be analyzed by the means of their spectra. The spectra are presented in Fig. 7.

Comparing the spectra in Fig. 7 with the spectra of the engine without the defect, presented in Fig. 4, reveals that the spectra show more frequency components when the engine is defected. However, the components probably are subject to small frequency fluctuations, usually referred to as frequency smearing [8]. The smearing makes a particular frequency component to appear as an increased variance of the spectrum. Averaging of the spectra can flatten such variance, and using more averaged segments can make such frequency components unrecognizable.

More differences between the ICE in idle run and the ICE with the defect are visible in the zoomed part of the spectra, showing the frequency range between 0 and 150 Hz. When the engine was running without the defect (i.e. in Fig. 4), four sinusoidal components were clearly visible in this frequency range. On the other hand, when the engine experiences one cylinder misfire, between eight and ten sinusoidal components can be recognized. Particularly clearly visible, and with high level, are the three sub-harmonics of the main frequency – the components with frequencies 8.19, 15.36, and 23.55 Hz. What is important, these components are visible in both the reference and the MEMS

accelerometer signals. This proves that the MEMS accelerometer can provide with signals of quality sufficient for condition monitoring of the engine.

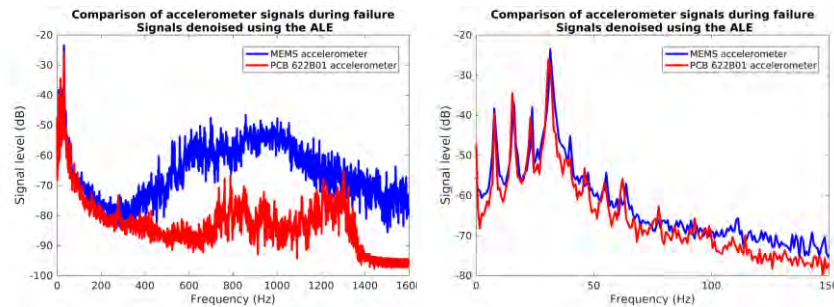


Figure 7. Spectra of the accelerometer signals during ICE one cylinder misfire, after de-noising using the ALE (left). Zoomed part of the spectrum (right)

## 5. Conclusions

The research reported in this paper considered the application of MEMS accelerometers to condition monitoring of internal combustion engines. MEMS accelerometers are recently very popular sensors, with application in many areas of life, including cellphones, smart watches, toys, and others. For this purpose, MEMS accelerometers are generally cheap, especially comparing to high-quality piezoelectric devices. However, the price is correlated to the frequency range and signal quality.

In this research, the ADXL 345 MEMS accelerometer was compared with the PCB 622B01 device. The former comes at a price of less than \$3, the latter costs around \$350. The ADXL 345 accelerometer has a frequency range from 0 to 1.6 kHz. Its overall performance in the tests was good. The signals acquired with both the accelerometers had comparable quality in the frequency range up to 200 Hz. Above this frequency, the MEMS accelerometer produced a higher level of wide-band signal, which was probably measurement noise. Nevertheless, both the signals allowed to recognize an increased level of harmonics when a small defect of the engine, in the form of one cylinder misfire, was introduced.

The signals acquired from both the accelerometers during engine's run-up and run-down were processed by Wigner-Ville spectrum estimate as well. In both cases the Wigner-Ville spectrum presented similar level of details.

To conclude, the ADXL 345 MEMS accelerometer is certainly suitable for condition monitoring of an ICE, if the frequency range up to 200 Hz is sufficient. Further research is necessary to confirm its usefulness in higher frequency range.

## Acknowledgments

This research was supported by the state budget for science, Poland, in 2018 and 2019, under no 02/010/BK18/0102.

## References

1. M. S. Kan, A. C. C. Tan, J. Mathew, *A review on prognostic techniques for non-stationary and non-linear rotating systems*, Mechanical Systems and Signal Processing, **62–63** (2015) 1 – 20.
2. S. Delvecchio, P. Bonfiglio, F. Pompoli, *Vibro-acoustic condition monitoring of internal combustion engines: A critical review of existing techniques*, Mechanical Systems and Signal Processing, **99** (2018) 661 – 683.
3. I. Komorska, *A vibroacoustic diagnostic system as an element improving road transport safety*, International Journal of Occupational Safety and Ergonomics, **19** (2013) 3 – 17.
4. G. W. Elko, F. Pardo, D. López, D. Bishop, P. Gammel, *Capacitive MEMS microphones*, Bell Labs Technical Journal, **10** (2005) 187 – 198.
5. A. Frank, *Modul do rejestracji sygnałów w samochodach i przesyłania ich do stacji bazowej*, Master's thesis, Silesian University of Technology, Gliwice, 2018.
6. S. Haykin, *Adaptive Filter Theory, Fourth Edition*, Prentice Hall, New York, 2002.
7. L. Cohen, *Time-frequency distributions – a review*, Proceedings of the IEEE, **77** (1989) 941 – 981.
8. R. B. Randall, J. Antoni, *Rolling element bearing diagnostics—a tutorial*, Mechanical Systems and Signal Processing, **25** (2011) 485 – 520.

## **Scientific Legacy of Professor Andrzej Rakowski in Current Studies of Pitch Discrimination in Music**

Andrzej MIŚKIEWICZ

*Chopin University of Music, Department of Sound Engineering,  
Okólnik 2, 00-368 Warszawa, Poland, andrzej.miskiewicz@chopin.edu.pl*

### **Abstract**

This paper is an overview of experimental studies of pitch discrimination and pitch strength in music conducted in recent years at the Chopin University of Music. The studies were inspired by Professor Andrzej Rakowski's findings and ideas on the foundations of pitch perception in music. The measurements of pitch discrimination show that the ability to hear pitch differences markedly decreases below 200-Hz frequency so that the pitch discrimination threshold increases to about a semitone at very low frequencies. The auditory system's relatively poor ability of pitch discrimination of low-frequency tones also manifests itself in much less accurate identification of musical intervals and melodic patterns in the lowest octaves of the musical scale, comparing with higher octaves. The paper also discusses the results of an experiment which indicate that some percussion instruments of the indefinite pitch family produce a sensation of pitch strength comparable with melodic instruments.

**Keywords:** pitch discrimination, pitch strength, musical intervals, melody identification in music

### **1. Introduction**

During his impressive scientific career, which lasted for six decades, a substantial part of Professor Andrzej Rakowski's work was focused on pitch perception, an area in which his accomplishments were most remarkable. In his studies of pitch perception he investigated, among others, various phenomena concerned with pitch discrimination. His first notable achievement in that field, reported at the ICA Congress in Budapest, was the finding that the difference limen for pitch (pitch DL) of a 1-kHz tone, measured in extensively trained musician subjects, might be as low as 0.35 cents, which corresponds to nearly 1/300 of a semitone [1]. This finding was received with some disbelief at the congress as the lowest pitch DLs, known at that time, *e.g.*, [2], were about three times larger. The experiment reported in Budapest began Professor Rakowski's extensive explorations of pitch discrimination, focused on two main problems: (1) he measured pitch DLs of pure tones to estimate the limits of the auditory system's frequency resolution and (2) he sought to determine how accurately are discerned the pitches of musical tones.

Professor Rakowski's investigations of pure-tone pitch discrimination have deepened the understanding of the frequency discrimination and frequency coding processes in the auditory system. In his studies conducted with the use of recorded samples of musical tones he has given detailed evidence for an effect known in musical praxis, that the pitches of tones played on different instruments or in different registers of

an instrument are heard with different accuracy. Rakowski [3] introduced a novel concept of *pitch strength* referred to the accuracy or salience of the pitch sensation.

This paper gives an overview of research on pitch discrimination and pitch strength in music conducted in recent years by Professor Andrzej Rakowski's associates at the Chair of Musical Acoustics, Chopin University of Music. Special attention is given in the overview to the perception of low-frequency tones. The investigations discussed here were inspired by Professor Rakowski's studies of pitch perception and were carried out to pursue his scientific ideas further, within the current research areas in psychoacoustics.

## 2. Pitch discrimination of low-frequency tones

Pitch DL, a measure of auditory sensitivity to pitch changes and frequency-discrimination ability of the auditory system has been typically defined as the just detectable change in frequency,  $\Delta F$ , between two successive tone pulses. Reported data have shown that  $\Delta F$  increases with tone frequency and amounts to about 1 Hz at a 500-Hz frequency, 3 Hz at 1 kHz and 15 Hz at an 8-kHz frequency [4].

Pitch discrimination of low-frequency tones, in a range below 200 Hz, has been to a much lesser degree explored than the discrimination of higher-pitched tones. It should be noted here that the audio frequency range below 200 Hz encompasses about 40% of the musical pitch scale, down from note G3. In psychoacoustics, the ability to detect a change in a tone's frequency is usually specified as  $\Delta F$ , in hertz, or as a percentage of a reference frequency. In studies related to music pitch DLs are also expressed as the just detectable pitch interval, in cents. A few studies conducted with the use of very low pitched tones have shown that pitch DL, in cents, substantially increases with decreasing tone frequency below 200 Hz [1, 2, 4] and amounts to nearly a semitone at 25 Hz [5].

Published pitch DLs for frequencies below 100 Hz are sparse and show considerable variability across experiments and subjects. In a recent, more systematic study of pitch discrimination of low frequency pure tones Rogowski and Miśkiewicz [6] measured pitch DLs at six frequencies from 20 to 100 Hz and at two higher frequencies – 250 Hz and 1000 Hz. The filled symbols in Fig. 1 show the pitch DLs, in cents, averaged across six subjects. For comparison, Fig. 1 also shows pitch DLs determined in other studies, by an adaptive, up-down adaptive 2AFC method and by the method of adjustment.

The data from 2AFC experiments (Fig. 1) show that pitch DL substantially increases with decreasing tone frequency below 100 Hz and amounts to as much as 270 cents at 20 Hz [6]. It is also apparent that pitch DLs determined by adjustment are much lower than those measured in a 2AFC task. The difference in the size of pitch DL determined by different methods has long been known in the literature. Wier *et al.* [7] reported that the pitch DL estimated for a 1-kHz pure tone by an adaptive, 2AFC procedure, 70.7% correct, is by a factor of 1.6 larger than the DL from the adjustment procedure. This difference is a partial explanation for the very low pitch DLs obtained by Rakowski [1]. Another factor that contributed to a pronounced decrease of the pitch DL estimated in his experiment was the use of correct-answer feedback. The data replotted in Fig. 1 from Jaroszewski [8] indicate that pitch DLs measured by the adjustment method, with and without feedback, differ by a factor of 2 to 4 at various frequencies.

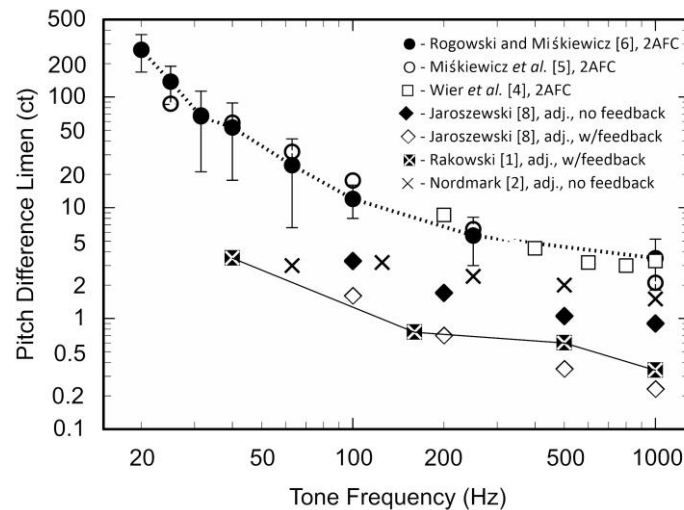


Figure 1. Pitch DLs, in cents, in a 20–1000 Hz frequency range. The data show the 70.7% correct point estimated by an adaptive, 2AFC procedure [4–6], the DLs determined by an adjustment procedure with no feedback [2, 8], and by adjustment with feedback [1, 8]

The method of adjustment involves various judgmental variables which are very difficult to control. One such a variable is the cue used by the subject for comparing the two tones during the adjustment task. Although it is assumed that the subject evaluates the identity of two tones by comparing their pitches, he/she may also, consciously or unconsciously, use other cues. For example, the judgment of identity might possibly be aided by the sensation of roughness, which is strongly pronounced at low frequencies.

To determine whether the pitch DL measured for low-frequency pure tones differs when the subject's task is to indicate which tone has a higher pitch and when the subject has to evaluate whether two tones are identical we shall present the results of an experiment conducted by Wolski [9] with the use of a procedure modelled after Sęk and Moore [10]. Pitch DLs were measured in two stimulus paradigms. In the first, two tone pulses were presented in each trial and the subjects were asked to indicate which pulse had a higher pitch. The discrimination threshold determined in such a way has been termed the *difference limen for frequency*, *DLF*, by Sęk and Moore [10]. In the second paradigm two successive pairs of tone pulses were presented; in one pair the tone frequencies were identical and in the other one different. The subjects were required to indicate in which pair the tone pulses were different. The threshold measured in such a paradigm has been called the *difference limen for change*, *DLC* [10]. In both conditions thresholds were estimated at the 79.4% correct point, by an up-down, adaptive, 2I, 2AFC procedure with feedback. Figure 2 shows the DLFs and DLCs averaged across four subjects [9], the DLFs and DLCs reported by Sęk and Moore [10] and the group DLFs determined by Rogowski and Miśkiewicz [6], replotted from Fig. 1. Open diamonds show the results of the best subject in the experiment of Rogowski and Miśkiewicz [6].

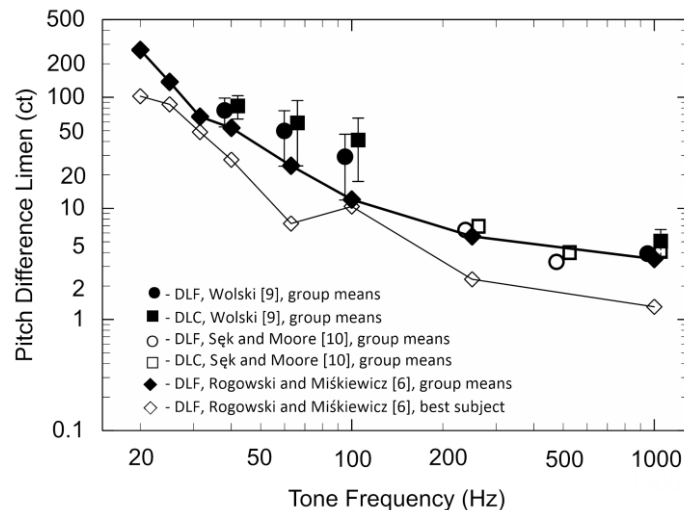


Figure 2. Difference limens for frequency (DLF) and difference limens for change (DLC) measured for pure tones in various experiments, in a frequency range of 20–1000 Hz. For clarity of presentation the symbols indicating the DLF and DLC values are slightly shifted to the left and to the right along the abscissa. The error bars show the standard deviation of individual means around the group mean in Wolski's [9] experiment

The results plotted in Fig. 2 show that DLFs and DLCs are similar to one another at very low frequencies. This finding indicates that the possibility of using auditory cues other than pitch for the judgment of perceptual identity of two tones does not improve the result of frequency discrimination measurement over the level obtained when the subject is required to detect the direction of the pitch change caused by a difference in frequency between two tones. It also should be noted that the data collected by Wolski [9] are in agreement with those of Sęk and Moore [9] who reported that DLFs and DLCs did not differ one from another at frequencies up to 2 kHz whereas at 4 kHz and higher frequencies DLFs were markedly larger than DLCs.

At this point, let us go back to Rakowski's experiment [1] which was the starting point for our discussion. It seems to be clear that the low pitch DLs obtained in his study resulted, at least in part, from using the method of adjustment with feedback. Apart from this, it should be noted that the pitch DLs measured in later experiments in Professor Rakowski's laboratory were very often lower than those obtained by other researchers with the use of the same methods. An explanation for this effect is Professor Rakowski's special approach to the measurements of pitch discrimination.

Depending on the purpose of the experiment, Rakowski's measurements of pitch discrimination were either subject-oriented or object-oriented experiments, *cf.* [11]. An experiment was subject-oriented when it was conducted to study the physiological limits of the auditory system's frequency resolution. When the measurements were made to determine how accurately are discerned the tone pitches in music, the experiment was object-oriented and the subject served in that case the role of a human instrument



for pitch measurement. Professor Rakowski gave much weight to the selection and training of his human measuring instruments. His investigations of pitch perception were usually conducted on musicians who underwent extensive training prior to the participation in the experiment. Sometimes Professor Rakowski managed to fish out an unbelievably gifted person for pitch discrimination from the student population at the Chopin University of Music.

An example from our recent study [6], of a subject with particularly good abilities for pitch discrimination, is shown in Fig. 2. The pitch DL of that subject, measured as the 70.7% correct point, is 2-3 times smaller at a 1-kHz frequency than in other experiments conducted on trained subjects with the use of a similar method. It is also apparent that the pitch DLs of the best subject are lower at frequencies below 1000 Hz.

### 3. Pitch strength of low-frequency tones in music

In a paper introducing the concept of pitch strength Rakowski [3] suggested that the measure of pitch strength be derived from the dispersion of multiple adjustments of the pitch of a pure tone to the pitch of the tone under investigation. Although other authors [12] opted for alternative methods of pitch strength assessment, Professor Rakowski strongly advocated the adjustment method and argued that it reflected, to a considerable extent, the conditions of pitch intonation in music. An unquestionable advantage of that method is that it provides both an estimate of pitch strength, inferred from the standard deviation of pitch adjustments of the test tone, and an estimate of the sound's pitch level, taken as the mean of adjustments. Such a method is also very useful for identifying individual pitch levels when a sound is perceived as having more than one pitch.

The measurements of pitch DL, discussed in Section 2, have shown that the pitches of low frequency tones are heard less accurately than at higher frequencies. From a perspective of music perception it is important to determine whether the poorer ability of discerning pitch changes of individual, low-frequency tones is manifested by less accurate identification of musical intervals and melodies in low pitch registers?

The influence of the pitch strength on musical interval identification accuracy was studied in an experiment conducted on a group of musicians by Rogala *et al.* [13]. The sound stimuli were pure-tone dyads constituting 13 intervals, from unison to octave. Pitch strength was controlled by presenting the dyads in different octaves and by varying their duration. Figure 3 shows the percentage of correct identifications of the intervals, as a function of tone duration, for dyads presented in the second octave (65.4–130.8 Hz), the fourth octave (261.6–523.3 Hz) and in the sixth octave (1047–2093 Hz). The data plotted in Fig. 3 show two effects: (1) due to the splatter of the sound spectrum, interval identification worsens when a brief dyad is shortened and the pitches of the tones are heard less accurately, (2) interval identification is much worse in the lowest octave due to poor frequency discrimination ability of the auditory system at low frequencies.

In an earlier study Rogala [14] used a method of pitch strength assessment that reflected, to a very large extent, the conditions of pitch perception in music. The stimuli were short melodies, made up of electronically transposed musical instrument sound samples presented in a sequence, in random order. The transposition enabled to produce

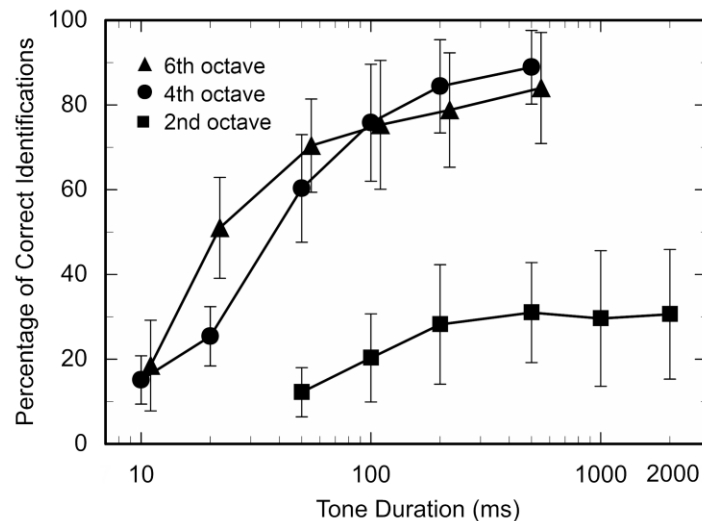


Figure 3. Percentage of correct identifications of musical intervals, as a function of duration of a pure-tone dyad, in different octaves of the musical scale. The error bars show the standard deviation of 14 individual means around the group mean. Data from [13]

melodic patterns from samples of non-melodic percussion instruments which make only one sound and are not pitched to any specific note in music. The stimuli also included melodies produced from sound samples of string instruments, woodwinds, brass instruments, and melodic percussion instruments. The subjects were asked to write down each melody presented in a series of trials.

Figure 4 shows the group mean percentage of correctly written down melodies for each instrument. The instruments are ordered along the abscissa according to ascending pitch of the original sound sample used for making a melody. The last two instruments – the bass drum and the triangle – belong to the class of indefinite pitch instruments. The data shown in Fig. 4 indicate that melodies made up of low-pitched tones, in the first and in the second octave, were identified less accurately than those in higher octaves. A noteworthy effect is that the melodies made up of a transposed sample of the triangle, an indefinite pitch instrument, yielded an identification score comparable to low-pitched tones of melodic instruments.

#### 4. Final remarks

The studies discussed in this paper have demonstrated that the ability of discerning pitch changes of pure tones substantially worsens with decreasing frequency, below about 200 Hz. The poorer ability of pitch discrimination of low-frequency tones causes the musical intervals and melodies to be identified less accurately in the lowest octaves of the musical scale than in higher octaves. However, one may have serious doubts as to whether the pitch DLs measured for pure tones reliably reflect the accuracy of pitch perception of musical instrument tones at low frequencies?

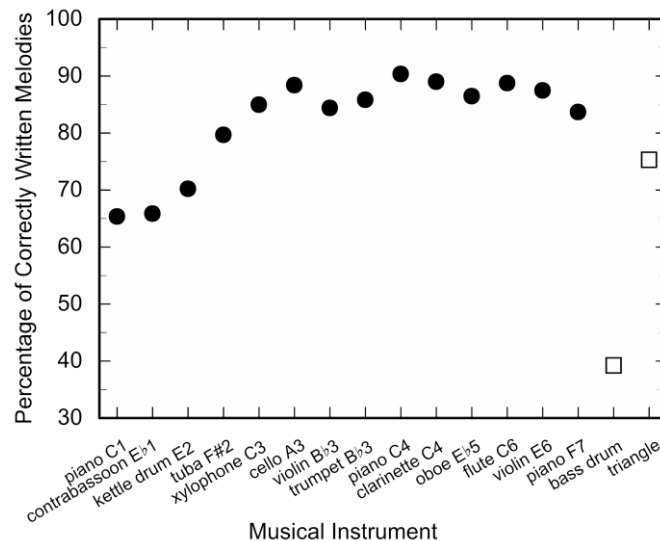


Figure 4. Group percentage of correctly written down melodies, made up of electronically transposed sound samples of various musical instruments. The instruments are ordered on the abscissa according to ascending pitch of the sample. Data replotted from Rogala [14]

It may seem that pitch DLs, to be representative of pitch discerning ability in music, should be measured with harmonic complex tones or with musical instrument samples, however, the use of such stimuli would be burdened with very difficult methodological problems. When tones with harmonic spectra are used in a psychophysical procedure of pitch DL measurement at very low frequencies it is easy to indicate which of the tones has a higher pitch in a trial by listening to the pitches of the higher partials, rather than to the low pitches, or nominal pitches of the tones. The auditory threshold curve and the equal-loudness curves are very steep in the low-frequency range therefore the higher partials of low-pitched musical tones are very much highlighted and easy to hear out.

Professor Rakowski paid very much attention to the relevance of his laboratory experiments to the sound perception conditions encountered in music, and distinguished absolute pitch DLs from operational DLs. Absolute pitch DL is the smallest frequency difference that a subject can discern in laboratory conditions. Operational pitch DL is a pitch difference threshold estimated in conditions relevant to the perceptual context meant to be studied in the experiment, for example, to music listening [15].

Professor Rakowski took a broad, cognitive perspective in his research on pitch perception and put forward a theory in which pitch was considered a multidimensional perceived attribute of sound, serving the role of a communication code in music. In today's terms we would call him a precursor of cognitive and ecological psychoacoustics in the scientific study of sound perception in music. The peak of his research activity fell on the period when explorations of pitch perception were focused mainly on purely sensory and methodological issues. Most studies of pitch perception were conducted with the use of pure tones and static complex tones at that time

and an attempt to bring an experiment closer to real life conditions was always a big challenge. The present, digital sound processing technology and computer-controlled procedures have removed most of those impediments. It also should be noted that much insight into the processes of pitch perception in music is now obtained from brain imaging studies in cognitive neuroscience. In the context of current advancements in auditory research Professor Andrzej Rakowski's scientific legacy is gaining a new perspective.

## References

1. A. Rakowski, *Pitch discrimination at the threshold of hearing*, Proc. of the 7th International Congress on Acoustics, Budapest, (1971) 373 – 376.
2. J. O. Nordmark, *Mechanisms of frequency discrimination*, J. Acoust. Soc. Am., **44** (1968) 1533 – 1540.
3. A. Rakowski, *Measurement of pitch*, Catgut Acoustical Society Newsletter, **27** (1977) 9 – 11.
4. C. C. Wier, W. Jesteadt, D. M. Green, Frequency discrimination as a function of frequency and sensation level, J. Acoust. Soc. Am., **61** (1977) 178 – 184.
5. A. Miśkiewicz, A. Rakowski, P. Rogoziński, M. Kocańda (2002), *Frequency difference thresholds vs. pitch difference thresholds for musical tones*, in: Formation and Perception of Musical Sound Sequences, A. Rakowski (Ed.), Chopin Academy of Music, Warsaw (2002) 125 – 154 (in Polish).
6. P. Rogowski, A. Miśkiewicz – *Pitch discrimination of pure tones at very low frequencies*, Archives of Acoustics (in preparation).
7. C. C. Wier, W. Jesteadt, D. M. Green, *A comparison of method-of-adjustment and forced-choice procedures in frequency discrimination*, Percept. Psychophys., **19** (1976) 75 – 79.
8. A. Jaroszewski, *Effect of feedback on pitch discrimination*, Acustica, **76** (1992) 137 – 141.
9. M. Wolski, *Audibility of pitch changes of low-frequency tones*, unpublished Master's Thesis, Department of Sound Engineering, Chopin University of Music (2009), (in Polish).
10. A. Sek, B. C. J. Moore, *Frequency discrimination as a function of frequency, measured in several ways*, J. Acoust. Soc. Am., **97** (1995) 2479 – 2486.
11. T. Letowski, *Guidelines for conducting listening tests in sound quality*, Proc. Noise-Con 94, Ft. Lauderdale (1994) 987 – 992.
12. H. Fastl, G. Stoll, *Scaling of pitch strength*, Hear. Res., **1** (1979) 293 – 301.
13. T. Rogala, A. Miśkiewicz, P. Rogowski, *Identification of harmonic musical intervals: The effect of pitch register and tone duration*, Arch. Acoust., **42** (2017) 591 – 600.
14. T. Rogala, *Melody identification as a method of measuring pitch strength*, Muzyka, **55** (2010) 65 – 76 (in Polish).
15. A. Rakowski, *Categorical perception of pitch in music*, The State College of Music, Warsaw (1978), (in Polish).

## Measurement of Perforated Panels at a Scaled Measurement Setup

Katarzyna BARUCH

*AGH University of Science and Technology,  
al. Mickiewicza 30, 30-059 Kraków, kbaruch@agh.edu.pl*

Tadeusz KAMISIŃSKI

*AGH University of Science and Technology,  
al. Mickiewicza 30, 30-059 Kraków, kamisins@agh.edu.pl*

Aleksandra MAJCHRZAK

*AGH University of Science and Technology,  
al. Mickiewicza 30, 30-059 Kraków, majchrzak@agh.edu.pl*

### Abstract

In the paper, the authors present an ongoing research on the absorption and measurement uncertainty of perforated panels made at different scales. Knowing the similarity criteria describing the relation between a full-size perforated panel and its scaled equivalent, it is possible to conduct the measurements of the elements of significantly reduced size – with an area not exceeding 0.2 m<sup>2</sup>. This procedure notably decreases the costs resulting from the production, transportation and storing the measurement samples. At the same time, the obtained values of sound absorption coefficient measured for the samples at 1:8 scale will characterize their full-size equivalents of geometry changed according to the derived similarity criteria. The paper discusses the possibilities of measurement of scaled samples.

**Keywords:** scale modelling, law of similarity, dimensional analysis, orifice, miniaturisation

### 1. Introduction

The methodology of sound absorption measurements is mainly based on the laboratory tests of full-size elements [1, 2]. However, only a limited number of institutions have required technical rooms, so the costs of such a measurement are usually high. An interesting alternative to the laboratory sound absorption coefficient measurements may be provided by so-called model tests, which require measurement samples made at scale. Creating models according to the required similarity criteria, may considerably decrease the costs of both making the samples and having them measured.

In acoustics, scale model tests are mainly used for the analysis of room acoustics phenomena [3, 4], transmission of sounds through the building partitions [5, 6] and sound absorption by materials and systems [7, 8]. Despite such a broad range of application, so far there has been no method which would allow scaling perforated panels and keeping their sound absorbing properties unchanged in the shifted frequency range. Therefore, the aim of the carried research is to propose such a methodology and in the following paper the possibilities of measuring the samples at a scaled measurement setup are discussed, together with the measurement uncertainty.

## 2. Subject of the research

The subjects of the research were acoustic perforated panels made at scales 1:4 and 1:8 in relation to their full-scale equivalent. The scales were chosen so as to enable the verification of the obtained results by measurement in a reverberation chamber made at 1:8 scale in relation to the full-scale test room of the Department of Mechanics and Vibroacoustics AGH in Cracow.

The measurement samples were made according to the similarity criteria derived by the authors [9], excluding the criterion regarding air viscosity. The following dimensions were scaled: thickness of the panel  $t_p$ , radius of the orifice  $r$ , distance between the centres of the orifices  $D$ , and distance between the panel and reflective surface  $d$ , according to the following relations, where  $f$  is the frequency, and  $c_0$  is the speed of sound.

$$\Pi_{t_p} = \frac{t_p f}{c_0} \quad (1)$$

$$\Pi_r = \frac{r f}{c_0} \quad (2)$$

$$\Pi_d = \frac{d f}{c_0} \quad (3)$$

$$\Pi_D = \frac{D f}{c_0} \quad (4)$$

Table 1 sets together the parameters describing the studied full-size panel and its scaled equivalents.

Table 1. Parameters of a full-size panel and its scaled equivalents used for the verification measurements

parameter	1:1 scale	1:4 scale	1:8 scale
scale factor	1	4	8
thickness of the panel	12.0 mm	3.0 mm	1.5 mm
radius of the orifice	4.0 mm	1.0 mm	0.5 mm
dimensions	-	400 x 450 mm	400 x 450 mm
perforation rate	12.56%	12.56%	12.56%
distance from the reflective surface	100.0 mm	25.0 mm	12.5 mm
measurement frequency range	100 – 5 000 Hz	400 – 20 000 Hz	800 – 40 000 Hz

### 3. The measurement of sound absorption coefficient at a scaled measurement setup

#### 3.1. Measurement procedure

The methodology for sound absorption coefficient measurements in a reverberation chamber made at 1:8 scale is based on the method described by the standard PN-EN ISO 354 [1]. Analogously to the methodology of sound absorption coefficient at the full-size measurement environment, the values of reverberation time T20 must be measured in two configurations: in an empty chamber and in a chamber with the sample under study. The standard recommends taking the measurements in at least 12 spatially independent combinations of sound source and microphone for each configuration. Based on the measured reverberation times it is possible to calculate equivalent sound absorption area of a tested sample, and then – sound absorption coefficient, using equation below:

$$\alpha = \frac{55.3V}{S} \left( \frac{1}{c_2 T_2} - \frac{1}{c_1 T_1} \right) - \frac{1}{S} 4V(m_2 - m_1), \quad (5)$$

where  $V$  is the volume of the chamber,  $S$  is the area of the tested sample,  $c, T, m$  are: sound speed in air, reverberation time in a chamber and intensity attenuation coefficient, respectively, and indexes 1 and 2 denote the configuration of the measurement: without and with the sample. The results should be given in 1/3-octave frequency bands, in the range of 800-40 000 Hz.

#### 3.2. Measurement setup

Since the measurement of sound absorption coefficient in a scaled reverberation chamber is based on the procedure described in the standard PN-EN ISO 354 [1], the miniaturized measurement room should meet the requirements of this standard, after adjusting the requirements to the scale factor of the chamber.

The design requirements for a 1:8 scale measurement chamber and the parameters of the chamber used for experiments are given in Table 2. In regular room atmospheric conditions, the chamber does not meet the requirement regarding the minimum reverberation time in entire frequency range. However, as it was shown in previous study [10] it is not necessary for the measurement of sound absorption coefficient.

The reverberation chamber used for the measurements is a model of the full-size reverberation chamber of the Department of Mechanics and Vibroacoustics AGH in Cracow. The Schroeder frequency resulting from the dimensions of the chamber is around 320 Hz. This frequency is a limit measurement frequency of the chamber – the measurements are only reliable above 320 Hz. This requirement is met for the measurement samples made at 1:8 and 1:4 scales, whereas the measurement samples made at bigger scale factors (for example 1:2) would require lower measurement frequencies – below the Schroeder frequency of the chamber.

Table 2. Requirements for the design of the reverberation chamber and 1:8 scale and the properties of the chamber used for experiments

	minimum recommended		at the measurement setup	
volume	0.29 m <sup>3</sup>		0.35 m <sup>3</sup>	
area of the sample under study	0.16-0.19 m <sup>2</sup>		0.18 m <sup>2</sup>	
reverberation time	frequency	T20	frequency	T20 (room atmospheric conditions)
	1 000 Hz	0.63 s	1 000 Hz	0.99 s
	2 000 Hz	0.63 s	2 000 Hz	0.96 s
	4 000 Hz	0.63 s	4 000 Hz	0.79 s
	8 000 Hz	0.56 s	8 000 Hz	0.56 s
	16 000 Hz	0.44 s	16 000 Hz	0.30 s
	32 000 Hz	0.25 s	32 000 Hz	0.14 s

The walls of the model and the additional reflective elements are made of plexiglass, in order to ensure minimum sound absorption and maximum insulation from the airborne sounds. A high voltage spark source is used at the setup, which generates signals of 400 Hz – 40 kHz; it is possible to register impulse responses of the chamber in this frequency range, which after scaling (scale factor 1:8) corresponds with the frequency range of 50 Hz – 5 kHz. Two ¼-inch free field microphones GRAS 46BE are used for the data acquisition. The microphones are connected to the measurement interface M-AUDIO FireWire 1814, through SV 06A. The acquisition of data regarding air temperature and relative humidity is realized through the thermo-higrometers Aosong AM2302. The measurement setup is shown in Figure 1.

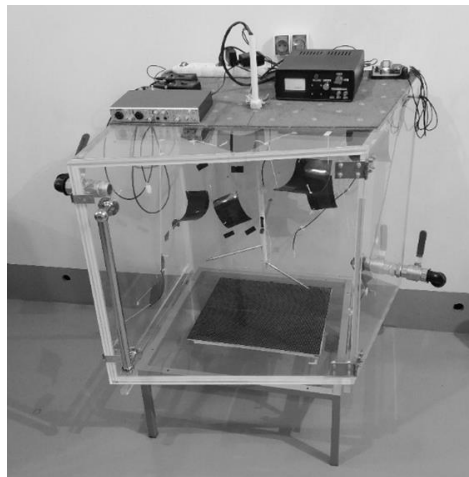


Figure 1. Setup for the measurement of sound absorption coefficient (reverberation chamber at 1:8 scale)



### 3.3. Measurement uncertainty

For a complete expression of the results of sound absorption coefficient measurements, it is necessary to determine the measurement uncertainty. According to PN-EN ISO 354 standard [1] only two factors comprise the total measurement uncertainty: reverberation time measurement uncertainty and reproducibility limits. However, studies on the measurement uncertainty in full-size measurements indicate that for the accurate assessment of sound absorption coefficient measurement uncertainty, other factors must be considered as well, such as the area of the measurement sample [11] or the atmospheric conditions [12]. The latter is especially important for high frequency bands. Since the measurement in a scale reverberation chamber is shifted towards higher frequencies, the atmospheric conditions should be considered in the total measurement uncertainty. Most commonly, for the determination of measurement uncertainty in case of indirect procedures, the law of uncertainty propagation is used. However, given the correlation of input parameters [13] and complexity of the relation between the output parameter – sound absorption coefficient  $\alpha$  and input parameters such as temperature and relative air humidity (involved in the intensity attenuation coefficient  $m$ ), the authors have chosen Monte Carlo method for the determination of the total measurement uncertainty.

In order to use Monte Carlo method for the determination of measurement uncertainty, the distributions of input parameters used for the determination of an output parameter – in this case sound absorption coefficient  $\alpha$ , must be known. They can be estimated using the obtained measurement results or based on the precision of measurement instruments. If the number of measurement results is less than 30, Student's t-distributions of  $n - 1$  degrees of freedom should be assumed [14]. This type of a distribution was assumed for generating the values of reverberation time T20. For the generation of relative air humidity values, temperature values and specimen size values, uniform distributions were used. If the maximum measurement error defined for an instrument is  $\Delta\epsilon$ , it should be assumed that the real value of the measured parameter may be situated at any point of the interval  $\pm\Delta\epsilon$  equally possibly. The distribution of the input value is then a uniform distribution of width  $2\Delta\epsilon$ . Having the distributions of input parameters,  $N$  values of each input parameter must be generated and the calculations of the output value (in this case – sound absorption coefficient) must be repeated the same number of times. The bigger the  $N$ , the more accurate the final result. For practical use it is agreed that  $N = 10^6$  gives satisfactory results [15]. Having  $10^6$  values of sound absorption coefficient in non-decreasing order it is possible to determine an interval which covers  $P\%$  of the obtained results. In case of a symmetric distribution the limits of the interval are given by the samples  $y_{max}$  and  $y_{min}$  of the indexed equal to  $N\frac{P}{2}$  and  $N\left(1 - \frac{P}{2}\right)$ , respectively. If  $P = 95\%$ , the value of  $0.5(y_{max} - y_{min})$  corresponds to the total measurement uncertainty of the output value.

The measurements of sound absorption coefficient were performed for nine independent combinations of sound source-microphone; it was repeated twice for each combination. The values of reverberation times obtained in the measurements were tested for gross errors, using Grubbs test [16]. The obtained statistics characterizing each measurement data set were used for the generation of the distributions of input parameters. For the reverberation time, Student's t-distributions of 17 degrees of freedom

(or less – if some of the values were rejected by Grubbs test) were generated. For the values of relative air humidity, temperature and sample size, uniform distributions were used, and the widths of these distributions were dependent on the measurement instruments. The accuracy of the temperature measurement was 1°C, for relative air humidity it was 2%, and for the size of a sample – 0.5 cm (this was connected not only with the measurement definition, but also inaccurate manufacturing).

#### 4. Results

The results of the sound absorption coefficient measurements of the specimens made at 1:4 and 1:8 scales are presented in the figures below, together with the measurement uncertainty. The samples were additionally verified in a finite element method model, created in COMSOL Multiphysics software. The samples and the models were created using the derived similarity criteria, neglecting the criterion regarding air viscosity. The results obtained in the numerical model were additionally transformed to obtain the statistical sound absorption coefficient to be compared with the measurement results [17]. The consistency between the obtained values is very good – for the sample made at 1:4 scale, root mean square difference between the results of the simulation and measurement is 0.028 for all the tested frequency bands (13.6% of the maximum value), and for the sample at 1:8 scale – 0.031 (11.6% of the maximum value). The curves obtained by measurement are slightly higher for the frequencies above 800 Hz which may be caused by the material properties of the manufactured samples – the simulations assume perfectly smooth and rigid surfaces.

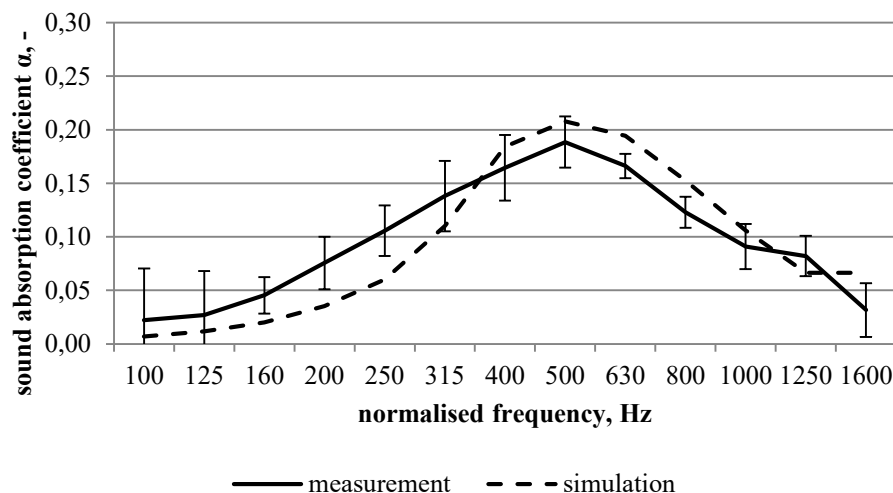


Figure 2. Sound absorption coefficient of a sample made at 1:4 scale as a function of frequency, obtained by measurement and in a numerical simulation (after necessary mathematical transformation)

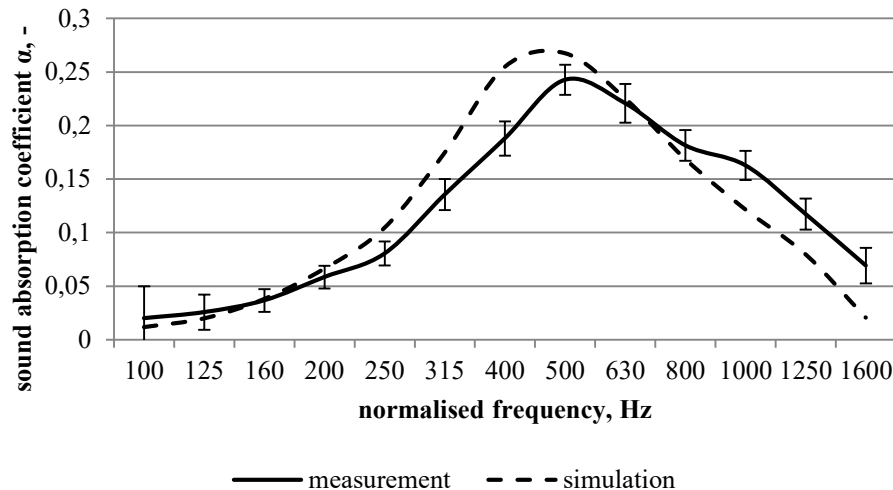


Figure 3. Sound absorption coefficient of a sample made at 1:8 scale as a function of frequency, obtained by measurement and in a numerical simulation (after necessary mathematical transformation)

## 5. Conclusions

In the paper, the authors present the verification of the previously proposed similarity criteria derived for perforated panels. The verification was performed by measurements; two samples were tested: a sample made at 1:4 scale and a sample made at 1:8 scale in relation to their full-size equivalent. The measurements, performed in a 1:8 scale reverberation chamber, show very good consistency with the numerical simulations. Also, the uncertainty of the measurement of sound absorption coefficient of scaled perforated panels was discussed. The values of uncertainty correspond to the full-size measurements, which proves that scale model measurements can be used for the verification of designed solutions without any aggravation of the measurement accuracy.

## Acknowledgments

This work was supported by AGH University of Science and Technology grant no. 16.16.130.942.

## References

1. PN-EN ISO 354, *Akustyka -- Pomiar pochłaniania dźwięku w komorze pogłosowej*, (2005).
2. PN-EN ISO 10534-2, *Akustyka -- Określanie współczynnika pochłaniania dźwięku i impedancji akustycznej w rurach impedancyjnych -- Część 2: Metoda funkcji przejścia*, (2003).
3. M. Barron, *Acoustic scale model testing over 21 years*, *Acoust Bull*, **22** (1997) 5 – 12.

4. H. Jang, J. Jeon, *Evaluation of the absorption by the orchestra in concert halls using scale model and computer simulation*, Int. Symp. Room Acoust., (2013).
5. A. Flaga, A. Szelać, *Dimensional analysis and similarity criteria for the model tests of sound transmission through simple partitions*, Environ Eff Build Struct People – Investig Stud Appl (2016).
6. F. Piekara, A. Szelać, K. Baruch, J. Rubacha, T. Kamisiński, *Badania modelowe izolacyjności akustycznej przegród budowlanych od dźwięków powietrznych*, Aktual Inżynierii Akust i Biomed (2016).
7. M/ Barron, S. Coleman, *Measurements of the absorption by auditorium seating - a model study*, J Sound Vib, **239** (2001) 573 – 87.
8. K. Baruch, T. Kamisiński, *Metodyka kompensacji wilgotności względnej powietrza przy pomiarze współczynnika pochłaniania dźwięku w fizycznym modelu komory pogłosowej*, Postępy Akust., (2015) 427 – 38.
9. K. Baruch, T. Kamisiński, *Analiza wymiarowa w badaniach pochłaniania dźwięku perforowanych ustrojów akustycznych*, Postępy Akust., (2017) 389 – 400.
10. K. Baruch, A. Majchrzak, B. Przysucha, A. Szelać, T. Kamisiński, *The effect of changes in atmospheric conditions on the measured sound absorption coefficients of materials for scale model tests*, Appl Acoust, **141** (2018) 250 – 260.
11. G. Wszolek, *Uncertainty Analysis for Determination of Sound Absorption Evaluation Index DLa*, 7th Forum Acusticum, (2014).
12. A. Iżewska , K. Czyżewski, *Niepewność pomiaru współczynnika pochłaniania dźwięku w komorze pogłosowej*. Pr Inst Tech Bud, **40** (2011) 3 – 13.
13. M. Müller-Trapet, M. Vorländer, *Uncertainty analysis of standardized measurements of random-incidence absorption and scattering coefficients*, J Acoust Soc Am, **137** (2015) 63 – 74.
14. W. Batko, P. Pawlik, G. Wszolek, *Sensitivity Analysis of the Estimation of the Single-Number Sound Absorption Evaluation Index DLa*, Arch Acoust, **42** (2017) 689 – 96.
15. ISO/IEC Guide 98-3:2008: *Uncertainty of measurement. Supplement 1: Propagation of distributions using a Monte Carlo method*, (1995).
16. A. Zięba, *Analiza danych w naukach ścisłych i technice*, Wydawnictwo Naukowe PWN; (2014).
17. C. Jeong, *Converting Sabine absorption coefficients to random incidence absorption coefficients*, J Acoust Soc Am, **133** (2013) 3951 – 62.

## **Sound Strength G Prediction in Orchestra Pit Based on Barron and Lee “Revised Theory”**

Tadeusz KAMISIŃSKI

*AGH University of Science and Technology,  
kamisins@agh.edu.pl*

Krzysztof BRAWATA

*AGH University of Science and Technology,  
kbrawata@gmail.com*

### **Abstract**

Musicians in the orchestra pit often report problems with too high sound levels during performances. At the same time, numerous surveys indicate frequent problems with the mutual hearing of musicians with each other, as well as themselves or a singer from the stage.

The structure of the orchestra pit causes the musicians to be exposed to strong reflections with low delay, which increases the overall sound pressure level. In the literature, one can find recommendations that the space of the orchestra pit should be treated with sound absorbing materials in a wide range and with sound-scattering materials without indicating their quantity or localization.

This paper focuses on the development of tools for assessing and predicting the value of energetic parameters such as sound strength G or clarity C80 in the orchestra pit based on its acoustic absorption and the volume of the hall. The sound propagation model in concert halls proposed by Barron and Lee was adapted for this purpose. The sound strength G can be used to predict the acoustic conditions in the orchestra pit, such as the sound pressure level, the mutual audibility of the musicians, and how the room supports playing musicians. The analysis covers several existing halls of different geometry and size, as well as the design of the barrier of the orchestra pit with the proposed modifications.

**Keywords:** sound strength, revised theory, orchestra pit

### **1. Introduction**

The most frequently reported by musicians working in orchestral pit is the problem of too high sound levels to which they are exposed. Too high sound levels cause problems with hearing individual groups of instruments, and often yourself. High concentration of music on one's own game causes a decrease in its ability to follow the conductor, which affects the difficulty of maintaining the balance between the soloists on the stage and the musicians in the orchestra pit [1].

Due to the limitations of the available surface, usually absorbing materials of low thickness are used, and thus, with high acoustic absorption for medium and higher frequencies (carpet on the floor [1], curtains hung on the walls [2], overhang treated with mineral wool boards, etc.). This results in a negative change in reverberation characteristics, resulting in some instruments being favoured and another part being masked. An additional difficulty in the correct design of the acoustic adaptation is

the fact that the shape, size, height, depth of the orchestra and the various acoustic defects resulting from it are of great diversity.

In the literature it is difficult to find a description of the tools and values of objective parameters that could be used to design and evaluate the quality of the orchestra.

## 2. Method

To assess the sound level in the room, the sound strength  $G$  is used. This parameter, divided into time intervals, is well suited to the assessment of the reverberant noise.

The sound propagation model proposed by Barron and Lee [3] for concert halls was used to predict the value of the sound pressure parameter  $G$  with the time distinction between early sound ( $G_{\text{early}}$ ) and late sound ( $G_{\text{late}}$ ). The calculation model is described by equations 1 a-d.

$$d = \frac{100}{r^2} \quad (1a)$$

$$e_r = (31200 T/V) e^{-0,04r/T} (1 - e^{-1,11/T}) \quad (1b)$$

$$l = (31200 T/V) e^{-0,04r/T} e^{-1,11/T} \quad (1c)$$

$$L = 10 \log(d + e_r + l) \text{ [dB]} \quad (1d)$$

where the acoustic energy is divided into three parts:

$d$  – the value of direct sound energy,

$e_r$  – the value of the energy of early reflections,

$l$  – value of late sound energy,

$T$  – reverberation time [s],

$V$  – room air volume [m<sup>3</sup>],

$r$  – distance between the source and the receiver [m].

Due to the occurrence of phenomena related to coupled rooms, such as double slope of the decay curve, the reverberation time and volume resulting from the separation of the orchestra pit from the room were used to predict the early energy, while reverberation time was used to predict the value of late energy ( $G_{\text{late}}$ ). absorbency of the entire room and its volume.


The influence on the observed parameters of the musicians and equipment present in the orchestra was also examined.

Obtained results from the Barron and Lee model were compared with the results obtained on the measurement path. The measurements were taken in the Krakow Opera.

## 3. Research

The research was carried out in the orchestra of the opera theater. The table 1 presents a short summary of the most important parameters of the room [4].

Table 1. General parameters value of Opera Krakowska hall

Opera Krakowska	
	
Open	2008
Architect	Romuald Loegler
Seating Capacity	735
Room air volume [m <sup>3</sup> ]	6504
Stage air volume [m <sup>3</sup> ]	3933
Reverberation time T <sub>500-1000Hz</sub> [s]	1.11
Early Decay time EDT <sub>500-1000Hz</sub> [s]	1.09
Sound strength G <sub>mid</sub> [dB]	3.8
Clarity C <sub>80</sub> [dB]	3.91

Below (Fig. 1) is a schematic diagram of the measurement points in the orchestra pit, which was used when examining objects for the purpose of this work, meeting the above-mentioned requirements.

Measuring points can be described as:

- P1 – location of the first group of violins,
- P2 – the left back of the orchestra pit,
- P3 – location of the cello / double-bass group.

All points were located at least 1 m from large reflective surfaces. Height of sources and microphones 1 m above the floor surface.

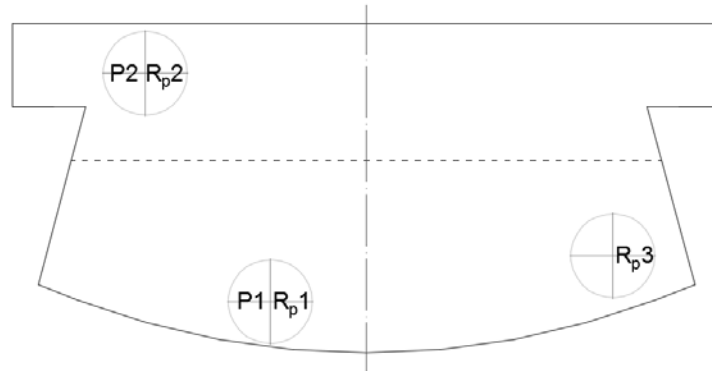


Figure. 1 Measurement layout in orchestra pit: P – source position, R – receiver position; sources and receivers 1 m from reflection surface, and 1 m above floor [5]

To calculate the time of the report with the equipment of the orchestra or with the musicians, the values of the acoustic absorption of individual elements were used based on the results made by J. Rubacha [6]. Values are shown in Table 2.

Table 2. Sound absorption [1m] of empty chairs, and musicians seating on chairs [6]

	1/1 oct [Hz]					
	125	250	500	1000	2000	4000
Empty chairs	0.05	0.15	0.4	0.6	0.9	1
Musician on chair	0.15	0.3	0.8	1.2	1.4	1.5

#### 4. Results

The tables 2-4 listing of the obtained from measurements and model values. The Root Mean Square Deviation (RMSD) was used to assess the quality of predicted values.

Table 2. Values of sound strength  $G_{mid}$  derived from measurement (M) and from model (B)

$G_{mid}$ [dB]	Empty		Chairs		Musicians	
	M	B	M	B	M	B
P1-Rp2	13.7	12.1	9.6	11.9	8.2	10.3
P1-Rp3	10.5	10.8	5.8	7.5	5.0	4.9
P2-Rp1	13.7	12.1	9.6	11.9	8.2	10.3
P2-Rp3	8.7	10.0	9.3	5.1	6.6	2.1
RMSD	1.35		2.77		2.71	



Table 3. Values of early sound strength  $G_{80\_mid}$  derived from measurement (M) and from model (B)

$G_{80\_mid}$ [dB]	Empty		Chairs		Musicians	
	M	B	M	M	B	M
P1-Rp2	13.7	11.8	11.9	9.4	7.9	10.3
P1-Rp3	10.4	10.1	7.5	4.6	3.5	4.9
P2-Rp1	13.7	11.8	11.9	9.4	7.9	10.3
P2-Rp3	8.6	9.0	5.1	8.5	5.5	2.1
RMSD	1.36		2.84		2.49	

Table 4. Values of late sound strength  $G_{late\_mid}$  derived from measurement (M) and from model (B)

$G_{late\_mid}$ [dB]	Empty		Chairs		Musicians	
	M	B	M	M	B	M
P1-Rp2	0.2	0.3	-4.3	-0.1	-4.4	-1.1
P1-Rp3	-0.6	2.1	-0.1	-0.9	-0.5	-2.0
P2-Rp1	0.2	0.3	-4.3	-0.1	-4.4	-1.1
P2-Rp3	-1.1	0.4	1.5	-1.4	-0.2	-2.6
RMSD	1.57		3.33		2.70	

## 5. Conclusions

The article presents the method predicting sound strength  $G$  values using computational models based on the Barron and Lee methods.

The method allows you to quickly and easily investigate to what extent sound absorption orchestra pit absorbed influence the value of sound power. In addition, it is possible to examine how much the degree of closure affects the conditions in the orchestra.

The proposed method can also be used to assess the quality of the interior of the orchestra and investigate the impact of other elements of the orchestra and its equipment on the distribution of sound levels in the areas of the audience and the stage.

## References

1. H. V. Fuhs, Applied Acoustics: Concepts, Absorbers, and Silencers for Acoustical Comfort and Noise Control: Alternative Solutions-Innovative Tools-Practical Examples, Springer Science & Business Media, 2013.
2. J. Meyer, Acoustics and the performance of music: Manual for acousticians, audio engineers, musicians, architects and musical instrument makers, Springer Science & Business Media, 2009.

3. M. Barron, L. Lee, Energy relations in concert auditoriums. I. The Journal of the Acoustical Society of America, **84.2** (1988) 618 – 628.
4. T. Kamisiński, M. Burkot, J. Rubacha, K. Brawata. Study of the effect of the orchestra pit on the acoustics of the Kraków Opera Hall, Archives of Acoustics, **34**(4) (2009) 481 – 490.
5. K. Brawata, Akustyczne aspekty konstrukcji fosy orkiestrowej w kontekście interakcji ze sceną, PhD Thesis, 2019.
6. J. Rubacha Aspekty akustyczne struktur widowni w salach widowiskowych, PhD Thesis, 2013.

## Noise Radiation from Circular Rods at Low-Moderate Reynolds Number

Joanna Maria KOPANIA

*Lodz University of Technology, 266 Piotrkowska Street,  
90-924 Lodz, Poland, joanna.kopania@p.lodz.pl*

Grzegorz BOGUSŁAWSKI

*Lodz University of Technology, 266 Piotrkowska Street,  
90-924 Lodz, Poland, grzegorz.boguslawski@p.lodz.pl*

Patryk GAJ

*Institute of Power Engineering - Thermal Technology Branch "ITC" in Lodz,  
113 Dabrowskiego Street, 93-208 Lodz, Poland, patryk.gaj@itc.edu.pl*

Kamil WÓJCIAK

*Institute of Power Engineering - Thermal Technology Branch "ITC" in Lodz,  
113 Dabrowskiego Street, 93-208 Lodz, Poland, kamil.wojciak@itc.edu.pl*

### Abstract

The well-known dominant sources of airframe noise are associated with unsteadiness of separated and/or vortical flow regions around the high-lift system (flaps, slats) and the aircraft undercarriage (landing gear). Current practical landing gear noise prediction models are individual component - based, which means that the various components are divided into groups according to the frequency range, in which they predominantly radiate noise. Since the far-field noise spectra are approximately Strouhal - based, the emitted frequency is assumed to be directly related to their size: the large elements are responsible for the low frequency region of the spectra, and the small components for the high frequency region. On the basis of such understanding of the noise generation mechanism, the special configurations that lead to considerable noise suppression were proposed. One element of these configurations are rods with different shape and cross section. In this work the situation when circular rods are in area of laminar-turbulent flow were analysed. The measurements were carried out for single circular rod with different diameters to study the noise effect depended on Reynolds number. Far field noise for broad range of Reynolds numbers was also examined depending on distance from the source of noise.

**Keywords:** aeroacoustical noise, circular rod, low Reynolds number

### 1. Introduction

The noise radiating from the outer parts of an aircraft is the dominant airframe noise source. Plenty of numerical and experimental studies are being conducted to identify the noise generated from these elements [1, 2]. Many components of a landing gear of aircraft (struts, cables, axles and wheels) can be modelled by rods of various lengths and cross-sections. Characteristic noise is radiated from single and multiple rods configurations. The subcritical, critical, and turbulent states of flow are of most interest in aspect of aircraft noise because of the Reynolds number range it

encompasses. For circular rod the Reynolds number is based on the rod diameter and the free-stream velocity of the uniform flow (especially when  $Re > 10^5$ ) [3, 4]. But so far, it is not known at which value of  $Re$  the fully turbulent state of flow starts [5, 6].

The noise generated by a circular rod in a uniform or turbulent flow has been studied through [7, 8]. Vortex shedding noise from single cylindrical rods were studied. Far field noise, surface pressure fluctuations and span-wise correlation lengths over a broad range of Reynolds numbers ( $2 \times 10^4$  to  $5.5 \times 10^5$ ) were examined [7]. The Aeolian tones from a single rod at very high Reynolds (namely  $2.5 \times 10^5$  to  $2 \times 10^6$ ) were related to the decrease and increase of the tone's amplitude in the super- and post-critical regimes to surface pressure fluctuation characteristics [9, 10]. The effect of free-stream turbulence on the vortex shedding noise from a single rod has also been studied and more broadband nature of the vortex shedding noise in the presence of free-stream turbulence was identified.

In these experimental studies acoustic and aerodynamic measurements were performed on a single circular rod with different diameters to study the effect of low-Reynolds number on the radiated noise. The test matrix included rods of diameter – 8 mm, 10 mm and 16 mm and low  $Re$  numbers were studied. Obtained flow physics and acoustics properties have been analysed to understand the effects of aerodynamically noise of circular rods. The present studies can be also referenced to the flow in ventilation system (elements like grills, diffusers, slats), because in these system the low Reynolds numbers are often used and velocity does not exceed 5 m/s.

## 2. Experimental method

The measurements were performed on the specially constructed test stand with the outlet to the anechoic room at the Institute of Power Energy in Lodz – Figure 1. Airflow was induced by a fan mounted on the inlet of the stand and regulated by the power inverter. The anechoic test chamber is cubic, approximately  $350 \text{ m}^3$  in volume and has got walls that are acoustically treated with foam wedges providing a reflection free environment.

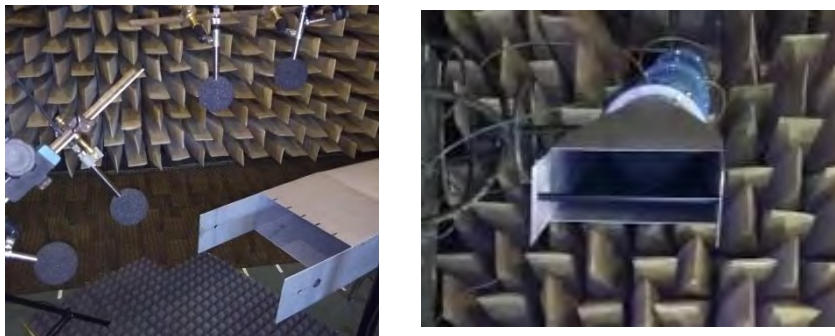


Figure 1. The construction of test stand with outflow to anechoic room and microphones

The rods were supported by two vertical side plates that were mounted to the short sides of the nozzle. The diameter of the cylindrical rods was 8 mm, 10 mm and 16 mm. The measurements of noise were made by using analyzer SVAN 958. The four

microphones (M1, M2, M3, M4) were located at a distance of 500 – 600 mm above the rods, with the distance between microphones presented on Figure 2. The microphones were calibrated before commencing the acoustic test. Microphones were used to measure noise in the flow field. When a microphone is located within an airflow field, it is recommended as best practice that a windscreen or nose-cone accessory should be used when taking acoustic measurements. In head-on laminar flow, the nose-cone accessory is the best choice. In all turbulent flow and parallel orientation laminar flow, the windscreen accessory is the best choice [11]. In these studies the microphones were located parallel orientation to flow streamline and were used in turbulent flow, so windscreens were placed on these microphones. In this studies, due to the large number of data, the results obtained only from microphone M1 were interpret.

The measurements were taken at range of flow velocities between 0.02 m/s and 17 m/s (for studied rods range of Reynolds numbers was between  $Re = 12 \div 18000$ ). The velocity distribution in outlet of the stand test was measured by using a wing anemometer connected with a HD 2103.2 instrument. Mean velocities were calculated by using the logarithmic Chebyshev method. Additional, detailed flow measurements at several points behind the rods have been performed using thermo anemometer measurement technique. For thermoanemometer the points of measurement were logged at locations  $z$ -axis = (-30 mm; -10 mm; 0 mm; 10 mm; 30 mm) and  $x$ -axis = (10 mm and 30 mm), relative to the surface of mounted rods.

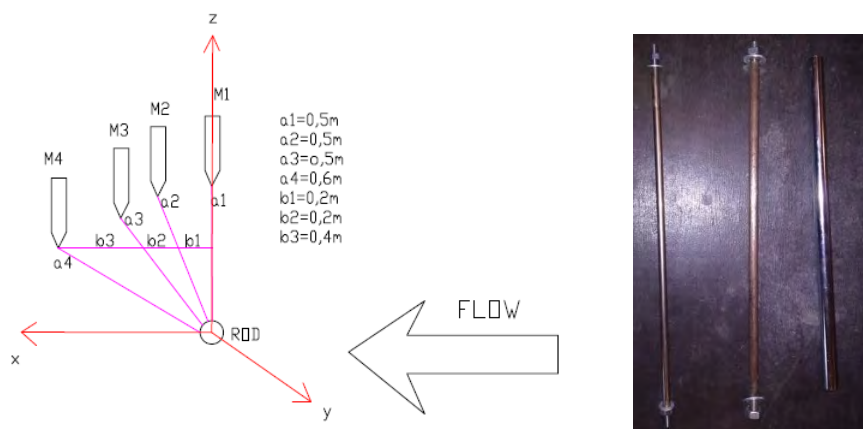


Figure 2. Localization of microphones above the rods (on the left) and rods used in these studies (on the right)

### 3. Aeroacoustical results

Aerodynamic and acoustic parameters of airfoil or other objects (e.g. circular or square rods) in low Reynolds number are important for learning its physical nature. Most of the flying animals (insects, birds), fly at Reynolds number of  $10^3 - 10^5$ , due to their low speed and small length scales and they have got varied physiology which accommodates

to those low velocities. Current trend of "inspiration on nature" to improve the aerodynamics of flight and reduce flight's noise parameters requires knowledge in this aspect. But very important is knowing the aerodynamical and acoustic parameters for simple objects like rods with different cross section, or flat or bend plate, which might be used for studying more complicated structures.

The first step in this study was determination of the distance between the rods and microphones. The preliminary studies allowed determining the maximum velocity in the outlet of constructed research stand at 17 m/s. Dependence between of the observed peak for rod with diameter of 10 mm (315 Hz at flow velocity 17 m/s) on the 1/3 octave sound pressure level spectrum (SPL), determined as the difference between the SPL spectrum of studied rod and background of the research stand, is presented at the Figure 4. The results show the measurements acquired with microphone M1 (Figure 2), for an observer located at a 90° elevation angle. The value of studied SPL difference increases with distance, and due to stable position of microphones as a studied farness from the rods, the 500 mm was chosen.

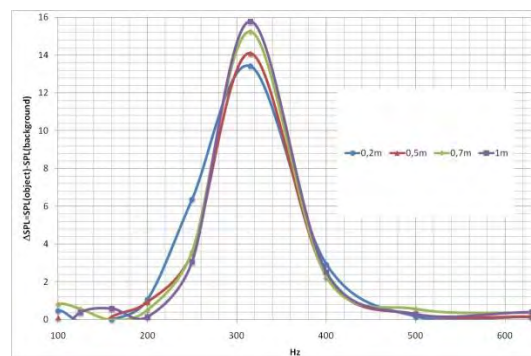


Figure 3. Dependence between the observed peak on the 1/3 octave sound pressure level spectrum (SPL), determined as the difference between the SPL spectrum of the studied rod with 10 mm diameter and background of the research stand

Acoustic measurements were performed on a simple single rod configurations to determine the 1/3 octave spectrum of the sound pressure level (SPL). This analysis will aid in the interpretation of the results and might be ground for future studies when this shape of rods will be used in more complex arrangements. The free stream flow was uniform, and the surface of the rods was smooth. The 1/3 SPL spectrums were measure for empty research stand and research stand with rods. The difference between these two 1/3 SPL spectrum was important in these studies. The spectrums for each of studied rods was only evaluated for frequencies and velocities where the level difference of these two measurements is above 2 dB. As is seen from Figure 4 and Table 1 the maximum peak at low range of frequencies is observed for studied rods. This peak move to higher frequencies along with increase the velocity of flow. For example, there is maximum at 125 Hz at 6.5 m/s for rod with diameter 8 mm, but when the velocity increase to 15.8 m/s the maximum peak move to 315 Hz.

Table 1. Differential SPL values of characteristic peaks the 1/3 octave spectrums depending on frequency and velocity of air flow (microphone M1)

m/s	rod with 8 mm diameter						rod with 10 mm diameter						rod with 16 mm diameter				
	Hz						Hz						Hz				
	100	125	160	200	250	315	100	125	160	200	250	315	80	100	125	160	200
4.8	1.99						2.22						5.19				
6.5		3.48						3.04						8.51			
8.2			6.47						6.46					9.56			
9.8				6.79						5.55					11.2		
11.4				5.94						5.57						13.61	
12.9					9.61						7.7					13.53	
14.4						10.5						10.65					12.24
15.8						13.78						14.09					13.09

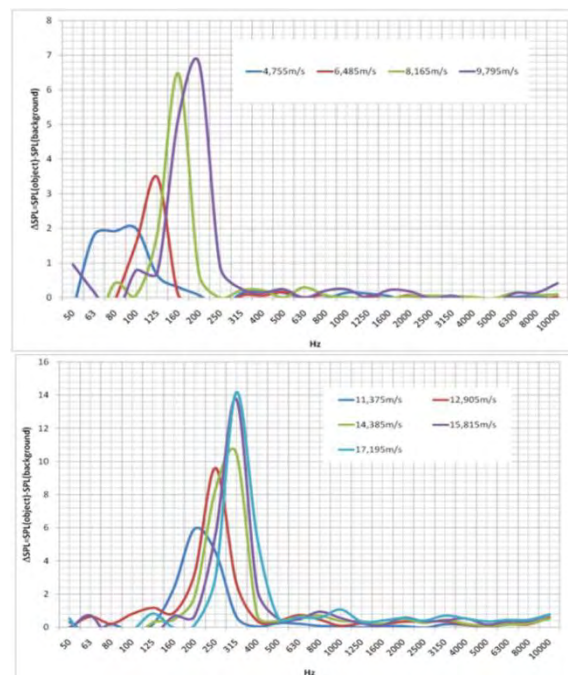


Figure 4. Differential 1/3 SPL spectrum between the SPL spectrum of studied rods with 8 mm diameter and background of the research stand dependent on velocity

If diameter of rod increases twice (from 8 mm to 16 mm), the maximum peak at the same velocity moves to lower frequency. For example, at velocity of 15.8 m/s, there is

maximum peak at 315 Hz for the rod with diameter 8 mm, but at 200 Hz for rod with diameter 16 mm.

To find the dependences between velocities, frequencies and observed peaks for studied rods, the colour-maps of differential 1/3 SPL were done. Sound field propagation depends on many factors like distance, air density, temperature, humidity, terrain, wind direction, etc. In addition to these simple factors there are issues relating to tonality of the noise source and also octave spreading. Additionally, noise is effected by refraction and reflection, so frequently does not travel in a straight line. The perception of noise is highly subjective and non-linear, and it must still be interpreted to have any meaning. In this context the pragmatic approach to noise calculation is needed. In this work the differential 1/3 SPL, determined as the difference between the 1/3 SPL spectrum of studied rod and background of the research stand, for studied rods is examined, what eliminates other phenomena, which could influence on 1/3 SPL spectrum. The 2 - D colour-maps were done by using bi-cubic interpolation depending on the obtained parameters for studied rods. The bi-cubic interpolation can achieve good performance for smooth regions because it assumes smoothness of obtained data. For each rods, the differential 1/3 SPL colour-maps were done for the range of frequency 63 – 400 Hz (Figure 5).

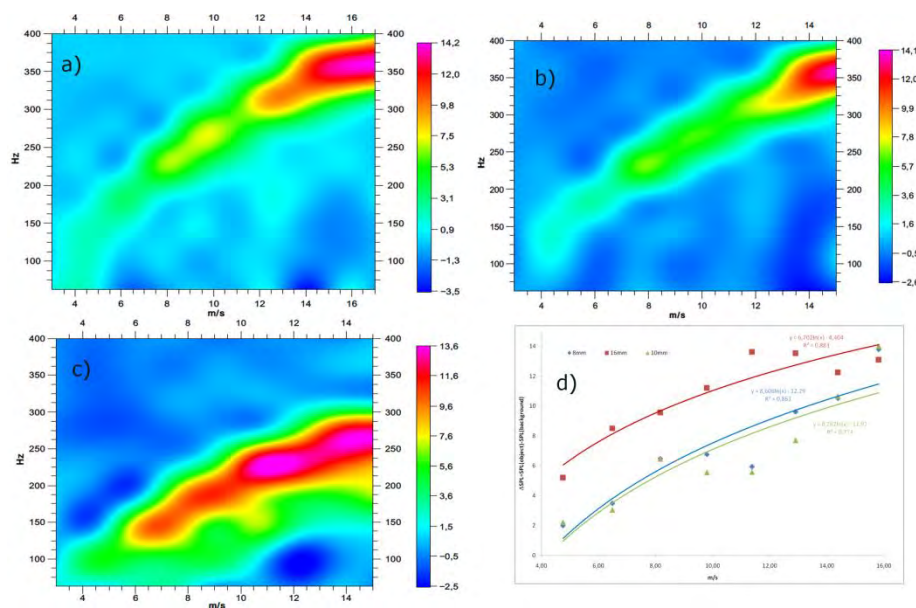


Figure 5. Differential 1/3 SPL colour-map dependent on velocities ( on the graph: a) for 8 mm rod (microphone M1); b) for 10 mm rod (microphone M1); c) for 16 mm rod (microphone M1); d) linear-log models for values of differential 1/3 SPL for observed peaks of rods dependent on velocities

As seen from these colour-maps, the maximum peak (from microphone M1) is moving along the curve could be described by the linear-log model, what is seen



on the Figure 5, graph d). Logarithmic functions are very helpful when working with phenomena that have a very wide range of values, because they allow you to keep the actually obtained values in a smaller range. Logarithmically transforming variables in a regression model is a very common way to handle situations where a non - linear relationship exists between the variables and such situation we often have got in acoustic measurements. However exact interpolation and get the better adjustment of data by these functions require more measurements, what will be continued.

The velocities of flow at several points behind the rods have been evaluated using thermo anemometer measurement technique. The points of measurement were chosen at locations  $z = (-30 \text{ mm}; -10 \text{ mm}; 0 \text{ mm}; 10 \text{ mm}; 30 \text{ mm})$  – perpendicular to stream flow and  $x = 10 \text{ mm}$  and  $30 \text{ mm}$  (according to flow stream), relative to the surface of mounted rods. In the Figure 6, the mean velocities of flow in chosen points behind the 10 mm rod and 16 mm rod are presented.

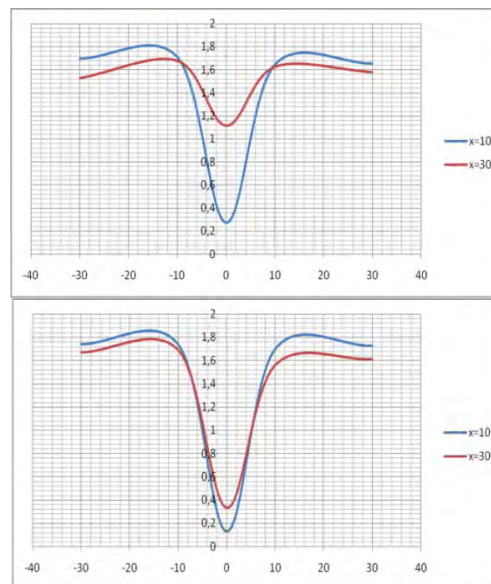


Figure 6. The mean velocities of flow behind the 10 mm rod (on the top) and 16 mm rod (on the bottom) in  $z$  and  $x$  axis at  $1.2 \text{ m/s}$

The points of measurement were chosen at locations axis  $z = (-30 \text{ mm}; -10 \text{ mm}; 0 \text{ mm}; 10 \text{ mm}; 30 \text{ mm})$  – perpendicular to stream flow and axis  $x = 10 \text{ mm}$  and  $30 \text{ mm}$  - according to flow stream, relative to the surface of mounted rods. At the Figure 6 the velocities of flow in chosen points behind the 10 mm rod and 16 mm rod are presented. Results show a similar velocity distribution along the  $x$ -axis. It can be seen that the minimum of the velocity profile behind the 10 mm rod is at 0.27, and behind the 16 mm rod is at 0.12, at the distance of  $x = 10 \text{ mm}$ . The velocity profile at distance of  $x = 30 \text{ mm}$  increases for 10 mm rod to 1.1 but a little for 16 mm rod to 0.33.

Figure 6 shows a different shape of the wake for the 16 mm rod as for the 10 mm rod. This means that the width of the wake behind the rods is predicted correctly.

#### 4. Conclusions

Experiments have been performed to investigate the flow around a circular rods with different diameter into uniform air stream. The vortex shedding frequencies and flow velocity were measured for Reynolds numbers between 12 and 18000, but the range  $Re$  above 3000 was emphasized. The sound pressure level as a 1/3 octave spectrum and position of the dominant noise is highly dependent on the velocity. For studied rods, the maximum peaks at low frequencies are observed (from 100 Hz to 315 Hz). Maximum peaks move to higher frequencies along with the increase of the velocity of flow. Finally, the aerodynamic parameters, as a flow velocity around cylinder, were performed. The velocity profile behind the rods is predicted correctly according to literature data.

#### References

1. M. G. Smith, L. C. Chow, *Validation of a prediction model for aerodynamic noise from aircraft landing gear*, AIAA paper 2002 – 2581.
2. W. Dobrzynski, H. Buchholz, *Full scale noise testing on airbus landing gears in the German-Dutch Wind Tunnel*, AIAA paper 97-1597.
3. M. M. Zdravkovich, *Flow around circular cylinder, Volume 1*, Oxford University Press, 1997.
4. M. M. Zdravkovich, *Flow around circular cylinder, Volume 2*, Oxford University Press, 2003.
5. O. Inoue, N. Hatakeyama, *Sound generation by a two-dimensional circular cylinder in a uniform flow*, Journal of Fluid Mechanics, **471** (2002) 285 – 314.
6. S. J. Park, C. W. Lee, *Flow structure around a finite circular cylinder embedded in various atmospheric boundary layers*, Fluid Dynamics Research, **30**(3) (2002) 197 – 215.
7. R. H. Schlinker, M. R. Fink, R. K. Amiet, *Vortex noise from non-rotating cylinders and airfoils*, AIAA paper 76 – 81.
8. H. Fujita, H. Suzuki, A. Sagawa, T. Takaishi, *The Aeolian tone characteristics of a circular cylinder in high Reynolds number flow*, AIAA paper 99 – 1849.
9. W. A. Olsen, *Noise generated by impingement of turbulent flow on airfoils of varied chord, cylinders, and other flow obstructions*, AIAA paper, 76 – 504.
10. M. R. Davis, N. H. Pan, *Noise generated by the interaction of turbulent jets with circular cylinders*, Journal of Sound and Vibration, **135**(3) (1989) 427 – 442.
11. A. R. Barnard, *Flow Induced Noise Reduction Techniques for Microphones*, Sound & Vibration, October (2014) 3 – 12.
12. L. McCormack, S. Delikaris-Manias, V. Pulkki, *Parametric acoustic camera for real-time sound capture, analysis and tracking*, Proceedings of the 20th International Conference on Digital Audio Effects (DAFx-17), Edinburgh, UK, September (2017) 5 – 9.

## Required Attenuation of Aircraft Noise in Buildings in the Light of Data from the Chopin Airport Monitoring

Krzysztof RUDNO-RUDZIŃSKI

*Department of Acoustics and Multimedia, Wrocław University of Science  
and Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland,  
krr@pwr.edu.pl*

### Abstract

The minimum attenuations of aircraft noise required by the Polish Standards concerning acoustic insulation and noise levels inside buildings have been compared. Data from the noise monitoring system of the Chopin Airport were used.

The minimum attenuation required by PN-B-02151-02:1987 differs from that required by PN-B-02151-3:2015-10. The highest required attenuation occurs when taking into account the maximum sound level  $L_{Amax}$  of aircraft noise. The requirements related to  $L_{Amax}$  are 2.0 to 7.8 dB higher than those associated with the equivalent sound level  $L_{Aeq}$ .

If the number of flight operations at night reaches the allowed maximum number of forty, the requirements related to  $L_{Aeq}$  may be decisive.

The requirements connected with  $L_{Aeq}$  according to PN-B-02151-02:1987 are 0.6 to 2.6 dB higher than coming from PN-B-02151-3:2015-10.

**Keywords:** building acoustic, aircraft noise, permissible noise level in buildings

### 1. Introduction

Aircraft noise is the third most dominant source of environmental noise in Europe in terms of the number of people exposed [1]. However, due to the adverse impact on people, the WHO recommends lowering aviation noise indexes to values lower than the recommended ones for road and rail noise [2].

The ICAO in a balanced approach to noise management [3] has identified instruments to mitigate noise impact, such as building codes and building noise insulation programs, as one of the principal noise reduction measures. The practical implementation of these instruments results from the conditions in individual countries [4]. However, it has generally been known for a long time that in the bedrooms the equivalent sound level of  $L_{Aeq}$  within 8 hours of the night should not exceed 30 dB and 45 dB  $L_{Amax}$  for a single acoustic event [5].

There are two standards in Poland regarding the protection against aircraft noise in buildings. PN-B-02151-3:2015-10 let calculate the sound reduction indexes required in buildings under the construction process [6], while PN-B-02151-02:1987 gives the permissible noise levels in residential rooms [7]. This study compares the requirements of noise protection in buildings resulting from the above-mentioned standards. The aircraft noise data were taken of the noise monitoring system of the Chopin Airport in Warsaw [8].

## 2. Required attenuation of a building envelope

The difference  $D_A$  of A-weighted sound level outside the building  $L_{A,out}$  and inside the building  $L_{A,in}$  is a measure of noise attenuation by the building envelope.

$$L_{A,out} - L_{A,in} = D_A \quad (1)$$

In order to the requirements for protection against aircraft noise to be met,  $D_A$  has to be at least equal to the minimum value resulting from the standards:

$$D_A \geq D_{A,min} \quad (2)$$

$D_{A,min}$  is the required attenuation of aircraft noise through the partitions of given room. The value of  $D_{A,min}$  equals to the difference of the level of external noise and the permissible noise level in the room. As a rule, higher noise protection requirements apply to the nighttime.

## 3. Aircraft noise levels

The data from the Chopin Airport noise monitors were used including the A-weighted sound exposure level  $L_{AE}$ , A-weighted maximum sound level  $L_{Amax}$ , date and time of the event, type of aircraft and type of flight operation

With regard to the airports where  $L_{Amax}$  is not measured, this noise index can be estimated using the methods described in the literature [6, 9].

Figure 1 shows locations of noise monitors of Chopin Airport. Only monitoring points located on inbound and outbound tracks were included.

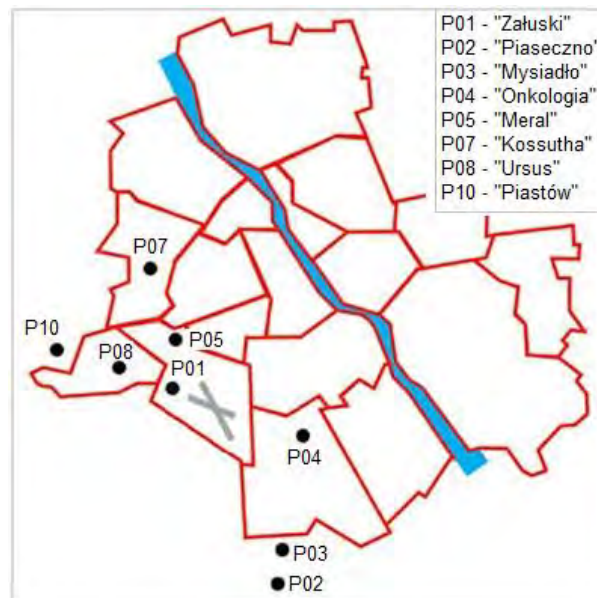


Figure 1. Layout of the Warsaw Chopin Airport noise monitors (based on [8])

The average value of the sound exposure level  $L_{AE}$  at monitors was calculated from the equation

$$L_{AE} = 10 \log \left( \frac{1}{n} \sum_{i=1}^n 10^{0.1 L_{AE,i}} \right) \quad (3)$$

where  $L_{AE,i}$  is A weighted sound exposure level of the  $i$ -th flight operation, and  $n$  is the number of flight operations during the whole year or in the three consecutive months with the highest number of flights at night.

The mean value of the maximum sound level  $L_{Amax}$  was analogously calculated

$$L_{Amax} = 10 \log \left( \frac{1}{n} \sum_{i=1}^n 10^{0.1 L_{Amax,i}} \right) \quad (4)$$

Calculated values are given in Table 1. The first two columns give the monitor number and name. Column 3 shows the average number of flight operations  $N$  at nighttime. The average  $L_{AE}$  and  $L_{Amax}$  are given in column 4 and 5, respectively.

Columns 6 and 7 shows respectively the average number  $N_L$  and the average level  $L_{Amax,L}$  of “loud events”. The loud events include the flight operations which A-weighted maximum sound level  $L_{AmaxF} \geq 70$  dB.

In column 8 the coefficient  $p$  is given, which is a ratio of maximum number of flight operations in a half hour at the nighttime to the overall number of flight operations at night.

Table 1. Noise parameters at the Chopin Airport monitors at the nighttime

Monitor	Name	$N$	$L_{AE}$	$L_{Amax}$	$N_L$	$L_{Amax,L}$	$p$
1	2	3	4	5	6	7	8
P01	Załużski	18.4	93.6	87.3	17.6	87.4	0.29
P02	Piasieczno	6.8	80.2	70.2	1.4	74.3	0.23
P03	Mysiadło	14.5	83.8	74.5	10.6	74.6	0.24
P04	Onkologia	6.0	88.8	79.3	5.3	79.8	0.32
P05	Meral	5.5	84.5	74.0	4.3	77.1	0.36
P07	Kossutha	1.3	82.0	72.0	0.7	74.5	0.35
P08	Ursus	16.8	83.5	73.3	10.1	74.6	0.28
P10	Piastów	13.9	78.5	66.9	0.7	72.2	0.25

The data refer to 2017 with the exception of P05 Meral, where the data from 2018 were included, when the number  $N_L$  of loud operations increased above 3.

The noise parameters were calculated for the whole year with the exception for P04, P05 and P07. Massive air traffic passes over these points only for a few weeks a year. For example at the P04 the flight operations are very unevenly distributed over individual months. In particular, in April 2017 a very large number of flights occurred (Figure 2). At night there were 472 flights, what corresponds to an average of 15.7 flights during one night. In the whole year the average number of flights at night was 1.7, while during of three consecutive most unfavorable months, the average was 6.0 flights per night.

As the annual number of takeoffs and landings in 2017 over P04 exceeded 6,000, according to PN-B-02151-3:2015-10 calculations should include all days and nights of the year. This means taking into account 1.7 flights during the night, while the average

in April 2017 was nearly ten times higher. This also means not taking  $L_{Amax}$  into account, because the number of flights at night did not reach  $N = 3$ , and consequently, the lack of the protection against awakening caused by loud flight operations repeated over one night.

As a compromise for the P04, P05 and P07 monitors the noise parameters have been calculated taking into account three months when the largest air traffic is observed, like in the airports with the number of start - landing pairs up to 3000 a year according to PN-B-02151-3:2015-10. For such airports averaging the number of operations throughout the whole year for the all runways would lead to understating the noise rating in relation to the real noise annoyance in certain areas around the airport. The selection of the averaging interval should apply to individual runways, not to the entire airport.

The noise monitors are located at distances from 1 to 9 km from runway thresholds. As can be seen from Table 1 the largest noise parameters are at P01 located closest to the airport. As the distance from the airport increases,  $L_{AE}$  and  $L_{Amax}$  decrease. The spread among the noise monitors equals to 15 dB for  $L_{AE}$  and over 20 dB for  $L_{Amax}$ .

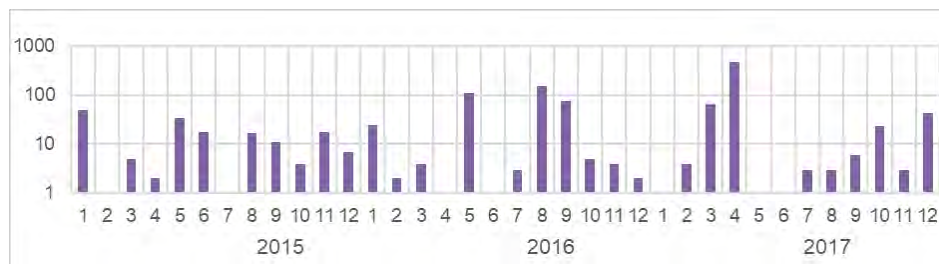


Figure 2. The nighttime flights number in months at monitor P04 - Onkologia

The average number  $N$  ranged from 1.3 to 18.4 during the 8 hours of the night. Close to the airport, the number of operations is greater because the individual inbound and outbound tracks are connected.

Between one quarter and over one third of all operations at night happens during the busiest half hour. Values of the coefficient  $p$  are the highest in the directions on which the number of flights is the smallest.

The mean value of the  $L_{Amax,L}$  exceeds 70 dB at all sites. The number of loud operations is at least 3, with the exception of P02, P07 and P10, which are outside the branching of outgoing tracks.

External noise indexes were calculated from the Table 1 data. The A-weighted equivalent sound level  $L_{Aeq, out, 8h}$  (according to PN-B-02151-3: 2015-10) was calculated from the formula

$$L_{Aeq, out, 8h} = L_{AE} + 10 \lg \frac{N}{8 \cdot 3600} \quad (5)$$

The A-weighted equivalent sound level for half an hour with the largest number of operations  $L_{Aeq, out, 1/2h}$  (according to PN-B-02151-02: 1987) is given by

$$L_{Aeq,out,\frac{1}{2}h} = L_{AE} + 10\lg \frac{pN}{0,5 \cdot 3600} \quad (6)$$

#### 4. Internal noise level

Noise requirements have been collected for rooms (bedrooms, great rooms, multi-purpose rooms and the like) in residential buildings at nighttime, resulting from Part 2 and Part 3 of PN-B-02151 in Table 2.

Table 2. Noise evaluation criteria in rooms in residential buildings at nighttime

Standard	Criterion
PN-B-02151-3: 2015-10 („Part 3”)	reference level $L_{Aeq,in,8h} = 25$
PN-B-02151-3: 2015-10 („Part 3”)	reference level $L_{Amax,in} = 50$ if $3 \leq N_L \leq 5$ reference level $L_{Amax,in} = 45$ if $N_L > 5$
PN-B-02151-02: 1987 („Part 2”)	permissible level $L_{Aeq,in,1/2h} = 30$

#### 5. Required attenuation of aircraft noise

The required attenuations  $D_{A,min}$  to fulfill the requirements from Table 2 are calculated below. It was assumed that aircraft noise is the dominant noise in the environment.

Using the formulas (5) and (6) and criteria of internal noise from Table 2, the required attenuation has been expressed as follows:

$$D_{Aeq,8h} = L_{AE} + 10\lg \frac{N}{28800} - L_{Aeq,in,8h} \quad (7)$$

$$D_{Amax} = L_{Amax,L} - L_{Amax,in}(N_L) \quad (8)$$

$$D_{Aeq,\frac{1}{2}h} = L_{AE} + 10\lg \frac{pN}{1800} - L_{Aeq,in,1/2h} \quad (9)$$

where:

$$\begin{array}{ll} \text{for } N_L < 3 & L_{Amax,in}(N_L) \text{ is not applicable (n/a)} \\ \text{for } 3 \leq N_L \leq 5 & L_{Amax,in}(N_L) = 50 \\ \text{for } N_L > 5 & L_{Amax,in}(N_L) = 45 \end{array}$$

The results of the required attenuation calculations for monitoring sites are given in Table 3 where also the differences in requirements are shown.

The typical value of noise attenuation by an external envelope for residential buildings is approximately 30 dB. As shown in Table 3, the required attenuation is about 30 dB at the monitoring points near the airport. In P01, it exceeds 40 dB.

It is seen that the equivalent sound level according to the Part 2 creates higher requirements than the long-term equivalent sound level according to the Part 3. The difference in the required attenuation is from 0.6 to 2.6 dB.

Table 3. Required attenuation of aircraft noise in buildings at monitoring sites

site	$D_{Aeq,8h}$ [dB]	$D_{Amax}$ [dB]	$D_{Aeq,1/2h}$ [dB]	$D_{Aeq,1/2h} - D_{Aeq,8h}$ [dB]	$D_{Amax} - D_{Aeq,1/2h}$ [dB]	$D_{Amax} - D_{Aeq,8h}$ [dB]
P01	36.7	42.4	38.3	1.6	4.1	5.7
P02	18.9	n/a	19.6	0.6	n/a	n/a
P03	25.8	29.6	26.6	0.8	3.0	3.8
P04	27.0	34.8	29.0	2.0	5.8	7.8
P05	22.3	27.1	24.9	2.6	2.2	4.8
P07	26.2	29.6	27.6	1.5	2.0	3.4
P08	20.3	n/a	21.4	1.1	n/a	n/a
P10	18.9	n/a	19.6	0.6	n/a	n/a

Requirements concerning the attenuation of external noise resulting from the Part 2 of PN-B-02151 and from the Part 3 can be compared analytically. The difference between the requirements given by (9) and (7) indicates which standard is more demanding:

$$\delta_{Aeq} = D_{Aeq, \frac{1}{2}h} - D_{Aeq, 8h} = 10 \lg(16p) - 5 = 10 \lg(p) + 7 \quad (10)$$

If a number of flights in each interval half-hour long is the same, then  $p = 1/16$ . In this case  $\delta_{Aeq} = -5$  dB. If all flights take place in only one half hour, then  $p = 1$  and  $\delta_{Aeq} = 7$  dB.

For  $p = 0.20$ , the requirements of the both standards are equal. For  $p$  like in Table 1 the requirements of the Part 2 are 0.6 to 2.6 dB higher than those of the Part 3.

When the number of loud events is minimum 3, the requirements concerning  $L_{Amax, out}$  are bigger than these concerning with  $L_{Aeq}$ . Values of  $D_{Amax}$  are higher by 2.0 to 5.8 dB than  $D_{A, \frac{1}{2}h}$  and by 3.4 to 7.8 dB higher than  $D_{A, 8h}$ .

For  $N > 5$  the difference  $D_{Amax} - D_{A, 8h}$  equals to

$$\delta_{A(max-eq)} = D_{Amax} - D_{Aeq, 8h} = L_{Amax, L} - L_{AE} - 10 \lg(N) + 24,6 \quad (11)$$

The difference  $\delta_{A(max-eq)}$  diminishes when number of flights grows. When  $N$  will increase to the maximum predicted number 40,  $\delta_{Amax/eq}$  can decrease to zero. In this case, the required attenuation connected with  $L_{Amax}$  will no longer matter.

## 6. Conclusions

The study compares the minimum attenuation of aircraft noise by the building envelope which results from the standards that currently apply [10]. The comparison was carried out for bedrooms, multi-purpose rooms and the like and concerned the nighttime. Data concern the locations of monitors belonging to the noise monitoring system of the Chopin Airport Warsaw

The minimum of noise attenuation by the building envelope, resulting from of PN-B-02151-02:1987, differs from the minimum according to PN-B-02151-3:2015-10.



The highest required attenuations occur in places where the average number of operations at night is not less than 3. In such case required attenuation is determined by the long-term maximum sound level  $L_{Amax}$ . The requirements related to  $L_{Amax}$  are 2.0 to 7.8 dB higher than those associated with  $L_{Aeq}$ . This is in line with the principle of preventing waking up during the night caused by noisy acoustic events [11].

If the number of aircraft operations increases to the maximum allowed number 40 during one night, the difference between required attenuation related to  $L_{Amax}$  and related to  $L_{Aeq}$  may decrease to zero. In this case, the requirements connected with  $L_{Amax}$  will no longer be relevant and the requirements related to  $L_{Aeq}$  will be decisive.

Between the requirements related to  $L_{Aeq}$  a bit higher requirements are set by PN-B-02151-02:1987 than by PN-B-02151-3:2015-10. With regard to the data examined, the replacement of PN-B-02151-02:1987 by PN-B-02151-3:2015-10 reduces the required attenuation by 0.6 to 2.6 dB. It should be remembered, however, that according to the Part 3, the equivalent sound level of aircraft noise is averaged over three months or a whole year, whereas the equivalent sound level according to the Part 2 may be referred to the busiest half-hour which happened only in one night. This is not a problem on tracks where the number of flights does not change significantly overnight. However, in the case of the tracks with an unequal distribution of the number of flights on a timeline, the problem arises as to whether the protection of buildings against noise occurring for several weeks a year should be as good as if the noise occurs every night. On the other hand, if aircraft noise occurs only for a month in a year, but its equivalent level is calculated in relation to the whole year, the average annual noise level will be lower by 10.8 dB. Lowering the required attenuation in buildings by such size will undoubtedly reduce the quality of rest at night.

However, it should also be noted that the requirements of PN-B-02151-3:2015-10 are only valid during the construction process. If transport noise will increase during the use of building, the lack of PN-B-02151-02:1987 will mean the lack of formal and legal possibility of assessment of this noise in the building. In particular, this problem will occur for buildings on limited-use areas, for which the requirements for protection against aviation noise are formulated as the permissible noise level in a room. This applies to almost all the biggest areas of limited use of civil and military airports in Poland. The only Polish Standard which gives the permissible noise level in rooms is PN-B-02151-02:1987. The alternative standard PN-B-02151-3:2015-10 is assigned only to the building process. Besides, it does not specify the permissible noise level in rooms, but gives so called reference levels, i.e. numerical values to be used in the calculation of the required acoustic insulation of partitions, but they cannot be used to assess the noise occurring in rooms.

## References

1. [https://www.eea.europa.eu/data-and-maps/daviz/number-of-people-exposed-to-6#tab-googlechartid\\_chart\\_21](https://www.eea.europa.eu/data-and-maps/daviz/number-of-people-exposed-to-6#tab-googlechartid_chart_21), (2019.06.05).
2. *Environmental Noise Guidelines for the European Region*, WHO Regional Office for Europe, ISBN 978 92 890 5356 3, World Health Organization 2018.

3. *Guidance on the Balanced Approach to Aircraft Noise Management*, Doc. 9829 AN/451, second edition, ICAO 2008.
4. *BAFU Progress report on aircraft noise abatement in Europe v3* Berlin, Maastricht, July 2015.
5. B. Berglund et al., *Guidelines for community noise*, World Health Organization (1999).
6. PN-B-02151-02:1987, *Protection against noise in rooms in buildings - Permissible values of sound level in rooms* [in Polish].
7. PN-B-02151-3:2015-10, *Building acoustics - Protection against noise in buildings - Part 3: Requirements for sound insulation of partitions in buildings and building elements* [in Polish].
8. <https://www.lotnisko-chopina.pl/pl/monitoring-halasu.html>, (2018.12.17).
9. K. Rudno-Rudziński, *Determination of the maximum noise level of F-16 aircraft based on the exposure levels* [in Polish], Progress in Acoustics, Polish Acoustical Society, Wrocław, 381-392 (2015).
10. *Regulation of the Minister of Infrastructure of 12 April 2002 on technical conditions to be met by buildings and their location* [in Polish], (Dziennik Ustaw 2002 nr 75, poz. 690).
11. Night Noise Guidelines for Europe, World Health Organization Europe 2009.

## Simulation of a Single Mode Wave Generation in Cylindrical Systems Applying Numerical Methods

Łukasz GORAŻD

*AGH University of Science and Technology, Faculty of Mechanical Engineering  
and Robotics, Krakow, Poland, lukasz.gorazd@agh.edu.pl*

Anna SNAKOWSKA

*AGH University of Science and Technology, Faculty of Mechanical Engineering  
and Robotics, Krakow, Poland, anna.snakowska@agh.edu.pl*

### Abstract

The paper presents simulations of a selected single mode generation in systems containing duct-like elements applying the finite element method (FEM). Simulations were carried out for values of the Helmholtz number exceeding the plane-wave propagation, *i.e.* for a multimode wave. The presented results constitute the first step to analyse propagation of the multimode wave through more complicated systems such as mufflers, even in a form of a cascade. Propagation of the incident wave in the form of a single mode greatly simplifies solutions of many problems, to mention only derivation of the transmission or the scattering matrices.

The results obtained can be applied to analyse the effectiveness of attenuation of acoustic silencers or elements of heating, ventilation and air conditioning systems (HVAC) at the design stage. Results of simulations are compared with already published experimental data obtained in a measurement set-up containing the self-designed single mode synthesizer.

**Keywords:** cylindrical duct, multi-mode wave propagation, single mode generation

### 1. Introduction

Acoustic waveguides are systems in the form of pipes in which an acoustic wave can propagate. They are used in heating, ventilation and air conditioning systems (HVAC) where the duct outlet is often a source of noise inside rooms or the outdoor environment [1, 2]. The waveguide theory is also applying in the analysis of noise coming from the jet engines placed in the cylindrical housing.

The purpose of this paper is numerical simulations related to the generation of a single selected mode inside a rigid cylindrical infinite waveguide using the finite element method (FEM). Simulations enable analysis of the acoustic phenomena occurring in the duct at the design stage without the necessity of construction time-consuming and expensive measuring setup.

When the frequency of the acoustic wave propagating in the duct exceeds the cut-off frequency of the Bessel mode (1,1), (indexing according to the Wejnshitejn monograph [3]) the sound pressure field is the result of the sum of component amplitudes of individual modes that can propagate without attenuation. Determination of the scattering matrix  $\mathbf{S}$  [4] for multimode wave in the case of acoustic muffler is associated with the determination of  $4N^2$  coefficients ( $N$  - is the number of propagating modes in inlet and outlet pipes [5]). In the case of a wave in the form of a single mode, only  $4N$  coefficients of the scattering

matrix must be determined [5]. The ability to generate a single mode greatly simplifies the analysis of the sound pressure field and its comparison with theoretical calculations or measurements.

The paper presents the first step, which is simulation of single mode generation in a cylindrical infinite waveguide with a constant cross-section. The required acoustic pressure field distribution was obtained by means of a set of point source models located on the cross-section of the waveguide. Verification of calculations was carried out using the previously published results for analogical acoustic measurements of real waveguide [6]. The results of the work can then be used to analyse more complex acoustic systems and elements of ventilation systems, such as acoustic mufflers.

## 2. Theory

Propagation of sound in the air is quite well described by the wave equation

$$\Delta p(\vec{r}, t) - \frac{1}{c^2} \frac{\partial^2 p(\vec{r}, t)}{\partial t^2} = 0 \quad (1)$$

where  $p$  is the acoustic pressure and  $c$  is the speed of sound.

The calculations in the frequency domain were performed using the Helmholtz equation being a Fourier transform from the wave equation.

$$\Delta p(\vec{r}) + k^2 p(\vec{r}) = 0 \quad (2)$$

where  $k = \frac{\omega}{c} = \frac{2\pi f}{c}$  is the wave number,  $\omega$  is the angular frequency,  $f$  is the frequency.

In theoretical considerations, taking into account the solution of equation (2) in the cylindrical coordinate system and the Neumann boundary condition, we obtain the expression for the sound pressure of a single mode

$$p_{mn}(\rho, \varphi, z) = A_{mn} e^{im\varphi} J_m\left(\frac{\mu_{mn}\rho}{a}\right) e^{i(\gamma_{mn}z - \omega t)} \quad (3)$$

where  $(\rho, \varphi, z)$  is the coordinates of the cylindrical system,  $a$  - waveguide radius,  $A_{mn}$  - amplitude of single mode sound pressure,  $J_m$  - Bessel functions of first kind,  $\mu_{mn}$  -  $n$ -th root of derivative of the Bessel function  $J'_m(\cdot)$ ,  $\gamma_{mn} = \sqrt{k^2 - \left(\frac{\mu_{mn}}{a}\right)^2} = \frac{1}{a} \sqrt{(ka)^2 - \mu_{mn}^2}$  - longitudinal wave number. In the case when a larger number of modes propagate within the waveguide, the total sound pressure of the waveform is a superposition of the allowed due to the parameter  $ka$ , modes

$$p(\rho, \varphi, z) = \sum_{m,n} p_{mn}(\rho, \varphi, z) \quad (4)$$

### 3. Waveguide geometry and simulation conditions

The simulations were carried out for a waveguide model with a radius of  $a = 103.25$  mm and 1.75 m length (Fig. 1). The geometrical dimensions are consistent with the real object on which previously published measurements were carried out. For a such waveguide geometry, the cut-off frequency of the first Bessel mode (1,1) is 976 Hz. That means that only a plane wave can propagate below this cut-off frequency. The simulations included in the article concern the generation of the first four modes, namely (0,0), (1,1), (2,1) and (0,1). The excitation frequency above the cut-off frequency of mode (0,1) for which the Helmholtz number is  $ka = 3.83$  is selected and is equal 2116 Hz. The mesh size of finite elements corresponds to 1/6 of the wavelength of the excitation frequency and is 0.02 m. For more accurate calculations, the element size was ten times smaller. A 3D mesh has been created from  $2.13173 \times 10^6$  tetrahedral elements included 372219 nodes for which calculations were carried out.

In order to fulfill the assumptions of the infinite waveguide on the surface of the outlet on both sides of the waveguide, the boundary condition was set as totally absorbing acoustic waves, creating so-called anechoic terminations. The surface of the waveguide was treated as rigid, therefore the Neumann boundary condition is associated with this surface. This means that the normal component of the acoustic velocity on the surface of the waveguide is zero  $v_n|_{\Sigma} = 0$ .

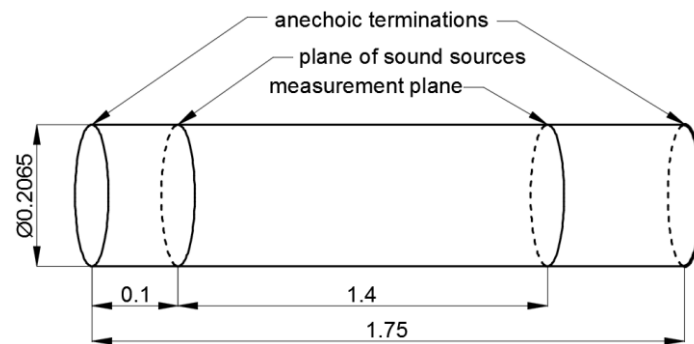


Figure 1. Geometry of the waveguide (dimensions in [m])

The location of the point sound sources in the FEM model corresponded to the distribution of sources as in the measurement set-up. The sources were located on a constant cross-section at a distance of 0.1 m from the anechoic termination on an equilateral triangle plan. Using the Green function and the shape function of a individual mode (0,0), (1,1), (2,1), (0,1), complex amplitudes of each sound sources were determined. The location of the sound sources is shown in Figure 2. Table 1 presents the determined amplitudes of sources for the generation of acoustic wave in the form of a single selected mode.

The distribution of sound pressure field was analysed on a cross-section at a distance of 1.4 m from the plane of the point sound sources (Fig. 2). For practical reasons, sound pressure values (in Pa) are presented in the form of sound pressure levels (SPL) in dB.

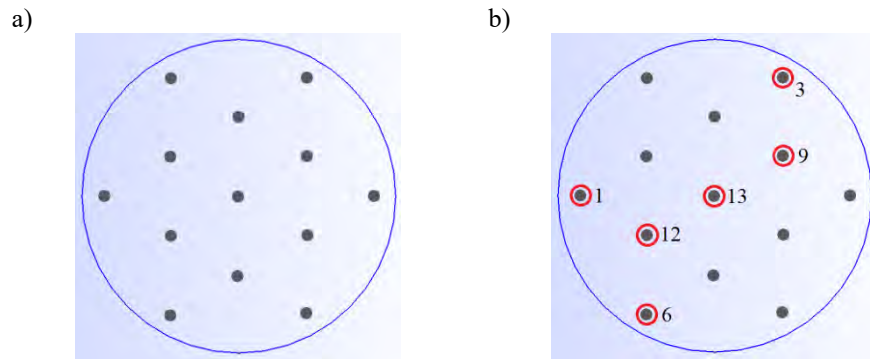


Figure 2. Location of point sound sources on the cross section of the duct; a) all sources, b) sources selected for the generation of single modes (marked with a red envelope)

Table 1. Amplitudes of individual sound sources located on the cross-section inside the waveguide

Point	amplitude of the modes [Pa]			
	mode (0,0)	mode (1,1)	mode (2,1)	mode (0,1)
1	$0.46+0i$	$0.00+0i$	$0.05+0i$	$-0.46+0i$
6	$0.46+0i$	$-1.00+0i$	$0.51+0i$	$-0.46+0i$
12	$-1.00+0i$	$0.80+0i$	$-1.00+0i$	$1.00+0i$
13	$0.57+0i$	$0.00+0i$	$0.87+0i$	$0.03+0i$
9	$0.10+0i$	$-0.80+0i$	$-0.89+0i$	$-0.10+0i$
3	$0.00+0i$	$1.00+0i$	$0.46+0i$	$0.00+0i$

The grid of points in which the sound pressure level was determined was obtained by rotating points distributed along the waveguide radius with a step of 5mm in the range of full angle 360 with a step equal to 15. The total number of calculation points is 580 (Fig. 3).

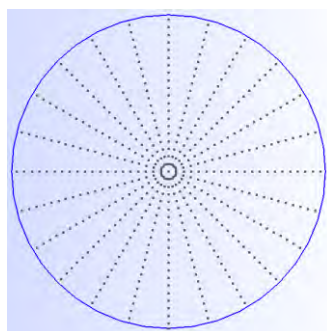


Figure 3. Arrangement of calculation points on the cross-section of the waveguide at distance 1.4 meters from the plane of the sound sources

#### 4. Results and their interpretation

For the above-described rigid model of the infinite cylindrical waveguide, calculations of sound pressure level in points were carried out. It can be seen that only in the case of mode (2,1) all six point sound sources are involved in the generation of acoustic single-mode wave. In other cases, all sources are not needed. The results obtained are presented below (Fig. 4).

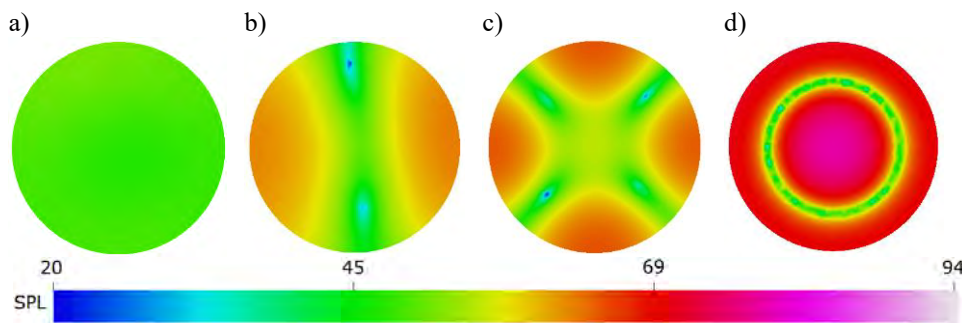


Figure 4. results of simulation single-mode generation using FEM: a) mode (0,0), b) mode (1,1), c) mode (2,1), d) mode (0,1)

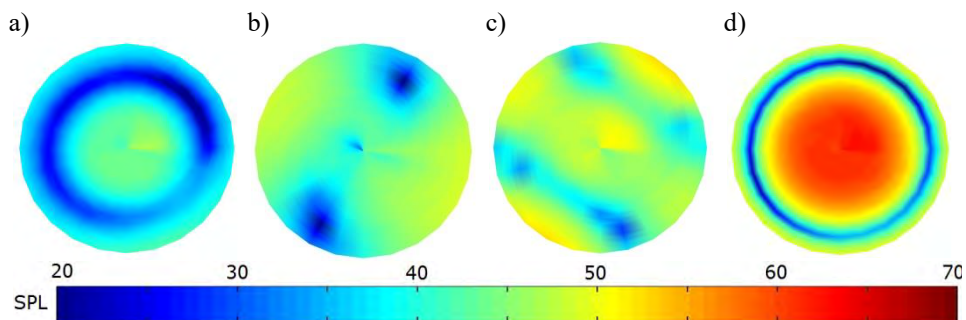


Figure 5. Results of single mode generation measurements using a rotary matrix of sound sources [6]: a) mode (0,0), b) mode (1,1), c) mode (2,1), d) mode (0,1)

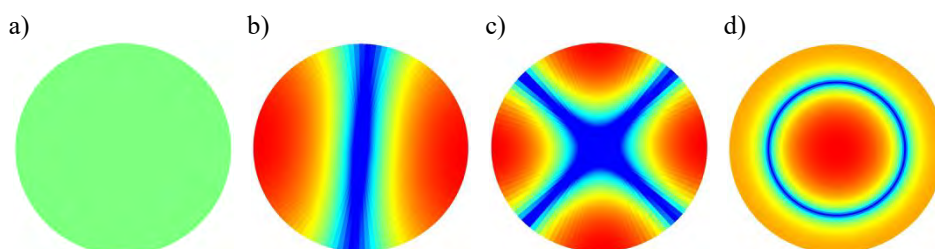


Figure 6. Theoretical shape functions of modes: a) mode (0,0), b) mode (1,1), c) mode (2,1), d) mode (0,1)

In the figures 4 – 6, it can be seen that the shape of the sound pressure level distribution obtained by the FEM calculations (Fig. 4) shows similarity to the theoretical shape functions (Fig. 6). Figure 5 shows the measurements results on a logarithmic scale. Fig. 5a) shows a significant contribution of mode (0,1), so it can be concluded that for  $ka = 4$  it was not possible to generate a plane wave without the participation of the other permitted modes.

## 5. Conclusions

The presented simulation results of the single mode generation show that the shape of the sound pressure level distribution obtained by the FEM method is similar to the shape of the modes obtained by means of measurements. Even better similarity is visible comparing simulations with theoretical calculations. For quantitative analysis, decomposed of the modes should be carried out, *i.e.* their complex amplitudes should be determined applied the Fourier-Lommel [7] transform and the properties of the shape function. Finally, it can be concluded that these preliminary FEM calculations are accurate enough and the results can be used to analyse more complex acoustic systems such as silencers or flow regulators.

## Acknowledgments

The study reported in this paper was financed by the Dean's of the Faculty of Mechanical Engineering and Robotics AGH University of Science and Technology – research project number 16.16.130.942.

## References

1. P. Joseph, P. A. Nelson, M. A. Fisher, *Active control of fan tones radiated from turbofan engines. I. External error sensors*, J. Acoust. Soc. Am., **106** (1999) 766 – 778.
2. A. Snakowska, J. Jurkiewicz, *Efficiency of energy radiation from an unflanged cylindrical duct in case of multimode excitation*, Acta Acustica united with Acustica, **96** (2010) 416 – 424.
3. L. A. Vainshtein, *The theory of diffraction and the factorization method*, Golem Press, 1969.
4. A. Sittel, J. M. Ville, F. Foucart, *Multiloading procedure to measure the acoustic scattering matrix of a duct discontinuity for higher order mode propagation conditions*, J. Acoust. Soc. Am., **120** (2006) 2478 – 2490.
5. A. Snakowska, K. Kolber, Ł. Gorazd, J. Jurkiewicz, *Derivation of an acoustic two-port scattering matrix for a multimode wave applying the single-mode generator*, Acoustics 2018 proceedings of joint conference, IEEE (2018), 294 – 298.
6. A. Snakowska, Ł. Gorazd, J. Jurkiewicz, K. Kolber, *Generation of a single cylindrical duct mode using a mode synthesizer*, Applied Acoustics, **114** (2016) 56 – 70.
7. J. M. Auger, J. M. Ville, *Measurement of linear impedance based on the determination of duct eigenvalues by a Fourier-Lommel's transform*. J. Acoust. Soc. Am., **88**(1) (1990) 19 – 22.



## Animal Mimicry for Covert Communication with Arbitrary Output Distribution: Beyond the Assumption of Ignorance

Krzysztof Władysław ZUBER

*Wrocław University of Science and Technology, Wybrzeże Wyspiańskiego 27,  
50-370 Wrocław, Poland, krzysztof.zuber@pwr.edu.pl*

Krzysztof J. OPIELIŃSKI

*Wrocław University of Science and Technology, Wybrzeże Wyspiańskiego 27,  
50-370 Wrocław, Poland, krzysztof.opielinski@pwr.edu.pl*

### Abstract

The paper describes a new method of embedding human communication in acoustic sequences mimicking animal communication. This is done to ensure a low probability of detection (LPD) transfer of covert messages. The proposed scheme mimics not only individual sounds, but also the imitated species' communication structure. This paper presents a step forward in animal communication mimicry – from pure vocal imitation without regard for the plausibility of communication's structure, through Zipf's law-preserving scheme, to the mimicry of a known communication structure. Unlike previous methods, the updated scheme does not rely on third parties' ignorance of the imitated species' communication structure beyond Zipf's law – instead, the new method enables one to encode information in a known zeroth-order Markov model. The paper describes a method of encoding an arbitrary message in a syntactically plausible, species-specific sequence of animal sounds through evolutionary means. A comparison with the previous iteration of the method is also presented.

**Keywords:** animal mimicry, covert communication, hidden Markov model

### 1. Introduction

Successful transmission of covert messages depends not only upon an appropriately robust coding method, but also upon decreasing the probability of their interception – for instance, in the case of underwater communication, the mere fact of a submarine sending a message can lead to its detection, regardless of whether the message is decoded or not. To limit the probability of detection of message transmission, a nature-inspired method of mimicry has been proposed [1]. This method relies on sending the hidden message as a string of an animal species' vocal signals.

However, even using animal sounds to transmit a message is not sufficient for a successful mimicry. This is because animals in general, and especially aquatic mammals (such as dolphins and whales), use complex structures in their communication [2,3,4]. In fact, any intelligent communication tends to distribute its units according to Zipf's law [3,5,6] which is not evident in superficial animal mimicry schemes. Thus, even a simple, automatic rank-frequency plot of received vocals will invalidate these schemes.

To address this problem, a mimicry method based upon generating a syntax for encoding covert messages, which gives the output units a Zipfian distribution while maximising both the information throughput and output distinctness was devised and described [7]. However, even that method has serious limitations – the most important

being that it pre-supposes that third parties are ignorant of the mimicked species' communication structure beyond the very general fact that it obeys the Zipf's law.

Therefore, a new method of mimicry was designed with the aim of encoding and decoding arbitrary messages composed of symbols as sequences of animal vocal units with a specified distribution. The method relies on constructing hidden Markov models. The scheme has been tested and compared to the alternative, Zipfian method, and proved to be the better option when it comes to successful mimicry, while having comparable information throughput rate.

## 2. Hidden Markov Model

To achieve the desired operation of encoding – namely, the ability to encode data expressed as **symbols** with arbitrary frequencies as animal species' vocal units with a specified probability distribution (for instance, the probability distribution of syllables in birdsong [8] or of temporal spacing between dolphin pulses [9]) – the hidden Markov model (HMM) approach was chosen.

Hidden Markov models rely fundamentally on transitions within hidden, unobservable states, which can probabilistically generate **observations** or **outputs** [10]. Transitions between hidden states are described by the stochastic matrix  $A$  (called the **state transition matrix**), and observation generation is described by matrix  $B$ , called the **output(emission) matrix**, where  $A_{ij} = P(s_{t+1} = x_j | s_t = x_i)$ ,  $B_{ij} = P(o_t = y_j | s_t = x_i)$ ,  $X = \{x_1, x_2, \dots, x_N\}$  are the  $N$  possible hidden states (**state alphabet**),  $Y = \{y_1, y_2, \dots, y_M\}$  are the  $M$  possible observations (**output alphabet**),  $S = \{s_1, s_2, \dots, s_T\}$ ,  $S \in X$  is the sequence of hidden states at step  $t$  out of  $T$  (**state sequence**), and  $O = \{o_1, o_2, \dots, o_T\}$ ,  $o \in Y$  is the sequence of observations in  $T$  steps (**observed sequence**) ( $N, M, T \in \mathbb{N}^+$ ). In addition, the probability of the system being initially in any given hidden state (**initial state distribution**) is represented by  $\pi = \{\pi_1, \pi_2, \dots, \pi_N\}$ ,  $\sum_{m=1}^N \pi_m = 1$ .

For the purpose of encoding, the hidden states are assumed to correspond to the characters of the message to be sent, and the output states – the emitted vocal units (which, it must be remembered, do not necessarily correspond to vocal syllables, but can also represent separations between units in communication dependent on timing or the number of repetitions in communication dependent on unit repetition [2]).

For successful mimicry, the probability distribution of output units must follow the probability distribution of units in the mimicked species' vocal communication,  $\forall t P(o_t = y_i) = F(i)$ ,  $F(i) \in [0, 1]$ ,  $\sum_{i=1}^M F(i) = 1$ . Similarly, for the input symbols,  $\forall t P(s_t = x_i) = G(i)$ ,  $G(i) \in [0, 1]$ ,  $\sum_{i=1}^N G(i) = 1$ . Thus, it is assumed that both symbols' and observations' probabilities are independent.

A trivial solution of the problem of constructing  $A$ ,  $B$  and  $\pi$  would be to set them to  $A_{ij} = G(j)$ ,  $B_{kl} = F(l)$ ,  $\pi = v$ , where  $v$  is the left eigenvector of  $A$  with the eigenvalue of 1 (i.e. the stationary distribution of the Markov chain described by matrix  $A$ ), which ensures

that the system starts at steady state and does not need time to settle for the desired output distribution.

However, doing so, while ensuring the proper input and output distributions, prevents any encoded message from being decoded with any certainty. To understand why this is the case, and devise steps to be taken to remedy the problem, it is illustrative to consider the decoding algorithm used most broadly for decoding HMMs – the Viterbi algorithm.

### 3. Viterbi algorithm

Viterbi algorithm for decoding HMMs (i.e. establishing the most probable sequence of hidden states,  $\hat{S} = (\hat{s}_1, \hat{s}_2, \dots, \hat{s}_t)$  from a sequence of observations  $O$ ) relies on recursively identifying the most probable state at every step given the most likely hidden state at the previous step (this entire exposition is based on [10]).

The algorithm's operation relies on calculating two tables:

$$C_1(j, t) = \max_n (C_1(n, t-1) \cdot A_{nj} \cdot B_{j o_t}) \quad (1)$$

which describes, given previous observations, the highest probability of the hidden state transitioning to  $x_j$  and emitting observation  $o_t$ , and

$$C_2(j, t) = \arg \max_n (C_1(n, t-1) \cdot A_{nj}) \quad (2)$$

which describes the most likely (given previous observations) state transitioned from ( $\hat{s}_{t-1}$ ) at step  $t$  given  $\hat{s}_t = x_j$ . The algorithm is initialised with

$$C_1(j, 1) = \pi_j \cdot B_{j o_1}, C_2(j, 1) = 0 \quad (3)$$

and run until the tables  $C_1$  and  $C_2$  are filled. Then, the hidden state corresponding to the greatest element in  $C_1(j, T)$  is identified, so

$$\hat{s}_t = \arg \max_j (C_1(j, T)) \quad (4)$$

is the most likely ending state given all the observations, and the most likely previous hidden states are recursively chosen as  $\hat{s}_{t-1} = C_2(\hat{s}_t, t)$ .

Equations (1), (2), (4) show the conditions of reliable functioning of the algorithm (from (2)):

$$\forall t \in (1 \dots T), j \in (1 \dots N) \exists! \arg \max_n (C_1(n, t-1) \cdot A_{nj}) \quad (5)$$

as multiple maxima in any column of  $C_1(n, t)$  make decoding ambiguous, because the state sequence cannot be chosen uniquely, and, from (4):

$$\forall t \in (1 \dots T), j \in (1 \dots N) \exists! \arg \max_j (C_1(j, t)), \quad (6)$$

which is a stronger version of the same requirement, but for  $C_1(j, T)$  dictated by the necessity of being able to unambiguously establish the state most likely responsible

for the final output. As we want to be able to decode sequences of arbitrary length, this requires us being able to terminate decoding at any step.

Let us consider why the trivial construction of  $A$  and  $B$  guarantees the failure of decoding. Substituting into equations (1), (2), (3) we obtain:

$$\text{a) } C_1(j,1) = G(j) \cdot F(o_1),$$

$$\text{b) } C_1(j,2) = \max_n (G(n) \cdot F(o_1) \cdot G(j) \cdot F(o_2)) = \max_n G(n) \cdot (F(o_1) \cdot G(j) \cdot F(o_2)),$$

so, by induction, every column of  $C_1$  will be a scalar multiple of  $G$ . If there are multiple maxima in  $G$ , the algorithm fails at this stage by condition (5),

$$\text{c) } C_2(j,2) = \arg \max_n (G(n) \cdot F(o_1) \cdot G(j)) = \arg \max_n (G(n)) \text{ – the outputs are always}$$

decoded as originating from the state with the highest probability.

Thus, we can describe another mode of failure, where regardless of observed outputs, the predicted state does not change. This is brought about by the outputs being independent of states and consecutive states being independent of each other, as is the case if the matrices are constructed in a trivial way.

In addition to the absolute proscriptions on the construction of the HMM described above, there are also several risk factors increasing the possibility of the derailment of the method: (a) multiple hidden states sharing the same steady-state probability (i.e. multiple identical elements of  $\pi$ ), especially combined with other risk factors, (b) multiple identical large elements in columns of  $A$  and  $B$  increase the probability of multiple maxima in columns of  $C_1(j,t)$ , (c) multiple identical large elements in rows of  $A$  may lead to multiple maxima in  $C_1(j,t)$  and failure as described by equation (6).

In addition, it must be borne in mind that the operation of the method relies on successive multiplications of probabilities, so the elements of  $C_1$  will rapidly decrease, thus the possibility of two elements being considered equal due to computation precision when it is not the case increases with the length of the observation sequence. Therefore, for practical purposes, every instance of the word “identical” in this section can be replaced by “similar”.

Using these findings, the method of HMM construction was developed.

#### 4. Constructing the HMM

First, one has to consider the message to be encoded. If each symbol has a distinct probability (for instance, if the symbols to be transmitted are letters in a natural language) the situation is simpler, as we only have to deal with repeated elements in columns of  $A$ . However, if the symbols have equal probabilities (for instance, if they represent bits), the problem of repeated elements in rows of  $A$  arises.

The number of hidden states can be increased by splitting one symbol into its combinations – for instance, introducing 01 and 00 instead of 0 when encoding digital message. This approach ensures the preservation of the capacity of the HMM to encode any signal, and can split the large maxima in  $A$  and increase symbols throughput. However, there is a limit on the relative number of hidden states and emitted states, which will be discussed later.

The solution of this problem is pre-seeding the message to be encoded with blank characters, with the probability of inserting the blank dependent on the current and next symbol – this can be represented by the matrix  $D_{ij}, i, j = 1 \dots N$ , which represents the probability of inserting a blank signal if the current symbol is  $x_i$  and the next one is  $x_j$ . Then, the size of the original  $A$  is increased by one row and one column to accommodate the new blank symbol, and the new transition matrix  $A^*$  is constructed:

$$A_{ij}^* = \begin{cases} A_{ij} \left(1 - D_{ij}\right), j = 1 \dots N, i = 1 \dots N \\ G(i) \sum_j A_{ij} D_{ij}, j = N+1, i = 1 \dots N \\ \left( \sum_i G(j) G(i) D_{ij} \right) / \left( \sum_i \left( G(i) \sum_j A_{ij} D_{ij} \right) \right), j = 1 \dots N-1, i = N \end{cases} \quad (7)$$

with the remaining element  $A_{N+1, N+1}^*$  equal to 0. The elements of  $D$  are selected to ensure acceptable distinctness of transitions of hidden states by Matlab® optimisation packages (under requirement of each row and column composed of distinct elements).

The emission matrix, just like the transition matrix, must be reconstructed from the trivial construction. Here, the distinctness of elements in columns is paramount in the current iteration of the method – the distinctiveness within rows would be important if we relaxed the assumption of perfect reception, or if we wanted to use the output from this stage as an input to another matrix.

The construction of the emission matrix also requires a software optimiser – for this paper, Matlab® optimisation package was used. The problem is set up thus: (a)  $\sum_l b_{kl} = 1$  (emission at each hidden state), (b)  $\sum_k \pi_k b_{kl} = F(l)$  (desired observation distribution),  $\max_l \sum_k b_{kl} \cdot \log_M b_{kl} = \min$  (maximum distinctness of elements in columns), (c) each column composed of unique elements, and optimised using Matlab® optimisers. This problem involves  $N \cdot M$  variables (the size of  $B$ ), and (excluding pathological cases)  $N + M - 1$  independent equations ( $N - 1$  from condition a),  $M$  from condition b)). Therefore, for the problem to have a non-trivial solution,  $NM > N + M - 1$ , or  $NM \geq N + M$ . This equation describes the practical limit of splitting hidden states and addition of synonymous states. Naturally, the greater the degree of inequality, the more possibilities of optimisation exist, and the more the matrix elements can be tweaked. The initial distribution is selected as the left eigenvalue of  $A^*$ . It is assumed that we can perfectly distinguish between observed states.

## 5. Results

For the sake of comparison with a mimicry encoding method which relies on encoding messages in permutations of animal signals to give the output a Zipfian distribution [7], the hypothetical scenario of encoding 3 binary syllables (1, 01, 00 – 1.5 bits) in six vocal

units (referred to as **1 2 3 4 5 6**) was explored. The frequency of units returned by the Zipfian method is presented in Figure 1 as compared to a purely Zipfian distribution.

Next, a HMM was constructed to mimic the same scenario: 3 binary syllables and the filler syllable (last row and column of  $A^*$ ), and an exactly Zipfian distribution of six output units (frequencies of 0.408, 0.204, 0.136, 0.102, 0.0816, 0.068 – squares in Figure 1). The model definition was then selected:

$$D = \begin{bmatrix} 0.9963 & 0.1814 & 0.3918 \\ 0.0442 & 0.9970 & 0.1099 \\ 0.0112 & 0.0868 & 0.9939 \end{bmatrix} \quad A^* = \begin{bmatrix} 0.0018 & 0.2046 & 0.1521 & 0.6415 \\ 0.4779 & 0.0008 & 0.2225 & 0.2988 \\ 0.4944 & 0.2283 & 0.0015 & 0.2758 \\ 0.5513 & 0.1947 & 0.2540 & 0 \end{bmatrix} \quad (8)$$

$$B = \begin{bmatrix} 0.0004 & 0.5406 & 0.0970 & 0.2805 & 0.0807 \\ 0.5497 & 0.1061 & 0.0040 & 0.0145 & 0.0015 \\ 0.0027 & 0.0018 & 0.5952 & 0.0172 & 0.2316 \\ 0.9893 & 0.0034 & 0.0021 & 0.0026 & 0.0020 \end{bmatrix} \quad \pi = \begin{bmatrix} 0.3414 \\ 0.1707 \\ 0.1707 \\ 0.3171 \end{bmatrix}$$

Afterwards, tests of reliability of encoding and decoding were undertaken – observation sequences between 2 and 2000 characters long were generated and decoded, with their error rate (frequency of errors in state decoding) averaged over 20 runs for each sequence length. The results are presented in Figure 2. In addition, a test of decoding a 2000-observation sequence split into blocks of length 10 is presented in the same figure.

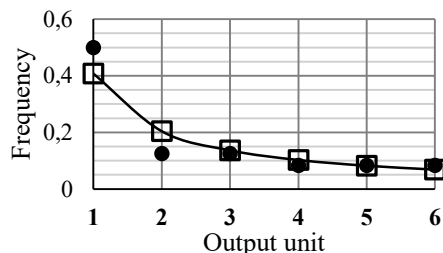


Figure 1. Zipfian method output unit frequency (black circles) and Zipfian distribution for six elements (black squares)

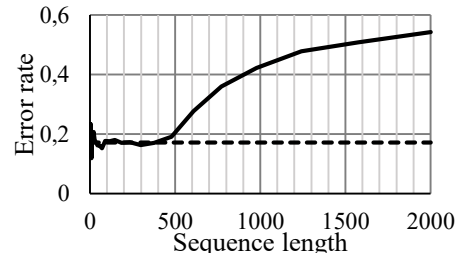


Figure 2. Decoding error rate for HMM for varying sequence length (solid line) and error rate for decoding a long sequence in blocks of 10 units (dashed line)

## 6. Discussion

Based upon the results (Figure 2), one notices that the HMM method is inherently error-prone – we can see two regimes of its operation: for a range of decoded observation sequence lengths, the error rate is stable and relatively low, transitioning to a much higher value for longer sequences.

The fact that any error can be explained thus: the Viterbi algorithm gives the most probable sequence of hidden states, but sometimes the improbable transition will happen,

throwing off the decoding process. This risk is not present in the Zipfian mimicry method, as each vocal unit permutation can only correspond to one message block.

The jump in error rate is due to one of the problems described in section 0 – namely, the successive multiplication of probabilities leading to the method not being able to establish unique maxima at consecutive steps of calculations, eventually decoding every observation as emitted from the most likely hidden state. However, this behaviour of the method does not preclude it from being used to form and decode long chains of units, such as those present in dolphin bursts [9], since the observation sequence can always be decoded in short blocks, which guarantee operating in the nominal, low-error rate value. This is because the initial state distribution is defined as the steady-state distribution of states, which eliminates any dependence of the decoded message on the chosen exact starting point. This also means that a portion of the message will always be accurately decoded even if not all vocal units are received.

Both errors can be decreased by constructing the HMM in a way that maximises the distinctness of state transitions – at the moment, this is done intuitively, a more in-depth mathematical treatment of the problem would be highly beneficial. However, there is a practical limit to this improvement, as eventually, for long sequences, computational precision will make decoding impossible – even a relatively large probability of 0.5 taken to the 500<sup>th</sup> power (a qualitative analogue of decoding the 500<sup>th</sup> observation) is of the order of  $3 \cdot 10^{-151}$  – well below the computational precision of Matlab® (32 digits by default).

The quality of mimicry described by both methods is very different. The Zipfian method can only approximate one (Zipfian) distribution, always trading the number of redundant vocal units for the closeness of the distribution to the desired one. In addition, in the Zipfian method, the individual meaning-carrying permutations of units must be separated either by a pause or by a non-coding vocal unit, thus either introducing a hard limit on sequence length or radically reducing the information throughput (the separator unit is guaranteed to be the one with the greatest frequency, as it appears in every permutation). In contrast, the HMM method can operate on arbitrary, even non-Zipfian vocal unit distributions, and can reliably generate a vocal sequence of arbitrary length. While it is true that a some symbols will be decoded incorrectly, this problem can be alleviated by using error-correcting codes (such as convolutional codes) to pre-code the message.

The pre-seeding of the message with blanks, while necessary for improving the robustness of encoding and decoding, decreases the bitrate (in the simulated HMM – to 1.02 bits/vocal unit from 1.5 in the Zipfian method). It would be beneficial to investigate the exact trade-off between bitrate and distinctness of the transition matrix further and introduce a measure of weighing bitrate against distinctness, as was done in [7]. It must also be borne in mind that the scenario of encoding bits was chosen in this paper precisely because it tends to be challenging due to multiple hidden states appearing with the same probability – the method is general enough to encode other signals, e.g. letters of alphabet or Morse code, which are far less challenging due to the each signal, having a different probability of occurring in most cases.

The question of whether transition and (especially) emission matrices meeting the criteria of uniqueness of elements can be constructed for a general distribution of message symbols and output vocal units is still unanswered, and would benefit from

further mathematical study – to be frank, we know how to make the method not fail immediately and intuit how to make it robust, but there is no mathematical clarity as to what optimisation strategy would be the best or even if optimisation will always succeed.

Finally, the HMM approach to mimicry looks like a promising avenue of research, and it appears very likely that further stacking of HMMs will enable one to simulate not only desired vocal unit distributions, but also entire conditional structures of animal species' communication. Also, the possibility of combining the Zipfian method's approach of introducing synonyms will be investigated, as will the relaxation of the assumption of perfect distinction between vocal observations/emissions.

## 7. Conclusions

The HMM-based method of mimicry for covert communication can reliably encode and decode messages while mimicking the desired vocal unit frequency distribution.

The HMM-based method of mimicry can operate reliably even when decoding long or incomplete chains of observations.

## References

1. S. Liu, G. Qiao, A. Ismail, *Covert underwater acoustic communication using dolphin sounds*, J Acoust Soc Am, **133**(4) (2013) EL300 – EL306.
2. A. Kershenbaum et al., *Acoustic sequences in non-human animals: a tutorial review and prospectus*, Biol Rev, **91**(1) (2016) 13 – 52.
3. B. McCowan, S. F. Hanser, L. R. Doyle, *Quantitative tools for comparing animal communication systems: information theory applied to bottlenose dolphin whistle repertoires*, Anim Behav, **57** (1999) 409 – 419.
4. R. Ferrer-i-Cancho, B. McCowan, *A law of word meaning in dolphin whistle types*, Entropy, **11** (2009) 688 – 701.
5. L. R. Doyle. *Animal Communications, Information Theory, and the Search for Extraterrestrial Intelligence (SETI)*. [Online]. <https://www.seti.org/seti-institute/animal-communication-information-theory-and-seti>.
6. J. Kanwal, et al., *Zipf's law of abbreviation and the principle of least effort language users optimise a miniature lexicon for efficient communication*, Cognition, **165** (2017) 45 – 52.
7. K. W. Zuber, K. J. Opieliński, *Animal mimicry in covert underwater communication application of syntax generation and simulated genome method*, in 2018 Joint Conference - Acoustics, Ustka, 2018, 1 – 5.
8. K. Katahira et al., *Complex sequencing rules of birdsong can be explained by simple hidden Markov processes*, PLoS ONE, **6**(9) (2011) 1 – 9.
9. A. R. Luis, M. N. Couchinho, M. E. dos Santos, *A quantitative analysis of pulsed signals emitted by wild bottlenose dolphins*, PLoS one, **11**(7) (2016) 1 – 11.
10. L. R. Rabiner, *A tutorial on hidden Markov models and selected applications in speech recognition*, Proc. of the IEEE, **77**(2) (1989) 257 – 286.



## Estimating the Population Exposed to Transportation Noise: a Case Study on Poznań City

Anna PREIS

*Institute of Acoustics, A. Mickiewicz University, Poznań, Poland,  
apraton@amu.edu.pl*

Jan FELCYN

*Institute of Acoustics, A. Mickiewicz University, Poznań, Poland,  
janaku@amu.edu.pl*

Piotr KOKOWSKI

*Akustix sp. z o.o., Przeźmierowo, Poland*

Michał GAŁUSZKA

*Akustix sp. z o.o., Przeźmierowo, Poland*

Paweł LIBISZEWSKI

*Akustix sp. z o.o., Przeźmierowo, Poland*

### Abstract

European Union Directive 2002/49/EC relating to the assessment and management of environmental noise (named as END) in article 6 paragraph 3 states that “harmful effects may be assessed by means of dose-effect relations referred to in ANNEX III”. In this still unfinished ANNEX III there are formulas which present how to calculate the number of people affected by a given noise. The dose-effect relations have been recently presented in WHO document “Environmental Noise Guidelines for the European Region”. These Guidelines allow to predict the percentage of people who will be affected by a specific kind of noise. E.g. it is shown how to calculate the percentage of highly annoyed people for a given value of noise index,  $L_{den}$ . In our paper we propose how to calculate the total number of people affected by noise in their living conditions and discuss the implementation of methods recommended in ANNEX III in Poland.

**Keywords:** ANNEX III, number of people affected by noise

### 1. Introduction

The main purpose of the Environmental Noise Guidelines for the European Region [1] is to provide recommendations for protecting human health from exposure to environmental noise originating from various noise sources. EU Member States agreed that recommendations from these guidelines should be observed in constructing the revision of ANNEX III of the Environmental Noise Directive (END). The recommendations introduced in the working version of ANNEX III [2] refer only to noise produced by three kinds of transportation sources: road traffic, railway and aircraft. The recommended  $L_{den}$  values for these three kinds of environmental noise sources could be a preventative measure for the occurrence of four specific types of harmful health effects associated with high noise levels: ischaemic heart disease, stroke, high annoyance and high sleep

disturbance. Such harmful effects may be assessed by means of dose-effect relations presented in WHO document, as well as, in the ANNEX III (Equations (3) to (20)).

The basic aim of ANNEX III is to define the best assessment method for calculating the number of people affected by noise produced by various transportation noise sources. The proposed methods are quantitative and are designed to calculate separately the number of inhabitants exposed to each of the three kinds of transportation noise. These calculated values can later be ranked according to the relative importance of a particular noise source.

In sum, ANNEX III contains methods allowing to calculate the number of people affected by three kinds of noise sources with respect to four types of harmful effects. However, the guidelines neglect the data informing about the population structure of the considered urban area. Therefore it is not possible to calculate the relationship between the number of affected people and their living conditions. For example, the regulations in ANNEX III do not allow to relate a given place in a city to highly annoyed people exposed to road traffic. Such data are important in constructing an efficient action plan which recommends means to reduce noise in a given area in a city under consideration. The higher the number of highly annoyed people in a given area the more vital is the action towards reduction of noise. Since ANNEX III is still unfinished it should be decided if it should focus on a general or a more specific solution. The first, general approach will consist in proposing methods allowing to calculate the total number of highly annoyed people in a whole city. In such a case, one has to answer the question how these calculations can be used in constructing an efficient action plan? The second, specific approach will demand much more work, but would be easier to implement.

## 2. Methodological problems in a working version of the Annex III document

Before ANNEX III will be established as a law in noise management some important methodological problems should be addressed. We will discuss one of them.

The document assumes that harmful effects of noise causing an increase in the risk of ischaemic heart disease (*IHD*), as well as an increase in the risk of stroke (*STR*), can be quantified by the relative risk. This assumption can be presented in the form of the following equation:

$$RR = \frac{P_{\text{event when exposed}}}{P_{\text{event when not exposed}}} \quad (1)$$

where  $RR$  is the probability of an event occurring (e.g. developing a disease) in the exposed group to the probability of the event occurring in the non-exposed group. The  $RR$  values are calculated for each kind of noise source separately.

The calculations of  $RR$  for a given harmful effect and noise source require insertion in the equation of (1) the specific values of  $L_{den}$  at which risk of *IHD* or *STR* starts. An example of such specified equation for the incidence rate of the *IHD* for road traffic is:

$$RR_{IHD,i,road} = \begin{cases} \exp((\ln 1.08/10) \cdot (L_{den} - 53)) & \text{for } L_{den} \text{ greater than 53 dB} \\ 1 & \text{for } L_{den} \text{ equal or smaller than 53 dB} \end{cases} \quad (2)$$

Values for other kinds of noise sources and harmful effects could be calculated without problems. The problem arises when we want to calculate the number of people affected by given harmful effect caused by particular noise source -  $N_{x,y}$

$$N_{x,y} = PAF_{x,y,i} \cdot L_y \cdot P \quad (3)$$

$$N_{x,y} = PAF_{x,y,i} \cdot M_y \cdot P \quad (4)$$

where  $PAF_{x,y}$  is the population attributable fraction,  $P$  is the total population of the area under assessment (the sum of the population in the different noise bands). To perform these calculations we need data of incidence rate,  $L_y$  and mortality rate,  $M_y$ . These data should be obtained from statistics on health region or country where the area is located. However, in some countries (e.g. in Poland) such data are difficult or even impossible to obtain.

This problem does not occur, when two other harmful effects: high annoyance ( $HA$ ) and high sleep disturbance ( $HSD$ ) are taken into consideration. These harmful effects can be quantified by the absolute risk ( $AR$ ) defined as: "Occurrence of the harmful effect in a population exposed to a specific level of environmental noise". There are well known equations (see: (7), (8) and (9)) for three different kinds of noise sources which can be used for the quantification of the number of highly annoyed ( $HA$ ) people. Similar equations (see: (10), (11) and (12)) could be applied in calculations of the number of highly sleep disturbed ( $HSD$ ) people.

For  $HA$  and  $HSD$  the total number  $N$  of people affected by the harmful effect  $y$  (number of attributable cases) due to the source  $x$  is derived (for each combination of noise source  $x$  - road, railway, aircraft, and harmful effect  $y$  -  $HA$ ,  $HSD$ ) by the following equation:

$$N_{x,y} = \sum_j (n_j \cdot AR_{j,x,y}) \quad (5)$$

where  $AR_{x,y}$  is the absolute risk of specific harmful effect ( $HA$ ,  $HSD$ ) and  $n_j$  is the number of people that are exposed to the  $j$ -th exposure band of noise. What needs clarification is the decision which approach to choose: general or specific. What should be calculated: the number of affected people for the whole city or the number for a given area in the city?

### 3. The current, legally recommended in Poland, method of calculation of the number of people affected by noise

Noise maps in Poland are calculated for five kinds of noise sources separately: tram, railway, road, aircraft and industry noise. In addition, the map for each kind of noise source has several layers with different noise level recommendations. Within each layer the places (expressed in a number of people in a given layer -  $m$ ) where the recommended  $L_{den}$  value is exceeded are identified. For these places the difference ( $\Delta L$ ) between the actual  $L_{den}$  value and the recommended one is calculated. These parameters are input values for the equation for index  $M$ , which is related to the number of people exposed to a given kinds of noise source:

$$M = 0.1m(10^{0.1\Delta L} - 1) \quad (6)$$

The difficulty in calculating index  $M$  is that there are no rules how to relate its value to the different areas. As can be seen from equation (6) neither the number of people exposed to a given noise source nor the dose –response relationship for different harmful effects are calculated.

#### 4. Environmental noise guidelines implementation in Poland. A case of Poznań

In our opinion, the environmental noise guidelines can currently be implemented in Poland in a limited form, only for two harmful effects: high annoyance ( $HA$ ) and high sleep disturbance ( $HSD$ ). Our calculations are performed for three kinds of noise sources: road traffic, rail noise (rails + trams) and aircraft noise.

To assess the actual number of ( $HA$ ) or ( $HSD$ ) people in the studied area it is recommended to use local exposure response function ( $ERF$ ) (the exposure response function was earlier called dose-response relationship). However, in Poland such data are not available. In such case the generalized  $ERFs$  (cf. questions 7 to 13) can be applied.

We decided to divide the whole city of Poznań into squares 200x200 m. There are 6902 squares which cover the whole city  $\sim 262$  km<sup>2</sup>. Each square represents one point in the maps of  $HA$  and  $HSD$  people. The results are presented in a form of two different maps (for % $HA$  and % $HSD$  respectively) and as two numbers: % $HA$  and % $HSD$ , for the whole Poznań.

The new acoustic map for Poznań constructed along these lines has already been calculated. The results of these new calculations will be compared to the  $M$  values calculated earlier.

##### 4.1. Calculations of the number of highly annoyed people due to road, railway and aircraft noise

The input data to this calculations are the number of residents of the given square in Poznań exposed to each kind of noise source, in 5 dB classes ranging from 55 dB to at least 75 dB  $L_{den}$ . Using the  $ERF$  for high annoyance, for each kind of noise source separately, the absolute risk ( $AR$ ) of  $HA$  people for each  $L_{den}$  value was calculated. By multiplying this  $AR$  value by the total number of people exposed to a given  $L_{den}$  we obtain the number of  $HA$  people. Having those two numbers: the number of people exposed to a given value of  $L_{den}$  and the number of  $HA$  people, it is possible to calculate the % $HA$  people for each  $L_{den}$ .

There are three equations which should be applied to calculate the  $AR$  value for road traffic, railway and aircraft noise:

$$AR_{HA,road} = (78.9270 - 3.1162 \cdot L_{den} + 0.0342 \cdot L_{den}^2)/100 \quad (7)$$

$$AR_{HA,rail} = (38.1596 - 2.05538 \cdot L_{den} + 0.0285 \cdot L_{den}^2)/100 \quad (8)$$

$$AR_{HA,air} = (-50.9693 + 1.0168 \cdot L_{den} + 0.0072 \cdot L_{den}^2)/100 \quad (9)$$

These equations are valid under the assumption that the harmful effects of noise causing high annoyance start at:  $L_{den} = 40$  dB for road traffic,  $L_{den} = 34$  dB for railway noise and  $L_{den} = 33$  dB for aircraft noise. To obtain the number of % $HA$  people in a given square

area we have to divide the number of all people exposed to a given kind of noise source (for all  $L_{den}$  values in a given square area), by the number of  $HA$  people in this same area and multiply it by 100. The same procedure should be repeated for each specific noise source. The maps showing the percentage of  $\%HA$  people and the  $M$  index for three kinds of noise sources are presented in Fig. 2(a), (b), (c) respectively.

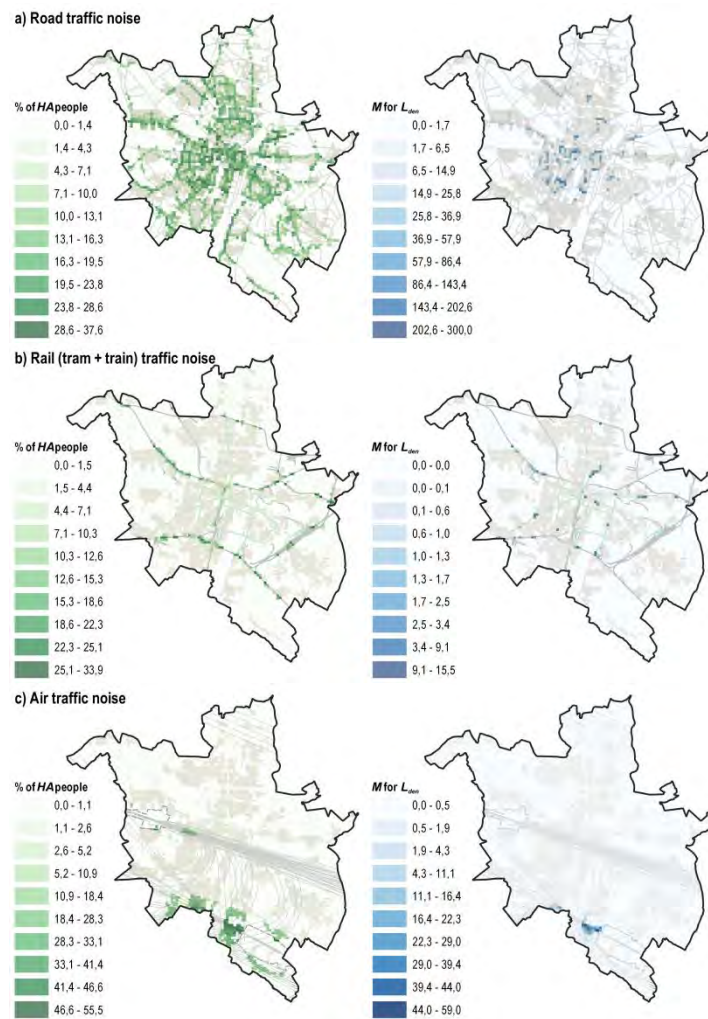


Figure 1. Noise map of Poznań representing  $\%HA$  people and index  $M$ , exposed to the (a) road traffic noise, (b) rail noise and (c) aircraft noise

As can be seen from the Figure 1, the method of calculations proposed here is much more precise than predictions based on the  $M$  index. The number of highly annoyed people

is higher than predicted with  $M$  index and the endangered areas are delimited more accurately (see also Table 1.).

#### 4.2. Calculations of the number of highly sleep disturbed people due to road, railway and aircraft noise

The same procedure was applied when calculating the number of highly sleep disturbed ( $HSD$ ) people. There are three equations which should be applied to calculate the  $AR$  value for road traffic, railway and aircraft noise:

$$AR_{HSD,road} = (19.4312 - 0.9336 \cdot L_{night} + 0.0126 \cdot L_{night}^2)/100 \quad (10)$$

$$AR_{HSD,rail} = (67.5406 - 3.1852 \cdot L_{night} + 0.0391 \cdot L_{night}^2)/100 \quad (11)$$

$$AR_{HSD,air} = (16.7885 - 0.9293 \cdot L_{night} + 0.0198 \cdot L_{night}^2)/100 \quad (12)$$

These equations are valid under the assumption that harmful effects of noise causing sleep disturbance start at:  $L_{night} = 43$  dB for road traffic,  $L_{night} = 43$  dB for railway noise and  $L_{night} = 33$  dB for aircraft noise. The maps presenting %HSD people for three kinds of noise sources are presented in Fig. 2(a), (b), (c) respectively. Analyzing the diagrams in Fig. 2 a similar conclusions can be drawn. The number of highly sleep disturbed people is higher than predicted with  $M$  index and the endangered areas are delimited more accurately. In the case of aircraft noise the endangered areas in the maps are similar to those delimited with the index  $M$ , but the number of affected people is larger than index  $M$  predicts.

#### 4.3. Calculations of the number of highly annoyed and highly sleep disturbed people for the whole city of Poznań

In a generalized approach to  $HA$  and  $HSD$  harmful effects of noise for the whole city we calculated one value for each kind of noise source separately. These values can be compared to the values obtained with the  $M$  index. Table 1. shows that the use of index  $M$  results in the underestimated number of  $HA$  and  $HSD$  people.

Table 1. The % $HA$  and % $HSD$  people, the absolute number of  $HA$  and  $HSD$  and index  $M$  calculated for three kinds of noise sources for the whole city of Poznań

Source	% of $HA$	$N$ of $HA$	$M$ for $L_{den}$	% of $HSD$	$N$ of $HSD$	$M$ for $L_{night}$
Road noise	10.92	55 632.4	11 014.1	3.26	16 601.0	8 843.1
Rail noise	0.27	1 360.2	119.5	0.73	3 714.0	147.5
Air noise	1.08	5 504.0	747.2	0.04	219.2	112.3

#### 4.4. Calculations of the number of highly annoyed and highly sleep disturbed people due to two or more concurrent noise sources

It is well known that in everyday life people are exposed to noise that is produced by several different sources. The question is: how can we calculate the overall harmful effects

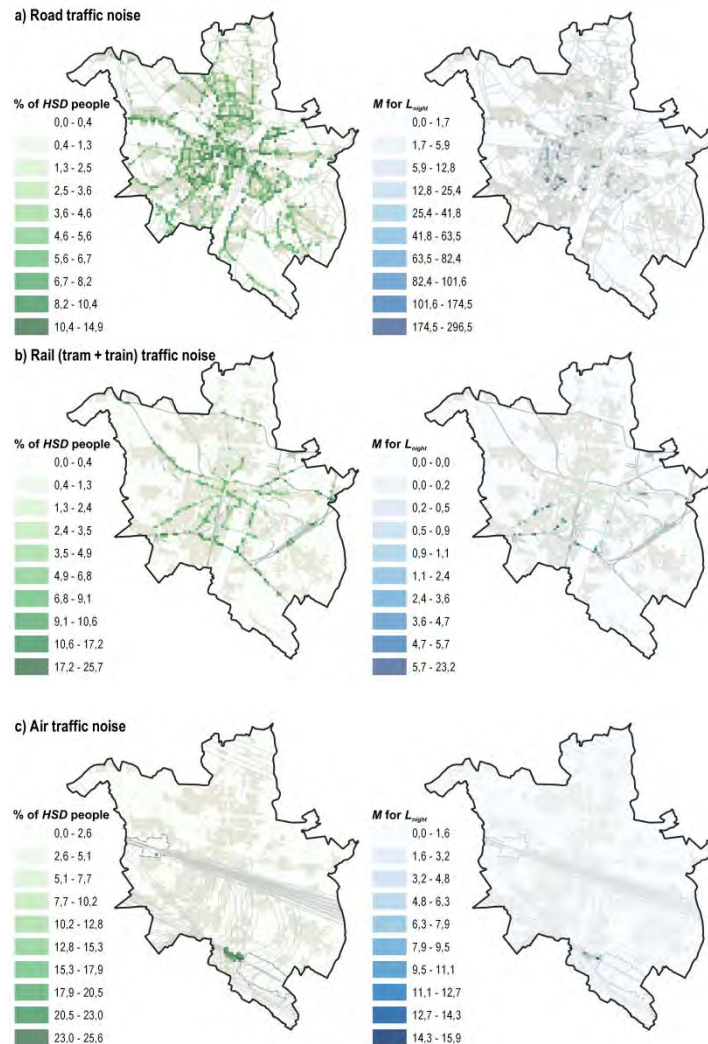


Figure 2. Noise map of Poznań representing percentage and index  $M$  of highly sleep disturbed people exposed to the (a) road traffic noise, (b) rail noise and (c) aircraft noise

generated concurrently by two or three noise sources? Based on the results of our preliminary study [3] we propose the summation of the number of people exposed to different kinds of noise sources at the same time with a special weighting. The weighting depends on deciding which noise source is dominant. Generally, we take into account 100% of people for the dominant sound source and 50%, 25% or 0% for the other noise sources. Such summation was performed for the  $HA$  harmful effect. For the  $HSD$  effect we simply summed up, with equal weight, the number of people  $HSD$  from all three kinds

of noise sources. The map representing the *HA* and *HSD* people from all noise sources is presented in Fig. 3.

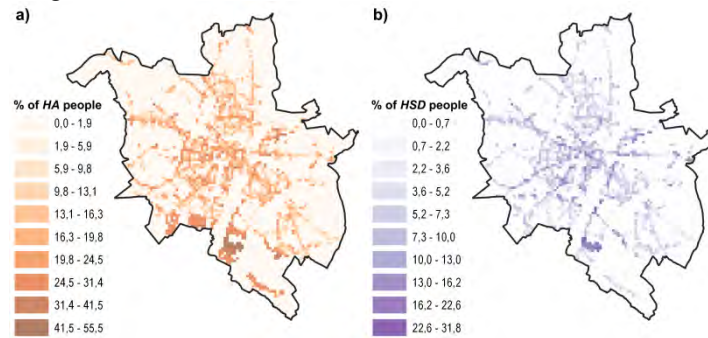


Figure 3. Noise map of Poznań representing the percentage of highly annoyed (a) and highly sleep disturbed (b) people

## 5. Conclusions

In this study we presented a case study of an implementation of the assessment methods for harmful effects of noise proposed in ANNEX III. Our study allows for the following conclusions:

- Two out of four harmful effects recommended in ANNEX III can be considered in the assessment process in Poland. These are: high annoyance (*HA*) and high sleep disturbance (*HSD*). The results of our calculations are presented both in the form of maps displaying the endangered areas, and generally as the number of the *HA* and *HSD* people in the whole city of Poznań
- The numbers of highly annoyed and highly sleep disturbed people were calculated based on the well documented exposure response functions. The hitherto calculated  $\Delta L$  value (equation (6)) which is used for calculation of the index *M* has no reference to any harmful effects.
- We have shown that calculated number of *HA* and *HSD* people significantly differs from the one calculated with index *M*
- Finally, we propose a new method of summation of the number of people affected by two or more concurrent noise sources.

## References

1. *Environmental Noise Guidelines for the European Region*. World Health Organization 2018, ISBN 978 92 890 5356 3.
2. *ANNEX III on the establishment of assessment methods for the harmful effects of noise according to Directive 2002/49/EC* -working version of the document, 2019.
3. J. Felcyn, *Model of total annoyance due to combined transportation sound sources in simulated noise scenarios 2019*, PhD Thesis, UAM Poznań.



## Problem of Placing the Organ Pipes on the Windchest

Damian WĘGRZYN

*Polish-Japanese Academy of Information Technology,  
Koszykowa 86 St., 02-008 Warsaw,  
damian.wegrzyn@pj.edu.pl*

Piotr WRZECIONO

*Warsaw University of Life Sciences,  
Nowoursynowska 166 St., 02-787 Warsaw,  
piotr\_wrzeciono@sggw.pl*

### Abstract

This paper presents research showing the problem occurring in the construction of a pipe organ, related to the placement of the organ pipes on the windchest. The close location of the organ pipes to each other influences the parameters of the sound generated by the pipes. It causes an intonation problem, namely the detuning of the organ pipes if they are located too close to each other on the windchest. The presented measurements show the influence of a distance between pipes of various types on basic sound parameters, such as frequency or volume level. The research carried out shows that in extreme cases the detuning reaches a temperate halftone. This has undoubtedly an impact on the tuning of organ pipes, especially in the case of a table organ or pipe organ built in a small space. In the future, the outcomes of the presented research can be applied in the windchest design.

**Keywords:** pipe organ, windchest, organ pipe distance

### 1. Introduction

An organ is a musical instrument built of many pipes, placed on a windchest, i.e. a device used to distribute air between particular pipes [1]. While making the instrument, there is not much space for a large distance between these numerous pipes on a windchest. In the small instruments, e.g., table organs, the dimensions of the windchest are relatively small. As a consequence, intonation problems occur because the neighbouring pipes can influence each other's sound. The main reason for this kind of issues is the way of sound generation, because the necessary condition of proper working of the organ pipe (i.e. generating the sound) is a turbulent flow through the so-called mouth of the pipe [1]. Changes in the flow significantly affect the sound parameters.

This paper presents research on the impact of placing organ pipes (on the windchest) on the basic sound parameters. This problem is a critical issue from the viewpoint of organbuilders, as well as musicians playing small organs. The final goal of this research is to create a model of the behaviour of organ pipes during tuning and intonation.

The results presented in this paper are new in the area of research on the acoustics of pipe organs. Previous publications relate primarily to the work of the organ pipe as a sound source without taking into account its environment [2, 3, 5, 6, 8-10].

## 2. Measurements

The presented research was carried out on the intonation table in an organbuilder's workshop. The air pressure in the windchest was set to 67 mm water gauge. In the organbuilding community, the measurement of relative pressure is commonly applied, using the u-tube manometer filled with water. The air temperature in the workshop was 19°C.

We performed the sound measurements using the sound recording at various distances from the obstacle. For each distance, the measurement was performed at least five times. The obstacles were set at distances from 0 mm to 80 mm, in 5 mm steps. At a distance equal to or greater than 80 mm, we did not notice changes in the parameters of the recorded signal.



Figure 1. Measuring point and sample measurement for Bourdon 8-foot:  
a) without obstacles, and b) with an obstacle

The measurements (Figure 1) were made for labial pipes representing various types: Principal 4-foot (open, metal, 75% tin and 25% lead, English bay leaf mouth with 29.4 mm width, 37.9 mm internal diameter, 530 mm wavelengt, the theoretical frequency: 323.16 Hz), Bourdon 8-foot (open, oakwood, external mouth with 41.1 mm width, 41.9 x 51.9 mm internal dimensions, 580 mm wavelengt, the theoretical frequency: 295.3 Hz) and Bourdon 16-foot (stopped, oakwood, external mouth with 41.8 mm width, 42.4 x 54.5 mm internal dimensions, 590 mm wavelengt,

the theoretical frequency: 145.15 Hz). They are basic voices in almost every pipe organ, so we consider them representative for our research.

The sound was recorded using two measuring microphones at the same time with omnidirectional characteristics, the sensitivity of 10 mV/Pa, and an equivalent noise level of 20 dBA. The microphone no. 1 was positioned at a distance of 58 mm from the top of the pipe, and microphone no. 2 at a distance of 49 mm from the mouth of the pipe. During the recordings, Class 1 acoustic calibrator was also used (1 kHz, 94 dB).

### 3. Analysis and results

During the measurements, significant changes in the pitch of the pipe were noticed in the case of a small distance between the pipe mouth and the obstacle. As a consequence, the decision was made to analyse the following sound parameters: the fundamental frequency ( $F_0$ ) depending on the distance of the obstacle from the mouth of the pipe, and the sound level of the first harmonic. The frequency analysis was performed in two stages. In the first stage, the fundamental frequency was found using the FFT [4] in GNU Octave programming language. The accuracy of the transform was 0.8 Hz. Due to the insufficient accuracy of this transform, the DTFT [4] in a program written by the authors in the Java environment was used in the next stage. The DTFT calculations were made in the range of  $F_0 \pm 5$  Hz, for  $F_0$  calculated in the first stage. The assumed DTFT accuracy was 0.01 Hz. Due to the usage of the calibration signal, the first harmonic's sound level analysis for the recordings was also performed (Figure 3).

The measurement uncertainty was calculated as a standard deviation from the obtained results. For the  $F_0$  frequency, the following formula was used:

$$\sigma_{F_0} = \sqrt{\frac{1}{N-1} \sum_{k=1}^N (F_0[k] - \mu_{F_0})^2} \quad (1)$$

where:  $\mu_{F_0}$  - average  $F_0$ ,  $N$  - number of measurements,  $k$  - measurement number.

Analogously to (1), the measurement uncertainty of the first harmonic's sound level was also calculated. Tables 1-3 present the results of the analysis. The dependencies between the  $F_0$  frequency of the pipe and the distance from the obstacle are presented in Figure 2.

Distances between sounds in the equal temperament are described by intervals. Interval is the difference between two sounds. Cent is a logarithmic unit of interval measurement [7]. For each interval, the difference  $W_{F_0}$  between the measured sound and the sound without the obstacle was calculated using the following formula:

$$W_{F_0} = 1200 \cdot \log_2 \left( \frac{F_0(x)}{F_0(80)} \right) \quad (2)$$

where  $x$  is the distance from the obstacle.

Table 1. The list of measurements for the organ pipe - Bourdon 16-foot stopped, c sharp

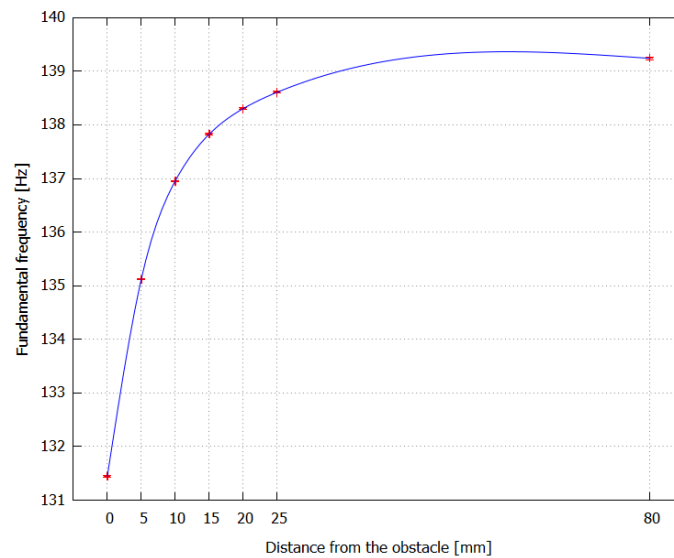
x	$\mu_{F0}$	$\sigma_{F0}$	$W_{F0}$	Microphone no. 1		Microphone no. 2	
				F <sub>0</sub> Sound level L <sub>Zeq</sub>	Measurement uncertainty $\sigma_{LZeq}$	F <sub>0</sub> Sound level L <sub>Zeq</sub>	Measurement uncertainty $\sigma_{LZeq}$
[mm]	[Hz]		[cent]	[dB]			
0	131.44	0.02	-99.8	86.32	0.03	103.46	0.02
5	135.12	0.01	-52.0	87.86	0.03	105.51	0.03
10	136.95	0.01	-28.7	89.32	0.02	105.61	0.03
15	137.82	0.02	-17.7	88.91	0.03	105.88	0.02
20	138.30	0.02	-11.7	88.30	0.06	106.02	0.05
25	138.60	0.02	-7.9	88.02	0.03	105.74	0.02
...							
80	139.24	0.02	0.0	87.01	0.04	104.33	0.04

Table 2. The list of measurements for the organ pipe - Bourdon 8-foot open, b

x	$\mu_{F0}$	$\sigma_{F0}$	$W_{F0}$	Microphone no. 1		Microphone no. 2	
				F <sub>0</sub> Sound level L <sub>Zeq</sub>	Measurement uncertainty $\sigma_{LZeq}$	F <sub>0</sub> Sound level L <sub>Zeq</sub>	Measurement uncertainty $\sigma_{LZeq}$
[mm]	[Hz]		[cent]	[dB]			
0	238.14	0.01	-103.1	109.15	0.04	106.61	0.04
5	246.02	0.03	-46.8	111.73	0.03	109.83	0.02
10	249.22	0.03	-24.4	112.18	0.09	109.63	0.09
15	250.61	0.04	-14.8	112.37	0.04	110.08	0.05
20	251.30	0.02	-10.0	112.41	0.06	110.18	0.07
25	251.76	0.03	-6.9	112.55	0.08	110.13	0.08
30	251.99	0.01	-5.3	112.60	0.04	110.20	0.04
35	252.19	0.06	-3.9	112.54	0.06	109.92	0.07
40	252.30	0.04	-3.1	112.58	0.03	109.89	0.03
45	252.38	0.02	-2.6	112.63	0.04	109.69	0.04
...							
80	252.76	0.02	0.0	111.55	0.04	106.99	0.04

Table 3. The list of measurements for the organ pipe - Principal 4-foot, e'

x	$\mu_{F0}$	$\sigma_{F0}$	$W_{F0}$	Microphone no. 1		Microphone no. 2	
				F <sub>0</sub> Sound level L <sub>Zeq</sub>	Measurement uncertainty $\sigma_{LZeq}$	F <sub>0</sub> Sound level L <sub>Zeq</sub>	Measurement uncertainty $\sigma_{LZeq}$
[mm]	[Hz]		[cent]	[dB]			
5	320.59	0.03	-35.3	95.72	0.02	97.45	0.04
10	324.14	0.02	-16.3	97.77	0.03	99.80	0.05
15	325.28	0.01	-10.1	98.04	0.05	100.70	0.07
20	325.81	0.02	-7.3	98.20	0.04	101.34	0.05
25	325.90	0.01	-6.9	98.22	0.02	101.67	0.03
30	326.16	0.01	-5.5	98.19	0.03	101.94	0.04
35	326.28	0.01	-4.9	98.34	0.36	102.01	0.03
40	326.37	0.01	-4.4	97.98	0.03	101.90	0.04
45	326.44	0.01	-4.0	97.99	0.02	101.90	0.03
50	326.49	0.02	-3.7	97.93	0.03	101.80	0.03
55	326.52	0.01	-3.6	97.90	0.02	101.67	0.04
...							
80	327.19	0.04	0.0	98.32	0.02	100.92	0.03

Figure 2. The dependence of the fundamental frequency  $F_0$  of the organ pipe on the distance from the obstacle - Bourdon 16-foot stopped, c sharp

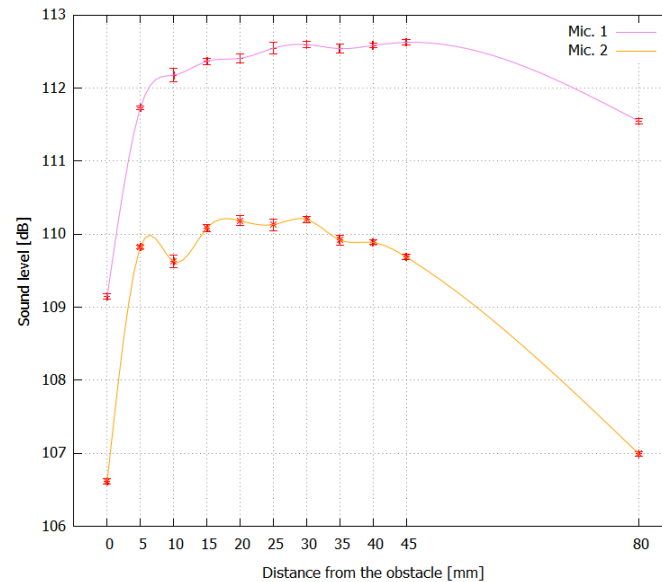


Figure 3. The sound level depending on the distance from the obstacle - Bourdon 8-foot open, b

The spectra shown in Figure 4 present changes in the harmonics of the pipe sound. All harmonics have a significantly lower sound level at a close distance of the obstacle from the pipe. Odd harmonics (i.e. frequencies  $kF_0$ , where  $k = 1, 3, 5, 7, \dots$ ), in the case of an open pipe, have a much lower sound level than without any obstacle. This phenomenon is typical for stopped pipes, but not for the open ones.

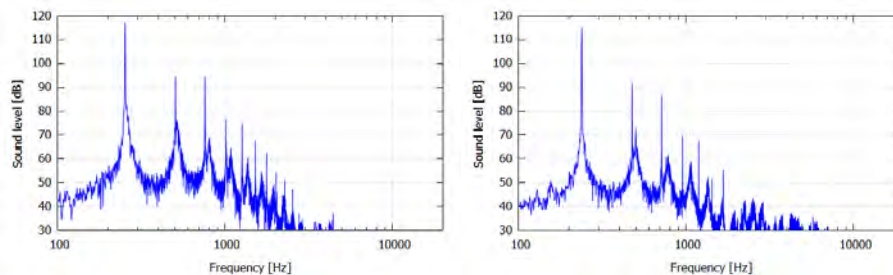


Figure 4. The spectrum of the organ pipe without the obstacle (left) and with the obstacle at 0 mm distance (right) – Bourdon 8-foot open b, microphone no. 1

#### 4. Musical consequences

Based on the measurements, it was found that the presence of an obstacle near the pipe has a significant impact on the pitch of the generated sound. These changes can be

as much as circa -100 cents (Figure 5), which in the case of equal temperament means lowering sound pitch by a semitone. An additional consequence of this phenomenon is the time-consuming tuning and intonation of organs, if the pipes are located close to each other. The most considerable change in  $F_0$  frequency in terms of the distance was observed for wooden pipes with a rectangular profile.

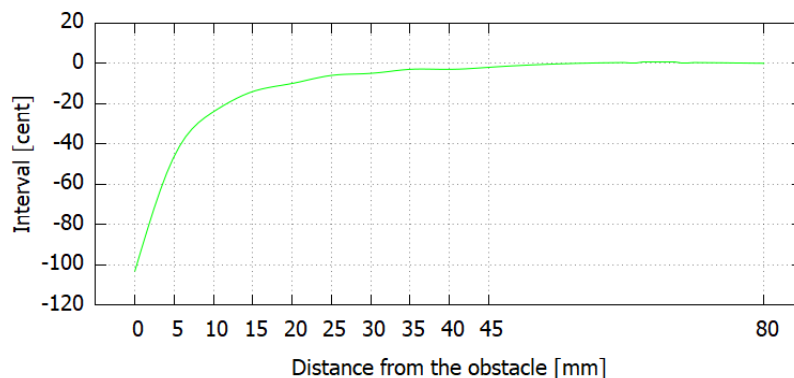


Figure 5. The interval between the tone of the organ pipe without obstacle and with the obstacle - Bourdon 8-foot open h

There was also a change in the sound level depending on the distance observed (Figure 3). For the studied pipes, these changes were about 4 dB. Such a change may also cause intonation problems.

## 5. Conclusions

The results of the presented measurements show an evident effect of the distance between the pipe and the obstacle on the parameters of the generated sound. In the case of pipe organs, such obstacles are other pipes placed on the windchest. These changes are significant for organ building. It can be expected that in the case of the dense arrangement of the pipes, the instrument may not be able to tune.

We found that the impact of the obstacle on sound generation by the organ pipe exists even when the distance between the obstacle and the mouth of the pipe is relatively large. For this reason, knowledge about the behaviour of the organ pipe near the obstacle is very important in the process of designing a windchest.

## Acknowledgments

The authors thank the organbuilder Władysław Cepka for providing the workshop and organ pipes for research.

## References

1. J. Angster, A. Miklos, P. Rucz, *Acoustics of Organ Pipes and Future Trends in the Research*, *Acoustics Today*, **13** (2017) 10 – 18.

2. H. Außerlechner, T. Trommer, J. Angster, A. Miklos, *Experimental jet velocity and edge tone investigations on a foot model of an organ pipe*, The Journal of the Acoustical Society of America, **126** (2009) 878 – 886.
3. S. Kokkelmans, M. P. Verge, A. Hirschberg, A. P. J. Wijnands, R. Schoffelen, *Acoustic behavior of chimney pipes*, The Journal of the Acoustical Society of America, **105** (1999) 546 – 551.
4. B. P. Lathi, *Linear Systems and signals*, Oxford University Press, New York 2010.
5. E. de Lauro, S. de Martino, E. Esposito, M. Falanga, E. Tomasini, *Analogical model for mechanical vibrations in flue organ pipes inferred by independent component analysis*, The Journal of the Acoustical Society of America, **122** (2007) 2413 – 2424.
6. A. W. Nolle, T. L. Finch, *Starting transients of flue organ pipes in relation to pressure rise time*, **91** (1992) 2190 – 2202.
7. E. Prout, *Harmony: Its Theory and Practice*, Cambridge University Press, New York 2011.
8. P. Rucz, F. Augusztinovicz, J. Angster, T. Preukschat, A. Miklos, *Acoustic behaviour of tuning slots of labial organ pipes*, The Journal of the Acoustical Society of America, **135** (2014) 3056 – 3065.
9. I. Vaik, G. Paal, *Flow simulations on an organ pipe foot model*, The Journal of the Acoustical Society of America, **133** (2013) 1102 – 1110.
10. M. P. Verge, B. Fabre, W. E. A. Mahu, A. Hirschberg, R. R. van Hassel, A. P. J. Wijnands, J. J. de Vries, C. J. Hogendoorn, *Jet formation and jet velocity fluctuations in a flue organ pipe*, The Journal of the Acoustical Society of America, **95** (1994) 1119 – 1132.



## Timbre Solfege and Auditory Profile Analysis

Jan ŻERA

*Warsaw Institute of Technology, Faculty of Electronics and Information Technology,  
Institute of Radioelectronics and Multimedia Technology, Nowowiejska 15/19,  
00-665 Warszawa, j.zera@ire.pw.edu.pl*

### Abstract

A course in auditory evaluation of sound, called *Timbre Solfege*, was developed by a team of researchers headed by Professor Andrzej Rakowski, at the Music Acoustics Laboratory, Chopin University of Music. A large part of the course, taught at the Department of Sound Engineering, has been focused on the detection and identification of timbre changes produced by formants and by other kind of sound spectrum modifications. Detecting formants in sound recordings is an auditory task that has much in common with auditory profile analysis, an area of research initiated and developed in psychoacoustics by Professor David M. Green, exploring the fundamentals of detection of changes in the sound spectrum envelope shape, independently of the differences in loudness between the sounds. The purpose of this study is an attempt to relate the results of the *Timbre Solfege* sound evaluation drills to the theory of the auditory profile analysis.

**Keywords:** sound spectrum, auditory profile analysis, formants, detection and discrimination

### 1. Introduction

The purpose of this paper is to compare detection thresholds of a formant created in the spectrum of noise (just noticeable formant amplitude) and discrimination thresholds of formant amplitude (just noticeable change in formant amplitude) with postulates of auditory profile analysis [1]. The data for such a comparison were taken from research studies oriented towards the development of auditory skills needed for professional practice in sound recording. The studies were part of a research program concerned with *Timbre Solfege*, a course aimed at perceptual and analytical evaluation of sound taught at the Chopin University of Music [2]. Auditory profile analysis is an area of research in psychoacoustics focused on the ability of discerning changes in the spectral envelope shape of complex acoustic signals. Thus, the measurements of formant detection and discrimination conducted within the *Timbre Solfege* research studies are both auditory tasks that can be examined with the use of the profile analysis methodology.

### 2. The *Timbre Solfege* course

The *Timbre Solfege* course (in Polish: *solfeż barwy*) was developed at the Music Acoustics Laboratory, Chopin University of Music, by a team of researchers led by late Professor Andrzej Rakowski. The course was introduced to the sound engineering curriculum in 1977 and since then its program and teaching methodology have been constantly expanded and modified. The course became a major four-semester practicum offered to students to develop various auditory skills, including detection

and identification of formants, discrimination of pitch and loudness, recognition of various dimensions of timbre, and assessment of sound quality. The course program also includes exercises intended to develop the skill of analysing auditory spatial information and understanding the principles of auditory image formation. Detailed description of the course program and teaching methods can be found elsewhere [2–7]. It has been proved in progress tests conducted throughout the course that *Timbre Solfege* markedly improves the students' sensitivity to timbre changes and memory for timbre [7].

Along with teaching, the *Timbre Solfege* course creates a basis for research in timbre perception, such as investigations of the detection [8] and discrimination [9] of formant amplitude, discrimination of the formant frequency [10] and bandwidth [11]. Among these research areas, investigations of formant detection and discrimination reported by Letowski and Rogala [8], and Rogala and Śliwka [9] will be considered later in this paper as compatible with the scope and structure of auditory profile analysis [1].

Letowski and Rogala [8] measured detection thresholds for a single 1/3-octave formant created in white noise, and similarly modified the spectra of two natural sounds – a recording of tubular bells and a sample of pop music recording (Figure 1a). Results of this experiment, shown in Figure 1b, present the formant levels corresponding to 75% correct responses obtained with the use of paired comparison method [8]. The formant level needed for formant detection in white noise (solid line in Figure 1) was large at low frequencies, amounting to nearly 11 dB at a 63-Hz frequency of the formant band. In high frequency range, the level of the just detectable formant decreased to about 2 dB. The results obtained for natural sounds are non-monotonic, and their discussion will be postponed as irrelevant to the purpose of current study.

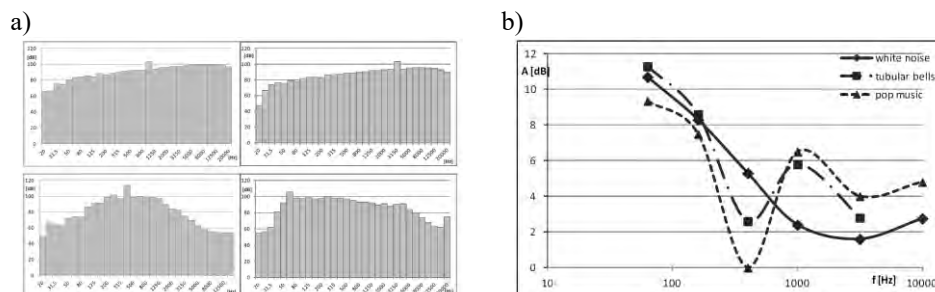


Figure 1. a) Upper panels: long-time average spectrum of white noise with added 1/3-octave formants at 1000 Hz (left) and 3150 Hz (right); lower panels: 400-Hz formant in tubular bells (left) and 63-Hz formant in pop music (right) spectra. b) Detection thresholds for formants created in spectra shown in a). Figures from ref. [8]

In the second experiment to be discussed here, reported by Rogala and Śliwka [9], formant discrimination thresholds were measured as a function of initial formant level and formant frequency. The initial level of a 1/3-octave formant imposed on wideband pink noise was either 3 or 12 dB. Using the *oddtity* stimulus presentation method (modified *3-alternative forced choice* paradigm), a psychometric function for the detection of a change in the formant level was determined. The discrimination

threshold was defined at 67% correct responses (middle point of the psychometric function). Average formant level discrimination thresholds, obtained for trained musicians and non-musicians, taken from reference [9], are presented in Figure 2.

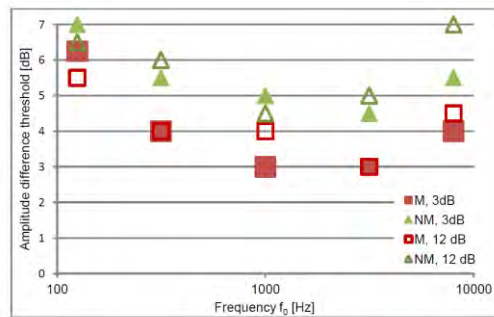


Figure 2. Discrimination thresholds of formant levels for musicians (M) and non-musicians (NM). Initial formant levels: 3 and 12 dB. Figure from reference [9]

The thresholds plotted in Figure 2, measured for centre noise band frequencies of 125, 400, 1000, 3150 and 8000 Hz, show a clear bowl shape, regardless of the formant initial level (3 or 12 dB) and the subjects' experience level. This bowl shape of threshold, favouring the 1-kHz frequency region, is characteristic of some results of auditory profile analysis and will be discussed in section 4. It is also apparent that discrimination thresholds for 3- and 12-dB formants are very similar, and differ by less than 1 dB. A constant discrimination threshold, in decibels, obtained for different formant levels indicates that Weber's Law holds for formant discrimination. This finding also suggests that the listeners may not perform spectral profile analysis, but rely on successive intensity comparisons in a frequency channel corresponding to the formant frequency. The thresholds measured for inexperienced subjects (NM group) are about 2–3 dB higher than those of experienced subjects (M group).

The results shown in Figures 1 and 2 will be compared in section 4 with the postulates of classical auditory profile analysis described by Green [1]. This will be done to examine, whether in the listening strategy in formant detection and discrimination tasks is based on successive signal comparisons in a single frequency channel or on simultaneous, across-channel comparisons.

### 3. Auditory profile analysis

Auditory profile analysis is a research direction introduced by David M. Green and his associates as a formalized approach to the auditory analysis of the spectral envelope profile of sound (for summary of early but substantial work see ref. [1]). The idea is to extend traditional intensity discrimination between successive sound samples by simultaneous comparison of intensity changes in different parts of the sound spectrum. The latter is simply discrimination of changes in the spectral envelope and thus the timbre of sound. The basic task consists in detection of a change in the intensity of

a single component in a multicomponent complex. With this respect, profile analysis is similar to the detection of 1/3-octave formant in noise, earlier described in this paper. During a single presentation of a sound sample (*e.g.*, the signal) in a profile analysis task the subject simultaneously compares the energy level resulting from an amplitude change in a given frequency channel with energy levels in all the other channels, centred at different frequencies. Subsequently, he/she makes a similar simultaneous comparison during the presentation of a second sound sample (*e.g.*, the standard). Detection of an amplitude increment is done by a comparison of different frequency regions instead of comparison of an energy change in a single frequency channel, in successive presentations, as is the case during loudness discrimination.

Although in everyday listening to spectrally complex sounds both successive and simultaneous comparisons are performed at the same time, proper measurement in laboratory settings requires to separate these two factors one from another. Roving of the overall level is used in profile analysis experiments to force the simultaneous across frequency channels comparisons by making successive level comparisons ineffective. From trial to trial, the overall signal level is randomly set with equal probability within 10-, 20- (a typical value) or even 40-dB range, usually in 1-dB steps. Such a roving of level prevents the subject from performing successive intensity comparisons, and forces him/her to use only across-spectral comparisons as a cue for detection of a spectral change in a signal. The roving range is directly related to the expected threshold and results from statistical estimation of how a random distribution in overall level disables the strategy of using a loudness cue in consecutive comparisons (for details see ref. [1], pp. 19-21).

Experiments on profile analysis [1, 12-14] were devoted to investigating several signal parameters that might influence the discrimination process. These included a number of components and the frequency spacing between them, the spectral range occupied by stimuli components, the effects of masking of neighbouring components, in contrast to the importance of components remotely positioned on the frequency scale, the effect of pedestal (*i.e.*, an initial increment in amplitude, relative to the amplitudes of other components). Although some studies were done for harmonic signals [15], most experiments on profile analysis were conducted with logarithmically spaced components in frequency what provided almost equal component distribution along the auditory filter bank. Within the scope of this study, this also corresponds to formants created in white or pink noise stimuli, using a 1/3-octave spectrum equalizer.

#### **4. Is the formant detection/discrimination a real life realisation of the profile analysis task?**

Laboratory experiments on profile analysis represent a quite theoretical construct of stimuli purposely designed for investigating the properties of the auditory system in its ability to analyse spectrally complex sounds. In this section an attempt is undertaken to directly compare formant detection, as it was done in the timbre solfege tasks, with some major results of profile analysis published by Green and Mason [1, 12], and Green and Kidd [14]. Actually, the two tasks will be compared. Firstly, the formant detection task in pink noise shown in Figure 1b will be compared with the profile analysis task

involving the detection of an amplitude increment of a single sinusoid in a 21-component logarithmic complex [1, 12]. This comparison is shown in Figure 3. Secondly, the formant level discrimination task shown in Figure 2 will be compared with a profile analysis task involving detection of an increment in a 950-Hz component with pedestal, as this task corresponds to the discrimination of formant level [14]. This comparison is shown in Figure 4.

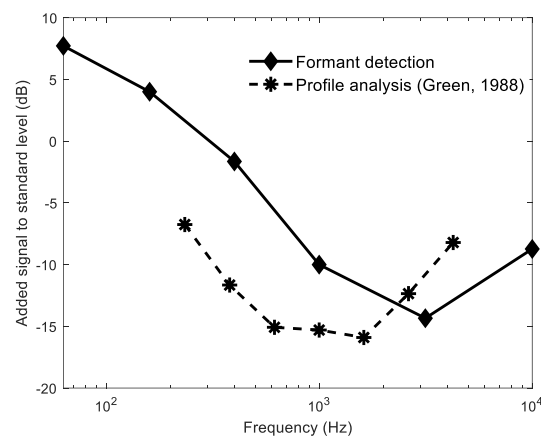


Figure 3. Comparison of a 1/3-octave formant detection task (solid line, diamonds) for white noise after Letowski and Rogala [8] with standard bowl of 21-component logarithmically spaced components in a range of 200–5000 Hz, in a profile analysis task after ref. [1] (dashed line, asterisks)

The solid line in Figure 3 presents the detection thresholds for 1/3-octave formants created in white noise, taken from Figure 2b and recalculated to the variable commonly used in profile analysis. In these studies, an increment  $\Delta A$  in amplitude added in phase to the initial amplitude  $A$  of the component is a signal in threshold measurements. For example, a 0-dB signal level corresponds to  $\Delta A = A$ . Thus an addition results in doubling the signal component amplitude or a 6-dB increase in level. A signal level of -20 dB corresponds to  $\Delta A = 0.1A$ , causing the amplitude of the signal component to be increased by 10% or only 0.8 dB, relatively to the amplitudes of other spectral components. This is why thresholds in profile analysis tasks are usually negative in decibels, except for very large amplitude increments. The detection thresholds seen in Figure 1b, in a range from 10.7 dB at 63 Hz down to 1.5 dB at 3150 Hz, are plotted in new coordinates of  $\Delta A$  level in Figure 3, in range from +7.7 dB down to -14.3 dB. The profile analysis data are taken from refs. [1, 12], and represent the well-established bowl for 21-component logarithmically spaced sinusoids (frequency ratio 1.1746 [12]). Unlike in the case of the data presented later in Figure 4, a direct comparison is possible, as both thresholds are defined for points located close to each other on the psychometric function: 75% correct for the formant detection task and 70.7% correct (adaptive procedure with a 2-down/1-up decision rule) in the profile bowl.

The variability of threshold with formant frequency, seen in Figure 3 for formant detection (solid line) is somewhat similar to the thresholds for amplitude change in the 21-component complex (dashed line). Both curves show a minimum of about  $-16$  to  $-14$  dB, although formant detection threshold at 3150 Hz, and profile analysis at 600–1700 Hz. The larger threshold increase for formant detection in low frequency range may be attributed to the energy decrease in narrowing the 1/3-octave band and thus an increase in masking by neighbouring bands. In contrast, log-spaced components in the auditory profile task produce approximately constant energy per critical band in the entire frequency range.

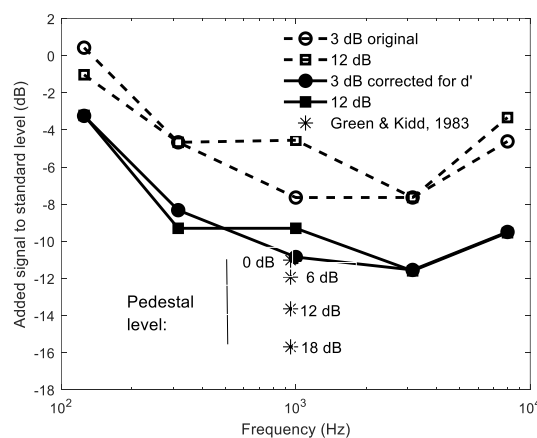


Figure 4. Comparison of 1/3-octave formant discrimination task (solid lines) and pedestal experiment by Green and Kidd [14] (asterisks). Dashed lines show experienced subjects' (M) original discrimination thresholds defined at 67% correct discrimination on the psychometric function. Solid lines show similar thresholds defined at 39% correct discrimination on a psychometric function corresponding to  $d' = 0.74$ , the value identical to that of 2AFC task at 70.7% correct responses in the pedestal experiment

The analysis of discrimination thresholds shown in Figure 4 refers to the data presented in Figure 2, restricted to the group of experienced musicians. These data are compared with the pedestal experiment by Green and Kidd [14] (asterisks in Figure 4). The dashed line in Figure 4 shows discrimination thresholds taken from Figure 2, expressed, as before, as  $20\log(\Delta A/A)$ . Due to differences in the experimental procedures, the values from experiments by Rogala and Śliwka [9] and Green and Kidd [14] had to be corrected to equalize the detectability index  $d'$  [16] before any comparison could be made. The thresholds in the pedestal experiment [14] correspond to  $d' = 0.74$  as they were defined at 70.7% correct responses, as a result of an adaptive procedure with a two-down/one-up decision rule and a two-alternative forced choice paradigm of stimuli presentation. The formant experiment [9] used an *oddball* procedure for which  $d' = 0.74$  occurs at 39% correct responses. Thus, using psychometric functions given in ref. [9], thresholds at 39% correct were determined and presented with solid lines in Figure 4.

The thresholds at 39% correct (solid lines) are on average by about 4 dB lower than those at 75% correct (dashed lines).

The corrected formant discrimination thresholds can be directly compared with the results of the pedestal experiment [14], of detecting a 950-Hz component in a 21-component complex (Figure 4, asterisks). Four pedestal levels, of 0, 6, 12 and 18 dB, were selected from ref. [14]. The average of thresholds for pedestals of 0 and 6 dB closely represent the initial formant level of 3 dB whereas the 12-dB pedestal corresponds to the 12-dB initial format level, to which an increment in level is added. It should be noted that formant discrimination thresholds closely correspond to 0 and 6-dB pedestals but are higher than the thresholds for the 12-dB pedestal. In the formant data reported by [9], the formant level discrimination thresholds are similar values for the 3- and 12-dB formants which was earlier interpreted as a finding indicating, that Weber's law holds for formant amplitude discrimination, which is not the case for the pedestal data. Overall, the data discussed in this paper suggest that the subjects concentrate in a formant discrimination task on the single frequency channel in which the change occurs and use sequential comparison, not simultaneous across-channel profile analysis. Such a listening strategy was very likely as the roving of signal level was not used in the experiments on formant discrimination.

## 5. Summary

This study was an attempt to compare the data on formant detection and discrimination obtained within the research studies concerned with the *Timbre Solfège* training program, developed by a team of researchers headed by Professor Andrzej Rakowski at the Musical Acoustics Laboratory, with a series of experiments within a research line called auditory profile, conducted by Professor David. M. Green and his associates. An added value of this comparison lies in the fact that the studies compared were, to a large extent, complementary in their research scope: the experiments with formants were conducted with close reference to auditory tasks commonly encountered in the praxis sound recording whereas the experiments on profile analysis used carefully designed stimuli aimed at studying specific properties of signal processing in the auditory system.

The major finding of this short study is that formant detection is based on inter-channel comparisons made independently for each of the stimuli compared in a trial. In contrast, formant discrimination consists in single channel comparison of amplitude at the formant frequency to detect changes in the formant amplitude and disregarding, as much as possible, the rest of the sound spectrum. This conclusion comes from the finding that the discrimination threshold for formant amplitude follows Weber's law which is not the case in auditory tasks based on profile analysis, involving across-channel comparisons. To verify these preliminary conclusions further experiments should be conducted at the Chopin University of Music with the use of some elements of profile analysis methodology.

## Acknowledgments

Work supported by the grant 504/04064/1034/40.00 from the Warsaw University of Technology.

## References

1. D. M. Green, *Profile analysis*, Oxford University Press, New York – Oxford 1988.
2. T. Letowski, A. Miśkiewicz, *Timbre Solfege: A course in perceptual analysis of sound*, In: *Signal Processing in Sound Engineering*, J. Adamczyk (Ed.), Warszawa: IPPT-PAN (2013) 83 – 96.
3. A. Rakowski, T. Letowski, K. Szlifirski, B. Okoń-Makowska, *Developing sensitivity to the timbre of sound at the Department of Sound Engineering*, Academy of Music in Warsaw (in Polish), in: *Studia z Teorii Przekazu Dźwięku*, Wydawnictwa Radia i Telewizji, Warsaw, Poland (1982) 183 – 195.
4. T. Letowski, *Development of technical listening skills: Timbre solfeggio*, J. Aud. Eng. Soc., **33** (1985) 240 – 244.
5. A. Miśkiewicz, *Timbre solfege: A course in technical listening for sound engineers*, J. Aud. Eng. Soc., **40** (1992) 621 – 625.
6. T. Letowski, A. Miśkiewicz, *Developing of technical listening skills for sound quality assessment*, Proceedings of *Inter-Noise '95*, Newport Beach, FL (1995) 917 – 920.
7. T. Rościszewska, A. Miśkiewicz, *Timbre Solfege: Development of auditory cues for the identification of spectral characteristics of sound*, Audio Engineering Society 138th Convention, May 7-10, Warsaw (2016).
8. T. Letowski, T. Rogala, *Formant perception: single formant*. In: *Sztuka słuchania (The Art of Listening)*, Chopin University of Music, Warszawa (2015) 45 – 63.
9. T. Rogala, *Discrimination of formant frequency in pink noise*. Audio Engineering Society 140th Convention, June 4-7, Paris (2016) Paper #9583.
10. T. Rogala, *Pink noise bandwidth discrimination*, Audio Engineering Society 142th Convention, May 20-23, Berlin (2016).
11. T. Rogala, P. Śliwka. *Discrimination of formant amplitude in noise*, Audio Engineering Society 138th Convention, May 7-10, Warsaw (2016) Paper #9282.
12. D. M. Green, Ch. R. Mason, *Auditory profile analysis: frequency, phase, and Weber's law*, J. Acoust. Soc. Am., **77** (1985) 1155 – 1161.
13. D. M. Green, Ch. R. Mason, G. Kidd Jr., *Profile analysis: critical bands and duration*, J. Acoust. Soc. Am., **75** (1984) 1163 – 1167.
14. D. M. Green, G. Kidd Jr., *Further studies of auditory profile analysis*, J. Acoust. Soc. Am., **73** (1983) 1260 – 1265.
15. J. Žera, Z. A. Onsan, Q. T. Nguyen, D. M. Green, *Auditory profile analysis of harmonic signals: critical bands and duration*, J. Acoust. Soc. Am., **93** (1993) 3431 – 3441.
16. N. A. Macmillan, C. D. Creelman, *Detection theory: a user's guide*, Cambridge University Press, Cambridge 1991.



## A Design of an Acoustic Coupler for Calibration of Hydrophones at Low Frequencies

Karol LISTEWNIK

*Central Office of Measures, Elektoralna 2 Str., 00-139 Warsaw,  
karol.listewnik@gum.gov.pl*

### Abstract

The purpose of this article is to present a coupler developed for the calibration of hydrophones at low frequencies in the Central Office of Measures (GUM). Due to the growing demand for marine environment research in the field of underwater noise, implemented in accordance with Directive 2008/56/EC, many models of autonomous underwater noise recorders have been developed. Ensuring the reliability of recorded data begins with the reliable calibration of hydrophones and/or the entire recorder. The choice of the calibration method was made on the basis of a detailed and broad analysis of calibration methods and similar solutions of coupling constructs worldwide. The analysis mainly took into account the documents of the last few years. On the basis of this analysis, the vibrating water column method was chosen and the construction of the coupler was developed. Regular calibration of hydrophones will contribute to reliable underwater noise monitoring in the Baltic region.

**Keywords:** acoustical coupler, calibration of hydrophone, low frequencies, receiving sensitivity

### 1. Introduction

The development of calibration methods and constructional solutions of calibrators at low frequencies from 1 Hz to 2 kHz is very important. There are many different types of autonomous underwater noise recorders in the world [1]. These recorders are calibrated by the manufacturer. Typically, the calibration of autonomous underwater noise recorders is limited to the calibration of hydrophones and internal electronics without testing the influence of the housing on their sensitivity. The growing number of recorders means that the task of periodic calibration requires more and more attention and time from their users, among others due to the time of delivery to the manufacturer, waiting time for calibration and delivery time to the consumer. It was investigated how hydrophones produced by manufacturers B&K and Reson change the receiving sensitivity between calibrations but only from 4 kHz (B&K) and 5 kHz (Reson) to ca. 200 kHz [2, 3]. Manufacturers rarely provide accurate calibration of hydrophones in the frequency range of 1 Hz to 2 KHz. It is assumed that, in the calibration process the receiving sensitivity of the hydrophone in the low frequency range is constant up to the first resonance. An analysis of the results of recalibrations of the same hydrophone receiving sensitivity has shown the results may differ significantly from the previous values given in the earlier calibration chart. The reason is the poor time stability of the receiving sensitivity reaching 1.6 dB (B & K – 3 recalibrations in 10 years) and 2.2 dB (Reson 1 recalibration in 3 years). In addition, the changes are non-linear.

In accordance with Directive 2008/56/EC, the quantity of acoustic recorders has grown; therefore a framework for Community action in the field of marine environmental policy (Marine Strategy Framework Directive) was established, which is the basis for monitoring underwater noise. The quality of the collected data is very important. Guidelines for the calibration of hydrophones and recorders are included in [4]. It is recommended that a full laboratory calibration is undertaken before and after every major deployment or sea-trial according to IEC 60565 2006 [5], ANSI S1.20 2012 [6]. The hydrophone calibrations are typically performed using couplers. Work on the division of the IEC 60565 2006 standard into two parts are underway; the second part exclusively concerns procedures for low frequency pressure calibration: IEC FDIS<sup>1</sup> 60565-2 2019 [7].

Increasing the availability, functionality and the need to shorten the time for calibration are the basis for the development of new and improved old calibration methods for hydrophones at low frequencies.

There is currently no standardization of the methods used to calibrate autonomous recorders.

Taking the above into consideration, the main objective of developing a calibration coupler at low frequencies at the GUM is to participate in improving the metrological capacity in underwater acoustic calibration for low acoustic frequencies.

## 2. Methods of calibrating hydrophones at low frequencies

In general, we can distinguish two calibration methods: primary methods that do not require reference to any acoustical standard, and secondary methods that require the calibrated reference standard to be used. Regardless of the method applied, the calibration includes at least the determination of the pressure sensitivity or pressure sensitivity level in accordance with one of the methods specified in the standard [5, 6] and estimation of measurement uncertainty shall be determined in accordance with [8]. Recently published papers present a method for an initial calibration of magnitude and phase of hydrophone sensitivity at frequencies in the range of 1 Hz to 2 kHz [9].

Different methods for calibrating hydrophones are well documented [10]. One of the most important projects carried out in recent years in the field of underwater acoustics is the EURAMET EMPIR project UNAC-LOW "Underwater Acoustic Calibration Standards for Frequencies Below 1 kHz" [11]. Goals of this project are the development of measurement consistency for the calibration of hydrophones and autonomous underwater noise recorders covering the frequency range of 20 Hz to 1 kHz. This frequency band covers the EU MSFD guideline requirements - Part I-III – which established one-third octave centre frequencies: 63 Hz and 125 Hz [4], also developing a multi-faceted and long-term strategy for each participant of the project, coinciding with the European Metrology Strategy and increasing the research potential in this area included upgrading standard IEC 60565.

The review of the latest calibration solutions for hydrophones at low frequencies will start with calibrators produced and tested as part of the EURAMET EMPIR UNAC-LOW

---

<sup>1</sup> Final Draft International Standard (FDIS)

project mentioned above. In the project description, the NPL reports that it has the ability to calibrate hydrophones for frequencies up to 1 kHz by the following methods:

- the air pistonphone method,
- the absolute method by a laser pistonphone,
- the vibrating column method.

The first method used by TUBITAK was to place the tested hydrophone with a calibrated reference receiver and a sound source in a closed chamber, so that the source interacted with both receivers at the same time. The reference receiver can be both a microphone and a hydrophone, depending on whether the chamber is filled with air or water. The estimated uncertainty amounted to  $\pm 1$  dB.

The standing wave method is used by the Swedish Defense Research Agency (FOI) using a calibration system (model C100, Underwater Science, Research & Development, Inc., USA) and is included in the standing wave tube method [12]. Calibration was performed entirely with an autonomous underwater noise recorder (Wildlife Acoustics SM2M). Tests for frequencies between 100 Hz and 1 kHz were carried out by comparison with a B&K 8104 reference hydrophone. The results of the tests showed a good agreement within  $\pm 1$  dB for frequencies from 100 Hz up to 700 Hz. Unexpectedly large deviations, from 3 dB up to 8 dB, were observed for frequencies between 800 Hz and 1 kHz. An explanation for this large discrepancy should be given.

Another method demonstrated in the EURAMET UNAC-LOW project was the method of air calibration using a pistonphone. It is generally recommended to use air pistonphone to check hydrophones before (and after) each dipping (pick up) during marine research. Over the research, the hydrophone was calibrated together with an autonomous underwater noise recorder in the frequency range of 20 Hz to 315 Hz [10]. The positive result of the research determined that this method be included in the project. The advantage of this method is the lack of liquid in the chamber, which makes the whole system lighter and easier to move. The disadvantage of this method is (as all methods using a calibration chamber) that the influence of the autonomous recorder housing on the obtained calibration results cannot be determined. For this method, the measurement uncertainty that occurred during the hydrophone calibration was not given.

In the National Physical Laboratory, an experimental hydrophone calibration was carried out using an interferometer and pistonphone [11]. The beam of the interferometer laser examined the vibrations of the piston. The purpose of this experiment was to investigate measurement uncertainty in the 20 to 158 Hz frequency range, which includes the required acoustic measurements for the 1/3 octave frequency of 63 Hz and 125 Hz.

At the VNIIFTRI (the All-Russian Research Institute of Physical-technical and Radiotechnical Measurements), the Russian Federation examined low frequency hydrophone calibration using a tensometric pressure sensor [13]. The acoustic chamber is a thick-walled chamber mounted to lattice on the floor where the hydrophone is located, and the reference transducer is screwed to the inner vertical wall in the form of a tensometric pressure sensor. The range of the tested frequencies ranges from 2 Hz to 3.15 kHz and the obtained uncertainty of measurement increases towards the upper frequency limit reaching a maximum of  $\pm 0.6$  dB.

In addition to Europe, in the US (Department of Sensors and SONAR, Naval Undersea Warfare Center), the recently developed primary method deserves special attention, which

is characterized by the combination of 1 Hz and 0.1 dB re 1 V  $\mu\text{Pa}^{-1}$  and  $\pm 1^\circ$ , respectively. The calibration takes place in a coupler reciprocity chamber depending on changes in frequency, temperature and pressure [14].

Ocean Networks Canada (ONC) measurements of near-field earthquake energy occur primarily below 60 Hz. The same calibrator can be used at low frequencies to calibrate hydrophones usually used to measure underwater noise. The VLF calibrator has a small volume chamber (50 mm diameter) filled with water or light oil (1325 m/s) [15]. The chamber is equipped with a tested hydrophone, a reference hydrophone, a pressure sensor and a temperature sensor. A piston with a diameter of 4 cm is moved by an actuator made of a piezoelectric stack with a maximum pitch of 10  $\mu\text{m}$ . The total measurement error includes error due to chamber dimensioning, ADC converter error, reference hydrophone error and investigated hydrophone error. The total error for 1 Hz is 0.12% and increases towards higher frequencies, reaching 0.35% for 1 kHz.

Another example of the use of a vibrational water column for the calibration of hydrophones are interlaboratory comparisons carried out by the Russian Metrological Institute of Technical Physics and Radio Engineering (VNIIFTRI) and Hangzhou Applied Acoustics Research Institute (HAARI – China), which took place in 2015. The research was based on the IEC 60565 2006 standard and the previous experience of F. Schloss et al. [16]. Before preparing the coupler for calibration of hydrophones, mathematical simulation of various sizes of calibration chamber was used. This enabled the resonance frequencies of the calibration chamber to be determined and, as a consequence, avoid resonance during calibration. The results of the calculations meant that the planned primary method was switched to a comparative method. Based on these simulations, a calibration chamber made of aluminum with a diameter of 30 cm and a wall thickness of 2 cm was made. The design allows the liquid level in the chamber to be modified (change of the resonant frequency). The authors report that the design allows the hydrophones to be calibrated from 30 to 1 kHz, with measurement uncertainties dependent on the hydrophone being tested, e.g. less than  $\pm 1$  dB for the Bruel & Kjaer type 8103 hydrophone [17].

At the National Metrology Center, the Science and Research Agency (NMC, A\* STAR), the national metrology institute (NMI) in Singapore, a hydrophone calibration system based on the vibrating column method was developed. It is used to determine the sensitivity of hydrophones for low frequencies in the range of 30 Hz to 2 kHz. The test chamber consists of a very rigid cylinder containing a water column with a volume of approximately 30 liters and an internal diameter of 30.6 cm. In the comparative method, two hydrophones were placed inside: a standard reference hydrophone and a tested hydrophone on an adjustable frame. The system provides calibration hydrophone with measurement uncertainty of less than 1.6 dB in the given frequency range. The system authenticates the measurement results of hydroacoustic sensors in applications related to oil and gas exploration, defence technologies and oceanic seismic-acoustic surveys [16].

### 3. Coupler design description

After studying the standards and analyzing the literature and test reports, a decision was made to choose the vibrating water column method. This approach has been extensively

described, and is relatively simple and economical to make. It is divided into two basic procedures. In the first, only one hydrophone is calibrated (the calibrated hydrophone can be calibrated first and then the reference one) and reference is made to the absolute value of the liquid pressure at the point of the geometric centre of the sensitive part of the calibrated hydrophone (the value of the modulus of the hydrophone sensitivity only). This approach does not require a reference hydrophone, although repeating the test with a reference hydrophone will make the test more reliable. The second option involves simultaneous immersion in a liquid at a selected reference depth and taking into consideration the test hydrophone and a comparison of the measurement results.

The first approach was chosen. This will allow you to achieve the ability to absolute calibration measurement of underwater sound. The acoustic coupler in this method is made in the form of a cylindrical calibration chamber filled with water. The hydrophone being tested is submerged immobile and stable in the water column. The calibration chamber is mounted on a sinusoidal vibration exciter that causes vertical sinusoidal vibration of the liquid column in the chamber. The design of the acoustic coupler for calibrating hydrophones via the vibrating water column method is presented in Figures 1 and 2.

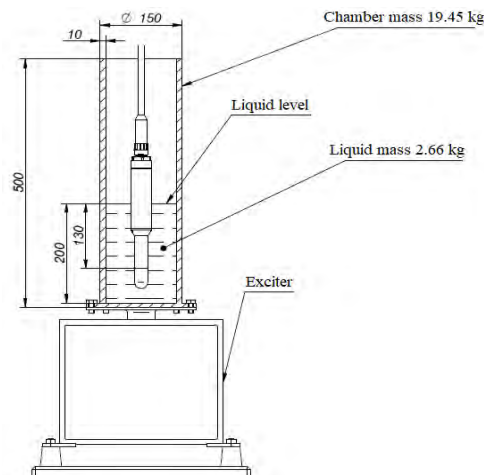


Figure 1. Design of an acoustic coupler for calibrating hydrophones via the vibrating water column method

The modulus of the pressure sensitivity of the hydrophone  $|M_p|$  shall be calculated as the ratio of the amplitudes of the sinusoidal signals of the open-circuit voltage  $U_0$  and pressure  $p_0$  (The height of the water column above the center of the sound-sensitive part expressed in pascals):

$$|M_p| = \frac{U_0}{p_0} \quad (1)$$

The depth of immersion of the centre of the hydrophone ( $d$ ) should be measured in the centre of the sound-sensitive part responsible for sound reception (which can be very difficult to achieve), or it can be omitted if the hydrophone sensitivity measurements are taken at two different depths. This depth difference can be established without knowledge of the location of the acoustic centre of the hydrophone, in this case we use formula (2) [4, 6]:

$$|M_p| = \frac{\Delta U_0}{\Delta d} \frac{1}{\rho_f x_0 \omega^2} \quad (2)$$

where:

- $\Delta U_0$  – difference of the peak values of the output voltages at the hydrophone for two different hydrophone immersion depths,
- $\Delta d$  – depth difference for 2 different measurements,
- $\rho_f$  – density of fluid,
- $x_0$  – vibration amplitude,
- $\omega$  – angular frequency.

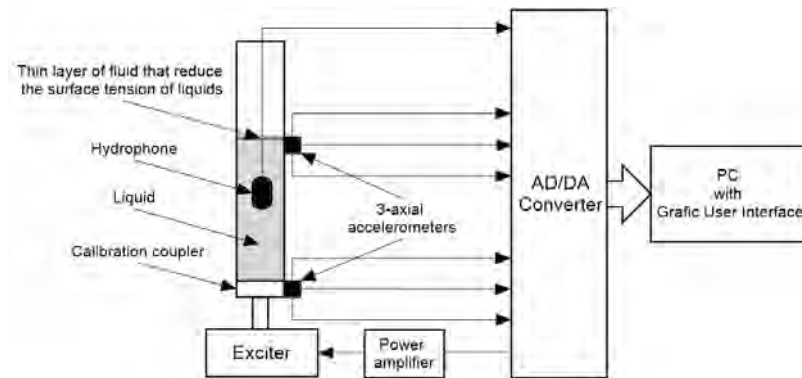


Figure 2. Measurement scheme

For check-ups the vibrating column method by alternately calibration the hydrophone under test and the reference hydrophone  $H$ , the following formula for calculating the voltage sensitivity of an unknown hydrophone can be used from [12]:

$$M_T = M_H + 20 \log[U_T/U_H] \quad (3)$$

where:

- $U_T$  – output voltage of the hydrophone under test  $T$ ,
- $U_H$  – output voltage of the reference hydrophone  $H$ ,
- $M_H$  – sensitivity level of the reference transducer (dB re V/ $\mu$ Pa),
- $M_T$  – sensitivity level of the hydrophone  $T$  (dB re V/ $\mu$ Pa).

#### 4. Conclusions

In recent years, it can be observed that autonomous underwater noise recorders have become increasingly popular. This is probably the reason why the number of projects, interlaboratory tests and individual solutions connected with ensuring the calibration of hydrophones at low frequencies has been growing over the last several years. In addition to previously known and improved calibration methods in the free field, there was interest in developing a method based on calibration in a chamber filled with water or air.

“... developing a multi-faceted and long-term strategy for each project participant, consistent with the European Metrological Strategy and increasing research potential in this area included updating the IEC 60565 standard.”

The above sentence accurately reflects the current view on the calibration of hydroacoustic devices and hydrophones. This article shows that there is a broader aspect: What about the reliability of calibration during the period of validity of the recalibration function, are all the functions cited available that users may have measurement uncertainty is less than 1 dB, and test results that are calibrated hydrophones is not time stable? Room changes between recalibrations reach 2.2 dB. What real uncertainty does the hydrophone have at the end user? As in the case when one of the solutions is to go to the period of recalibration of hydrophones, it excludes sending hydrophones abroad and for use in applications with hydrophone calibrators (unless additional hydroacoustic devices / hydrophones are available). In this article, based on the analysis of current solutions, the method of calibrating hydrophones by vibrating water column in an acoustic coupler was selected. It is a simple method, often used and the costs of its technical implementation and maintenance are economically justified and smaller compared to other methods.

At present, a calibrator project has been drafted and presented in this paper. The construction of a hydrophone calibration station with a vibrating water column is the first stage in the development of metrological infrastructure in the area of underwater acoustics in Poland (in the Central Office of Measures) [19].

#### References

1. R. S. Sousa-Lima, T. F. Norris, N. O. Oswald, D. P. Fernandes, *A Review and Inventory of Fixed Autonomous Recorders for Passive Acoustic Monitoring of Marine Mammals*, *Aquatic Mammals*, **39**(1) (2013) 25 – 53.
2. S. Schael, *Adulteration Of Underwater Acoustic Measurements*, Proceedings of the 2<sup>nd</sup> International Conference and Exhibition on Underwater Acoustics, (2014) 359 – 366.
3. K. Buszman, I. Gloza, R. Jozwiak, K. Listewnik, *The Calibration Of Hydroacoustic Channel Of Mobile Measurement Module*, Proceedings of the 2<sup>nd</sup> International Conference and Exhibition on Underwater Acoustics, (2014) 367 – 372.
4. R. P. A. Dekeling, M. L. Tasker, Van der Graaf, et al, *Monitoring Guidance for Underwater Noise in European Seas, Part I-III: Executive Summary*, JRC Scientific and Policy Report EUR 26557 EN, Publications Office of the European Union, Luxembourg 2014.

5. IEC 60565:2006, *Underwater Acoustics – Hydrophones – Calibration in the Frequency Range 0.01 Hz to 1 MHz*, Geneva, Switzerland: International Electrotechnical Commission.
6. ANSI/ASA S1.20:2012, *Procedures for Calibration of Underwater Electroacoustic Transducers*, New York: American National Standards Institute.
7. IEC FDIS 60565-2 2019, *Underwater acoustics – Hydrophones – Procedures for low frequency pressure calibration*, Geneva, Switzerland: International Electrotechnical Commission.
8. ISO/IEC Guide 98-3, *Uncertainty of measurement –Part 3: Guide to the expression of uncertainty in measurement*, International Electrotechnical Commission. Geneva 2008.
9. W. H. Slater, S. E. Crocker, S. R. Baker, *A primary method for the complex calibration of a hydrophone from 1 Hz to 2 kHz*, *Metrologia*, **55** (2018) 84 – 94.
10. G. Hayman, S. P. Robinson, P. A. Lepper, *The Calibration and Characterisation of Autonomous Recorders used in Measurement of Underwater Noise*, *Advances in Experimental Medicine and Biology*, **875** (2015) 441 – 445.
11. S. Buogo, S. Mauro, J. F. Borsani, S. Curcuruto, A. Biber, A. C. Çorakçi, A. Golick, S. Robinson, G. Hayman, S. Barrera-Figueroa, M. Linné, P. Sigray, P. Davidsson, *Underwater acoustic calibration standards for frequencies below 1 kHz: current status of EMPIR UNAC-LOW project*, *ACTA IMEKO*, **7**(2) (2018) 32 – 38.
12. J. Ward, S. Robinson, *Low Frequency Hydrophone Calibration With A Laser Interferometer*, *Conference Proceedings of the 4<sup>th</sup> International Conference and Exhibition on Underwater Acoustics*, (2019) 33 – 40.
13. A. Isaev, A. Shchelkunov, *Low frequency hydrophone calibration with using tensometric pressure sensor*, *The Journal of the Acoustical Society of America*, **123** 3348 (2008).
14. W. H. Slater, *Extending The Calibration In The Underwater Sound Reference Division (USRD) Reciprocity Coupler To Incorporate Phase*, Master of Science In Engineering Acoustics Thesis, Naval Postgraduate School, Monterey 2016.
15. D. T. Dakin, N. Bailly, J. Dorocicz, J. Bosma, *Calibrating Hydrophones At Very Low Frequencies*, *Proceedings of the 2<sup>nd</sup> International Conference and Exhibition on Underwater Acoustics*, (2014) 373 – 378.
16. R. J. Bobber, *Underwater Electroacoustic Measurements*, Peninsula Publication, Los Altos 1990.
17. A. Malarkodi, G. Latha, PSSR Sridhar, Niels V. Bøgholm, *Design Considerations and Establishment of a Low Frequency Hydrophone Calibration Setup using the Principle of Vibrating Water Column*, *International Journal of Acoustics and Vibration*, **23**(1) (2018) 44 – 48.
18. S. Cui, D. W. Y. Khoo, *Underwater Calibration of Hydrophones at Very Low Frequencies from 30 Hz to 2 kHz*, *J. Phys.: Conf. Ser.*, **1065** 072015 (2018) 1 – 4.
19. K. Listewnik, D. Dobrowolska, *Development of Metrological Infrastructure in the Field of Underwater Acoustics in Poland*, *Joint Conference - Acoustics, Ustka 2018*, 1 – 5.



## Influence of Noise on Children's Sense of Hearing

Adam KACZOROWSKI, Maria KOMASA

*Adam Mickiewicz University, Institute of Acoustics, Faculty of Physics,*

*ul. Uniwersytetu Poznańskiego 2, Poznań*

*akaczor96@gmail.com, maria.komasa@gmail.com*

### Abstract

Children are at risk of developing hearing problems due to their exposure to high sound pressure levels for excessive amounts of time. The aim of this project was to establish the relationship between sound levels in the most hazardous fields and their effect on children's sense of hearing. Measuring sound pressure levels in schools, analysing earphone and headphone usage and comparing them to audiometry results enabled basic dependencies to be found. Some of the examined children fall outside the safe exposure times to certain sound pressure levels (according to WHO directives). The results analysed in frequency bands show that there is a clear correlation between a shift in hearing threshold and intensity levels in music. Noise during lessons is not very high, but the noise during breaks and of bell rings can influence children's hearing to a certain extent. Due to improper usage of earphones and headphones, the young people we examined may develop hearing loss at an early age.

**Keywords:** hearing loss, acoustic reflex, school noise, headphones, earphones, music spectrum

### 1. Introduction

Hearing is one of the most important senses and it's health should be protected. The purpose of our study was to find out whether and to what extent the sound and noise that surround children every day affect their hearing. It is well known, that long-term exposure to loud sounds can negatively influence the sense of hearing. It is especially dangerous for children, because their body is still evolving, which makes them more vulnerable to hearing problems that could affect their entire future life. The sources of sound that are the most dangerous were divided into 3 groups: noise in schools, noise generated by public transport, and music listened to through earphones and headphones.

The acoustic reflex is an important phenomenon these studies. The sound pressure level (SPL), that triggers the stapedius muscle contraction ranges from 70 to 90 dB SPL [1]. The purpose of this reflex is to protect our hearing organs from the high energy sounds, that could potentially have a negative impact on our hearing. When a sound coming to our ears is above the acoustic reflex threshold (ART), the stapedius muscle contracts, decreasing the energy that comes to the cochlea [2, 4]. The stapedius muscle is very small (6.3 mm) and is unable to work for a long time. It's role is to protect our hearing when a short, high energy level sound appears [3, 4].

When the sound level is above the ART and it lasts longer than a few dozen seconds (depending on sound frequency), the muscles simply get "tired" and no longer help us deal with noise [1]. This is the reason why the SPL, which triggers the acoustic reflex has been chosen as a reference for the measurements.

The sounds and noise that surround children and which we have chosen for measurement are partially over ART, and certainly last longer than a few seconds. This means that children's hearing can be harmed. The children that took part in our studies spend from 28 to 31 hours a week at school in lessons and another 5 hours during breaks. On their way to and from school, they sometimes travel by public transport, which can also be noisy (but this depends on the vehicle), and while travelling they also use earphones and headphones. Taking these factors into consideration, exposure to high level sounds might have a serious effect on their hearing, and this constitutes the subject of our research.

## 2. The research group

The research was carried in two schools: the 66 Primary School and the 62 Primary School in Poznań, Poland. They are both district schools, which increases the probability that the children have neither serious hearing problems (which could be qualified as disabilities) nor extraordinarily responsive hearing (which could be possible in music schools or classes). 35 children aged 10-13, from two classes (5th and 6th grade) were examined. None of the participants has ever taken part in any similar project before. We assumed that the level of knowledge and experience concerning the subject of our research is similar in the group as a whole. Nobody reported hearing problems before participating in the project. There were representatives of both head-phone and earphone users among the participants. We treated 12 children that do not use headphones or earphones as a control group.

## 3. Methodology and equipment

To specify the noise sources that affect children's hearing to the greatest extent, three main sources of noise and sound were chosen: schools, public transport and earphones/headphones.

In schools there were three areas to be explored: the noise of bell rings, noise during lessons, and noise during breaks. All these types were analysed in relation to exposure time. When it comes to listening to music, WHO directives might be exceeded, in terms of both loudness and listening time. The SPL in different means of public transport were measured. The influence of this high background SPL on higher music volumes was also investigated.

The first field of research was primary schools (SP). We measured the acoustic background in empty classes and corridors when children were absent. We measured sound levels using SVAN945 and SVAN945A sound level meters. The second part of our study involved measuring sound levels during classes and breaks. We measured the equivalent sound pressure level during lessons (45 minutes) and during breaks (10 minutes). We used the same meters as previously.

The next part of the project was to check the volume at which children listen to music. In order to do this, we used mp4 players (Lenco Xemio-668), in-canal earphones (Pioneer SE-C1T(R)) and over-ear headphones (Motorola Pulse Max). There were music files saved on the devices (Brotherswing by Caravan Palace). Mp4 players were chosen because it is an easy volume adjusting system. There were discontinuous values from 0 to 27

assigned to the actual loudness. The children participating in the project were divided into 3 groups – the ones that use earphones, the ones that uses headphones and the ones not using either. During tram journey, children were given an mp4 player, played a the song and adjusted the music volume as they felt comfortable. The settings - the volume level - for each child were written down. After each listening session we changed the earphone buds and cleaned the headphones, for hygienic purposes. The measurements were made on Poznan's Fast Tram (Poznański Szybki Tramwaj - PST). We chose public transport as a place to take sound volume measurements because most of the children use headphones while traveling.

The third field of the research concerns children's hearing health. A pure tone audiometry test was performed for each child.

The forth part of the research involved measuring how the mp4 settings correspond to the pressure level in dB. To check this, we took a Neumann KU 100 artificial head, and using Bruel & Kjaer Pulse LabShop, calculated how loud the sound on headphones and earphones really was.

The last part of the research involved establishing the sound pressure level of the noise in public transport. We travelled using different types of trams and routes, and buses as well. We measured the equivalent sound pressure level ( $L_{AeqT}$ ) using SVAN945 and SVAN945A sound level meters.

Finally, we took all of our data and copied them into MS Excel. We also used SvanPC and SvanPC++ to read results from the sound level meters and MATLAB to perform some calculations.

## 4. Results and discussion

### 4.1. Schools

The sound levels measured in schools are featured in Tables 1 and 2.

Table 1. Background noise level in SP 62

Class	Background noise level [dB A]
Computer lab	41.3
Chemistry lab	23.4
Classroom on the ground level	31.2
Classroom on the first floor	31.0
Classroom on the first floor close to a street (lights turned off)	40.3
Classroom on the first floor close to a street (lights turned on)	28.4

Table 2. Background noise level in SP 66

Class	Background noise level [dB A]
Class on ground level (lights turned on)	37.2
Class on the first floor (lights turned on)	35.5
Class on the second floor (lights turned on)	36.3
Class on ground level (lights turned off)	25.6
Class on the first floor (lights turned off)	24.5
Class on the second floor (lights turned off)	29.2

Table 3.  $L_{AeqT}$  during classes and breaks

School	Class	$L_{AeqT}$ [dB A]
SP 62	Mathematics	64.8
SP 62	Biology	64.2
SP 62	Break	88.9
SP 66	Mathematics	61.4
SP 66	Substitution	75.4
SP 66	Break	97.3

Table 4.  $L_{AeqT}$  during breakes when the bell rang

School	$L_{AeqT}$ [dB A]
SP 66	98.3
SP 66	100.1

The first conclusion concerns schools and the noise inside them. We are positively surprised, that the background noise is low - usually under 35 dBA. The exceptions are only the computing class, which is probably caused by the fans inside and classes with old, halogen light bulbs – their sound pressure level is high enough to cause an audible difference [Tables 1, 2]. Another conclusion is that  $L_{AeqT}$  during classes is not as high as we had expected – it is lower than the acoustic reflex threshold, so lessons are not hazardous for children's hearing. The only exception was substitutions, but these are a minor part in the weekly teaching schedule, so they should not affect hearing [Table 3]. Unfortunately, breaks and bell rings in schools are dangerous. With sound pressure levels over the acoustic reflex threshold, they can cause serious problems and lead to hearing loss in the future [Table 4]. Children spend approximately 5 hours a week on breaks, which is longer than WHO describes as safe. Unfortunately, due to the limitations of our small research group we can not be sure how it actually affects children, but the average absolute hearing threshold (AHT) is higher than we would expect, considering the age of the children was 10-12. This may be caused by school or public transport noise, but there may also be other threats we did not take under consideration.

#### 4.2. Public transport

To find out how noisy it actually is inside trams and buses, we measured the equivalent sound pressure level using an SVAN 945A sound level meter. We travelled by old and new trams on the PST route and by the 168 and 198 buses, both articulated and ordinary ones. We collected data from 5 minute periods [Table 5].

Table 5.  $L_{AeqT}$  in trams and buses

Model	Line	$L_{AeqT}$ [dB A]
Konstal 105Na	14	83.9
Solaris Tramino S105p	12	79.0
Solaris Tramino S105p	15	77.5
Solaris Urbino 18	198	76.7
Solaris Alpino	168	69.5

#### 4.3. Type of headphones

We examined how the type of headphones used affects hearing. To do so, we measured sound pressure level using artificial head in ear canals with in-canal earphones and under over-ear headphones. Then we compared those values and searched for a correlation between type of headphones and absolute hearing threshold.

Table 6. The relation between the loudness on mp4 players and  $L_{pA}$  on headphones

Volume	$L_{pA}$ [dB A]
10	63.8
12	70.2
15	74.5
16	77.7
20	85.0
22	88.6
24	91.5
27	98.7

Table 7. The relation between the loudness on mp4 players and  $L_{pA}$  on earphones

Volume	$L_{pA}$ [dB A]
8	72.3
12	81.1
13	84.4
15	86.0
17	90.2
18	93.2
19	93.9
21	100.3
27	108.0

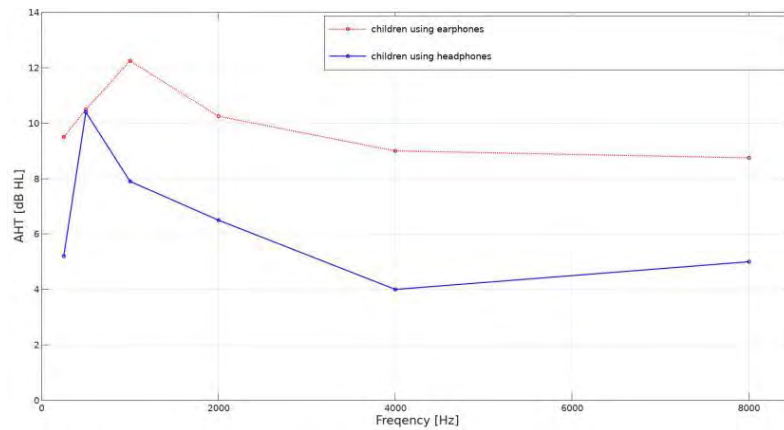


Figure 1. Comparison between headphones and earphones influence

The conclusion is clear - we can see that children that use in-canal earphones have slightly, but significantly higher AHT than children using headphones. As it turns out,  $L_{pA}$  in earphones is much higher than in the headphones, when we set the mp4 loudness to the same volume level, and this may cause a difference in absolute hearing thresholds [Table 6 and 7] [Figure 1].

#### 4.4. Connections between noise and hearing

In our study we wanted to know if and how listening to music using earphones and headphones can affect children's hearing. To do so, we performed pure tone audiometry and then we were looking for correlations.

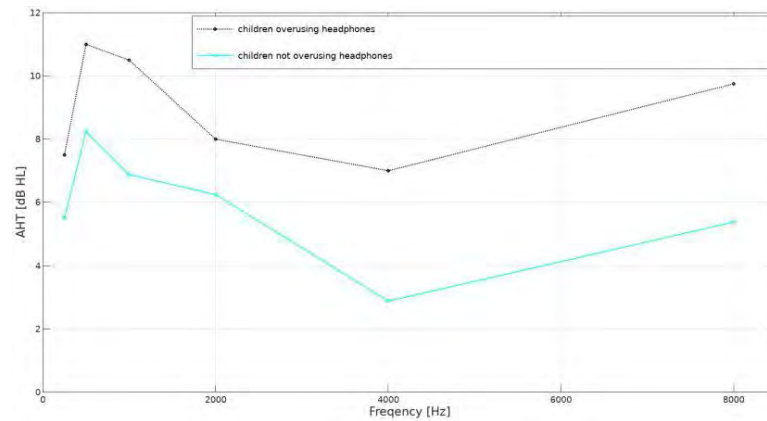


Figure 2. Thresholds

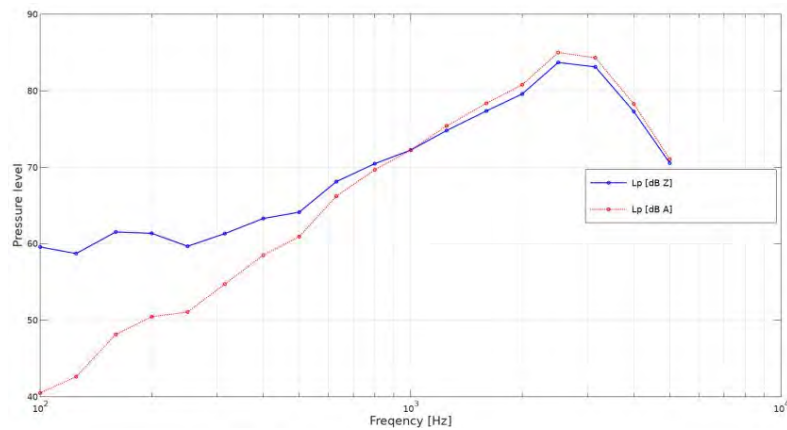


Figure 3. Music spectrum

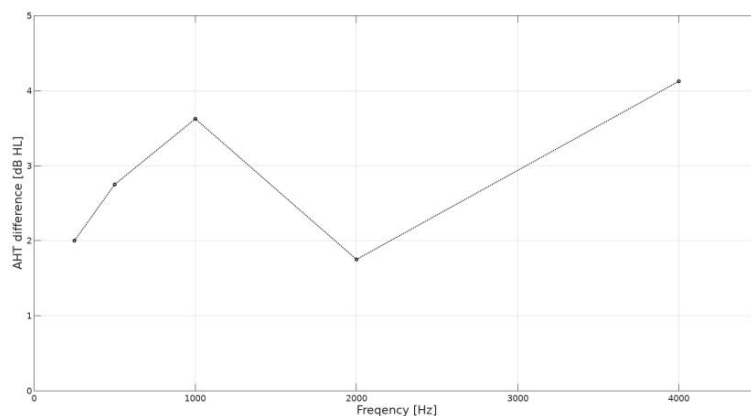


Figure 4. Difference in Absolute Hearing Thresholds

The last conclusion is the influence of earphone and headphone usage. As we can see in Figure 2, children that use earphones and headphones for longer periods or at louder settings (we will say that they overuse headphones) than WHO recommends have a higher absolute hearing threshold. We can see a correlation between a shift of AHT and the music spectrum - with the exception of 2000 Hz [Figure 3 and 4]. It is difficult to tell why this is so. The biggest difference is at around 4000 Hz, where the maximum of music energy is – and although our results might be uncertain, due to the small research group and the difficulties associated with pure tone audiometry, they are backed by theory. It is very worrying that the AHT of such young people is so high. If nothing changes, these children might become hearing-impaired in a few dozen years just because of extensive and improper usage of headphones and earphones.

## 5. Conclusions

The research group was large enough to show basic dependencies, although to gain a better understanding and perform a more thorough examination the group should be bigger. The schools investigated are district schools located in rather quiet surroundings. The outcomes may differ if the circumstances change - for example, if the schools are situated near a loud road, airport or level crossing, it would negatively affect background noise level. The number of children in classes may also play a role in the sound pressure level during lessons and breaks. Public transport can also be very loud, but due to the modernization of trains and buses this problem might be eliminated in a few years, because newer vehicles are quieter. We have only analysed 5 pieces of music of different genres. To get a more precise relationship between music and the absolute hearing threshold, more genres of music should be examined.

We believe that thanks to our study and the workshops that we conducted for children, they will be more aware of possible threats to their hearing. What is more, the schools where we have conducted our research receive feedback from us, so that they can improve the acoustic conditions in schools and make learning more effective and the time spent during breaks safer.

## Acknowledgments

We would like to thank dr Roman Gołębiewski and dr Andrzej Wicher for helping us with this research. We wouldn't be able to do this without their help and knowledge - we are really grateful for all what they have done for us. We appreciate the children, teachers and schools that took part in our study - they have made the research possible. We also would like to say thank you to Stephen Darsley for correction. At the end we would like to greatly acknowledge our reviewers, for effort and time spent on this paper.

## References

1. T. Gierek, A. Slaska-Kaspera, *The stapedius muscle – the present opinions about anatomy and physiology*, Otolaryngologia, Polska 2007.
2. E. Hojan, *Protetyka Słuchu*, UAM, Poznań 2014.
3. World Health Organization, *Toolkit for safe listening devices and systems*, 2019.
4. B. Moore, *Wstęp do psychologii słyszenia*.



## Automatic Recognition of Artificial Reverberation Settings in Speech Recordings

Krzysztof KACHNIARZ\*

*Warsaw University of Technology, Faculty of Electronics and Information Technology,  
Institute of Radioelectronics and Multimedia Technology, Nowowiejska 15/19,  
00-665 Warsaw, [kkachni1@mion.elka.pw.edu.pl](mailto:kkachni1@mion.elka.pw.edu.pl)*

*\*Promity Sp. z o.o., Wiejska 14/25, 00-490 Warsaw*

Marcin LEWANDOWSKI

*Warsaw University of Technology, Faculty of Electronics and Information Technology,  
Institute of Radioelectronics and Multimedia Technology, Nowowiejska 15/19,  
00-665 Warsaw, [marcin.lewandowski@ire.pw.edu.pl](mailto:marcin.lewandowski@ire.pw.edu.pl)*

### Abstract

The aim of this study is to create the method for automatic recognition of artificial reverberation settings extracted from a reference speech recordings. The proposed method employs machine-learning techniques to support the sound engineer in finding the ideal settings for artificial reverberation plugin available at a given Digital Audio Workstation (DAW), i.e. Gaussian Mixture Model (GMM) approach and deep Convolutional Neural Network (CNN) VGG13, which is a novel approach. Training set and data set are 1885 speech signals selected from a EMIME Bilingual Database which were processed with 66 artificial reverberation presets selected from Semantic Audio Labs's SAFE Reverb plugin database. Performance of the proposed automatic recognition method was evaluated using similarity measures between features of reference and analysed speech recordings. Evaluation procedure showed that a classical GMM approach gives 43.8% of recognition accuracy while proposed method with VGG13 deep CNN gives 99.94% of accuracy.

**Keywords:** artificial reverberation, machine learning, digital audio signal processing

### 1. Introduction

Artificial reverberation is one of the most common digital audio effect used in sound, music or video production. There are three different ways that reverberation can be created and added to a signal: convolution-based, delay-networks and physical modelling. Artificial reverberation algorithms have been under development beginning with [1]. Review of this more than 50 year development process can be found in [2] and during this time, uncountable artificial reverberation algorithms have been implemented. These algorithms running as a software plugins are equipped with numerous presets, which are the combination of various reverberation plugin settings. To efficiently create a desired room impression, the sound engineer must be familiar with all of these settings, which are different for each available plugin. Thus finding the best set of reverberation plugin parameters that identifies desired room acoustic features is time-consuming and non-trivial task. Over the years several techniques to simplify workflow with artificial reverberation plugins have been proposed [3-6]. Recently Reiss et al. [7] proposed a design of an adaptive digital audio effect for

artificial reverberation, controlled directly by desired reverberation characteristics, that allows it to learn from the user in a supervised way.

The aim of this study is to create the method for automatic recognition of artificial reverberation plugin settings (preset) extracted from a reference speech recordings. Based on how precisely the system recognizes proper plugin preset (with known low-level reverberation settings), the user can then fine tune individual parameters of the plugin to create desired acoustic impression. This approach is similar to [6], but instead of using GMM-based (Gaussian Mixture Model) system historically used in speaker recognition [8] the proposed method employs deep Convolutional Neural Network (CNN) VGG13-based technique, which was proposed in [9, 10].

## 2. Methodology

The proposed method for automatic recognition of artificial reverberation settings uses VGG13 neural network proposed in [10] and technique for signal preprocessing and feature extraction proposed in [9]. Prior to training phase, a long sequences of silence were removed from input signals by segmenting each audio file into frames and thresholding RMS energy of the frames. The silence threshold was chosen to be -56 dBFS. After silence-removal each file was transformed into log-scaled mel-frequency spectrograms with STFT window size of 1024 samples, hop size of 512 and 64 mel bands. Following feature extraction, each feature vector was split into chunks of a fixed size. The next step was data augmentation called mixup [11] as described in [9]. The VGG13 neural network structure was the same as proposed initially in [10].

The reverberation recognition efficiency of the VGG13 neural network approach was compared with efficiency of the GMM-UBM (Gaussian Mixture Model - Universal Background Model) method. This method has been used successfully for speaker recognition systems over the years and in the work [6] it was used for the artificial reverberation recommendation system. In this work the GMM-UBM method from [12] Matlab Toolbox was used. MFCC features were extracted from input signals instead of log-scaled mel-frequency spectrograms in neural network approach. From each input audio file, 20 MFCC cepstral coefficients were extracted. The window size was set to 25 ms and frame intervals were 10 ms. UBM training parameters were set according to [12].

## 3. Data Set and Reverb Presets

Speech recordings used to train and test VGG13 and GMM-UBM models were obtained from EMIME Bilingual Database [13] i.e. 145 sentences recorded in semi-anechoic chamber by 7 females and 6 males in English language, which makes a total of 1885 several-seconds audio files. The files were sampled at 22 kHz and 16 bit.

Training of VGG13 and GMM-UBM models were conducted with audio files processed with artificial reverberation VST plugin SAFE Reverb [14]. It's an open source software with an open API so there is a lot of various presets made by users. Many of them have parameters that are very similar to each other so based on low-level reverb preset settings cosine similarity was calculated between all of them and 66 presets were selected. To put all 66 reverb presets in context to each other, a classic

[illegible]

Figure 1. Similarity of the reverb presets displayed after multidimensional scaling

- Standard set - randomly chosen 90% of audio files for a training set and 10% of audio files for a test set,
- Sentence independent set (to check if the model performance is independent of the sentence) - 15 specific sentences were selected for a test set and remaining 130 for a training set,
- Speaker independent set (to check if the model performance is independent of the speaker) - one specific male and one specific female voices were selected for a test set.

## 4. Results

To measure the performance of the VGG13 and GMM-UBM models, two metrics were chosen: accuracy and confusion matrix. Accuracy defines how often the model correctly recognizes reverberation settings in the reference audio files (taken from test sets). A confusion matrix summarizes the classification performance of a model with respect to test set. It is a two-dimensional matrix, indexed in one dimension by the true (actual) class of an object, e.g. the reverb setting and in the other by the reverb setting that the model assigns (predicts) from the test set audio files. Reverb settings are labelled as preset names from the SAFE Reverb plugin database and confusion matrix is normalized by class support size (number of elements in each class i.e. each reverb preset).

#### 4.1. Testing with Standard Data Set

First of all, the training of the two models was performed with 90% of the standard data set. Testing was performed with remaining 10% of the standard data set. Model GMM-UBM achieved an accuracy score of 43.8% and VGG13 achieved an accuracy score of 99.94%. Confusion matrix for two models are shown in Fig. 2 and Fig. 3.

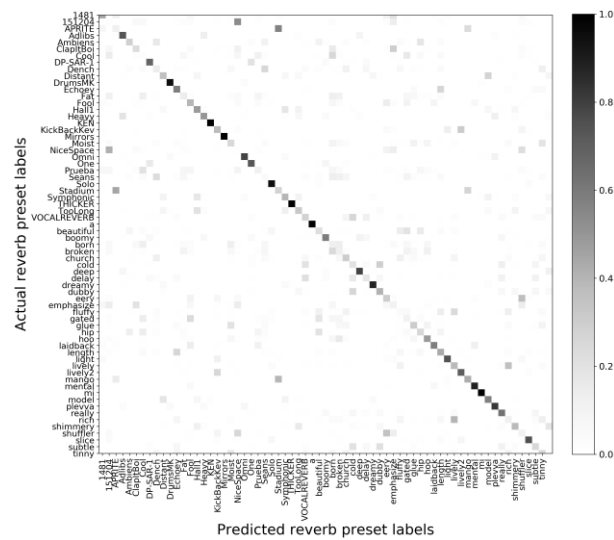


Figure 2. Confusion matrix of GMM-UBM model test with standard data set

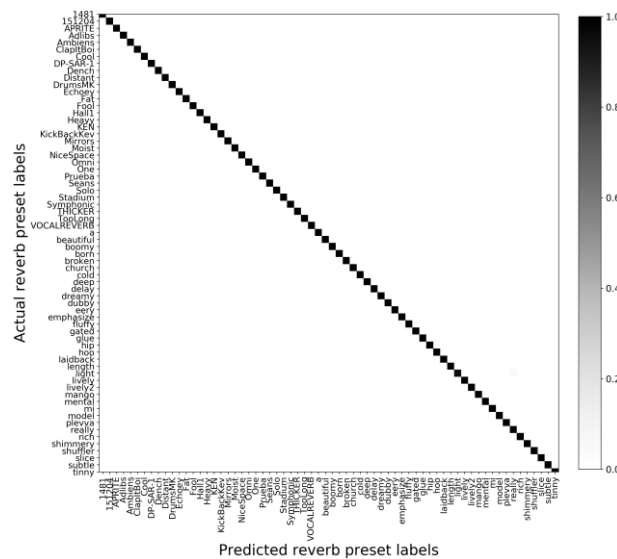


Figure 3. Confusion matrix of VGG13 model test with standard data set

#### 4.2. Testing with Sentence Independent Data Set

Training of the two models was performed with 90% of the sentence independent data set. Testing was performed with remaining 10% of the data set. Model GMM-UBM achieved an accuracy score of 65.05% and VGG13 achieved an accuracy score of 99.88%. Confusion matrix for two models are shown in Fig. 4 and Fig. 5.

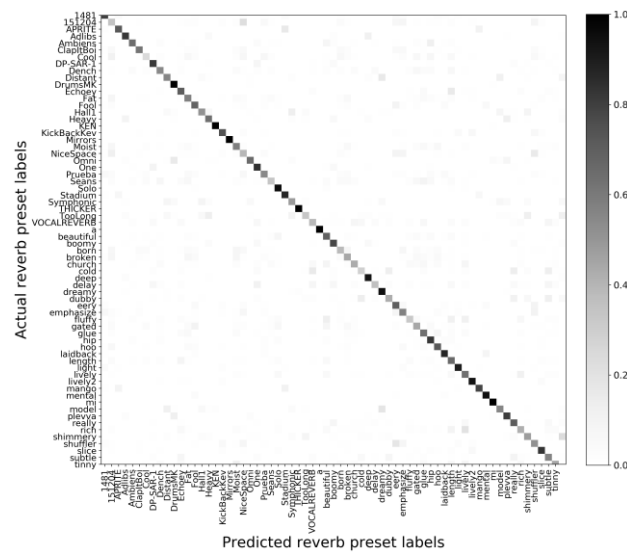


Figure 4. Confusion matrix of GMM-UBM model with sentence independent data set

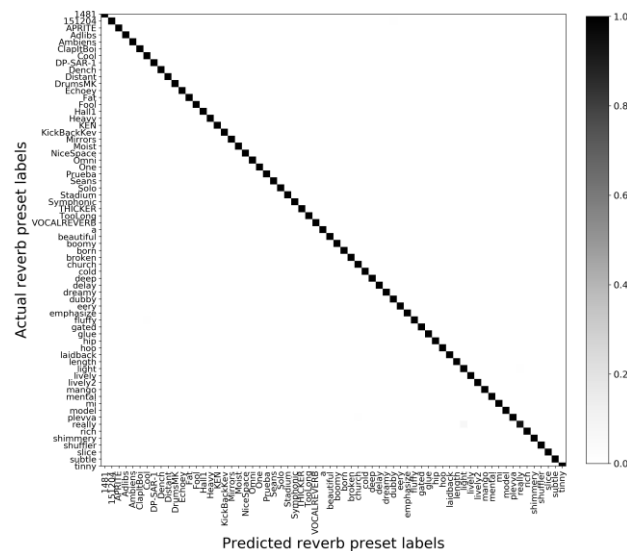


Figure 5. Confusion matrix of VGG13 model test with sentence independent data set

#### 4.3. Testing with Speaker Independent Data Set

Training of the two models was performed with 90% of the speaker independent data set. Testing was performed with remaining 10% of the data set. Model GMM-UBM achieved an accuracy score of 47.88% and VGG13 achieved an accuracy score of 99.36%. Confusion matrix for two models are shown in Fig. 6 and Fig. 7.

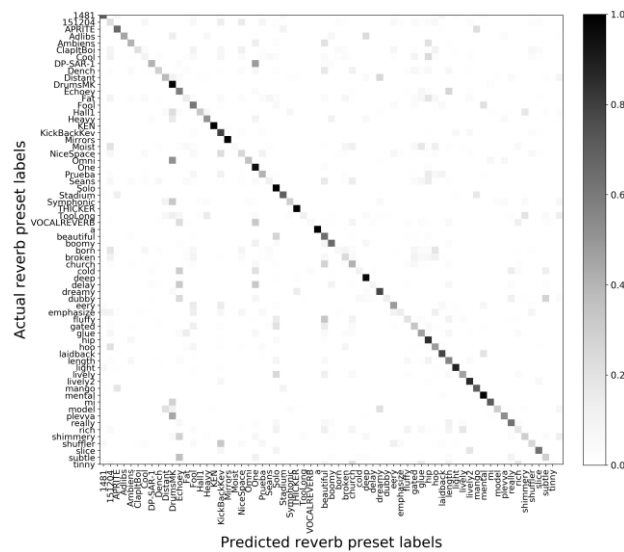


Figure 6. Confusion matrix of GMM-UBM model with speaker independent data set

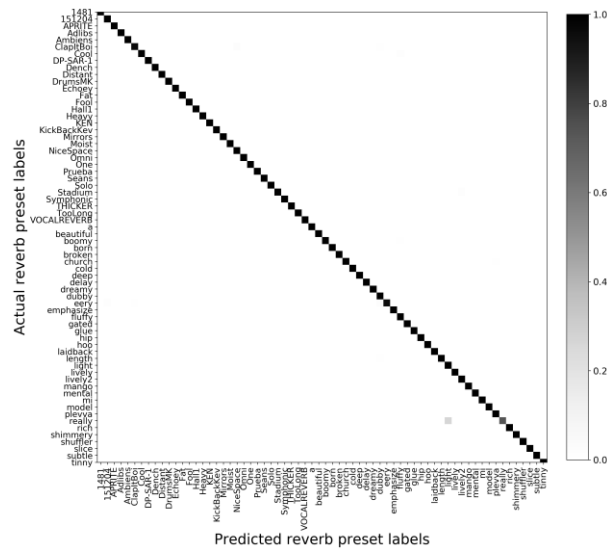


Figure 7. Confusion matrix of VGG13 test model with speaker independent data set

## 5. Conclusions

The new method for automatic recognition of artificial reverberation settings extracted from a features of the reference speech recordings was presented. This method is based on convolutional neural network trained in a supervised way. For comparison, the previously proposed recommendation system based on Gaussian Mixture Models was tested.

Application of that system could be numerous. The main goal of this system is to support the user in finding a reverberation preset that best matches desired room impression. Therefore, that system could support workflow of audio engineers in audio-video postproduction studios, consumer and professional sound studios or even be one of the features in Digital Audio Workstations. Another application of the proposed system could be support for dereverberation algorithms, which are important especially for speech intelligibility improvement. Algorithms of this kind require information about room acoustic's parameters in which recording has been made. The proposed system, based on recognized reverberation preset, could decode low-level reverb information and then help to suppress reverb level in the recording.

Evaluation tests have been conducted for three different data sets. In each case the accuracy obtained with model of convolutional neural network was much higher than accuracy obtained with model based on Gaussian Mixture Models. The recommendations of our model show that the system is almost always able to suggest similar reverb preset.

Future work will focus on exploring accuracy of the system in case of larger amount of presets and plugins. Also, it could be useful to integrate this model into some dereverberation algorithm and test it.

## Acknowledgments

This work was supported by the statutory grant 504/04064/1034/40.00 from the Warsaw University of Technology.

## References

1. M. R. Schroeder, B. F. Logan, *Colorless artificial reverberation*, IRE Transactions on Audio, **6** (1961) 209 – 214.
2. V. Valimaki, J. D. Parker, L. Savioja, J. O. Smith, J. S. Abel, *Fifty years of artificial reverberation*, IEEE Transactions on Audio, Speech, and Language Processing, **20**(5) (2012) 1421 – 1448.
3. J. Jullien, E. Kahle, M. Marin, O. Warusfel, *Spatializer: a perceptual approach*, 94<sup>th</sup> Convention of the Audio Engineering Society, Preprint, **3465** (1993).
4. M. F. Zbyszynski, A. Freed, *Control of VST plug-ins using OSC*, Proc. of the International Computer Music Conference, Spain 2005, 263 – 266.
5. Z. Rafii, B. Pardo, *Learning to control a reverberator using subjective perceptual descriptors*, 10<sup>th</sup> International Society for Music Information Retrieval (2009).

6. N. Peters, J. Choi, H. Lei, *Matching Artificial Reverb Settings to Unknown Room Recordings: a Recommendation System for Reverb Plugins*, 133<sup>rd</sup> Audio Engineering Society Convention, USA 2012.
7. E. T. Chourdakakis, J. D. Reiss, *A machine-learning approach to application of intelligent artificial reverberation*, Journal of the Audio Engineering Society, 2017.
8. D. Reynolds, T. Quatieri, R. Dunn, *Speaker verification using adapted gaussian mixture models*, Digital Signal Processing, **10**(1-3) (2000) 19 – 41.
9. T. Iqbal, Q. Kong, M. Plumbley, W. Wang, *Stacked Convolutional Neural Networks For General-Purpose Audio Tagging*, Centre for Vision, Speech and Signal Processing, University of Surrey 2018.
10. K. Simonyan, A. Zisserman, *Very deep convolutional networks for large-scale image recognition*, 3<sup>rd</sup> ICLR 2015.
11. H. Zhang, M. Cisse, Y. N. Dauphin, D. Lopez-Paz, *mixup: Beyond empirical risk minimization*, 6th ICLR 2015.
12. S. O. Sadjadi, M. Slaney, L. Heck, *MSR Identity Toolbox*, Microsoft Research 2013.
13. M. Wester, *The EMIME Bilingual Database*, The University of Edynburg: Centre for Speech Technology Research 2012.
14. Semantic Audio Labs, [Online], <http://www.semanticaudio.co.uk/>, [Access: 10.07.2019].



## Classes of Tonality of Signals in the Aspect of Active Elimination of Tonal Components

Michał ŁUCZYŃSKI

*Wrocław University of Science and Technology,*

*Chair of Acoustics and Multimedia,*

*michal.luczynski@pwr.edu.pl*

### Abstract

The aim of this paper is to analyse various types of signals defined as tonal where energy is concentrated in a narrow band of the spectrum. Not all tonal noises could be reduced using narrowband active noise control systems with cancelling signal generated based on source parameters. The author proposes the following classes of tonality: periodical (e.g. sinusoidal) signals, sinusoidal signals modulated by random function, sinusoidal signals with increased/decreasing amplitude and/or frequency, a combination of the two previous ones, and narrowband noises. For each type of the analysed signals active elimination of the tonal component using synthesized cancelling signal was carried out. Depending on the type of the signals different results were obtained which indicates the need to clarify the definition of a tonal signal, or use not one concept of a tonal signal, but several classes discussed above.

**Keywords:** tonality, signal processing, Active Noise Control

### 1. Introduction

The aim of this paper is to analyse different types of signals which can be described as tonal signals [1]. The author indicated that the concept of tonality should be clarified especially in the aspect of narrowband active noise control systems [2, 3] with cancelling signal generated based on source parameters [4, 5]. Not all tonal noises that meet the definition of tonal signals could be reduced efficiently using such systems [6]. Therefore the author introduces the classes of signal tonality.

Tonal signals is defined as signals where energy is concentrated in a narrow band of the spectrum [1, 7]. This definition applies to signals containing one or more tonal components. Therefore, it can be assumed that the tonal character of the signal means that in the spectrum there are frequency components (one or more) whose energy is concentrated around discrete frequency values (tonal components). This definition includes both pure tone and narrowband noise.

In the further part of the paper analysed tonal components were the only frequency component of the signal or one of many components observed in the signal spectrum. An individual analysis of a single component was possible due to the use of bandpass filters. The impact of the signal-to-noise ratio was not analysed.

### 2. Classes of tonality of signals

Fragments of analysed signal are shown at figure 1.

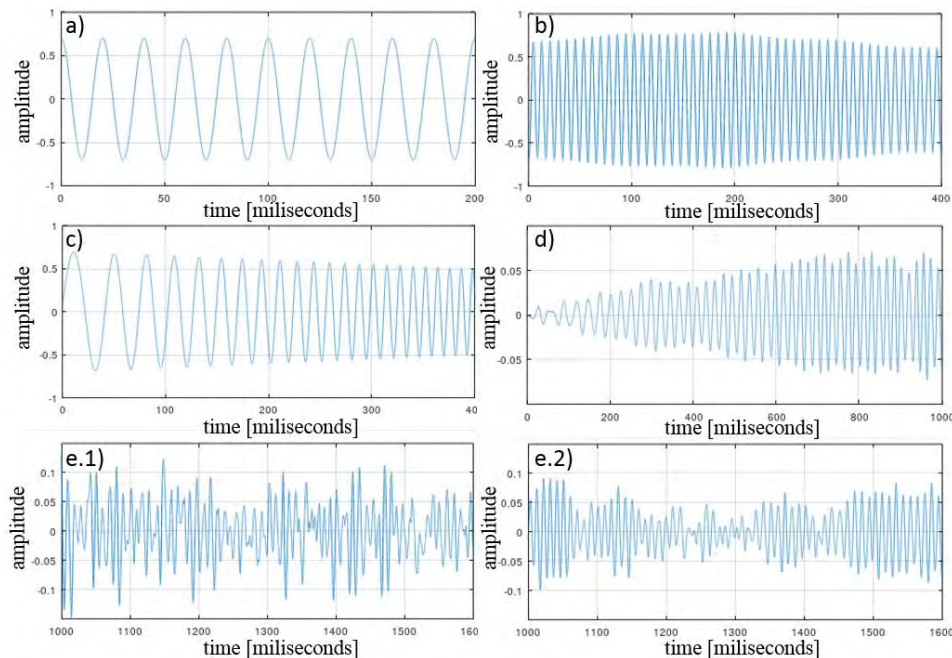


Figure 1. Chosen fragments of analyzed signals: a) stationary periodic signals, b) periodic signals modulated by a random function, c) periodic signal with increasing/decreasing amplitude and/or frequency, d) combination of b and c, e.1) narrowband noise  $Q = 10$ , e.2) narrowband noise  $Q = 50$

The types of tonal components depending on the variability of parameters (amplitude, frequency) are presented and listed below:

a) Stationary periodic signals (figure 1. a):

The simplest class of analysed signals. A single tonal component is a sinusoidal signal. The problem of the elimination of the tonal component is based on the correct detection of the parameters (amplitude, frequency, initial phase), synthesis a cancelling signal and addition to the primary signal.

Quasi-stationary signals:

A signal described in the previous point are an idealized model. Real signals are characterized by a variability of parameters over time. The following sections describe possible changes, both small random modulation of parameters and changes their values, such as increasing (or decreasing) their amplitude and/or frequency.

b) Periodic signals modulated by a random function (figure 1. b):

In order to present this class of signals, a sound of two-stroke engine at constant revolutions of about 7200 RPM was analyzed. The recording was carried out under controlled conditions [8, 9] in Acoustic Chamber of Department of Acoustics and Multimedia at Wrocław University of Science and Technology. The Acoustic

Chamber has one sound reflected surface (floor). Other surfaces are sound absorbing. Separate tonal component was obtained using frequency filtration. The signal is not an ideal pure tone with constant parameters.

- c) Periodic signals with increasing/decreasing amplitude and/or frequency (figure 1. c):

This class of signals is characterized by the change of parameters in different way than random variations. Increasing or decreasing of parameters (both amplitude and frequency) could be caused i.e. by changing the number of revolutions of motor or fan. In order to present this class of signal a frequency sweep was analysed. Short time analysis of sweep indicates that it is a narrowband signal, however using long time analysis it is supposed to be a broadband signal. The author suggests that frequency sweep should be considered as tonal signal.

- d) Combination of the two previous ones (periodic signals with increasing/decreasing amplitude and/or frequency modulated by a random function (figure 1. d):

While analysing tonal components of real signals, such as a sound of engine with changing the number of revolutions, it could be observed both increasing or decreasing parameters and some random variation described as frequency and amplitude modulation with non-deterministic function. The analysis of separated single tonal component of sound of combustion engine while increasing number of revolution was performed to define this class of tonal signals.

- e) Narrowband noise (figure 1. e)

The last type of signal that is included in the definition of tonal signals (energy is focused around discrete frequency values) is the signal obtained by frequency filtration of noise using a narrow band pass filter (quality of filter  $Q = 10$  figure 1.e.1 and  $Q = 50$  figure 1.e.2). Analysis was performed for two types of narrowband noises: digitally and analogue generated.

### 3. Experiment

For each of the types of signals, an experiment was carried out using MATLAB environment. A signal tonal component was eliminated using active noise reduction algorithms with generated cancelling signal (pure tone synthesis). The parameters of cancelling signal were obtained based on parameters of the primary signal.

The stages of the experiment are:

- a. Preliminary time and frequency analysis,
- b. Extraction of a single tonal component,
- c. Analysis of the variability of its parameters,
- d. Synthesis of a cancelling signal based on parameters of tonal component of primary signal,
- e. Active elimination of the tonal component (using different length of window - frame),
- f. Analysis obtained results.

The block diagram of the algorithms used in the experiment is shown in Figure 2.

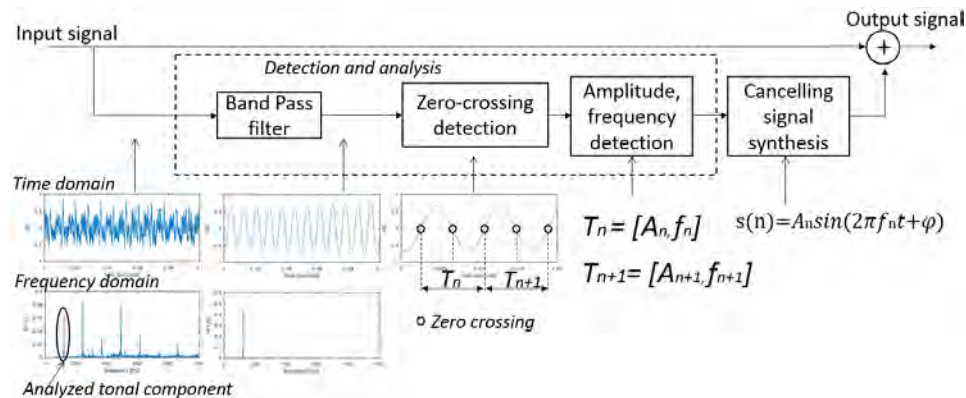


Figure 2. The block diagram of the algorithms used in the experiment. Under the algorithm there are shown an example of signals at subsequent stages of processing

The use of this processing method is possible if the differences between the frequencies of tonal components are sufficient to separate the components. The signals have been selected to obtain desirable results. Using the zero-crossing detection allows to divide a signal of tonal component into periods and determine amplitudes and frequencies in subsequence periods. These parameters were used for generation of a cancelling signal. Then an elimination using active noise cancellation was performed. Different length of time window (frame length) were used. The shortest frame is the length of one period and cancelling signal is a pure tone with a given frequency and amplitude such as that detected in a given period of primary signal (as in Figure 2). In the case of longer frame the parameters of cancelling signal are the average frequency and amplitude in the range of given frame of primary signal or the amplitude and frequency of the first period in a given frame (for signals with increased frequency). Phase mismatch problems were minimized due to zero-crossing detection. The process of elimination was initialized when primary signal crossed zero. The evaluation of a given class of signal was based on the observation of the output signal. The shape of output signal depended on difference between the original signal and the cancelling signal. This implies the differences between individual classes.

#### 4. Results

##### a) Stationary periodic signals

The signal was generated using MATLAB environment. The pure tone had given and amplitude, frequency and initial phase. The frequency was 50 Hz. The sampling rate was 44.1 kHz. There was no need for frequency filtration. Zero-crossing points were detected and then amplitude and frequency. The parameters were constant in the subsequent periods. The cancellation signal was synthesized. The length of a frame does not affect the results. The output signal obtained a level close to zero. Non exact zero values were caused by calculation resolution.

## Quasi-stationary signals

## b) Periodic signals modulated by a random function:

The analysed signal of separated tonal component of the sound of combustion engine contains 485 periods. Amplitudes and frequencies are noticed in subsequence periods. It turned out that the values are not constant as in the case of pure tone. This values are presented at histograms (figure 3).

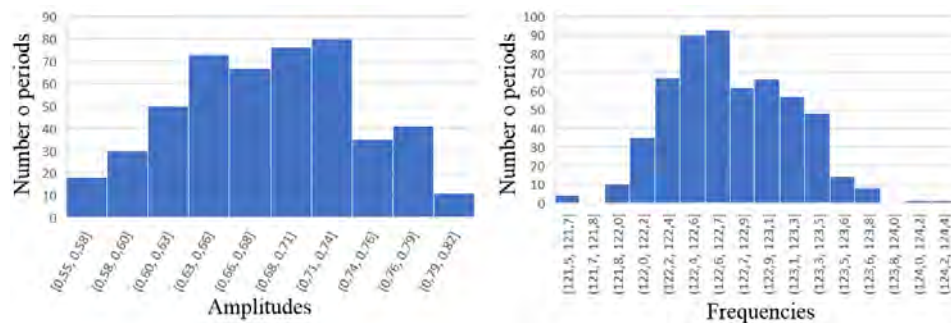


Figure 3. Histograms of amplitudes (left) and frequencies (right)

Amplitude and frequency values are some kind of statistical distribution. It was assumed that the parameters are constant mean values modulated by a non-deterministic signal. The amplitude/frequency can be defined not as a single number, but as a random variable. The average value of amplitudes is 0.68, maximum value is 0.82 and minimum is 0.55. Standard deviation is 0.06. The average value of frequencies is 122.7 Hz, maximum value is 124.3 Hz and minimum is 121.5 Hz. Standard deviation is 0.4 Hz.

The tonal component was eliminated for different length of time windows. Figure 4 presents the example results of experiments for chosen fragments of signal. Figure 4a shows the elimination result for window length (frame) is 1 period (for each period a cancelling signal was generated pure tone with the frequency and amplitude read for a given period of primary signal). Figure 4b shows the elimination result for window length of 10 periods. The parameters of cancelling signal (pure tone) were the average amplitude and frequency values in the range of 10 periods.

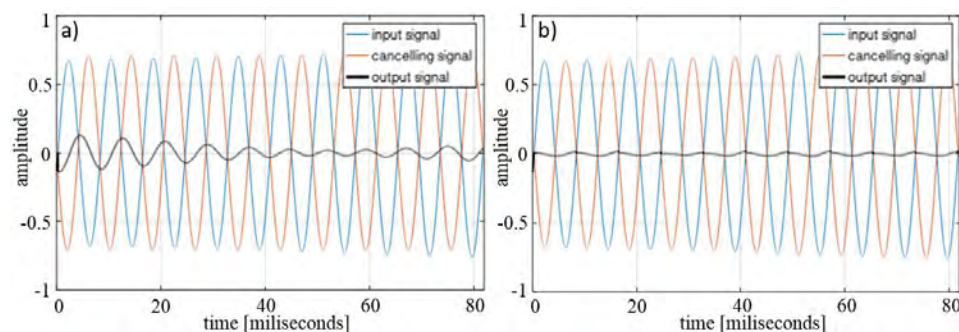


Figure 4. Results of the experiment for periodic signal modulated by a random function:  
a) frame length of 1 period, b) frame length of 10 periods

In both cases, it could be observed that there is no complete elimination of the given component. This is caused by the mismatch of the primary and cancelling signal. As well it could be seen that for the shorter frame the amplitude of output signal is lower (an errors of elimination are smaller). However, in most cases, the smaller number of signal samples causes higher detection errors, especially when using discrete Fourier transform based detection algorithms. For this class of signals, the efficiency of tonal components elimination is limited not only by detection errors as in the case of ideal periodical signals but also by random modulation variability.

c) Periodic signals with increasing/decreasing amplitude and/or frequency

Analysed frequency sweep with decreasing amplitude starts at 20 Hz and ends at 200 Hz after 1 second. Figure 5 shows six periods of an input (primary) signal with cancelling signal and results of the elimination (output signal). The parameters of the cancelling signal (pure tone) are the amplitude and the frequency detected in the first period of the primary signal.

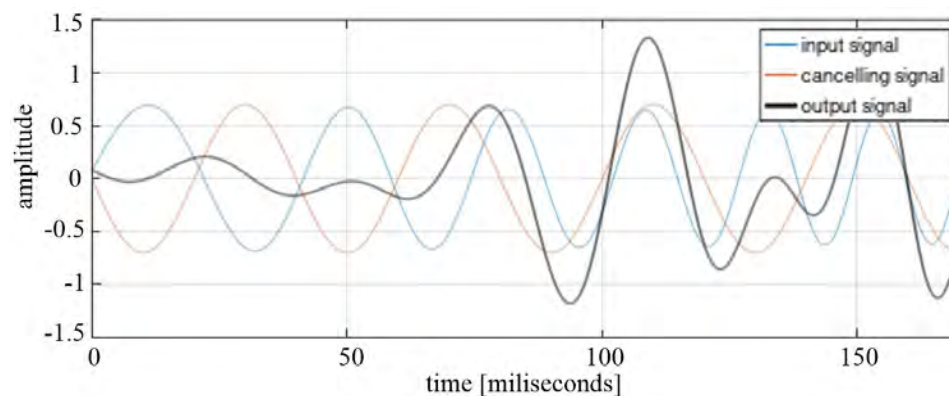


Figure 5. Results of the experiment for periodic signal with decreasing amplitude and increasing frequency (first six period of primary signal)

The mismatch of the primary and cancelling signal increased in subsequent periods so the tonal component is not eliminated efficiently. In the first period the level of tonal component is reduced but then the level of output signal is even higher than the input signal. The effective reduction of a tonal component level might be possible if a frame length is corresponded to changes of parameters.

d) Combination of the two previous ones (periodic signals with increasing/decreasing amplitude and/or frequency modulated by a random function

The sound of an internal combustion engine while increasing the number of revolution was analyzed to present this class of tonal signal. Both the change in amplitude and frequency resulting from increasing the number of engine revolutions and the modulation of these values by non-deterministic function are observed. Frequency filtration by bandpass filter with a variable middle frequency was used to separate

the selected frequency component. For the selected fragment where the number of revolution increase the values of the amplitude and frequency in subsequent periods were detected. The results are shown in Figure 6.

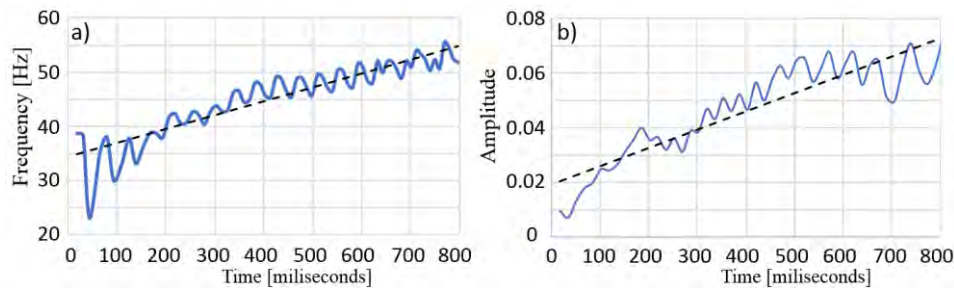


Figure 6. a) frequency and b) amplitude detected in subsequent periods of the primary signal of the engine while increasing number of revolutions with trend line marked

The two types of value changes described in the previous paragraphs (b and c) could be seen. This affects the tone component elimination efficiency. Elimination will be ineffective for longer frames (windows) such as for the sweep signal. If the frame is short there is no complete elimination because the changes take place even in single period.

#### e) Narrowband noise

The last class defined in this paper is a narrowband noise. It is non-deterministic, random signal. No significant differences between analogue and digitally generated noise were observed. However the parameter whose impact is highly important is bandwidth. A few experimental elimination were performed for noises obtained by filtering with filters with different quality [8, 9].

Quality of filters was 10, 50, 100, 200 and 500. Middle frequency was 1 kHz. For two first noises the level of output signal was higher than input signal. For noises with narrower bands the level decreased (even 12.1 dB for noised filtered by filter with  $Q = 500$ ).

## 4. Conclusions

Depending on the type of narrowband signal, one of the above, different results were obtained. This indicates the need to clarify the definition of a tonal signal or the use of several classes as presented above instead of single concept of tonal signal. Pure tone (periodical signal) is an ideal model of tonal signal. Other signals differ from the ideal model. For pure tones modulated by a non-deterministic (random) function, parameters for each subsequent period of the signal can take any value in the modulation range. However, these values are similar. The extreme case of these signals is narrowband noise. Changes of values of parameters depend on the noise bandwidth. The model of the signal with decreasing / increasing amplitude and/or frequency is an idealized model of signal. In the subsequent periods of a signal, a certain tendency (trend) of parameter



changes is noticeable. Such signals have their representation among real signals if there is simultaneous modulation of the parameters with a random function.

The author of this paper does not deny the definition of tonality as energy concentrated around the discrete frequency values. However, in the aspect of Narrowband ANC systems, the author recommends to define total signal more precisely. It could help to choose the parameters of an algorithm such as length of a frame and predict the effectiveness of the noise reduction. Signal interpretation in accordance with the presented classes may also have potential application in the case of predictive systems. Understanding the nature of signal variability helps in predicting the future values of parameters and minimize errors resulting from the mismatches primary and cancelling signals caused by processing latency.

### Acknowledgments

I would like to thank my main supervisor Prof. Andrzej Dobrucki and second supervisor Ph.D Stefan Brachmański for support.

### References

1. ISO 1996-2:2017, *Acoustics -- Description, measurement and assessment of environmental noise -- Part 2: Determination of sound pressure levels*.
2. S. M. Kuo, D. R. Morgan, *Active Noise Control: A Tutorial Review*, Chapter III. NARROWBAND FEEDFORWARD ANC, Proceedings of the IEEE, **87**(6) 1999 943 – 973.
3. S. Ryu, Y. J. Park, Y.-S. Lee, *Active Suppression of Narrowband Noise by Multiple Secondary Sources*, Journal of Sensors, Volume 2016, Article ID 6276828, 9 pages.
4. X. Qiu, X. Lia, Y. Aib, C. H. Hansen, *A waveform synthesis algorithm for active control of transformer noise: implementation*, Applied Acoustics, **63** (2002) 467 – 479.
5. P. Górski, W. M. Zawieska, *System aktywnej redukcji hałasu o przebiegu okresowym*, Bezpieczeństwo pracy 9/2002, 5 – 6 (in Polish).
6. M. Łuczyński, S. Brachmański, A. Dobrucki, *Active elimination of tonal components in acoustic signals*, 2018 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), Poznan, 2018, 355 – 360.
7. Z. Engel, G. Makarewicz, L. Morzyński, W.M. Zawieska, *Metody Aktywnej Redukcji Hałasu, Chapter 8.9 Aktywna redukcja hałasu transformatorów energetycznych*, Warsaw 2001 (in Polish).
8. M. Luczynski, S. Brachmanski, *Mathematical Model of the Acoustic Signal Generated by the Combustion Engine*, AES Convention: 142 (May 2017) Paper Number: 9717.
9. M. Łuczyński, *Model matematyczny sygnału akustycznego generowanego przez silnik spalinowy (eng. Mathematical model of the acoustic signal generated by the combustion engine)*, Master thesis, Wrocław University of Science and Technology, 2016.



## Prediction of Low-Frequency Sound Field in Rooms with Complex-Valued Boundary Conditions on Walls

Mirosław MEISSNER

*Institute of Fundamental Technological Research, Polish Academy of Sciences,  
Pawińskiego 5B, 02-106 Warsaw, Poland, mmeissn@ippt.pan.pl*

### Abstract

A modal representation of a room impulse response has been used to formulate expressions for low-frequency sound field in rooms of arbitrary shape. Based on theoretical results, a simulation program has been developed to predict a sound pressure distribution and a room transfer function for rectangular enclosure having walls covered by a material of complex impedance. Calculation results have shown that changes in the wall reactance entail a substantial modification of a sound pressure distribution. Furthermore, an influence of wall reactance on the room transfer function was investigated and it was discovered that a change in a reactance sign causes a shift in frequencies of modal vibrations excited in the room.

**Keywords:** room acoustics, modal vibrations, room impulse response, complex wall impedance

### 1. Introduction

The main objective of theoretical room acoustics is to investigate the steady-state and transient acoustic behaviors of enclosed spaces. There are many theoretical methods for modeling a sound field inside enclosures and among them are diffusion-equation models, geometrical approaches, wave-based methods and modal expansion methods. The diffusion-equation models [1] are an extension of the statistical theory to spatially varying sound field. Geometrical approaches are suitable for high sound frequencies and the ray tracing method [2], the beam tracing algorithm [3] and the mirror source technique [4] are the most popular methods for geometric modeling. In contrast to geometric approaches, wave-based methods provide a complete description the sound field because they solve the wave equation after suitable space discretization. The most common among these numerical techniques are the finite element method [5], the boundary element method [6] and the finite-difference time-domain method [7]. Modal expansion methods yield the acoustic modes of pressure vibrations inside enclosures and the sound field is expressed as a linear combination of these modes [8]. Modal expansion approaches are more difficult to apply for irregularly shaped rooms [9], but they fully describe a wave nature of the sound field like a diffraction and a creation of standing waves. They also enable to identify typical modal effects such as a modal degeneracy [10] and a localization of modes [11].

In the paper, the modal expansion method is used to predict a low-frequency sound field in rooms with complex-valued boundary conditions on walls. A theoretical modeling is accompanied with a numerical simulation performed for a rectangular enclosure with uniform and frequency constant impedance on room walls. Based on calculation results, the effect of complex wall impedance on a distribution of a pressure

amplitude is investigated and changes in the room transfer function with this impedance are analyzed.

## 2. Theoretical model

In a low-frequency range, room dimensions are comparable with a length of sound wave and the method, which is most appropriate for determining an interior sound field, is a modal analysis. According to this method, the room response can be described as a superposition of responses of acoustic modes excited in a room by a sound source, i.e.,

$$p(\mathbf{r}, t) = \sum_{m=1}^{\infty} p_m(t) \Phi_m(\mathbf{r}), \quad (1)$$

where  $\mathbf{r} = (x, y, z)$  is the position coordinate of a receiver, the functions  $p_m$  determine a temporal behavior of the sound pressure and  $\Phi_m$  are the real-valued eigenfunctions which satisfy the orthonormal property in the volume  $V$  of the room. In a theoretical model it is assumed that room walls are covered by a sound absorbing material with a complex impedance, thus, the pressure  $p$  fulfills the following boundary condition

$$\nabla p \cdot \mathbf{n} = -\frac{1}{c\zeta} \frac{\partial p}{\partial t}, \quad (2)$$

where  $\nabla$  is the nabla vector operator,  $\mathbf{n}$  is the unit vector normal to the walls, which is directed away from the room volume, and  $c$  is the sound speed. In Eq. (2) the quantity  $\zeta$  represents the specific impedance of a wall material and  $\zeta = \zeta_r + j\zeta_i$ , where  $\zeta_r$  and  $\zeta_i$  are a wall resistance and a wall reactance, respectively. It is assumed that the magnitude  $|\zeta|$  of the wall impedance is much larger than unity because typical materials covering room walls are characterized by a small sound absorption in the low-frequency range [12].

The procedure for finding the function  $p_m$  relies on a suitable solution of a wave equation and using the method presented in [13], it can be shown that for small sound damping in a room the function  $p_m$  is a solution of the following equation

$$\frac{\partial^2 p_m}{\partial t^2} + 2\xi_m \frac{\partial p_m}{\partial t} + \omega_m^2 p_m = c^2 \int_V q(\mathbf{r}', t) \Phi_m(\mathbf{r}') d^3\mathbf{r}', \quad (3)$$

where  $\omega_m$  is the natural eigenfrequency,  $q$  is the volume source term in the wave equation,  $d^3\mathbf{r}' = dx'dy'dz'$  is the volume element and the parameter  $\xi_m$  is given by

$$\xi_m = r_m + j\varphi_m = \int_S \frac{(\zeta_r - j\zeta_i) \Phi_m^2(\mathbf{r})}{|\zeta|^2} d^2\mathbf{r}, \quad (4)$$

where  $S$  is the surface area of room walls and  $d^2\mathbf{r}$  is the surface element. Since the wall resistance  $\zeta_r$  is non-negative, the parameter  $r_m$  represents the modal damping factor.

The impulse room response corresponds to the case of an impulsive temporal excitation of a room by a sound source located at a point. Thus, assuming that the volume source term in Eq. (3) has the form  $q(\mathbf{r}', t) = \delta(\mathbf{r}' - \mathbf{r}_0) \delta(t - t_0)$  one can obtain

$$\frac{\partial^2 p_m}{\partial t^2} + 2\xi_m \frac{\partial p_m}{\partial t} + \omega_m^2 p_m = c^2 \delta(t - t_0) \Phi_m(\mathbf{r}_0), \quad (5)$$

where  $\mathbf{r}_0 = (x_0, y_0, z_0)$  is the source position and  $t_0$  is the time of impulse generation. A method for resolving Eq. (5) was described in [14] and the obtained result is as follows

$$p_m(t) = \frac{c^2 e^{-\xi_m(t-t_0)} \sin[\Omega_m(t-t_0)] \Phi_m(\mathbf{r}_0)}{\Omega_m}, \quad (6)$$

where  $\Omega_m$  is the complex eigenfrequency for damped modal vibrations given by

$$\Omega_m = a_m + j b_m = \sqrt{\frac{c_m + \sqrt{c_m^2 + d_m^2}}{2}} + j \sqrt{\frac{-c_m + \sqrt{c_m^2 + d_m^2}}{2}} \quad (7)$$

where the quantities  $c_m$  and  $d_m$  are determined by

$$c_m = \omega_m^2 - r_m^2 + \varphi_m^2, \quad (8)$$

$$d_m = 2r_m \varphi_m. \quad (9)$$

A substitution of Eq. (6) into Eq. (1) leads to the function of the form

$$h(\mathbf{r}_0, \mathbf{r}, t) = c^2 \sum_{m=1}^{\infty} \frac{e^{-\xi_m t} \sin(\Omega_m t) \Phi_m(\mathbf{r}_0) \Phi_m(\mathbf{r})}{\Omega_m} \quad (10)$$

describing the room impulse response (RIR) between the sound source at the position  $\mathbf{r}_0$  and the receiver located at the point  $\mathbf{r}$ . Because of the causality condition, the RIR function  $h(\mathbf{r}_0, \mathbf{r}, t)$  is zero for  $t < 0$ . It satisfies also the reciprocity principle because the right side of Eq. (10) is a symmetric function of the source and receiver points coordinates.

The room impulse response is very useful in room acoustics because a knowledge of the RIR function enables to predict the room response to any sound source. In fact, when a volume source in the wave equation is described by the source function  $q$ , the pressure response to this excitation can be found from the following expressions [15]

$$p(\mathbf{r}, t) = \int_V q(\mathbf{r}', t) * h(\mathbf{r}', \mathbf{r}, t) d^3 \mathbf{r}' = \int_V \int_{-\infty}^t q(\mathbf{r}', \tau) h(\mathbf{r}', \mathbf{r}, t - \tau) d\tau d^3 \mathbf{r}', \quad (11)$$

where the asterisks denotes a convolution operation. The steady-state room response to a point source can be found assuming that in Eq. (11) the source function  $q$  takes the form

$$q(\mathbf{r}', t) = Q \delta(\mathbf{r}' - \mathbf{r}_0) e^{j\omega t}, \quad (12)$$

where  $\omega$  is the angular source frequency and the amplitude  $Q$  of a sound excitation is dependent on the source power  $W$  according to the formula  $Q = \sqrt{8\pi\rho c W}$ , where  $\rho$  is the air density. Thus, after performing the volume and time integrations in Eq. (11), a formula for the steady-state sound pressure is found to be as follows

$$p(\mathbf{r}, t) = P(\mathbf{r}) e^{j[\omega t + \phi(\mathbf{r})]}, \quad (13)$$

where  $P$  is the steady-state pressure amplitude

$$P(\mathbf{r}) = \sqrt{\left[ \sum_{m=1}^{\infty} \alpha_m \Phi_m(\mathbf{r}) \right]^2 + \left[ \sum_{m=1}^{\infty} \beta_m \Phi_m(\mathbf{r}) \right]^2}, \quad (14)$$

$\phi$  is the pressure phase given by

$$\phi(\mathbf{r}) = \arctan \left[ \frac{\sum_{m=1}^{\infty} \beta_m \Phi_m(\mathbf{r})}{\sum_{m=1}^{\infty} \alpha_m \Phi_m(\mathbf{r})} \right] \quad (15)$$

and the quantities  $\alpha_m$  and  $\beta_m$  are determined by

$$\alpha_m = \frac{Qc^2 [r_m^2 + a_m^2 - b_m^2 - (\omega + \varphi_m)^2] \Phi_m(\mathbf{r}_0)}{\left[ (r_m + b_m)^2 + (\omega + \varphi_m - a_m)^2 \right] \left[ (r_m - b_m)^2 + (\omega + \varphi_m + a_m)^2 \right]}, \quad (16)$$

$$\beta_m = \frac{-2Qc^2 [r_m(\omega + \varphi_m) + a_m b_m] \Phi_m(\mathbf{r}_0)}{\left[ (r_m + b_m)^2 + (\omega + \varphi_m - a_m)^2 \right] \left[ (r_m - b_m)^2 + (\omega + \varphi_m + a_m)^2 \right]}. \quad (17)$$

As it results from Eqs. (14), (16) and (17), the amplitude  $P$  is dependent on the source position  $\mathbf{r}_0$  and the source frequency  $\omega$  and, through the quantities  $a_m$ ,  $b_m$ ,  $r_m$ , and  $\varphi_m$  on the natural eigenfrequency  $\omega_m$  as well as the real and imaginary parts of the specific wall impedance  $\zeta$ . Thus, for constant  $\mathbf{r}_0$  and given source frequency  $\omega$ , Eq. (14) enables one to predict a spatial distribution the steady-state pressure amplitude for different values of  $\zeta_r$  and  $\zeta_i$ . On the other hand, when the source and observation coordinates  $\mathbf{r}_0$  and  $\mathbf{r}$  are specified, Eq. (14) makes possible to determine the low-frequency room transfer function for various values of  $\zeta_r$  and  $\zeta_i$  at a given point of the room.

### 3. Analysis of simulation results and conclusions

In the first part of a numerical study, an influence of the absorbing material impedance on a spatial distribution of the steady-state pressure amplitude will be investigated. In general, the developed theoretical model is valid for arbitrary room shapes. However, to perform numerical tests for wider frequency range, a rectangular enclosure is considered because modal vibrations in such a room are well understood and described. The room has the dimensions:  $l_x = 7$  m,  $l_y = 5$  m,  $l_z = 3$  m, and its walls are covered uniformly by an absorbing material with frequency constant complex impedance  $\zeta$ . Since a small sound damping on room walls is considered, the eigenfunctions  $\Phi_m$  occurring in Eqs. (14)–(17) were approximated by the mode shape functions for rigid room walls

$$\Phi_m(x, y, z) = \sqrt{\frac{\varepsilon_x \varepsilon_y \varepsilon_z}{V}} \cos\left(\frac{n_x \pi x}{l_x}\right) \cos\left(\frac{n_y \pi y}{l_y}\right) \cos\left(\frac{n_z \pi z}{l_z}\right), \quad (18)$$

where the modal indices  $n_x$ ,  $n_y$ ,  $n_z$  are non-negative integers and they are not simultaneously equal to zero (the trivial solution of the wave equation was excluded) and  $\varepsilon_s = 1$  if  $n_s = 0$ ,  $\varepsilon_s = 2$  if  $n_s > 0$ . The room was excited by the point source with the power  $W$  of  $1.5 \cdot 10^{-2}$  W located at the position:  $x_0 = 5$  m,  $y_0 = 3.5$  m,  $z_0 = 1.6$  m.

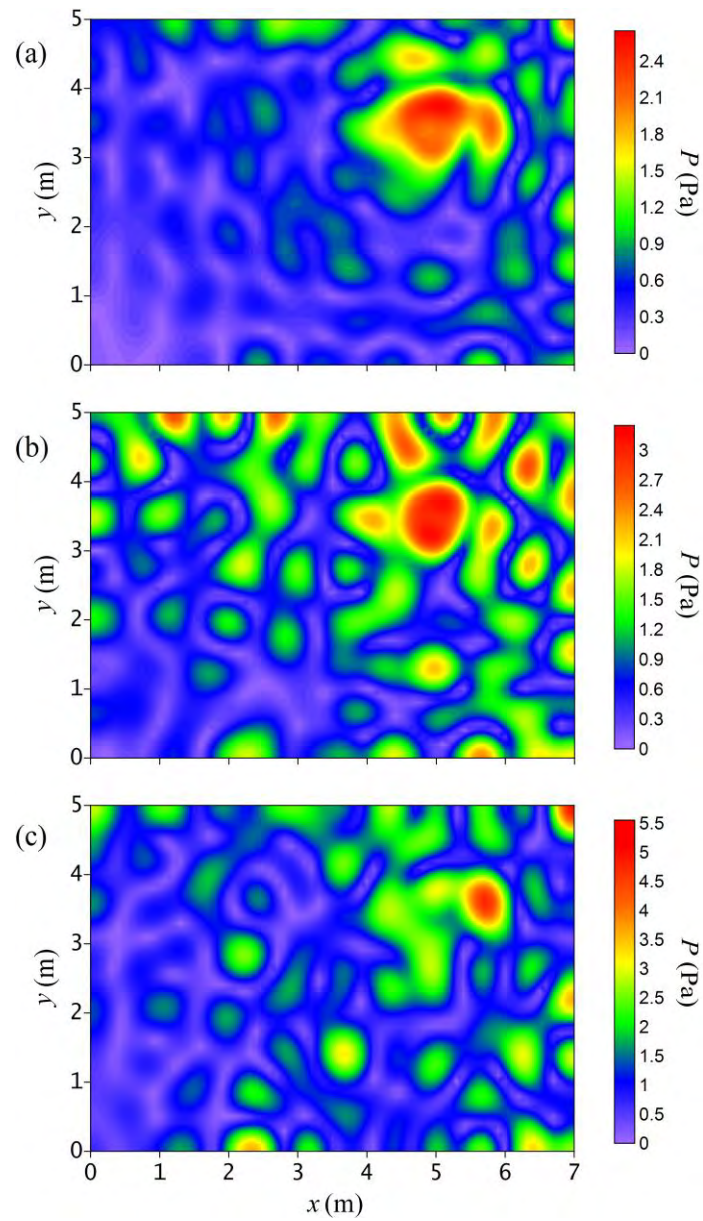


Figure 1. Mapped distribution of the steady-state pressure amplitude  $P$  on the  $(x, y)$  plane at a distance  $z = 1.2$  m from a room floor for the source frequency of 300 Hz and the wall impedance: (a)  $\zeta = 10$ , (b)  $\zeta = 10 + j15$ , (c)  $\zeta = 10 - j15$

In a numerical algorithm, the room volume was discretized by a mesh system that has uniform horizontal and vertical mesh sizes  $\Delta x = \Delta y = \Delta z = 0.05$  m and to predict the steady-state pressure amplitude, Eqs. (14), (16) and (17) were applied. The number of modes included in series in Eq. (14) is an important parameter in a numerical procedure because it greatly influences the simulation accuracy and calculation time. This number depends on room dimensions and it can be approximated from the formula [12]

$$M = \frac{4\pi V}{3} \left( \frac{f_c}{c} \right)^3 + \frac{\pi S}{4} \left( \frac{f_c}{c} \right)^2 + \frac{L}{8} \left( \frac{f_c}{c} \right), \quad (19)$$

where  $V = l_x l_y l_z$  is the room volume,  $S = 2(l_x l_y + l_x l_z + l_y l_z)$  is the surface of room walls and  $L = 4(l_x + l_y + l_z)$  is the sum of the lengths of the room edges. Equation (19) indicates that approximately  $M$  modal frequencies are located in the frequency range up to the cutoff frequency  $f_c$ . In a numerical simulation it was assumed that  $f_c = 600$  Hz, then after inserting this value into Eq. (19) one can obtain  $M = 2712$ .

Figure 1 shows exemplary simulation results obtained for the source frequency of 300 Hz. The graphs have a form of colored contour maps which are a two-dimensional representation of three-dimensional data. They illustrate a distribution of the steady-state pressure amplitude  $P$  on the observation plane  $z = 1.2$  m for three different wall impedances:  $\zeta = 10$  (Fig. 1(a)),  $\zeta = 10 + j15$  (Fig. 1(b)) and  $\zeta = 10 - j15$  (Fig. 1(c)). The data in Fig. 1 allow us to conclude that the use of absorbing material with non-zero reactance substantially modifies a spatial sound pressure distribution. This is mainly due to the influence of the reactance on damping properties of the absorbing material. These properties are best characterized by means of the random-incident absorption coefficient, which for the complex wall impedance, is expressed by the formula [12]

$$\alpha = \frac{8\zeta_r}{|\zeta|^2} \left[ 1 - \frac{\zeta_r}{|\zeta|^2} \ln(1 + 2\zeta_r + |\zeta|^2) + \frac{\zeta_r}{\zeta_i} \arctan\left(\frac{\zeta_i}{1 + \zeta_r}\right) \right]. \quad (20)$$

How to check it easily, the coefficient  $\alpha$  for the first wall impedance has the value of 0.489, and for two remaining impedances it amounts to 0.217. This explains why for non-zero reactance a maximal value of the amplitude  $P$  is larger than in the case of a real-valued wall impedance. It is surprising, however, that for the negative reactance this maximum is about 1.6 times higher than for the positive reactance. This is due to the fact that for the negative reactance the phases  $\varphi_m$  change the sign, thus affecting values of the parameters  $\alpha_m$  and  $\beta_m$  used to calculate the pressure amplitude  $P$  (Eqs. (16) and (17)).

The second part of the numerical study aims to investigate how the wall impedance affects the low-frequency room transfer function (RTF). Calculations of RTFs were carried out for the receiving point:  $x = 2$  m,  $y = 3$  m,  $z = 1.2$  m, for previously assumed wall impedances. The results are depicted in Fig. 2 and they illustrate frequency dependence of the level  $L$  of the pressure amplitude  $P$  in the frequency range 50–500 Hz. In Fig. 2(a) the RTFs obtained for  $\zeta = 10$  and  $\zeta = 10 + j15$  are presented and they prove that a modal structure of the sound field is better reproduced for non-zero reactance because in this case the room walls provide a smaller sound damping. In Fig. 2(b) the RTFs predicted for  $\zeta = 10 + j15$  and  $\zeta = 10 - j15$  are compared. The interesting thing is

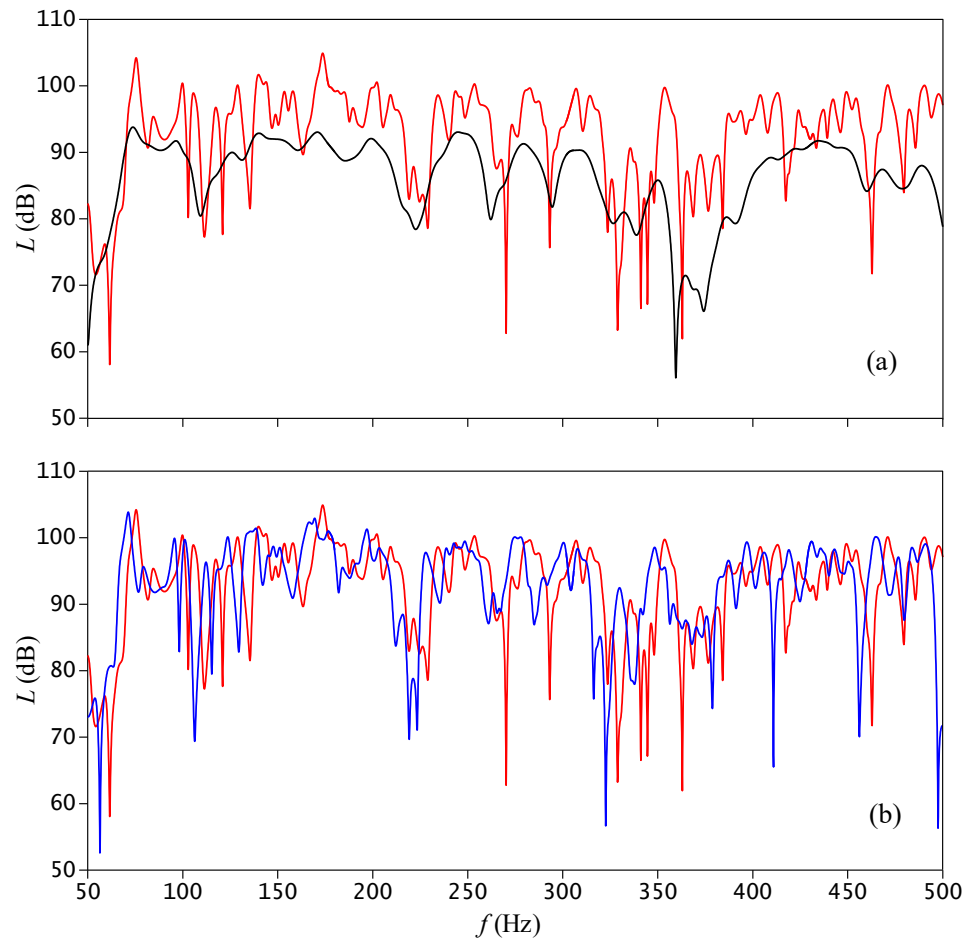


Figure 2. Comparison of the room transfer functions predicted at the observation point:  $x = 2$  m,  $y = 3$  m,  $z = 1.2$  m, for wall impedances:  $\zeta = 10$  (black line),  $\zeta = 10 + j15$  (red line),  $\zeta = 10 - j15$  (blue line)

that the RTFs are shifted about 4 Hz to each other. It results from the fact that in Eqs. (16) and (17) the phase  $\varphi_m$  is always summed with the sound frequency  $\omega$ , therefore the change of phase sign results in the observed frequency shift.

### Acknowledgments

This work was financially supported by the National Science Center, Poland under the project no. 2016/21/B/ST8/02427.

## References

1. V. Valeau, J. Picaut, M. Hodgson, *On the use of a diffusion equation for room-acoustic prediction*, Journal of the Acoustical Society of America, **119**(3) (2006) 1504 – 1513.
2. J. Summers, R. Torres, Y. Shimizu, B. Dalenbäck, *Adapting a randomized beam-axis-tracing algorithm to modeling of coupled rooms via late-part ray tracing*, Journal of the Acoustical Society of America, **118**(3) (2005) 1491 – 1502.
3. S. Laine, S. Siltanen, T. Lokki, L. Savioja, *Accelerated beam tracing algorithm*, Applied Acoustics, **70**(1) (2009) 172 – 181.
4. M. Aretz, P. Dietrich, M. Vorländer, *Application of the mirror source method for low frequency sound prediction in rectangular rooms*, Acta Acustica united with Acustica, **100**(2) (2014) 306 – 319.
5. T. Okuzono, T. Otsuru, R. Tomiku, N. Okamoto, *A finite-element method using dispersion reduced spline elements for room acoustics simulation*, Applied Acoustics, **79** (2014) 1 – 8.
6. T. Sakuma, Y. Yasuda, *Fast multipole boundary element method for large-scale steady-state sound field analysis. Part I: setup and validation*, Acta Acustica united with Acustica, **88**(4) (2002) 513 – 525.
7. D. Murphy, A. Southern, L. Savioja, *Source excitation strategies for obtaining impulse responses in finite difference time domain room acoustics simulation*, Applied Acoustics, **82** (2014) 6 – 14.
8. S. Dance, G. Van Buuren, *Effects of damping on the low-frequency acoustics of listening rooms based on an analytical model*, Journal of Sound and Vibration, **332**(25) (2013) 6891 – 6904.
9. K. Sum, J. Pan, *Geometrical perturbation of an inclined wall on decay times of acoustic modes in a trapezoidal cavity with an impedance surface*, Journal of the Acoustical Society of America, **120**(6) (2006) 3730 – 3743.
10. M. Meissner, *Computer modelling of coupled spaces: variations of eigenmodes frequency due to a change in coupling area*, Archives of Acoustics, **34**(2) (2009) 157 – 168.
11. M. Meissner, *Spectral characteristics and localization of modes in acoustically coupled enclosures*, Acta Acustica united with Acustica, **95**(2) (2009) 300 – 305.
12. H. Kuttruff, *Room acoustics*, 5th ed., Spon Press, New York, 2009.
13. M. Meissner, *Acoustic energy density distribution and sound intensity vector field inside coupled spaces*, Journal of the Acoustical Society of America, **132**(1) (2012) 228 – 238.
14. M. Meissner, *Prediction of reverberant properties of enclosures via a method employing a modal representation of the room impulse response*, Archives of Acoustics, **41**(1) (2016) 27 – 41.
15. S. Damelin, W. Miller, *The mathematics of signal processing*, Cambridge University Press, New York, 2012.



## **Coupled Speakers Directivity Measurements for Small Acoustic Omnidirectional Source Development**

Bartłomiej CHOJNACKI

*AGH University of Science and Technology,  
Mickiewicza Av. 30, 30-059 Cracow, Poland, bartlomiej.chojnacki@agh.edu.pl*

Tadeusz KAMISIŃSKI

*AGH University of Science and Technology,  
Mickiewicza Av. 30, 30-059 Cracow, Poland, kamisins@agh.edu.pl*

Klara JUROS

*AGH University of Science and Technology,  
Mickiewicza Av. 30, 30-059 Cracow, Poland, klara.juros@gmail.com*

Daniel KACZOR

*AGH University of Science and Technology,  
Mickiewicza Av. 30, 30-059 Cracow, Poland, dkaczor@student.agh.edu.pl*

### **Abstract**

Acoustic monopole construction is unsolved task, which engineers try to solve for many years. In the past they constructed many innovative solutions, such as spark or laser-gap sound sources, but those concepts caused many equipment troubles. It is impossible to select type of the measurements signal for this kind of sound sources, also any sparks and laser beams can provide strong electromagnetic distortions around the operation zone. In current state of art we are trying to provide non-standard solutions while traditional omnidirectional sound source, based on spatial configuration of electroacoustic transducers, is not described correctly and can be expanded.

Paper presents concept of acoustic monopole source based on coupled configuration for electrodynamic loudspeakers, similar to isobaric setup but non-constant volume between the speakers. By using FEM modelling we will present directivity patterns and project ideas, describing how overall results corresponds with geometrical parameters of speaker configuration and parameters. Presentation summary contain results of prototype device directivity patterns measurements. Received characteristics shows big potential in using cone-to-cone coupled speaker setup to reach acoustic monopole in frequency range defined by distance between transducers. Achieved knowledge allows to provide broadband, easy to build and small acoustic monopole with many possible applications.

**Keywords:** omnidirectional, point sound sources, coupled speaker set, sound radiation patterns

### **1. Introduction**

Proper sound propagation directivity patterns are important for any acoustic measurements because radiation characteristics of sound source used for measurements can affect them. For in situ measurements in highly diffused sound fields like industrial halls or other rooms with high reverberation, also for reverberation chamber, we can neglect influence of sound source directivity or compensate it with higher number of the measurement averages[1]. For the free field measurements like anechoic chamber we

should pursue for best possible omnidirectionality to avoid serious mistakes in measurements for example in HRTF measurements [2].

Acoustic monopole construction provides many problems as none universal method and design procedure exists [3]. Sound sources require different features for assumed measuring function, such as scale measurements, in-situ rooms measurements or near-field HRTF measurements – all those purposes requires different characteristics for sound sources.

Current approach provide coupled speaker solution similar to isobaric speaker, which is rarely used solution, invented in 1994 (patent granted in 1997), mainly for improving low frequency response of subwoofers [4]. It uses combination of two loudspeakers, facing each other and using the same signal (in-phase or out-of-phase). In this paper we investigated directivity patterns of loudspeakers in isobaric setup, deriving the dependencies between the driver parameters, box volume or distance between the drivers. Using isobaric speaker technology has huge potential in acoustic omnidirectional source design, allowing for development of small and efficient sound sources for multiple applications. Currently the only limitation seems to be narrow frequency range. Presented coupled speakers setup looks similar to isobaric configuration, but diaphragms are moving in opposite directions and the volume between the speakers is non constant. Paper discuss advantages and disadvantages of this solution and possible construction methods.

## 2. Omnidirectional speaker development and usage

Current state of the art contains many innovative solutions provided to reach correct, omnidirectional directivity patterns in wide frequency range and improve the accuracy of measurements performed in the requirements of point source simulation. It is pointed, for example in ISO 354 or ISO 17497 standards [5], that source directivity affects the measurements together with the measuring environment [6]. Despite the role they are designed for, omnidirectional sound sources have different sizes, working principles and are based on different types of transducers. Example of some previously developed device, such as commercial use Microflow low frequency source or developed by AGH spark gap sources are shown in Figure 1.



Figure 1. Microflow low frequency omnidirectional source (left) [7] and spark gap sound source for scale modelling [8]

Different constructions have one common feature which is possibly closest to the uniform, omnidirectional spatial response, but differ in the terms of reachable frequency ranges, powers and device dimensions, which also define the possible usages.

This paper is focused on development of small point source with electrodynamic microspeakers. Main application of this device could be loudspeaker arrays and remote controlled, variable directivity sources. This solution is easy to apply and can be expanded in many projects where point source simulation is required.

### 3. Isobaric loudspeaker configuration

The concept of isobaric speaker is known from 1997 [4], but so far there are very few scientific reports on this technology. This configuration was used mainly in subwoofers few decades ago, usually with the bandpass enclosures ( Figure 2).

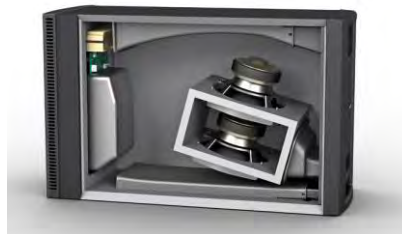


Figure 2. Typical isobaric loudspeaker based subwoofer [9]

Main features of this solution are: halving the  $V_{AS}$ ,  $R_e$  and  $L_e$  parameters and reducing or amplifying the efficiency by 3 dB (regarding the connection in parallel or series). Possible speaker connections are explained in Figure 3.

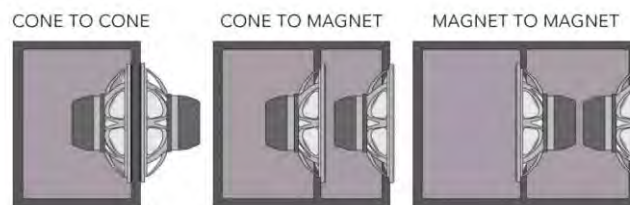


Figure 3. Possible isobaric speakers configuration [10]

In this research we focused on rarely considered feature and currently not measured for isobaric loudspeaker systems, which is directivity and spatial response. As in commercial applications we usually have the enclosure with vents (bandpass type) or closed enclosure assume omnidirectional sound propagation for low frequencies. On the other hand we also have known loudspeaker configuration in infinite baffle, considered omnidirectional in low frequencies and dipole in high frequencies, getting similar to piston radiation. To consider different types of propagation and compare them with isobaric speakers we used FEM modeling in COMSOL multiphysics (Figure 4).

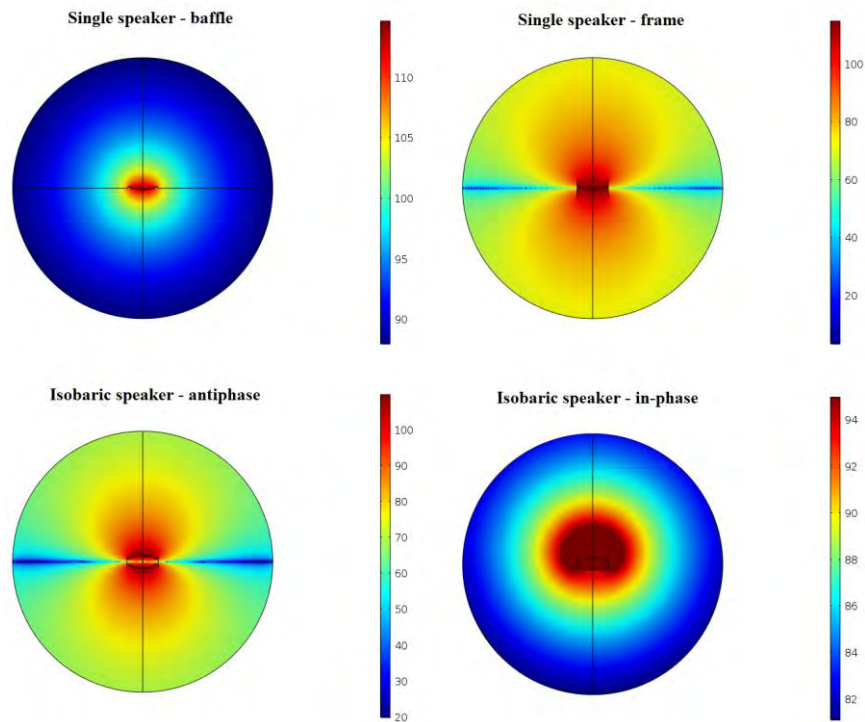


Figure 4. Simulated directivity for tested configurations – single speaker in infinite baffle (top left), single speaker in small frame (top right), isobaric speakers working in antiphase (bottom left) and isobaric speakers working in phase (bottom right)

Decreasing the frame size to minimum (Figure 4), basically, floating the speaker in the air, provide the dipole issue with the spatial speaker response. The same way occurs to be in the isobaric speaker configuration while the drivers are connected and driven with antiphase signal (one is moving forward while second is moving backward). Connection of the speakers in phase, so in the given time frame the diaphragms are moving to the centre and outside in the same way, allow us to avoid the dipole issue caused by lack of infinite baffle dividing the space and allowing to neglect the effect of back side propagation. This lead us to consider this as a possible point source and acoustic monopole. However, when the diaphragms of speakers are moving in opposite directions, the volume between them aren't constant anymore, so the name “isobaric” is no valid anymore. In this case we propose name “coupled speakers” to be used.

#### 4. Experimental verification of miniature omnidirectional source

At this part of the study we performed only preliminary tests in anechoic chamber. Sample set of isobaric speakers was used with Visaton K28 microspeakers, 28 mm of diameter. Figure of test samples is shown on Figure 5.



Figure 5. Test samples for the experiment – single speaker (left) and isobaric in-phase (right)

To verify performed FEM modelling we measured directivity patterns in 1/24 octave bands, using the rotating table in anechoic chamber and pink noise as a test signal. Spatial resolution was 10 degree in horizontal plane. As a omnidirectionality measure we used standard deviation calculated for each frequency band from 35 measurements, representing the full circle around the speaker (Figure 6), which is proved and simple way to assess the omnidirectionality [11].

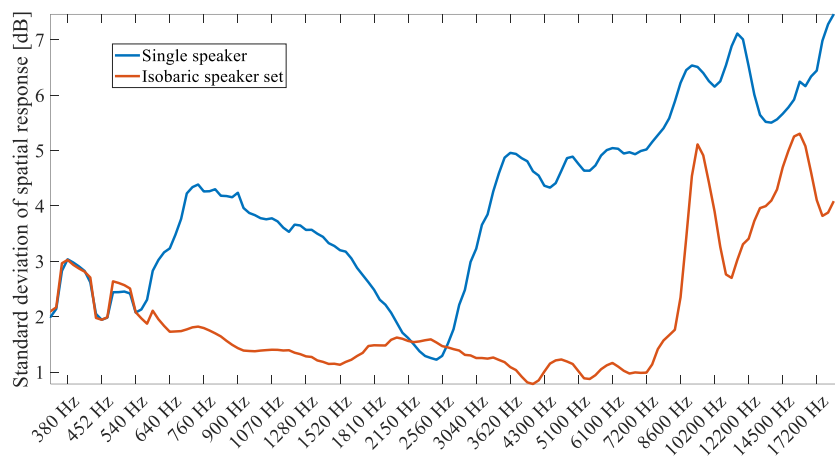


Figure 6. Standard deviation of spatial response for single driver and isobaric loudspeaker

Using this parameter we rate the result as closer to omnidirectional source if the standard deviation is smaller. Analysing Figure 6 we notice great increase in speaker omnidirectionality for isobaric in-phase configuration.

## 5. Conclusions

Paper presents the solution of coupled speakers set to provide small, omnidirectional sound source with narrow-band frequency response, limited by speaker size and distance between the speakers in coupled setup. Preliminary experimental verification proved initial concept and FEM modelling. There is a need to perform further estimation of isobaric speaker set concept in the terms of omnidirectional sources construction, such as geometrical dimensions of the inner volume and relevant speaker parameters for this type of speaker setup, however preliminary research prove the possible utility of this solution in further omnidirectional sources development.

## Acknowledgments

Research financed by research grant for PhD students from Dean of the Department of Mechanics and Robotics, AGH University of Science and Technology in Cracow.

## References

1. M. L. S. Vercammen, *Improving the accuracy of sound absorption measurement according to ISO 354*, Proc. Int. Symp. Room Acoust., 2010.
2. G. Z. Yu, B. S. Xie, D. Rao, *Directivity of spherical polyhedron sound source used in near-field HRTF measurements*, Chinese Phys. Lett., 2010.
3. A. Majchrzak, B. Chojnacki, M. Sobolewska, K. Baruch, A. Pilch, *The Measurement of Sound Scattering in a 1:8 Scale-Validation of the Measurement Stand and Procedure*, In: INTER-NOISE and NOISE-CON Congress and Conference Proceedings, (2018) 3287 – 94.
4. J. T. Larsen, J. R. Larsen, *Isobaric loudspeaker*, 1997.
5. C. Hak, R. H. C. Wenmaekers, J. P. M. Hak, L. C. J. van Luxemburg, *The source directivity of a dodecahedron sound source determined by stepwise rotation*, In: Proceedings of Forum Acusticum, 2011.
6. I. B. Witew, G. K. Behler, *Uncertainties in measurement of single number parameters in room acoustics*. Requirements Sources for Sound Data Analysis, 2291 – 5.
7. D. Shearer, *Measuring absorption below 100Hz with a P-U sensor*, 2016.
8. K. Baruch, A. Majchrzak, B. Przysucha, A. Szela, T. Kamiński, *The effect of changes in atmospheric conditions on the measured sound absorption coefficients of materials for scale model tests*, Appl. Acoust., 1412018, 150 – 260.
9. R. Alden, *Speaker Building 201: with 11 Completely Designed Speaker Systems including a 5.1 Home Theater System*, Audio Amateur Press 2004.
10. <https://www.vueaudio.com/isobaric-subwoofer-design/>
11. B. Chojnacki, M. Ziobro, J. Rubacha, *Piezoelektryczne wszechkierunkowe źródło dźwięku do akustycznych badań w skali w zakresie ultradźwięków*, In: A. Pilch, editor. Studium badawcze młodych akustyków. Kraków, Wydawnictwo AGH 2016.

## A Computer Model for Calculating the Speech Transmission Index Using the Direct STIPA Method

Paweł DZIECHCIŃSKI

*Wrocław University of Science and Technology, Department of Acoustics  
and Multimedia, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław,  
pawel.dziechcinski@pwr.edu.pl*

### Abstract

Computer models currently used for the simulation of the speech transmission index (STI) calculate the STI using the statistical method or are based on numerically determined impulse response of the transmission channel. The limitation of both these computational methods is that they do not allow to take into account the nonlinear properties of the transmission channel and fluctuating background noise. This paper presents a proposition of a model based on the direct STIPA method. This model allows computer simulations of STIPA for distributed sound systems, and enables analysis to include both changes in signal dynamics and fluctuating background noise. The paper presents the idea of the model and validation of its basic elements - the generator and the analyser. The possibilities of using the model for computer simulation of outdoor public address systems were also discussed.

**Keywords:** speech transmission index, STIPA, attenuation of sound by the atmosphere

### 1. Introduction

The most often used method of rating the speech intelligibility of sound systems is the STI (speech transmission index) method, which is the subject to Houtgast and Steeneken works [5, 11], Brachmański papers [1, 2] and IEC 60268-16 standard [6]. This method is based on the fact that the speech signal can be divided into two spectra – audible spectrum (octave bands from 125 Hz to 8 kHz) and modulation spectrum, including the phonemes (phoneme – the smallest unit of speech distinguishable for users of a given language) – from 0.63 to 12.5 Hz. When the signal passes through the transmission channel (electroacoustic system and chamber), the signal level in each octave band and the modulation depth for different modulation frequencies change, which is mainly associated with the presence of noises and reverberation in the channel. The single-number STI value is determined based on the modulation transfer function, which is the ratio of the respective modulation indexes at the output to the indexes at the input of the channel. Measurements of speech transmission index can be conducted using the direct method – analysing the modulation depth of a special measurement signal, or indirect – determining modulation indexes based on the impulse response of the transmission channel – by the Schroeder equation [10]. Computationally the speech transmission index can be determined by the statistical method or on the basis of the numerically determined channel impulse response [6]. The limitation of both calculation methods is that they allow including neither the nonlinear properties of the transmission channel nor the fluctuating

background noise. Although properly adjusted sound system should not be a source of large nonlinear distortion, but in practice commonly used nonlinear elements are e.g. dynamics processors. The paper proposes a computer model that allows taking into account in the STI determination both changes in the dynamic range of the signal and temporally fluctuating background noise.

The influence of various factors on speech quality (including intelligibility) was discussed, among others, by Brachmański [1]. In his book, however, he did not deal with sound systems, whose special feature is the transmission of sound over long distances. This is partly related to the problem of sound attenuation by the atmosphere. The proposed model allowed to assess the impact of this phenomenon on STI.

## 2. Model concept

The proposed model is based on a direct measurement method, a simplified STI version, which is STIPA (speech transmission index for public address systems). The general scheme of the model (Figure 1) assumes that the STIPA signal can come from  $n \geq 1$  number of loudspeakers. In general, the sound reaching the observation point will therefore be a superposition of signals coming from individual loudspeakers and noises occurring at this point.

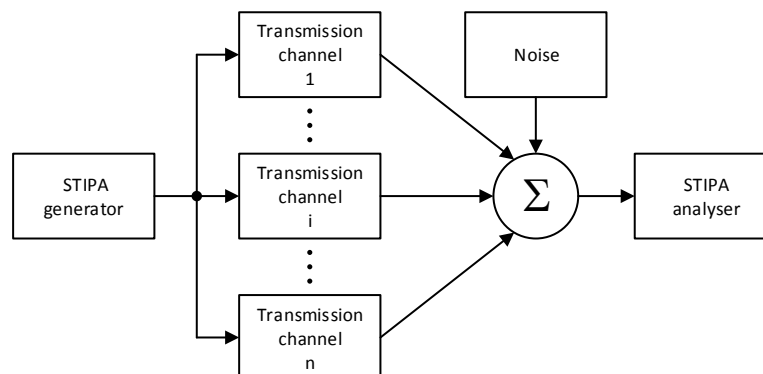


Figure 1. General structure of the proposed calculation model

In a typical electrical input sound system, the transmission channel consists of devices for processing the audio signal, loudspeakers, and sound propagation environment (Figure 2). The audio signal transmission chain can be divided into two parts:

- 1) A-chain – from the signal source to the main gain control of the loudspeaker channel,
- 2) B-chain – from the main gain control of a given loudspeaker channel to a loudspeaker.

In the A-chain, signal processing processes common to all transmission channels are carried out, such as dynamic compression or frequency characteristic equalisation. The B-chain is a loudspeaker processor and power amplifier. The audio processing path



can be modelled by applying signal processing procedures that correspond to those used in reality, or if their algorithms are not known, by generating a file with the STIPA signal and processing it in a real processor. A similar processing procedure can be used to process not only the STIPA signal, but also other audio samples. This possibility was designed for processing samples to be used in subjective tests.

Depending on the available data, the loudspeaker will be modelled as a filter with appropriate frequency characteristics or by convolution of a loudspeaker impulse response with a signal.

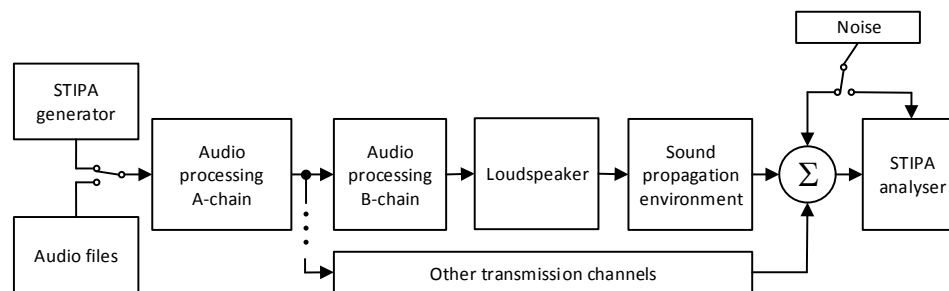


Figure 2. Detailed structure of the proposed calculation model

The sound propagation environment can be a room, an external environment or a free field. In the case of a room, it is assumed that the sound propagation environment will be modelled by convolution of the room impulse response with the signal. The impulse response can be obtained by means of measurements or as a result of modelling in an external programme. The algorithms from ISO 9613-1 [7] and ISO 9613-2 [8] standards will be used to model the external environment. The current version of the sound propagation model in the environment has been limited to the phenomenon of sound attenuation by the atmosphere. The simplest way to take into account the effect of this phenomenon on STIPA is to determine sound attenuation for octave bands from 125 to 8000 Hz and appropriate attenuation of the signal spectrum at the generation stage. Due to the fact that the model was also planned for processing signals other than STIPA, a different algorithm was chosen. It consisted of:

- 1) determining the frequency response of sound attenuation by the atmosphere on the basis of ISO 9613-1 standard,
- 2) designing a digital filter with determined frequency characteristics,
- 3) performing signal filtration.

The frequency characteristics of the filter resulting from sound attenuation by the atmosphere for a source-microphone distance  $d$  equal to 200 m, for two of the four analysed cases are shown in Figure 3.

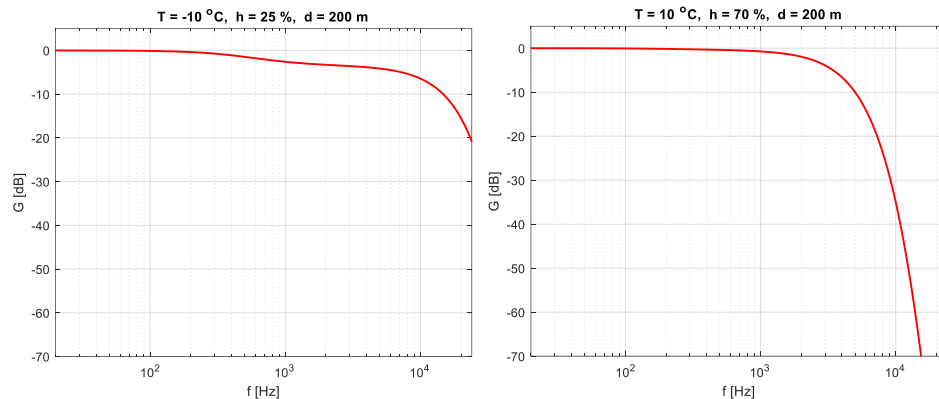


Figure 3. Frequency characteristics of the filter resulting from sound attenuation by the atmosphere for a source-microphone distance of 200 m

The impact of noise on STIPA can be included in the model in two ways. The first way is to take it into account computationally, in accordance with the algorithm specified in IEC 60268-16 standard. Taking into account this type of noise requires including in the model its sound pressure levels in the 1/1-octave bands. The second method is to sum the analysed measurement signal with real noise saved in the file and properly calibrated as to level.

It is assumed that in the model it will be possible to study the influence of individual transmission channel elements independently. An additional functionality of the model will be the ability to process any sound files that can be used, for example, for subjective research.

### 3. Validation of the STIPA generator and analyser

The IEC 60268-16 standard describes the algorithm for generating and analysing the STIPA signal, however, the requirements for the filters used in the generator and the analyser are usually described in terms of minimum requirements. For the generator,  $\frac{1}{2}$  octave filters with a stop-band attenuation of at least 60 dB/octave and the pass-band ripples not exceeding 0.3 dB are required. As an example, the standard proposes 12th order 0.2 dB ripple type I Chebyshev. In the analyser, the standard requires the use of class 0 or 1 filters in accordance with IEC 61260 standard with phase characteristics as close to linear as possible. The analyser should also use a low pass filter to obtain the signal envelope. The standard states only that its cut-off frequency should be approximately 100 Hz.

In the case of the generator, it is possible to use advanced filters requiring high computing power due to the fact that signal preparation is usually done only once (saving to a file or a carrier, etc.). In the case of the analyser, one should adapt to obtaining filters similar to those implemented in sound level metres. It is therefore important to optimise the computing power necessary for their implementation. The selection of optimal filters is made more difficult by the fact that preliminary

analyses have shown that different types of analyser filters are more suitable for different generator filters. Furthermore, due to the fact that the STIPA signal is a noise signal, the results obtained are variable in time. This variation also depends on the filters used in the analyser and generator. Finally, it was decided to perform a comparative analysis of STIPA results obtained using a generator realised with filters of adjacent band attenuation from 36 to 96 dB in 12 dB steps and acceptable irregularities in the 0.1, 0.2, and 0.3 dB pass-band. Type 1 Butterworth and Chebyshev filters were checked. In the case of the analyser, Butterworth filters of the orders 10, 12, 14, 16 and 18 were tested. As a smoothing filter, a Butterworth filters with a cut-off frequency of 100 Hz of orders of 1, 2 and 3 as well as 95 Hz cut-off frequency of order 2 were used. For each type of generator filter, 10 files with STIPA signal were generated with a length enabling to perform analysis for 15 seconds. Each of 720 combinations of filter types was tested. Including 10 replications, it gave 7200 results. The optimal configuration was chosen based on the following criteria:

- 1) the highest maximum of STIPA value (as close as possible to 1,000),
- 2) the highest minimum of STIPA value,
- 3) the highest average of STIPA value,
- 4) the lowest standard deviation of STIPA.

For a computer model, ensuring the generator's repeatability is not necessary because the selected excitation can be saved on a file that can be used later for analysis instead of the generator. Repeatability was also included in the analyses due to the fact that on the basis of the model, STIPA method was also planned to be implement on the platform of the sound level meter.

Out of 720 filter configurations analysed and 10 replications:

- 1) 81 combinations failed to obtain a STIPA value of 1.00,
- 2) for 226 combinations, the minimum STIPA value was lower than 0.97,
- 3) the average STIPA value for 293 combinations was less lower 0.99.

For 12 out of 720 tested filter configurations there were obtained:

- 1) the maximum STIPA value for 10 replications 1,000,
- 2) the minimum STIPA value for 10 replications higher than 0.97,
- 3) the average STIPA value for 10 replications higher than 0.99,
- 4) STIPA standard deviation for 10 replications lower than 0.007.

Of these 12 cases, the smallest order combination of filters was used in the model.

The validation of the computational part of the STI algorithm, i.e. determining the single-number final value based on the modulation transfer function, was performed using the data from the example in the IEC 60268-16 standard [6]. After simplifying the STI algorithm to STIPA, validation was based on comparing the results with the results of two measuring systems existing on the market.

#### 4. Impact of sound attenuation by the atmosphere on STIPA

The impact of sound attenuation by the atmosphere on STIPA will be discussed for the sound system emitting sound from the central cluster, not introducing linear and nonlinear distortions, and ensuring an even distribution of sound levels. The propagation concerns a free field. The analyses were carried out for four combinations of temperature

$T$  and relative humidity  $h$ . The expected effect of sound attenuation by the atmosphere is a decrease in sound pressure levels of the sound pressure level signal in the high frequency range. The effect of such a change in the signal spectrum is the increasing influence of the masking effect on the STIPA result and the decreasing signal-to-noise ratio. Figure 4 presents changes in STI as a function of the source-microphone distance  $d$ , assuming that the noise has an insignificant impact, and the equivalent continuous sound level of the  $L_{Aeq}$  measurement signal is 60 dB. For a signal of this level, the effect of masking is minimal, therefore changes in STI begin to be noticeable at distances above 60 m in the case of the combination of  $T$  and  $h$  providing the highest attenuation in high frequencies. The lowest of the analysed attenuations occurs for  $T = 10^\circ\text{C}$  and  $h = 25\%$ . For distances  $d$  up to 200 m such attenuation does not introduce any change to STIPA.

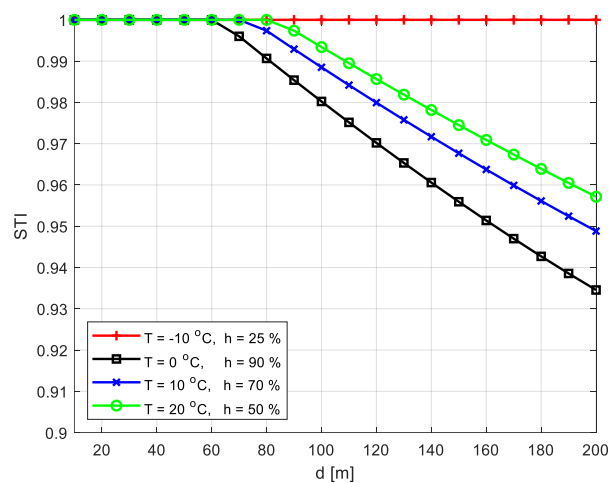


Figure 4. Impact of sound attenuation by the atmosphere on STIPA-signal level  
 $L_{Aeq} = 60$  dB, no noise

In practice, sound systems are rarely designed to work with sound levels close to 60 dB. Most often, the required STIPA value and the level of noise determine the sound level of the signal. In the case of a sound propagation environment with a small reverberation time or for a free field, to obtain an STI in the range of 0.50 - 0.60 the sound level of the signal is chosen to be about 10 dB higher than the noise level. Figure 5 shows the noise results with the sound level  $L_{Aeq} = 70$  dB and the "transport noise" spectrum [4, 9] and the signal with the level  $L_{Aeq} = 70$  dB. In this case, the effect of sound attenuation for long distances is already significant and for  $d = 200$  m STI for the value in the source proximity decreases by 0.08 - 0.13.

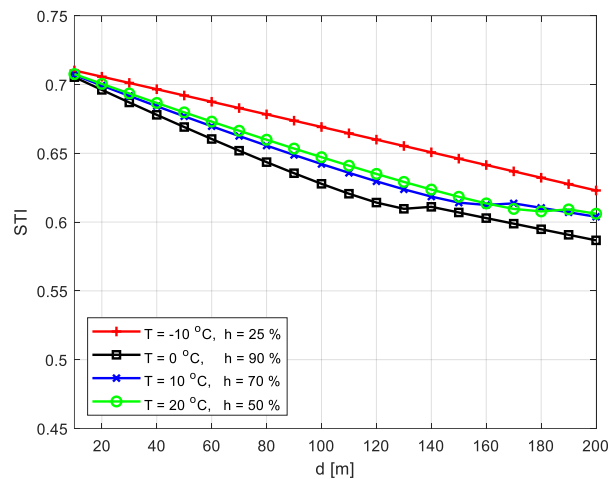


Figure 5. Impact of sound attenuation by the atmosphere on STIPA - signal level  
 $L_{Aeq} = 80$  dB, noise level  $L_{Aeq} = 70$  dB ("transport noise" spectrum)

Figure 6 shows a typical situation for a sound system in an environment with a high level of noise. In this case, the signal-to-noise ratio is fixed at about 6 dB. Compared to the previous case, the STI decrease is similar or slightly smaller, but the STI values obtained for large distances from the source are close to the minima required for many systems (0.45–0.50). It should be remembered that in reality the STI values obtained will be smaller due to factors that were not included in the analyses, i.e. linear distortions and reverberation.

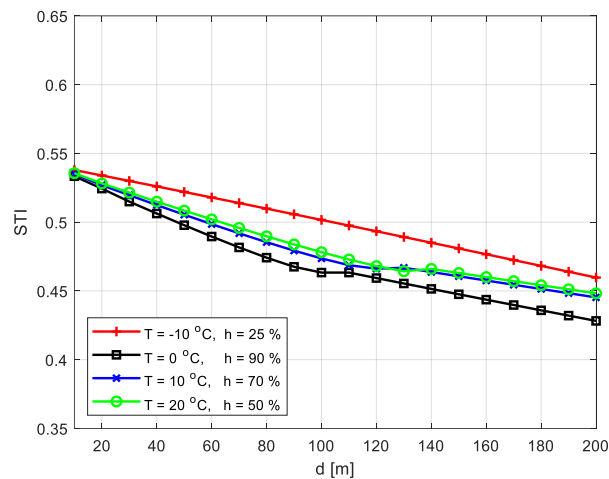


Figure 6. Impact of sound attenuation by the atmosphere on STIPA - signal level  
 $L_{Aeq} = 100$  dB, noise level  $L_{Aeq} = 94$  dB ("transport noise" spectrum)

## 5. Conclusions

The paper presents the idea of a computer model enabling the determination of the speech transmission index using the direct STIPA method. The advantage of this model is primarily the ability to take into account the presence of dynamics processors in the electroacoustic chain.

The analyses show that the selection of generator and analyser filters has a significant impact on the final result of the direct STIPA measurement method. Despite of the use of filters meeting the requirements of the standard IEC 60268-16, it is possible to obtain a dispersion of results in a series of 10 replications that are higher than 0.03 uncertainty of STIPA measurement.

The impact of sound attenuation by the air on STIPA values can be significant in the case of long distance sound transmission. It is particularly important for signals with a high sound level and a relatively small signal-to-noise ratio.

## References

1. S. Brachmanski, *Estimation of logatom intelligibility with the STI method for polish speech transmitted via communication channels*, Archives of Acoustics, **29**(4) (2004) 555 – 562.
2. S. Brachmanski, *Experimental comparison between speech transmission index (STI) and mean opinion scores (MOS) in rooms*, Archives of Acoustics, **31**(4) (2006) 171 – 176.
3. S. Brachmański, *Wybrane zagadnienia oceny jakości transmisji sygnału mowy*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2015.
4. P. Dziechciński, M. Hadyk, *Projektowanie systemów rozgłoszeniowych na otwartych peronach kolejowych*, Postępy akustyki 2017, red. D. Bismor, Gliwice: Polskie Towarzystwo Akustyczne Oddział Górnośląski, (2017) 197 – 208.
5. T. Houtgast, H. J. M. Steeneken, *The Modulation Transfer Function in rooms acoustics as a predictor of speech intelligibility*, Acustica, **28**(1) (1973) 66 – 73.
6. IEC 60268-16:2011 - *Sound system equipment - Part 16: Objective rating of speech intelligibility by speech transmission index*.
7. ISO 9613-1:1993, *Acoustics — Attenuation of sound during propagation outdoors – Part 1: Calculation of the absorption of sound by the atmosphere*.
8. ISO 9613-2:1996, *Acoustics — Attenuation of sound during propagation outdoors – Part 2: General method of calculation*.
9. ISO 717-1:2013, *Acoustics — Rating of sound insulation in buildings and of building elements – Part 1: Airborne sound insulation*.
10. M. Schroeder, *Modulation Transfer Functions: Definition and Measurement*, Acustica, **49** (1981).
11. H. J. M. Steeneken, T. Houtgast, *A physical method for measuring speech transmission quality*, J. Acoust. Soc. Am., **67**(1) (1980) 318 – 326.

## Model of Multipath Propagation of Ultrasonic Pulses in Soft Tissue Using Divergent Beam Tomography Method

Jędrzej BOROWSKI, Krzysztof J. OPIELIŃSKI  
*Faculty of Electronics, Wrocław University of Science and Technology,  
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland,  
jedrzej.borowski@pwr.edu.pl, krzysztof.opielinski@pwr.edu.pl*

### Abstract

The paper presents the model of calculating ultrasound waveform beam emitted inside the circular space of ultrasonic transducer ring array and propagated through a biological medium submerged in water. Each elementary transducer emits a burst signal, which then propagates through a medium and is received by a number of transducers on the opposite side of the ring array. The method allows for calculating runtime and amplitude of ultrasonic bursts while traveling from an emitter to a receiver through a specified soft tissue section geometry, having regard to the refraction and attenuation effects and directivity pattern of transducers. The soft tissue section geometry is constructed using circular shapes with given ultrasound speed and attenuation distribution. The elaborated software creates a set of received waveforms for each transmitting transducer. The presented results produced by the software can be used as a basis for further research on inverse problems in ultrasound waveform tomography.

**Keywords:** model of multipath propagation of ultrasonic pulses, ultrasonic transducer ring array, divergent wave beam, soft tissue, ultrasound tomography

### 1. Introduction

The concept of ultrasound waveform tomography shows much promise as a new method for medical imaging aiding in cancer detection and diagnosis [1, 2]. However, this imaging method is computationally complex, as the inverse problems have to be solved on large amounts of numerical data. We propose a method of generating exemplary data which can be an aid in developing, testing and evaluating algorithms of waveform tomography.

The method is designed to provide exemplary results simulating the behavior of the ultrasound tomography transducer ring array [3, 4], modeled after existing real world implementations. The generated results are simulating the propagation of ultrasonic pulses through biological media submerged in water. The model is highly flexible, allowing for modeling of results for media with various ultrasound speeds and internal composition of the soft tissue section. The elaborated method does not limit the shape of the section. However, the software implementation currently supports the geometry constructed from circular, non-intersecting shapes only. In order to simplify the generation of the results, only refractions and internal reflections of ultrasound beams are taken into consideration.

### 2. Model of ultrasonic pulse

The ultrasonic pulse is modeled using the equation for linear *burst* type signals [5]:

$$p = p_0 \sum_{n=0}^{\infty} \left( 1 - \left| \frac{2tf_0}{m_c} - 1 \right|^m \right) \sin(\omega_0 t - \phi_0) \quad (1)$$

Where  $p_0$  is the amplitude of acoustic pressure,  $t$  is the time,  $f_0$  is the pulse resonant frequency,  $m_c$  is the number of sine wave cycles,  $m$  is the exponent, which is used to shape the envelope, and  $\omega_0 = 2\pi f_0$ .

For the purposes of this paper, the generated pulses had the parameters  $f_0 = 2$  MHz,  $\phi_0 = 0$ ,  $m_c = 7$ , and the exponent  $m$  has the values of  $m = 2$  for the first half of the burst, and  $m = 0.5$  for the second half. An example of a generated burst is displayed on Fig. 5.

### 3. Model of transducer directivity

The directivity of each transducer element is modeled to simulate the attenuation of amplitude of the ultrasonic beam which is propagated or received in the transducer at a non-zero angle. The directivity pattern of a rectangular transducer can be calculated from the equation [6]:

$$K(\theta, \varphi) = \left| \frac{\sin \frac{ua}{2}}{\frac{ua}{2}} \right| \left| \frac{\sin \frac{wb}{2}}{\frac{wb}{2}} \right| \quad (2)$$

where  $u = (2\pi \sin \theta)/\lambda$ ,  $w = (2\pi \sin \varphi)/\lambda$ ,  $a$  and  $b$  are width and height of the transducer,  $\theta$  is the angle along the  $OX$  axis, and  $\varphi$  is the angle along the  $OY$  axis. Since the problem is considered in a 2-D environment, it can be assumed that  $\varphi = 0$ , and the factor containing the  $w$  variable can be disregarded:

$$K(\theta) = \left| \frac{\sin \frac{ua}{2}}{\frac{ua}{2}} \right| \quad (3)$$

In the case of narrow rectangular ultrasonic transducers (and so is in this case), better approximation of the directivity pattern gives the formula with the extra  $\cos \theta$  factor, which derives from the proper interpretation of the Rayleigh-Sommerfeld theory [7]:

$$g(\theta) = K(\theta) \cdot |\cos \theta| = \left| \frac{\sin \frac{ua}{2}}{\frac{ua}{2}} \cos \theta \right| \quad (4)$$

An example of a directivity pattern generated using the above formula can be seen on Figure 1.



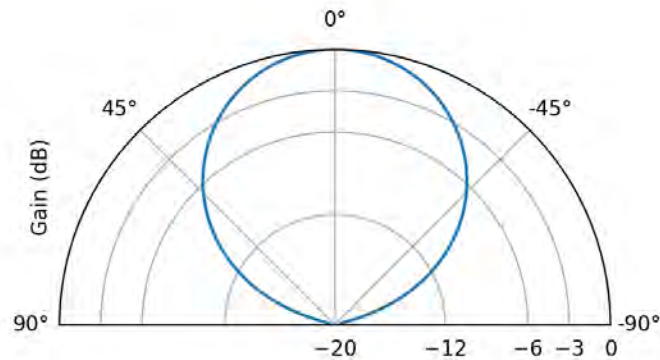


Figure 1. Simulated directivity pattern of a single rectangular transducer in the ring array

#### 4. Model of propagation

The soft tissue geometry is constructed using circular section elements, with given center point  $(x_0, y_0)$ , radius  $r$ , ultrasound speed  $c$ , material density  $\rho$ , and attenuation distribution  $a$ . The transducer ring array is constructed from  $n^2$  emitter-receiver elements and is approximated as a regular circle with the specified radius  $R$ . For the purposes of this paper, the geometry of the section placed inside the space of the ring array simulates a breast tissue submerged in water.

The ultrasonic beam is simulated by generating a number of rays from the source origin point, which is located in the middle of the emitting transducer. The rays are generated at angles in the range from  $-\alpha/2$  to  $\alpha/2$ , where  $\alpha$  is the angular width of the simulated beam. Together with each beam, a simulated pulse waveform is calculated using equation (1). The attenuation and delay of the pulse are applied to the waveform after the propagation is fully simulated. For evaluation, a value of  $\alpha = 45^\circ$  was chosen. For adequate simulation accuracy the recommended number of rays is greater or equal to the number of ring array transducers. Additionally, for each ray, the directivity gain  $g_e$  of the emitting transducer is calculated using equation (4).

Having computed the directional vectors of each ray, intersections with the nearest soft tissue geometry element can be easily found analytically. For each found point of intersection, the ray is recalculated according to Snell's law. The distance between a previous origin point (e.g. the emitter, or a previous material boundary) is saved together with the information about the ultrasound speed in the medium the ray is currently traversing.

The following equations are used to calculate the ray after colliding with material boundaries, where  $\mathbf{I}$  is the incident ray vector,  $\mathbf{N}$  is the normal vector of the geometry element,  $\mathbf{T}$  is the transmitted ray vector [8],  $\theta_1$  and  $\theta_2$  are the incidence and refraction angles. The vectors are normalized so that the lengths of both  $\mathbf{I}$  and  $\mathbf{T}$  are equal to 1.

The angle of incidence can be obtained by calculating the dot product of the incident vector and the inverse of the normal vector:

$$\cos \theta_1 = -\mathbf{N} \cdot \mathbf{I} \quad (5)$$

After applying Snell's law, it is possible to calculate the angle of refraction:

$$\sin \theta_2 = \left( \frac{c_2}{c_1} \right) \sin \theta_1 \quad (6)$$

This allows for calculation of the transmitted ray  $\mathbf{T}$  [8]:

$$\mathbf{T} = \frac{c_2}{c_1} \mathbf{I} + \left( \left( \frac{c_2}{c_1} \right) \cos \theta_1 - \sqrt{1 - \left( \frac{c_2}{c_1} \right)^2 \sin^2 \theta_1} \right) \mathbf{N} \quad (7)$$

If the calculated angle of refraction is real, then the incident ray is refracted by traversing the element boundary. In this case, it is necessary to calculate the weakness of the ultrasonic pulse. The amplitude transmitting coefficient can be calculated using the following formula [9]:

$$D_p = \frac{\cos \theta_1 - \frac{\rho_1 c_1}{\rho_2 c_2} \sqrt{1 - \left( \frac{c_2}{c_1} \right)^2 \sin^2 \theta_1}}{\cos \theta_1 + \frac{\rho_1 c_1}{\rho_2 c_2} \sqrt{1 - \left( \frac{c_2}{c_1} \right)^2 \sin^2 \theta_1}} + 1 \quad (8)$$

where  $\theta_1$  is the angle of incidence,  $\rho_1$  is the material density of the first environment,  $\rho_2$  is the material density of the second environment,  $c_1$  is the ultrasound speed of the first environment, and  $c_2$  is the ultrasound speed of the second environment.

If the calculated angle of refraction is a complex number, then internal reflection is occurring. In this case, the resulting vector is calculated with the following formula:

$$\mathbf{T} = \mathbf{I} + 2 \cos \theta_1 \mathbf{N} \quad (9)$$

The calculated transmitted vector  $\mathbf{T}$  can be used together with the intersection point to create a new, refracted ray, on which the process can be repeated for the rest of remaining collisions on the path of the ray. When there are no more soft tissue elements for the ray to collide with, an intersection with the transducer ring array is found, and the angle between the propagating ray and the receiving transducer element is calculated. Basing on this angle, the directivity gain for the receiving transducer  $g_r$  is calculated using equation (4). Half of the transducers opposite of the emitter are taken into consideration. If the ray collides with a transducer outside of this boundary, it is discarded. The waveform of the propagating ray is added to the input buffer of the receiver transducer.

## 5. Runtime and amplitude equations

The information gathered allows to calculate the complete waveform of the signal received by the transducer which is hit by the propagated rays. In order to obtain this waveform, the generated pulse must be attenuated and delayed accordingly. The attenuation factor is dependent on the length of the path that the ray has taken through the section geometry, ultrasound speeds of materials passed through, and the number of material boundaries traversed. The equation of the gain can be expressed as:

$$a_p = g_e g_r \prod_{i=1}^n s_i a_i \prod_{i=1}^k D_i \quad (10)$$

Where  $g_e$  is the directivity gain for the emitter transducer,  $g_r$  is the directivity gain for the receiver transducer,  $s_i$  is the distance and  $a_i$  is the attenuation distribution for  $i = 1 \dots n^{\text{th}}$  element passed by the ray,  $D_i$  is the transmitting coefficient for  $i = 1 \dots k^{\text{th}}$  material boundary passed by the ray.

The time delay of the received pulse can be expressed as:

$$t_p = \sum_{i=1}^n \frac{s_i}{c_i} \quad (11)$$

Where  $s_i$  is the distance and  $c_i$  is the ultrasound speed of  $i = 1 \dots n^{\text{th}}$  element passed by the ray.

## 6. Implementation and results

An implementation of this method has been elaborated in Python programming language, using the Shapely library for geometrical computations. The elaborated software is capable of generation and visualization of results of divergent beam tomography. The process can be parametrized by selecting the number of transducers in the ring array, the number of rays simulated in the beam, and the width of the beam. For each emitting transducer, the software generates a set of waveforms received by the receiving transducers.

Figure 2 presents examples of the modelling of the paths taken by the rays in the ultrasound beam. In this example, a beam of 32 rays with the width of  $45^\circ$  has been propagated through a medium submerged in water. In the left example, the element has a higher ultrasound speed than that in the water, and in the right example, the ultrasound speed in the element is lower than that in the water. It can be seen how the introduction of the tissue into the water is affecting the ultrasound beam.

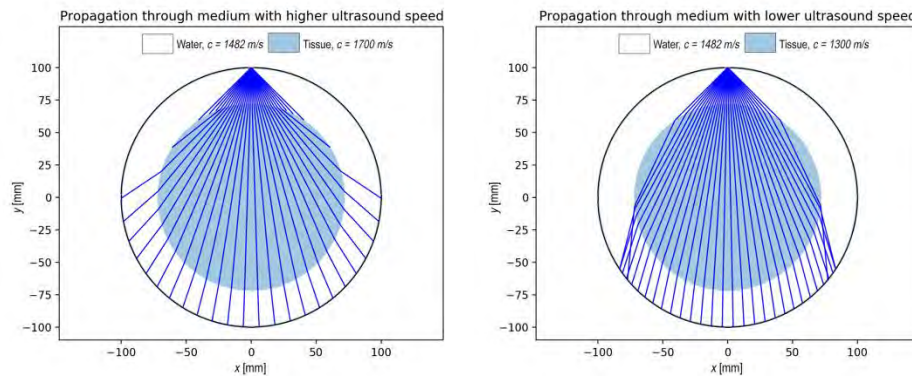


Figure 2. Examples of the modeled ray paths in the ultrasound beams passing through higher and lower sound speed media

Figure 3 presents a modeled propagation through a complex medium - a soft tissue section which contains elements with three different ultrasound speed values. The ultrasound speed of the main element is similar to the sound speed of tissue such as a female breast. The other ultrasound speed values have been chosen to visualize the focusing and scattering of the beams. The right side of the figure presents a selection of received waveforms from the aforementioned example. It can be seen how the angle of the beam, distance traveled by the ultrasound pulse, and the presence of tissue on its way affects the amplitude and time delay of the signal. The received waveforms are numbered from 1 to 25, and the locations of the transducers that received those waveforms are marked on the left plot. A number of the received waveforms, e.g. numbered 10, 18, and 21 are a result of multiple beams entering the same receiver transducer. In this case, it can be observed that the waveforms of the received signals are amplified or attenuated due to summing of signals incoming to the receiver.

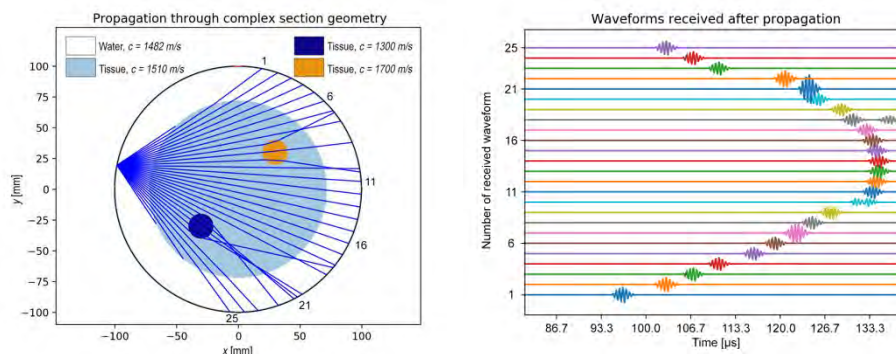


Figure 3. Example of the modeled ray paths in a complex section geometry, and a selection of waveforms received by the simulated transducer

Figure 4 shows another example of propagation through the complex geometry, where the differences between sound speeds of the elements are smaller than in the previous example. Focusing of the beams on the elements with lower sound speed (beams 18 to 20) and scattering of the beams on elements with higher sound speed (beams 9 to 11) can be observed. Beams 13 to 16 are passing through an object with lower sound speed first, and an object with higher sound speed second. This example presents that the model is capable of calculating many refractions for each beam.

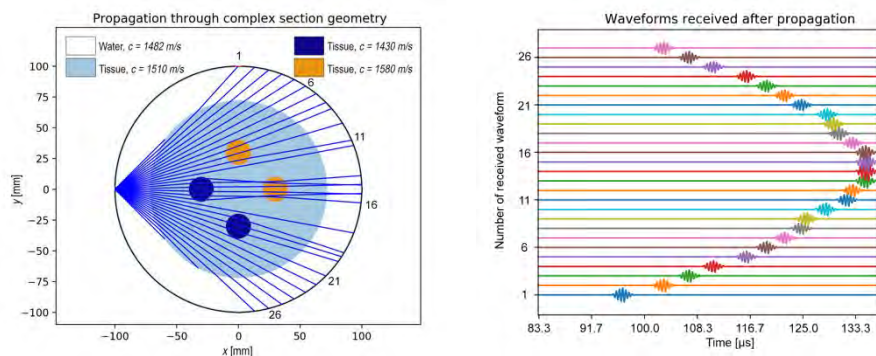


Figure 4. Example of propagation through a section with elements of varying sound speed, and the waveforms received by the simulated transducer

Figure 5 presents an example of an ultrasound pulse waveform before and after propagating through a complex section geometry. This example was generated using 1024 transducer elements, and 1024 rays for each propagation. Due to multipath nature of the elaborated method, the receiving transducer has been hit by many rays at different times, causing changes in the envelope of the received signal.

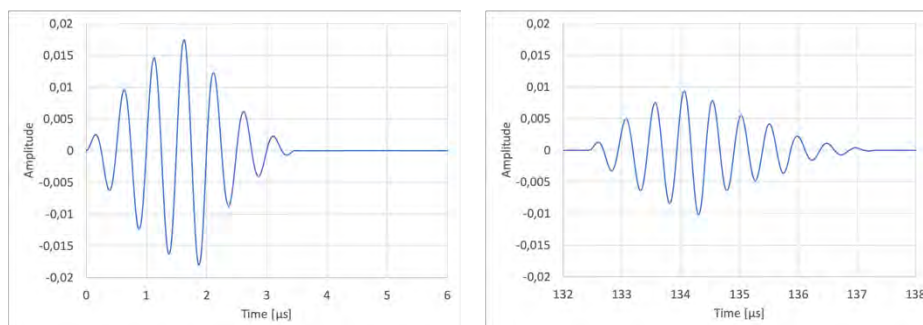


Figure 5. Examples of the generated pulse waveform before and after propagating through a system

## 7. Conclusions

The elaborated method allows for effective, and computationally inexpensive generation of exemplary results of ultrasound tomographic measurements, which can be used as reference measurements in the development of ultrasound waveform tomography algorithms. The flexible geometry and possibility of parametrizing results allows for analysis of factors such as the number of transducers, their directivity, and the size of the ring array on measurement results. The elaborated software due implementation can also be useful as a teaching aid to explain basic ultrasonic wave propagation phenomena using the implemented visualization methods.

## References

1. G. Sandhu, C. Li, O. Roy, S. Schmidt, N. Duric, *High-resolution quantitative whole-breast ultrasound: in vivo application using frequency-domain waveform tomography*, in [SPIE Medical Imaging], 94190D-94190D, International Society for Optics and Photonics (2015).
2. G. Sandhu, C. Li, O. Roy, E. West, K. Montgomery, M. Boone, N. Duric, *Frequency-domain ultrasound waveform tomography breast attenuation imaging*, in [SPIE Medical Imaging], 97900C-97900C, International Society for Optics and Photonics (2016).
3. T. Gudra, K. J. Opieliński, *The ultrasonic probe for the investigating of internal object structure by ultrasound transmission tomography*, *Ultrasonics*, **44**(1-4) (2006) e295 – e302.
4. K. J. Opieliński, P. Pruchnicki, P. Szymanowski, W. K. Szepieniec, H. Szweda, E. Świś, M. Jóźwik, M. Tenderenda, M. Bułkowski, *Multimodal ultrasound computer-assisted tomography: An approach to the recognition of breast lesions*, *Computerized Medical Imaging and Graphics*, **65** (2018) 102 – 114.
5. K. Opieliński, *Analysis and modelling of ultrasonic pulses in a biological medium*, *Archives of Acoustics*, **33**(4) (2008) 13 – 19.
6. H. F. Olson, *Acoustical Engineering*, D. Van Nostrand Company, New Jersey 1957.
7. A. R. Selfridge, G. S. Kino, B. T. Khuri-Yakub, *A theory for the radiation pattern of a narrow-strip acoustic transducer*, *Appl. Phys. Lett.*, **37**(1) (1980) 35 – 36.
8. A. S. Glassner, *An Introduction to Ray Tracing*, Academic Press, Palo Alto 1989.
9. J. Obraz, *Ultrazvuk v měřicí technice*, SNTL, Praha 1984.

## Coding Effects on Changes in Formant Frequencies in Japanese Speech Signals

Mateusz KUCHARSKI

*Wroclaw University of Science and Technology, Faculty of Electronics,  
Department of Acoustics and Multimedia, 50-370 Wroclaw,  
Wybrzeze Wyspianskiego 27, 223513@student.pwr.edu.pl*

Stefan BRACHMAŃSKI

*Wroclaw University of Science and Technology, Faculty of Electronics,  
Department of Acoustics and Multimedia, 50-370 Wroclaw,  
Wybrzeze Wyspianskiego 27, stefan.brachmanski@pwr.edu.pl*

### Abstract

This paper presents results of research on effects of lossy coding on formant frequencies for Japanese speech signals. Additionally changes in pitch of the voice were inspected. For this research four most popular lossy coding standards were chosen, MP3, WMA, AAC and OGG, and compared to original WAVE files. Audio files were created by the author based on ITU-T P.501 recommendation in two sampling frequencies, 16 kHz and 48 kHz, and converted into chosen codecs. To extract the data from audio files, open license software Praat was used. Due to discovered differences in time duration between original and encoded files, that also differed between individual codecs, only OGG and WMA standards were compared directly. MP3 and AAC standards were divided into Japanese syllables, averaged and then compared into also averaged WAVE files. Results were additionally compared to FLAC lossless codec.

**Keywords:** speech, speech coding, formant

### 1. Introduction

Researching coding influence on formant frequencies requires creating database of speech signals. Database of speech signal used for this research was created by the author as a part of BSc thesis (Tab. 1) [1, 4]. Sentences used were taken from ITU-T P.501 recommendations sentence list of Japanese language [2]. They were recorded in environment meeting the requirements of ITU-T P.800 using condenser microphone and open licence software [3]. Sampling rate during recording was 48 kHz, later additional 16 kHz audio files were created using those original files. All files were monophonic and coded in 16 bit PCM. Recordings of the same sentences were made twice, during two different sessions. Original WAVE files, in both 16 kHz and 48 kHz, were later converted into four chosen most popular lossy codec formats: MPEG Layer 3 (MP3), Windows Media Audio (WMA), Advanced Audio Coding (AAC) and OGG Vorbis (OGG), as well as lossless FLAC. Final database consists of 192 audio files.

During preparations for the experiment, several issues were discovered. Before data extraction begun, files' parameters were checked by additional software and during that author found out differences in time durations between original WAVE and encoded files. Those encoded with MP3, AAC and WMA codecs had longer time duration than

originals, while those with OGG and lossless FLAC were identical in that matter. The additional lengths were also different depending on the codec. Differences were smallest for the WMA codec, 5 ms for the first file, and noticeably larger for MP3, reaching 85 ms. The AAC codec was also problematic, even during database creation, and due to inconsistent bitrate used software was unable to identify its length. While partly responsible for the differences turned out to be codec starting in the beginning of the file, it was only about 15 ms, and the remaining 70 ms were in the file itself. The origin of these differences is unknown.

Table 1. Japanese sentences contained in database.  
The transcription of Japanese texts come from ITU-T.501 recommendation

Number of sentence	Notation	Sentence
1	Original	彼は鮎を釣る名人です。
	Hiragana	かれわあゆをつる めいじん です。
	Transcription	Kare wa ayu wo tsuru meijin desu.
2	Original	古代エジプトで十進法の原理が作られました。
	Hiragana	こだい えじぶと で じゅっしんほう の げんり が つくられました。
	Transcription	Kodai ejiputo de jussinhou no genri ga tsukurare mashita.
3	Original	読書の楽しさを知ってください。
	Hiragana	どくしょ の たのしさを しってください。
	Transcription	Dokusho no tanoshisa wo shitte kudasai.
4	Original	人間の価値は知識をどう活用するかで決まります。
	Hiragana	にんげん の かちわちしきを どう かつよう するか できまります。
	Transcription	Ningen no kachi wa chishiki wo dou katsuyou suruka de kimarimasu.
5	Original	彼女を説得しようとしても無駄です。
	Hiragana	かのじょ を せっとく しようとしても むだです。
	Transcription	Kanojo wo settoku shiyoutoshitemo mudadesu.
6	Original	その昔ガラスは大変めずらしいものでした。
	Hiragana	その むかし がらすわたいへん めずらしい ものでした。
	Transcription	Sono mukasi garasu wa taihen mezurashii monodeshita.
7	Original	近頃の子供たちはひ弱です。
	Hiragana	ちかごろ の こどもたちわひよわです。
	Transcription	Chikagoro no kodomo tachi wa hiyowa desu.
8	Original	イギリス人は雨の中を平気で濡れて歩きます。
	Hiragana	いぎりすじん わ あめの なかを へいきで ぬれて あるきます。
	Transcription	Igrisu jin wa ameno nakawo heikide nurete arukimasu.



Another problem that occurred was that software used for data extraction, program Praat, which did not support OGG, AAC and WMA files. To bypass this issue, files in these encodings were converted once again into WAVE. Because they were encoded with lossy codecs, the changes in parameters, that author was interested in, should have already impacted those files and thus be identical after another conversion using a lossless encoding. Also because AAC format is primarily designed to contain stereo data, it artificially created double tracks from original mono. Thus it was necessary to convert it back to mono using Audacity software.

## 2. The research

The problem of lengths was resolved by using two methods of comparing files. First method was used for OGG, WMA and FLAC codecs, as OGG and FLAC had the same number of points in which the formant measurements were taken and the WMA codec only differed by several points. The method was a simple side-by-side comparison. Using Praat software, a list of formant points was extracted, with a 0.00625 second interval between each point, for both the examined codec and original WAVE file. Then a simple subtraction for the first and second formant frequencies in each point between two files was conducted, and an absolute value of the result was taken.

The second method was a bit less straightforward. As the differences in formant points were too big, files could not be compared side-by-side. It was decided that each file was going to be divided into several parts, then data from each part would be extracted separately. A mean value would be taken and compared to mean value of corresponding part of the WAVE file. Fact, that Japanese language was being examined during this research turned out to be important in this part. Words might have been too long as chunks of data, and also a word can be really short as well as really long. On the other hand singular letters can be extremely hard to separate from each other. Fortunately, Japanese syllables are quite easy to separate. Also unlike western syllables, there is a very well defined number of them, due to writing systems functioning in Japan. There are three writing systems there:

- Kanji - Chinese logographic characters, that are used for the words' meanings (there are about 2200 of basic characters and many more specialistic and archaic ones);
- Hiragana - syllabary consisting 46 basic characters, 25 diacritics and 36 digraphs, used mostly for particles and word ending modifications;
- Katakana - syllabary consisting of the same syllables that hiragana, but used for foreign-borrowed words.

All of Japanese language can be written in both syllabaries, and it often is, for children and people learning it, so it was decided that Japanese syllables were going to be those needed chunks of sentence for the research.

The data was extracted using "view and edit" mode in Praat software. This is the mode that displays time course on a graph in the upper and spectrogram in the lower part of the screen, both of which share the same time axis. It is also possible to display formant points (red dots) as well as pitch (blue line) on the spectrogram. This display, with aid of playing selected parts of examined file, enabled the author to fairly

accurately separate syllables and thus extract wanted data. So, from each chunk, the list of formant points was extracted, and then mean values for first two formants were calculated. Additionally maximum and minimum values were noted.

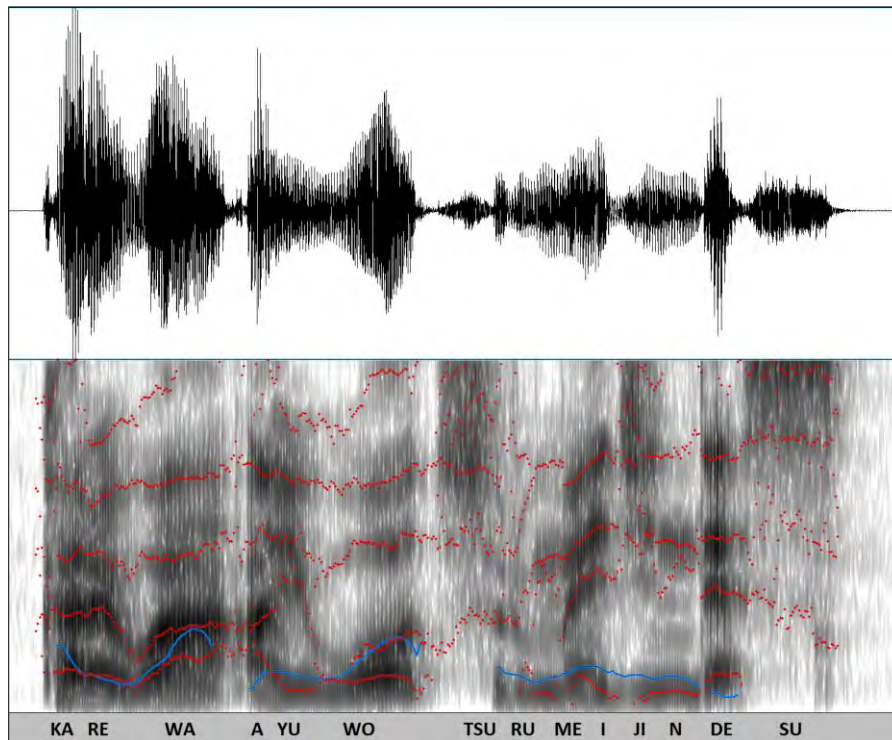


Figure 1. Time course (upper part) and spectrogram with formants and pitch (lower part) for the first sentence

In similar fashion data about pitch of the voice was collected. For OGG, WMA and FLAC codecs it was extracted as a list for the whole file. For the rest, it was extracted from the same parts (syllables), as formant data. There, the mean value as well as minimum and maximum were taken from software.

Data collected for each codec was then compared to corresponding data from original WAVE files. For the presentation, absolute values of differences were calculated. Results of this are presented in tables 2 and 3. There is one table containing mean values of differences in formant frequencies and their standard deviations for all lossy codec, as well as a similar table containing data about pitch, also for all codecs. Each table contains results for two versions of the sentence, recorded during two different sessions. Each version is also available in both 16 kHz and 48 kHz sampling frequency.

The comparison of lossless FLAC to original WAVE files resulted in exactly none differences - all formant points' values were the same for both. This result was expected, but its' point was not to check if there are any differences, but if the software used for

this research does not influence the outcome. Because differences were nonexistent, with respect to this standpoint, it is safe to assume that results achieved for the rest of researched files are accurate. As the results for this part were all zeros, it was assumed that presenting them in a table is not necessary.

Table 2. Mean values of formant differences with standard deviation for all codecs

Codec	Recording 1				Recording 2			
	16 kHz		48 kHz		16 kHz		48 kHz	
	F1	F2	F1	F2	F1	F2	F1	F2
OGG	35.51	40.27	15.68	17.43	34.18	35.12	13.85	16.51
Standard deviation	12.27	6.62	7.31	14.33	17.70	8.86	10.30	14.46
WMA	50.39	59.66	65.35	81.71	60.35	67.32	60.88	76.89
Standard deviation	12.06	13.28	17.43	23.18	29.75	24.88	11.59	18.85
MP3	53.27	61.19	48.43	53.59	34.78	52.83	29.49	35.70
Standard deviation	17.81	32.25	11.05	16.76	8.62	42.24	7.26	6.02
AAC	40.41	52.03	45.79	64.76	38.59	59.61	33.59	43.44
Standard deviation	18.26	31.56	20.11	41.01	11.00	50.19	8.52	8.71

Table 3. Mean values of pitch with standard deviation for all codecs

Codec	Recording 1		Recording 2	
	16 kHz	48 kHz	16 kHz	48 kHz
OGG	0.40	0.10	0.34	0.01
Standard deviation	0.60	0.25	0.35	0.01
WMA	2.59	3.83	5.09	5.88
Standard deviation	1.79	1.88	4.78	4.75
MP3	2.73	1.82	6.48	3.80
Standard deviation	3.74	0.86	9.10	4.86
AAC	2.92	4.51	2.31	3.36
Standard deviation	3.59	5.10	2.62	3.46

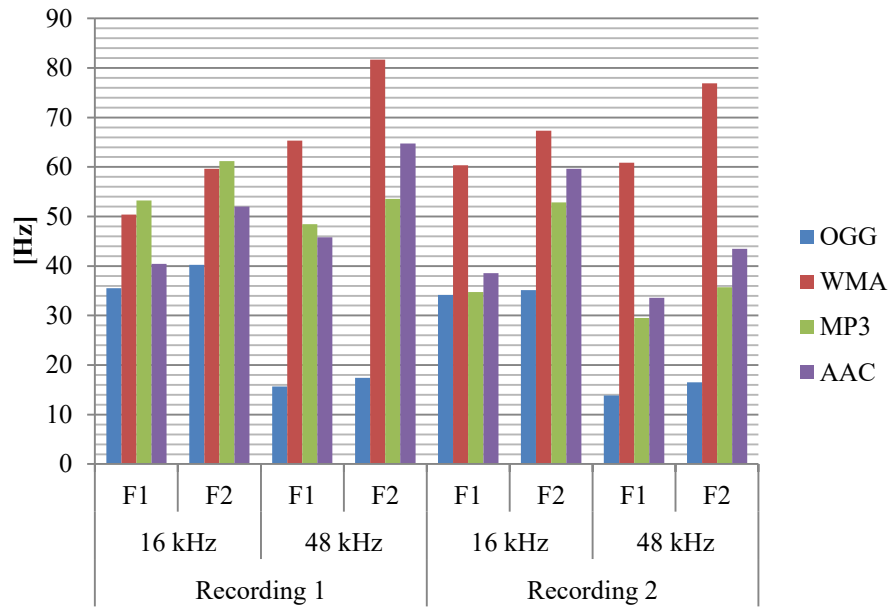


Figure 2. Mean values of formant differences for all codecs

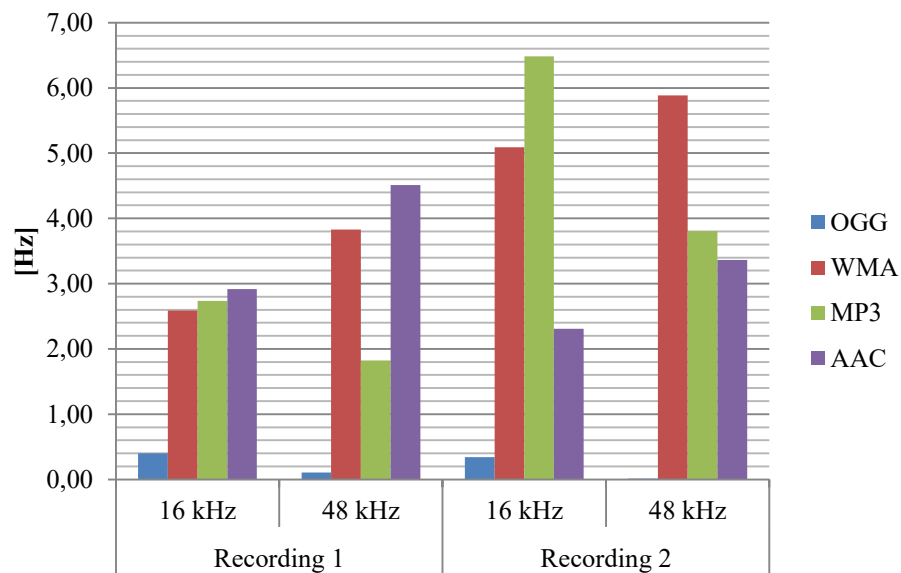


Figure 3. Mean values of pitch for all codecs

### 3. Results

Differences in formant frequencies fluctuate from about 15 Hz up to 80 Hz, depending on codec, sampling frequency and number of the formant. The lowest values were achieved for the OGG codec. Recording session did not matter that much in this case, however there is a significant difference between sampling frequencies. While for the original 48 kHz results are quite small, about 15 Hz, the converted 16 kHz versions have over twice as much. In WMA codecs' case the values are much higher, in fact they are the highest among all codecs. It might be caused by the method used for data extraction. It was examined the same way OGG was, but unlike OGG there are differences in lengths. They are not very big - highest noted difference in number of formant points was six and the mean value is about three. Also sentence number five has unusually high values, around and above 100 Hz difference, with 120 Hz being maximum for the first formant and 124 Hz for the second. MP3 codec was the first to be examined using the second method, that is by dividing file into parts. Mean values are lower than WMAs', but it also has some spikes over 100 Hz, reaching even 183 Hz in one point. Those points however are exceptions and the rest are close to mean value. The AAC codecs' results are very similar to MP3s', they even have spikes in the same places and with similar values.

Results of research in terms of pitch of the voice are also interesting, because there is almost no difference between original file and lossy encoded ones. As it can be seen in tables above, OGG has again the lowest differences, which are almost always below 1 Hz. The rest of the codecs give a few hertz differences with mean values usually not exceeding 6 Hz. Again, there are several files that have higher difference values, up to 14 Hz. In case of WMAs' higher values, it is also fault of uneven lengths. It can be seen that generally the values are also below 1 Hz, but there are additional points, just like formant points, that add values above 100 Hz. Because of that mean values are higher. There were also several unusual occurrences, when software detected pitch on much higher than normal frequencies, from 400 up to 500 Hz. These are however rare and do not last long, so it is safe to assume that in general lossy coding has almost no effect on pitch of the voice.

### 4. Conclusion

The objective of conducted research was to check, if popular lossy codecs have effect on certain voice parameters. The parameters were formant frequencies and pitch of the voice. Results presented in this paper show that for the formant frequencies there are in fact some differences. OGG codec, that was the easiest to examine, had differences of about 15 Hz for 48 kHz sampling frequency, and about 35 Hz for 16 kHz. The rest of the codecs, WMA, MP3 and AAC, had much higher values, with WMA having the highest, above 60 Hz mean, up to even 120 Hz. MP3 and AAC had slightly lower overall values, but in case of one file got almost up to 190 Hz. Unlike formant frequencies, pitch of the voice had almost no changes at all, again with several exceptions.

There are several potential perspectives for the future expanding of this research. It is possible to expand the database for other speakers, including natives. It is also possible

to check the results for another languages and research influence of the language itself on those parameters. Also another, more accurate methods might be developed.

## References

1. S. Brachmański, *Wybrane zagadnienia oceny jakości transmisji sygnału mowy*, Wrocław: Oficyna Wydawnicza Politechniki Wrocławskiej, 2015.
2. ITU-T Recommendation P.501, *Test signals for use in telephonometry*, 2017.
3. ITU-T Recommendation P.800, *Method for subjective determination of transmission quality*, 1996.
4. M. Kucharski, *Realization of Japanese sentences sets acoustical database for selected coding techniques*, Wrocław, BSc Thesis, Wrocław University of Science and Technology, 2017.

## Design of Control System for Active Vibration Suppression of Trapezoidal Plate

Konrad MIERZEJEWSKI

*University of Rzeszów, Pigońia 1, 35-310 Rzeszów,  
kmierzejewski@ur.edu.pl*

Lucyna LENIOWSKA

*University of Rzeszów, Pigońia 1, 35-310 Rzeszów,  
lleniew@ur.edu.pl*

### Abstract

An active vibration control system is proposed for suppressing the small amplitude plate vibration. The structure under study is a vibrating trapezoidal plate, having a constant thickness, to which MFC (Micro Fiber Composite) actuator is bonded. It was assumed, that the plate clamped at one edge is excited by a uniform periodic force generated by a loudspeaker. The control problem lies in using MFC actuator to reduce the plate vibrations. For the system under consideration the mathematical model obtained on the base of parametric identification method is constructed. This part of the research was done with the help of Polytec laser vibrometer. The apparatus is highly advanced tool that allows measurement of vibration of examined structure. With transfer function model obtained in identification process, using Matlabs Identification Toolbox, feedback control laws was created for changing response of the system in desired way. There are many ways to model controller having mathematical model of the object. In this article, authors propose approach to design an effective controller for vibration suppression of a trapezoidal plate with the use of the pole placement method in graphical SISOTool environment. This article describes concept, results of simulation tests and implementation for the designed controller.

**Keywords:** trapezoidal plate, active vibration control, MFC element, ARX model

### 1. Introduction

The issue of low frequency vibrations control is an extremely important problem that should be considered both at the design stage and later during their operation. The constant advancement of information technology allows real-time systems to be designed more and more effectively [2, 3, 13]. Working with efficient data acquisition cards, give the possibility to simulate, implement and test control systems. Rapid prototyping of control systems enables software such as Altium Designer, LabView, MATLAB/Simulink and many more. This article describes a study in which the xPC Target platform [8] was used, cooperating with the MATLAB environment, to design and implement a vibration control system for a trapezoidal plate using a collocated MFC sensor / actuator.

In general, vibration control systems can be divided into two groups: passive and active methods [7]. The first focuses on reducing adverse vibrations by modifying the parameters and structures of the systems. Active methods, unlike passive ones, based mainly on storage and energy dissipation, compensate vibrations through own production. Such solutions consist in connecting an energy source to the external system.

Such element can be placed anywhere in the device and its work in such a way as to generate or absorb vibroacoustic energy to reduce redundant vibrations.

The most important element of active vibration reduction systems (AVC) are sensors and actuators. They allow you to collect information about the subject of the test, so it can be used later to create a vibration compensation algorithm. Their construction should be considered in such a way that they will not negatively affect the system in which they will be mounted, for example by increasing stiffness or adding weight. Such element should be characterized by low mass, high efficiency, long life and small size. Piezoelectric elements come out against this challenge [1]. Smart Materials MFC are a new type of piezoelectric tools that sensors and actuators are made of. The mechanical and electrical properties of MFC have been described in many articles, confirming their effectiveness and reliability [10]. Researchers dealing with the issue indicate the placement of sensory and executive elements as extremely important. Improper deployment can cause lack of observability and controllability, which strongly affect the quality of the control system [12].

Easiest way to create a closed-loop system contains two steps. The first is to identify the mathematical model of the object being studied. The second step is to use the obtained model to design the controller. During the research, the authors determined the resonance frequencies, harmonics of the tested board and designed tuned controllers to suppress them. For this purpose, the generated ARX model was used by the standard LMS algorithm from MATLAB System Identification Tools. This article describes the method of creating an effective system of vibration control of a trapezoidal plate, using SISO Tool from MATLAB software. The research stand was based on the xPC Target platform, controlling the workstation equipped with two multifunction boards. At the end, the results of both simulations and preliminary results of the real application of the designed systems will be presented.

## 2. Construction of the stand

The object that was tested was an aluminum trapezoidal plate with dimensions as follows: longer base 63cm, shorter base 21 cm, height 44 cm [6]. It was attached for a longer base to iron-cast cuboid with a height of 90 cm and a weight of about 120 kg. A laser vibrometer measuring head was placed above the station. A rectangular MFC M-8557-S1 [11] element with dimensions of 85 x 47 mm was glued to the tested board.

In the first part of the experiment, its task was to transfer vibrations to the object and acquire data for the needs of the identification process. Later, after connecting to the station with xPCTarget, using the collocated sensor and actuator, it served as a vibration sensor of the board and transmitted to it the signal of the vibration damping regulator.

The location of the element on the plate was selected based on the vibrometer tests, in order to simulate the exemplary work and effectively vibrations reduction.



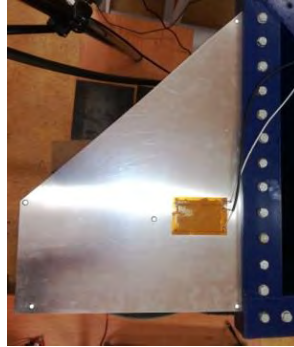


Figure 1. Research object with attached MFC sensor/actuator

All elements of the measurement system are shown on Fig. 2. The stand consisted of: 1 – vibrometer measuring head, 2 – LCD monitor, 3 – computer station with scanning software, 4 – cuboid body, 5 – trapezoidal plate, 6 – loudspeaker.



Figure 2. Testing stand

### 3. Measurement of vibration

To create a mathematical model of the tested object, the identification procedure should be performed. In order to understand the acoustic properties of the test panel, a laser vibrometer test was carried out. It involves the imposition of a virtual grid of measurement points on the object, and then inducing vibrations with a specific signal to collect data about the object's response. The plate was excited to vibration with a signal of constant amplitude and a changing frequency in the range of 20 to 500 Hz. The device's software generated an animation showing the disc response to the given stimulation. The result is shown below.

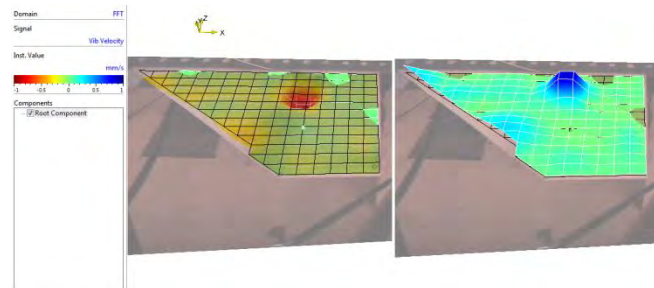


Figure 3. Resonant model shape. Maximum velocity amplitude

FFT analysis allowed determining the resonant frequencies of the plate. The reduction of vibration amplitude is the result of the operation of control signals determined by the designed controller.

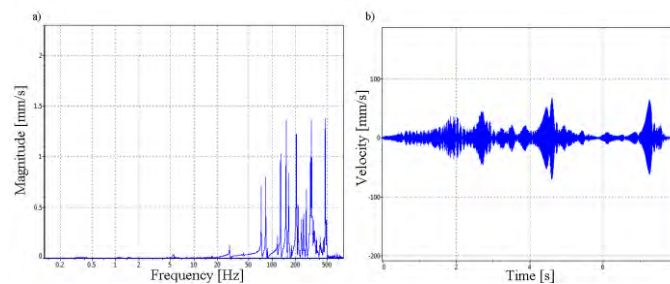


Figure 4. Diagrams: a) resonant frequencies of the plate;  
b) plate response for chirp excitation of 20–500 Hz

The reduction of the largest vibration amplitude at the first resonant frequency usually causes the reduction of other harmonics with smaller amplitudes occurring at different frequencies. With this in mind, the 73 Hz and 315 Hz frequencies were chosen for the digital controller design.

#### 4. Identification of the object

In order to find the parameters of the digital controller one should identify the object subject to control, i.e. determine its mathematical model. It defines the behavior of the object under certain conditions, which are described by the input and output of the object at present and in the past. Naturally, the model is a certain approximation of the actual object. Different types of models are used depending on the purpose of the model being created and the structure of the object being identified. Assuming that the sampled signal values can be related through the linear difference equation given by Eq.(1) [9]:

$$\begin{aligned} & y(k) + a_1 y(k-1) + a_{nA} y(k-nA) \dots \\ & = b_1 u(k-d) + b_{nB} u(k-d-nB+1) + e(k), \end{aligned} \quad (1)$$

where  $y(k)$  and  $u(k)$  are respectively the output and input in discrete time  $k = 1, 2, 3, \dots$ ;  $nA$  is the number of poles,  $nB$  is number of zeros,  $d$  indicates number of samples before input reacts to system output and  $e(k)$  stands for the noise.

Idea of active vibration suppression can be generally characterized by the diagram (Fig. 5). Plant is described by the transfer function  $G_0$ , the actuator-to-error transfer path. Used MFC element (controller) uses error sensor data to generate actuator drive signal to reduce plate vibration. Autoregressive model with exogenous input (ARX) which was shown above (1) can be represented using  $z^{-1}$  operator as Eq.(2):

$$y(k) = \frac{B(z^{-1})}{A(z^{-1})}u(k-d) + \frac{1}{A(z^{-1})}e(k), \quad (2)$$

where:

$$A(z) = 1 + a_1z^{-1} + \dots + a_{nA}z^{-nA}, \quad (3)$$

$$B(z) = b_1 + b_2z^{-1} + \dots + b_{nB}z^{-nB+1} \quad (4)$$

is considered as a base model of the vibrating planar trapezoidal plate.

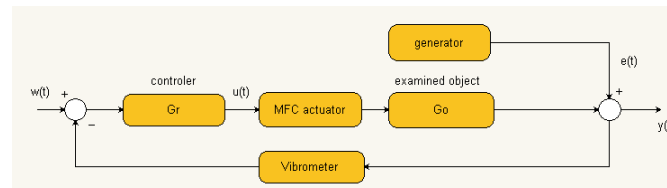


Figure 5. Control system diagram

In order to obtain polynomial coefficients in (3), (4) authors used sine signals with frequencies obtained in previous section (73 Hz and 315 Hz) and chirp signal which frequency changes from 20 to 500 Hz. Those three identification processes led to design three digital controllers, two tuned to reduce chosen resonant frequencies and one which reduce whole frequency band. The accuracy of the model derived was estimated according to the MATLAB criterion of *best fit* defined as:

$$BEST_{FIT} = (1 - |y - y_{model}| / |y - mean(y)|) * 100\%, \quad (5)$$

where  $y$  stands for measured output,  $y_{model}$  is simulated or predicted model output, and  $mean(y)$  designed mean or measured output. Results of identification process are shown below.

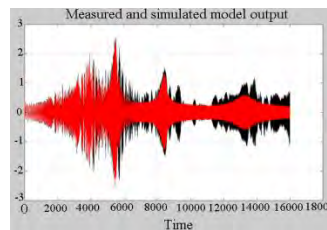


Figure 6. Identification results for chirp signal of 20-500 Hz

## 5. Controller design

Having transfer function obtained in identification process we can create a feedback control law to change the response of the system in a desired way [2, 4]. In order to fit the established criterion of our closed-loop system we were looking for a controller transfer function presented as the well-adjusted ratio of polynomials  $Q(z^{-1})/P(z^{-1})$ . Using this technique, authors have to bear in mind that the requirement of closed-loops systems is to keep polynomial roots in the left-hand plane.

Using SISO Designing Tool software pole placement was done by the designer. Finally, location of poles and zeros for the closed loop systems is shown in the figure below.

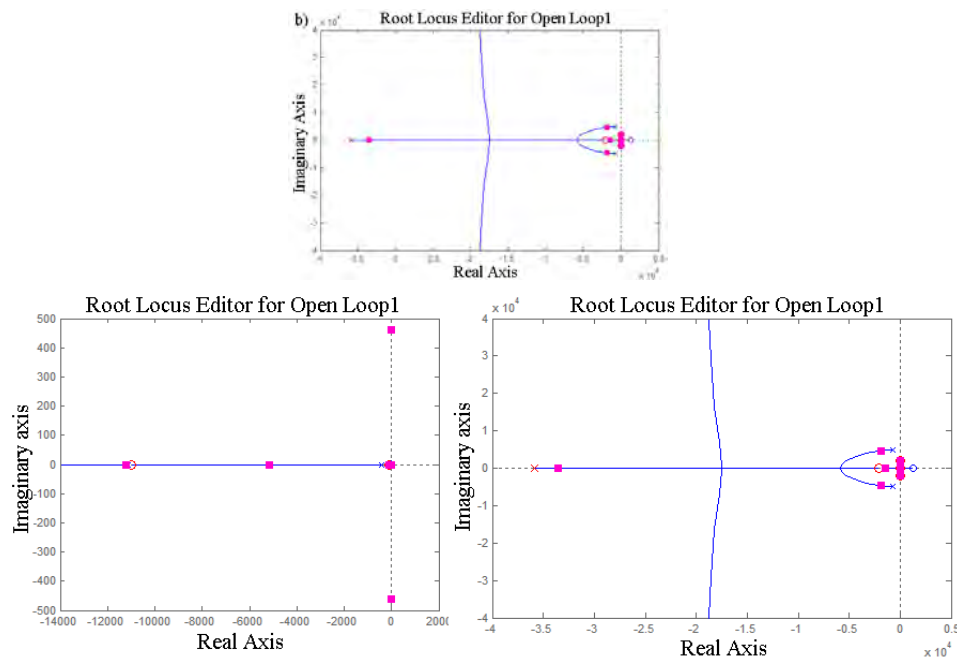


Figure 8. Location of zeros and poles of the system: a) identified 73 Hz, b) identified 315 Hz, c) identified chirp

Obtained transfer function of controller for control system based on identified 73 Hz sine signal presents as follow:

$$G_c = \frac{-4.634z^2 + 2.214z - 2.04}{z^2 - 1.992z + 0.9921} \quad (6)$$

Using MATLAB/Simulink software simulations of designed control systems was done. Acting of designed controller has been verified by testing behavior of the plate driven by the sinusoidal signal with frequency 73 and 315 Hz. Results are presented in Fig. 9.

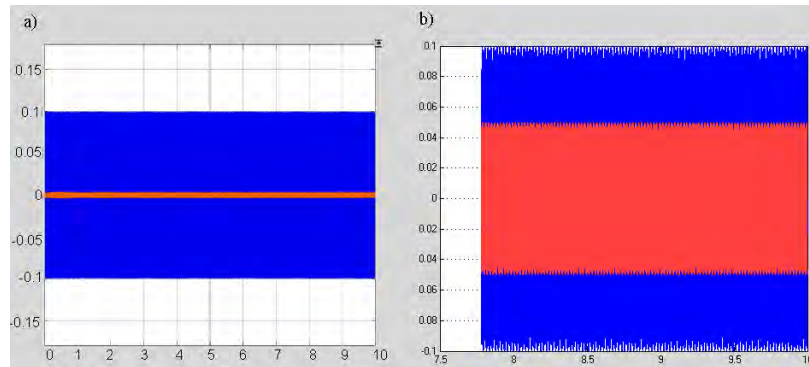


Figure 9. Simulation results for designed controller for: a) 73 Hz excitation, plate vibration reduced by about 90%, b) 315 Hz excitation, plate response reduced by about 50%. Both: blue color – open loop, red – closed loop system

## 6. Preliminary results of implementation

Designed controller has been implemented and tested on a real-time platform – PowerDAQ with xPC Target environment supervising action of two multifunctional boards. First card (PD2-AO-8/16) is the Analog Output Module used for controlling MFC actuators. The sampling rate at 100 kS/s per channel and 16 bit converter resolution were sufficient to completely control the operation of the actuators. To adjust the output voltage D/A cards to voltage level of the MFC actuator (from 1,500 V to -500V), it was necessary to use a high-voltage amplifier (HVA 1500 50/4), which processes the signal in the frequency range from 0 Hz to 16 kHz. The second card (PD2-MFS-8-500/14) contains analog inputs with 14-bit A/D transducer. One of the inputs of this card, which was connected directly to MFC element (M-8557-S1), provides voltage values within the range of  $\pm 10$  V and acceptable sampling rate [3]. Initial results of implementation are shown below. For the purpose of the experiment, plate was excited by sine wave signal, with frequency of 73 Hz and amplitude of 1 V. Obtained results shows substantial reduction, giving prospects for future research. Plate vibration was reduced to about 0.2 V i.e. 80% of initial excitation.

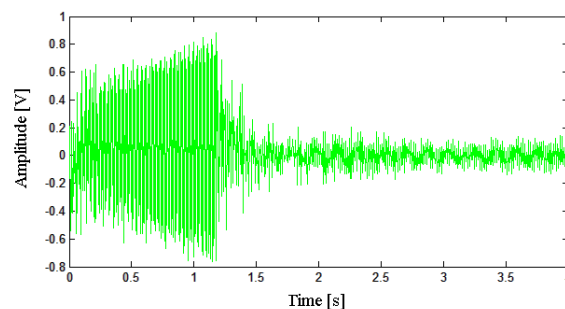


Figure 9. Initial results of tested control system

## 7. Conclusions

The main goal of the research was to investigate capability of MFC elements to suppress trapezoidal plate vibrations. The paper shows next steps in process of designing control system. Described stand is based on xPCTarget platform watching over two multifunctional boards. Feedback control is realized with rectangular MFC element attached to the plate. During research, the mathematical model was obtained and its transfer function was used for designing digital controller with the pole-placement method. Results of experiment, both simulation and real test confirmed effectiveness of chosen method and show that the plate displacement can be considerably suppressed.

## References

1. H. T. Banks, R. C. Smith, Y. Wang, *Smart Material Structures: Modeling, Estimation and Control*, Wiley-Masson, New York, 1996.
2. G. Ferrari, M. Amabili, *Active vibration control of a sandwich plate by non-collocated positive position feedback*, Journal of Sound and Vibration, **342** (2015) 44 – 56.
3. L. Leniowska, M. Sierżęga, *The vibration control of a circular plate by the use of a parametric controller with phase shift adjustment*, Mechatronics, **58** (2019) 39 – 46.
4. D. Mazan, L. Leniowska, *MFC Sensors and Actuators in Active Vibration Control of the Circular Plate*, Archives of Acoustics, **40**(2) (2015) 257 – 265.
5. K. Mazur, M. Pawełczyk, *Feed-forward equalization of sound radiation from a vibrating plate*, Mechanics and Control, **31**(1) (2012) 35 – 39.
6. K. Mierzejewski, L. Leniowska, *Active Vibration Control of Trapezoidal Plate using MFC Collocated Sensor/Actuator*, IEEE Proceedings of 2018 Joint Conference Acoustics, Ed. Jacek Marszał, Iwona Kochańska, 188 – 191. IEEE Conference Record: # 44813.
7. A. Preumont, K. Seto, *Active Control of Structures*, John Wiley & Sons, Ltd. 2008.
8. M. Sierżęga, L. Leniowska, *The application of xPC target platform for a circular plate vibration control*, Solid State Phenomena, **248** (2016) 135 – 141.
9. T. Söderström, P. Stoica, *System identification*, Prentice Hall Int., London, 1989.
10. T. D. Usher, K. R. Ulibarri, G. S. Camagaro, *Piezoelectric Microfiber Composite Actuators for Morphing Wings*, ISRN Materials Science, (2013), Article ID 189659.
11. W. K. Wilkie, G. R. Bryant, J. W. High et al., *NASA Langley Research Center Macro-Fiber Composite Actuator (LaRC-MFC): Technical Overview*, 2004.
12. S. Wrona, M. Pawełczyk, *Controllability-oriented placement of actuators for active noise-vibration control of rectangular plates using a memetic algorithm*, Archives of Acoustics, **38**(4) (2013) 529 – 536.
13. S. Wrona, M. Pawełczyk, *Shaping frequency response of a vibrating plate for passive and active control applications by simultaneous optimization of arrangement of additional masses and ribs. Part I Mechanical Systems and Signal Processing* 70, 682 – 698.

## Tests of Basic Voice Stress Detection Techniques

Piotr STARONIEWICZ

*Chair of Acoustics and Multimedia,  
Wrocław University of Science and Technology,  
Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland,  
piotr.staroniewicz@pwr.edu.pl*

### Abstract

The modern speech processing techniques enable new possibilities of potential applications. Besides speech and speaker recognition, also the information about speakers' physical condition, emotional state or stress can be detected in speech signal. Since emotional stress can occur during deception, its detection in speech could be used for law or security services. The paper presents the comparative tests of two voice stress detection techniques: one based on trials of microtremors detection relying on an iterative EMD method (Empirical Mode Decomposition) and the second one based on the statistical analysis of fundamental frequency and MFCC parameters. The preliminary tests were carried on the group of 12 speakers (6 males and 6 females) answering yes/no to the list of a few dozen personal questions. The presented research revealed the speakers' very high personal influence on the obtained results.

**Keywords:** Voice Stress Analysis, Empirical Mode Decomposition

### 1. Introduction

The speech processing technology makes it possible nowadays not only to recognise speech and speakers, but also to identify complex information about a speaker's state or condition. The techniques of speaker emotion recognition [1, 2] are dynamically developing and, in some cases, can be even more efficient in proper emotion classification than humans [3].

From the possible applications' point of view a very interesting method is the technique of stress detection in speech, also known as VSA (Voice Stress Analysis). The stress can be caused by external (i.e. physical) or internal (i.e. psychological) factors. Since deception is one of the possible internal factors, the detection of it could be a valuable application for law or security services. Using the classical polygraph encounters an important obstacle which is the necessity of physical connection to the subject. Hence on the one hand the techniques which do not have such a connection (e.g. face thermal vision or VSA) let us reduce other stress factors for the examined subject, but on the other hand, and can be more discreet or even unnoticeable by the subject. The basis of the numerous commercial VSA applications is the controversial Lippold [4] theory from the 1970s of microtremors, i.e. the reaction of muscle tension of vocal chords during the stress of around 8-12 Hz. At the same time, there are numerous commercial applications on the market, whose working algorithms are not disclosed for the understandable reasons and there are still very few scientific reports letting us assess the real usability of the VSA techniques.

## 2. Speech database

For the voice stress detection purposes a special speech database was designed. Similarly as in the polygraph tests it includes individuals' yes/no answers to the list of a few dozen questions. As the database speakers the couples which were in a relationship for a certain time (at least one year) were recruited. During the recordings the decisive part was played by the content of the asked questions. It was supposed to make the speaker abandon his/her personal "comfort zone" and force him or her to answer some questions deceitfully. During the recordings, the partners were sitting on the opposite sides and asking each other questions from the list. Each partner had a different list of questions. Three kinds of questions were used:

- *Relevant questions*, which were significant for obtaining the information from the object. These questions were asked directly and they concerned the relation between the partners. The purpose of those questions was to evoke stress, for example during the speaker's deceit.
- *Irrelevant questions* were used as a buffer between the relevant questions. The purpose of those questions was to introduce some break and the speaker's relaxation and they were not related to the topic of the discussion.
- *Control questions* reveal the truth and their purpose is to demonstrate the comparison to the relevant question.

The recordings were carried out in good acoustic conditions in a quiet room with the dynamic microphone Shure SM 58 SE, the acoustic mixer Behringer Eurorack MX 802A and the external analogue-digital converter Creative Sound Blaster Audigy 2 NX. The signals were recorded with the sampling rate of 44,100 samples per second with 16 bit resolution.

## 3. Tested stress detection techniques

Two stress detection techniques were applied. The first one was based on the trials of microtremors detection and the second one was based on the statistical analysis of chosen voice parameters.

The first technique relies on the iterative EMD method (Empirical Mode Decomposition), which was proposed by Huang in 1993 [5]. The EMD method allows to present the analysed, nonstationary signal as a sum of stationary signals called IMFs (Intrinsic Mode Functions).

The four methods were tested for that technique of stress detection. The simplest one (denoted as VSA1 below) extracts the microtremor from the signal. If the microtremor is in the range of 8-12 Hz, the result is "true", otherwise the application recognises the utterance as "false". This simple test was then developed into a method (VSA2) where the distribution of component frequencies was examined. The method recognises the "true" when the ratio of component frequencies inside the investigated band (8-12 Hz) to the sum of all component frequencies exceeds a chosen value. The third method (VSA3) is very similar but it demands the calibration process for each speaker, where the answers to the irrelevant questions are used. The fourth method (VSA4) also demands the calibration process with the answers to the irrelevant questions. During



the classification the algorithm compares the most significant component frequencies of the recognized sample to the most significant component frequencies of the reference samples.

The second applied technique was based on the analysis of voice parameters:

- Mean F0 value,
- Range of F0,
- Jitter,
- MFCC parameters (Mel Frequency Cepstral Coefficients).

Since the used parameters are from various domains, their values were normalized. The block scheme of the technique was presented in Figure 1. Two methods were tested for that technique: with mean (denoted as VSA5) and median (denoted as VSA6) values calculated for each parameter. The range of each parameter was then determined on the basis of standard deviations. If the value of the parameter for the answer to the relevant question was not in the determined range the “false” was recognized by the algorithm.

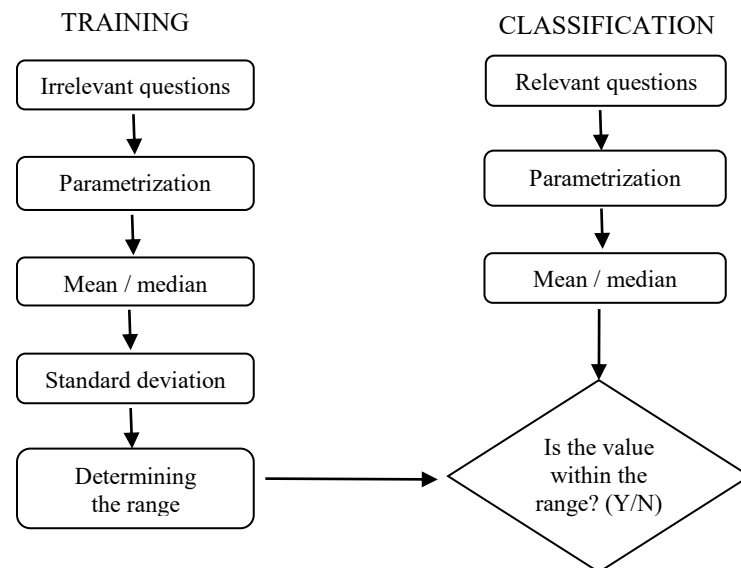


Figure 1. Scheme of stress detection based on statistical analysis of voice parameters

### 3. Results and discussion

The correctness of deception detection was defined as the ratio of the correctly identified negative and positive data samples to all the data samples. The singular results of the deception detections for each tested speakers were presented in Table 1, where the first four columns of the results (VSA1-VSA4) were obtained for the technique based on EMS detection and the last two columns (VSA5 and VSA6) for the technique based on the statistics of voice parameters. It can be noticed that the differences between

the results obtained for the singular subjects can be very substantial: from a very high detection correctness of around 80% (e.g. subject F1) to quite low, even around 40%, for the voices of some male subjects (e.g. subject M6).

Table 1. Detection of deception results for 12 speakers database and six tested stress detection techniques

Speakers		VSA1	VSA2	VSA3	VSA4	VSA5	VSA6
Male / female	Speaker's number						
Females	F1	70.7%	41.3%	70.7%	35.4%	64.8%	70.7%
	F2	81.4%	75.1%	75.1%	87.6%	76.6%	76.6%
	F3	47.2%	41.3%	64.8%	82.5%	64.8%	47.2%
	F4	58.9%	58.9%	47.2%	41.3%	70.7%	70.7%
	F5	56.4%	62.6%	68.9%	62.6%	58.9%	58.9%
	F6	47.2%	58.9%	52.9%	76.6%	76.6%	64.8%
Males	M1	47.2%	58.9%	70.7%	41.3%	58.9%	47.2%
	M2	47.2%	47.2%	64.8%	70.7%	56.4%	50.1%
	M3	41.3%	52.9%	47.2%	64.8%	64.8%	64.8%
	M4	58.9%	59.6%	58.3%	53.4%	81.4%	56.4%
	M5	58.9%	76.6%	64.8%	58.9%	52.9%	47.2%
	M6	37.6%	62.6%	43.9%	31.4%	73.4%	53.4%

The mean values of deception detection for male and female voices for all the tested methods are presented in Figure 2. All the obtained mean results exceed 50%, however, such results rather disqualify the tested techniques for serious law or security applications, at least as the stand-alone ones (a typical polygraph uses techniques from several domains at the same time). As it can be noticed, the best mean value much over 60% was obtained for considerably the simplest method with the statistical analysis of voice parameters (VSA5). Rather disappointing are the results of all the tested techniques based on the EMD method, especially that this method was considered as a very promising one by some researchers for VSA applications [6, 7].

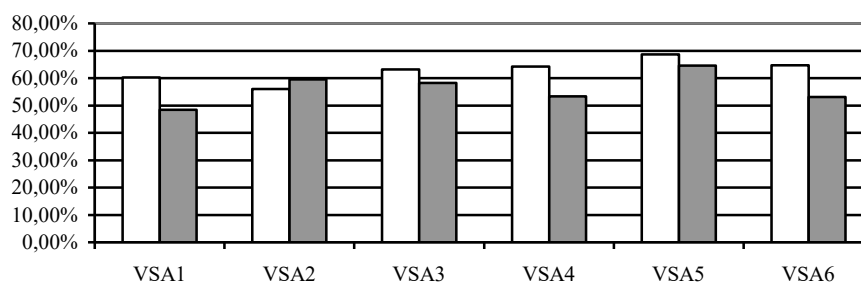


Figure 2. Mean values of deception results for female (white) and male (gray) voices for six tested stress detection techniques

### 3. Conclusions

The usage of VSA techniques in law or security applications is a very controversial problem. On the one hand the manufacturers of such software put pressure on introducing such applications. On the other hand, its effectiveness was not confirmed objectively so these techniques are not considered as reliable [8].

The results obtained from the carried out tests are significantly much lower than that which are declared by the producers of commercial devices for the detection of deception in the speaker's voice. Moreover, the scores depend very highly on the speaker's individual characteristics. It is evident that some subjects react to stress in the way that can be easier detected than others. The most promising results of over 60% were obtained for considerably simple techniques with statistical analysis of voice parameters, which still gives a promise that the scores of VSA techniques can be improved in the future with more sophisticated algorithms. The problem is even more difficult because of very short utterances that have to be used in such a detector (yes/no answers), which in speech emotion analysis using pitch changes can be a significant obstacle. Also, it is worth mentioning that in such experiments no database can be considered as a hundred percent reliable, even after an anonymous deception validation by the subjects, which was carried out in our case.

There exists a fundamental problem in all the techniques of detecting the deception. The stress is a human reaction to the deception, but it can also be caused by other stimuli. Therefore it would be beneficial to avoid intrusive techniques which demand physical connection of the subject to the detectors. It evokes additional stress which can potentially change the results. Other very big advantage of the non-intrusive techniques is that there is no necessity to make the subject aware of being tested.

### References

1. P. Staroniewicz, *Considering basic emotional state information in speaker verification*, Proc. 4<sup>th</sup> International Conference on Biometrics and Forensics (IWBF), Limmasol, Cyprus, 3-4 March 2016, IEEE 2016.
2. P. Staroniewicz, *Automatic recognition of emotional state in Polish speech, Toward autonomous, adaptive, and context-aware multimodal interfaces: theoretical and practical issues*, Lecture Notes in Computer Science, Springer, **6800** (2011) 347 – 353.
3. P. Staroniewicz, *Recognition of emotional state in Polish speech – comparison between human and automatic efficiency*, Lecture Notes in Computer Science, Springer, **5707** (2009) 33 – 40.
4. O. Lippold, *Physiological microtremors*, Scientific American, **224**(3) (1971) 65 – 73.
5. N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung, H. H. Liu, *The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis*, Proc. Roy. Soc. Land. A, (1998) 903 – 1005.

6. J. Z. Zhang, N. Mbitiru, P. C. Tay, R. D. Adams, *Analysis of stress in speech using Adaptive Empirical Mode Decomposition*, 2009, Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers, IEEE 2009.
7. N. Mbitiru, P. Tay, J. Z. Zhang, R. D. Adams, *Analysis of Stress in Speech Using Empirical Mode Decomposition*, Proceedings of The 2008 IAJC-IJME International Conference.
8. C. S. Hopkins, D. S. Benincasa, R. J. Ratley, J. J. Grieco, *Evaluation of voice stress analysis technology*, Proceedings of the 38<sup>th</sup> Hawaii International Conference on System Sciences, IEEE 2005.

## Comparison of the Image Compounding Methods for the Multi-Angle 3-D Ultrasound Imaging

Maciej SABINIOK

*Department of Acoustics and Multimedia, Faculty of Electronics,  
Wroclaw University of Science and Technology,  
Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland,  
maciej.sabinio@pwr.edu.pl*

Krzysztof OPIELIŃSKI

*Department of Acoustics and Multimedia, Faculty of Electronics,  
Wroclaw University of Science and Technology,  
Wybrzeze Wyspianskiego 27, 50-370 Wroclaw, Poland,  
krzysztof.opielinski@pwr.edu.pl*

### Abstract

The main goal of the paper is to compare the image compounding methods to minimise the artefacts arising in the Multi-Angle Conventional Ultrasound Imaging (MACUI) due to the system configuration. The MACUI method used for 3-D object imaging and the introduced imaging artefacts are described. Different ways of the image compounding by intensity averaging are presented in the work. Implemented methods of image compounding were tested for different types of objects mimicking soft tissue. The comparison allowed to determine the most appropriate method of intensity averaging in the compounding method. The method can be used to reduce the presence of image artefacts and enhance the quality of the resulting slices which are used to create 3-D volume of an object structure in MACUI method.

**Keywords:** Multi-Angle Conventional Ultrasound Imaging, 3-D ultrasound imaging, image compounding, soft tissue

### 1. Introduction

Breast cancer is one of the most common women's health issue worldwide. Different radiological methods are continuously being developed and researched to improve an early detection of malignancy in women's breasts. The extensive research is conducted using ultrasound as it is the non-invasive and non-ionizing method of breast imaging. Ultrasound tomography (UT) is a novel technique providing several new ultrasound modalities for breast tissue imaging. These modalities use transmission, reflection and scattering of ultrasound to obtain comprehensive information about tissue structure [1]. The prototype of the hybrid ultrasound tomography scanner developed in Poland by DRAMIŃSKI S.A. company in cooperation with Wroclaw University of Science and Technology team is currently at the stage of clinical evaluation [2]. Despite the advantages of using ultrasound tomography, its sensitivity and specificity are still worse in comparison to Magnetic Resonance Imaging (MRI), Mammography and Ultrasound [2]. Thus, the embedding a new method that uses B-mode ultrasound modality in the ultrasound tomography scanner is desirable. So far, Full Angle Spatial Compound Imaging (FASCI) method has been

designed to take advantages of B-mode modality [3]. FASCI utilizes specially designed transducer ring array for the ultrasound tomography device [5, 6]. The method benefits from use of existing hardware, however, the geometry of the system as well as not optimal transducers characteristics for B-mode imaging cause the limitation of the method. The study on a new method, which could be employed in ultrasound tomography device, was carried out by the authors. Multi-Angle Conventional Ultrasound Imaging (MACUI) method described in the next section has been designed and tested for potential use in UT device and for reducing the limitations of the FASCI.

MACUI method shows many advantages over FASCI, including direct acquisition of more informative transversal and sagittal sections, use of probe and electronics optimised for B-mode scanning and lack of circular geometry limitation such as backscattering and reflection distortion. Nonetheless, some artefacts are present in obtained images. Spatial Compounding (SC) is a known and widely used method of enhancing quality of ultrasound B-mode imaging. The authors aim to use the SC in MACUI method in order to minimise artefacts arising at the centre of the slices obtained from the combination of the images acquired for the  $\theta$  and  $(\pi - \theta)$  angles. Quantitative comparison of different ways of compounding by image intensity averaging was performed in terms of usability of the SC method for the MACUI.

## 2. MACUI method and artefacts

Multi-Angle Conventional Ultrasound Imaging involves the rotational imaging system to collect the set of data about examined structure under different horizontal angles  $\theta$ . For each angle a vertical section of an object is acquired. The probe of the conventional B-mode system can be placed in a different position in relation to the object under examination. Therefore, three basic configurations can be distinguished. Vertical top and vertical bottom configurations are equivalent to each other and depend on imaging system configuration. In the first one, the object slices are acquired by the probe moving over the top of the object, while in the second configuration, under the bottom. The third configuration is a vertical lateral one. In this configuration, the vertical slices are obtained from the probe moving around the lateral surface of the object. Vertical top and lateral configurations are schematically shown in Fig. 1. Probe is being rotated around Z-axis. Vertical projections of an object are acquired for the entire turnover with a predetermined angle step. The vertical bottom configuration with the centre of the ultrasonic probe moved aside the rotation axis can be potentially used as an additional, conventional ultrasound modality in UT device. In this configuration, the entire probe is placed outside the rotational axis. This provides large enough imaging area to be applicable for breast screening and allows using a breast holder needed for stretching the breast during examination procedure [7]. Considering that vertical top and bottom configuration are equivalent, the vertical top was used in the work in order to qualitatively evaluate the spatial compounding methods for MACUI.

The obtained images are then assembled into coronal projection of a structure. The rows taken at the specific heights from all of the images carry the data about specific coronal section.

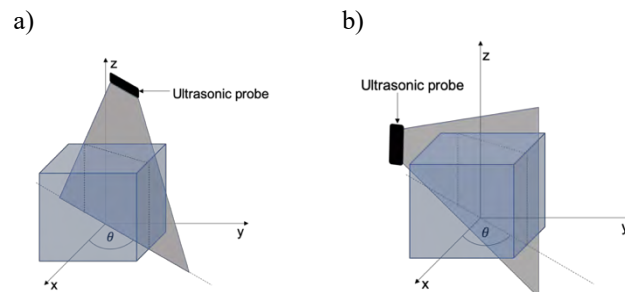


Figure 1. The idea of acquiring images of object's sections in the vertical top (a) and vertical lateral (b) configuration of the MACUI

The radial and angular position of the pixels in each row corresponding to the particular height are converted to the Cartesian coordinates to create the image of the structure's coronal section.

Tissue mimicking phantoms were submerged in the water tank and placed on the absorption material to minimise reflection from the bottom of the tank. The probe was mounted to the rotational mechanism that allowed for setting the required angles. SmartUs Telemed B-mode scanner was used as imaging device. The linear ultrasound probe with transducer array of 4 cm length with the centre frequency set to 7.5 MHz was used. Measurement set-up used to acquire test images for further evaluation of SC is presented in Fig. 2a. Test objects were scanned in two positions only. The second position was at angle  $(\pi - \theta)$ , while  $\theta$  stand for an angle of the first position. This is shown in Fig. 2a by presenting the second position of the ultrasonic probe in shadowed colours. It can be noticed that the imaging area of the probe at the initial position is overlapped by the imaging area for the probe at angle  $(\pi - \theta)$ . This region is the subject of interest in the work (Fig. 2b). It is important to combine both images in this area with, ideally, neither artefacts nor distortion. This influences the quality of a resulting slice of the object, which would be used later for reconstruction of the entire object's 3-D volume.

In the preliminary study on the MACUI method, the opposite images were compounded in the most straightforward and simple way. Both images were truncated at the rotation axis and concatenated together side by side. Despite the simplicity and implementation easiness of this method, it exhibits the image artefact in its centre. When the sectoral scanning is used for the imaging process in the described configuration, the ultrasound waves reach the object at the rotation axis at different angles and through different acoustic paths. That causes the object's shape distortion due to different average sound speed throughout an acoustic path and different angles of incident. As a result, truncated images are not well matched at the centre of the resulting slice. Figure 2b shows the example of the tissue mimicking phantom's slice obtained in MACUI method by image truncation. The area where the artefacts are the most visible has been cut and presented separately. It has to be mentioned that the phantom is not a very complicated structure. When more complicated tissue structure with smaller inclusions in this area would be considered, artefacts can significantly reduce their continuity or even prevent a viewer from differentiating them from surrounding tissue in reconstructed 3-D image.

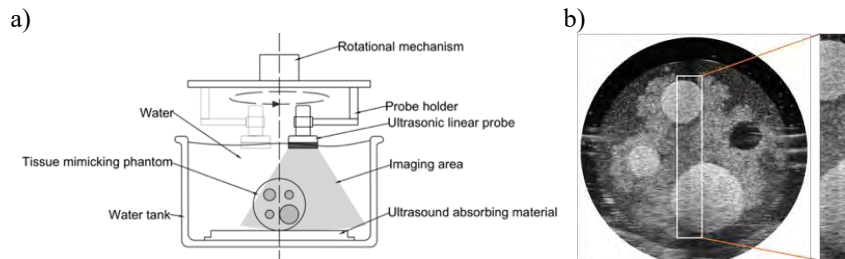


Figure 2. Measurement set-up for SC evaluation in MACUI method (a).  
Example of arising artefacts in MACUI utilizing image truncation

### 3. Materials and compounding methods

Multi-angle Compound Imaging (MACI) is a commonly used method to enhance the quality of ultrasound imaging. Usually, the method involves a single probe employing a beam steering system to alter the angle of the ultrasound beams during measurement. The object of interest is scanned under several angles (usually 3 to 9) and then images are combined to obtain a single image [3]. Spatial Compounding method minimises a number of artefacts presented in the conventional B-mode imaging. Speckle noise, reflection artefacts from specular reflectors, acoustic shadows and angle dependency reduction, as well as an increase in contrast resolution, boundaries visualisation, image understanding and general image quality over conventional B-mode ultrasound have been reported in a number of works [3-6]. Intensity averaging of images is a well-established way to obtain spatially compounded images in ultrasound B-mode scanning techniques. Because of this, the authors attempted to perform the qualitative evaluation of different types of intensity averaging as the first step in study to improve the quality of the images and reduce the artefacts arising in proposed MACUI method. What significantly differentiates the spatial compounding approach in MACUI method from MACI is a probe placement. In contrast to MACI, in the MACUI method two images are taken from opposite sides of object vertical symmetry axis (Fig. 1a). Therefore, the images are less correlated with each other and there are larger differences in the acoustic path of the ultrasound waves approaching the structure in the overlapped area. This results in higher amount of shape distortion differences between images.

The selected ways of compounding by image averaging are Arithmetic, RMS, Geometric and Weighted Average. In the Arithmetic method, the intensity of each pixel in the overlapping area is the simple arithmetic average of the intensity of corresponding pixels in the images that have been compounded. It can be expressed by the equation:

$$I_{Avg\_x,y} = \frac{1}{N} \sum_{n=1}^N I_{x_n,y_n} \quad \text{for } N=2 \quad = \quad \frac{I_{x_1,y_1} + I_{x_2,y_2}}{2} \quad (1)$$

In general, image intensity averaging can be performed for  $N$  images, however MACUI configuration presented in this paper involves two images in a compounding task. Indices  $x, y$  indicate the position of the pixel on the compound image, while indices  $x_1, y_1, x_2, y_2$  stand for the placement of the pixels in the corresponding images. RMS method uses



the root mean square to calculate the pixels intensity in the compounded image. This is described by the equation for the general case of  $N$  images involved in compounding process and specified for two images:

$$I_{RMS\_x,y} = \sqrt{\frac{1}{N} \sum_{n=1}^N I_{x_n,y_n}^2} \quad \text{for } N=2 \quad \sqrt{\frac{I_{x_1,y_1}^2 + I_{x_2,y_2}^2}{2}} \quad (2)$$

Similarly, the equation defines the geometric mean used in geometric averaging of the images' intensity pixels:

$$I_{Geo\_x,y} = \sqrt[N]{\prod_{n=1}^N I_{x_n,y_n}} \quad \text{for } N=2 \quad \sqrt{I_{x_1,y_1} \cdot I_{x_2,y_2}} \quad (3)$$

The last method utilizes specially adapted weighted average for spatial compounding in MACUI method:

$$I_{WAvg\_x,y} = \sum_{n=1}^N W_{n,x,y} \cdot I_{x_n,y_n} \quad \text{for } N=2 \quad W_{1,x,y} \cdot I_{x_1,y_1} + W_{2,x,y} \cdot I_{x_2,y_2} \quad (4)$$

In general case of  $N$  images  $W_{n,x,y}$  are the weighting coefficients for the specific pixel at  $(x, y)$  point in compounded image. Index  $n$  corresponds to the particular image for which the coefficient is used. This method is equivalent to the smooth images blending in the compounding area. The vertical line in the centre of the compounding area is the simple average of the pixels engaged in compounding. For each row in the compounding area, the weighted coefficients for the left-side image decrease when index  $x$  increases and increase when index  $x$  decreases. It happens with the step  $S$ . For the right-side image increasing and decreasing of the weighted coefficients take place in opposite order. Dependence between  $W_{1,x,y}$  and  $W_{2,x,y}$  and step  $S$  are defined accordingly:

$$W_{1,x,y} = 1 - W_{2,x,y} \quad (5)$$

$$S = \frac{1}{M_y + 1} \quad (6)$$

where  $M_y$  is a number of pixels in a row at  $y$  height inside the compounding region. The linear change in coefficient had been chosen arbitrary as the most straightforward method of image blending.

Two self-made agar gel phantoms were used as a tissue mimicking objects. Dimensions and view of the phantoms are presented in Fig. 3. The main part of both of the phantoms were made from 10% Petrygo radiator fluid and water solution, with addition of 3 g of agar powder per 100 ml of the mixture (No. 5 on Fig. 3a). Inclusions in each phantom were made to obtain different acoustic properties of a gel. In the first one, the cylindrical parts 1, 2 and 3 were made by changing the Petrygo radiator fluid concentration, which altered the speed of ultrasound in the gel. The proportion of the Petrygo radiator fluid in parts 1, 2 and 3 are 15%, 20% and 5% accordingly. In the second one, those parts were prepared by addition of a different portion of graphite

powder with the particle sizes range of 0 – 60  $\mu\text{m}$ . This allowed for changing the ultrasound wave attenuation coefficient of obtained gel. The amount of graphite powder added to parts 1, 2, 3 are 4.55 g/100 ml, 2.5 g/100 ml and 6.35 g/100 ml accordingly. In both phantoms, part 4 was left as an empty cylinder. The cylinders were filled with the water during measurements. Fig. 3b shows the exemplary photo of the real phantom with graphite powder inclusion.

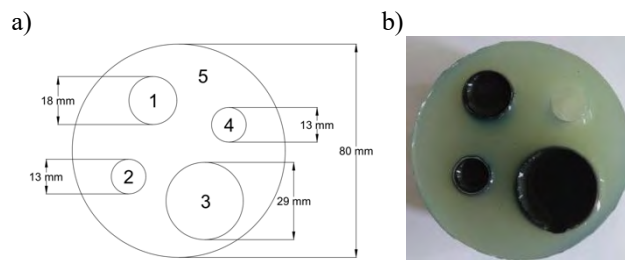


Figure 3. Dimensions of self-made agar gel phantoms (a) and the photo of the real made agar phantom with the graphite powder inclusion (b)

#### 4. Results and analysis

In Fig. 4, the phantom with addition of the graphite powder is presented. Numbers 1, 2, 3, 4 correspond to Arithmetic, RMS, Geometric and Weighted Average methods accordingly. No additional processing like filtering or denoising image had been used during imaging process of individual images used in spatial compounding. Unexpected agar gel nonuniformity in the inclusions' surrounding area (No. 5 in Fig. 3a) has come out during measurement. This unpredicted effect is beneficial from the images evaluation point of view, as such irregularity in the structure can reveal the artefact which could not be visible otherwise.

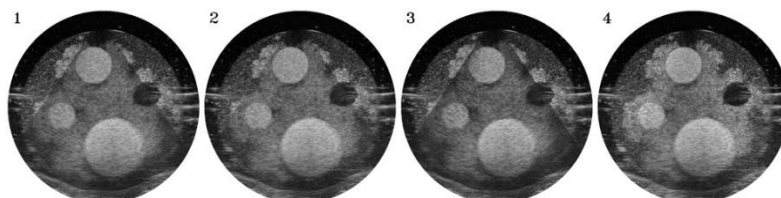


Figure 4. Result of the compound imaging of the phantom with addition of the graphite powder in the MACUI method without image normalization procedure

It can be seen that in the results of the first three methods the boundaries of compounding region are clearly visible. Due to this fact, the additional normalization was made to scale the images' intensity in both parts, inside and outside of compounding region. The normalization factor was 255, which is the maximum value of the 8-bit grey scale images used in compounding procedure. The results after normalization are shown in Fig. 5. The image obtained with the Weighted Average method is not affected by boundaries' intensity inhomogeneity as they are effectively blended with the appropriate

weighted factor. At the boundaries, data from the image for which the phantom was scanned closer the centre of B-mode device imaging area are of the greatest importance. The entire image was normalized for this reason. As can be seen, the normalization procedure significantly improves the resulting image in the Arithmetic, RMS and Geometric Average methods. However, in the Average and Geometric methods boundaries of the SC region are still highly visible, but the improvement over methods 1 and 3 is significant. Additionally, the normalization procedure slightly increases the overall brightness in SC region for tested phantoms.

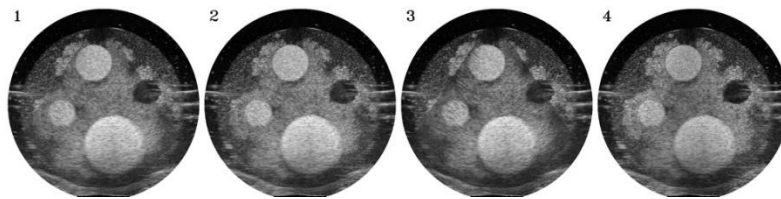


Figure 5. Result of the compound imaging of the phantom with addition of the graphite powder in the MACUI method with image normalization procedure

The best results, regardless of the normalization procedure usage in SC for the MACUI, were obtained with the Weighted Average method due to the lack of intensity differences at the SC region's boundaries.

The results of SC for phantom with different proportion of the Petrygo radiator fluid are shown in Fig. 6.

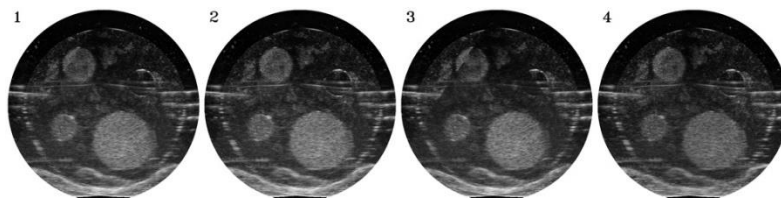


Figure. 6. Result of the compound imaging of the phantom with different proportion of Petrygo radiator fluid in the MACUI method image normalization procedure

Imaging results of the phantom presented in Fig. 6 obtained by the Arithmetic (1) and Geometric (3) average methods are not sufficient in terms of the consistency of the SC region's boundaries like in the results presented in Fig. 5. The result obtained by the RMS method significantly reduced the issue of boundaries' inhomogeneity, but not completely. The best result in this term was obtained with the use of the Weighted Average method. Moreover, the increase in inclusions' boundaries blurring can be noticed comparing to the phantom presented in Fig. 5. This is due to different average speed of ultrasound through different path in the phantom. This results in larger amount of shape distortion of the individual inclusion scanned under different perspectives. Therefore, the SC by the intensity averaging is not sufficient for structures in which the speed of ultrasound varies. Additional processing for shape differences reduction would be required in this case.

## 5. Conclusions

Examples discussed in the paper show that the spatial compound imaging without additional processing can be successfully used in the Multi-Angle Conventional Ultrasound Imaging approach unless the speed of sound does not vary in the structure. The best results were obtained with the proposed Weighted Average method. The compounded image is not affected by the intensity inhomogeneity at the boundaries of SC region. Relatively small amount of inclusions' shape distortion in tissue mimicking object can be effectively minimised with using the Weighted Average method, resulting in satisfactory imaging quality of the individual structure's slice obtained in MACUI method.

In case of sound speed differences across the structure, the significant blurring effect on inclusions' boundaries is exhibited. This reduces the possible usage of the MACUI method without additional processing aiming to reduce differences in inclusions' shape between images. In case of real tissue structure, both attenuation coefficient and speed of sound change. This minimise the usability of the simple SC methods to be successfully applied in MACUI method. Nevertheless, the most suitable for the MACUI method and possible further improvement is proposed Weighted Average method as it is not affected by artefacts at the SC region's boundaries. Future work will cover the study on the adequate method of shape differences reduction between images.

## References

1. N. Ozmen, R. Dapp, M. Zapf, H. Gemmeke, N. V. Ruiter, K. W. A. van Dongen, *Comparing Different Ultrasound Imaging Methods for Breast Cancer Detection*, IEEE Trans Ultrason Ferroelectr Freq Control, **62**(4) (2015) 637 – 646.
2. T. Milewski, M. Michalak, A. Wiktorowicz, K. Opieliński, P. Pruchnicki, M. Bułkowski, J. Gielecki, M. Jóźwik, *Hybrid Ultrasound Tomography Scanner – a Novel Instrument Designed to Examine Breast as a Breast Cancer Screening Method*, Biomed J Sci & Tech Res, **14**(4) (2019) 1 – 5.
3. R. Entrekin, P. Jackson, J. R. Jago, B. A. Porter, *Real Time Spatial Compound Imaging in Breast Ultrasound: Technology and Early Clinical Experience*, Medica Mundi, **43**(3) (1999) 35 – 43.
4. J. E. Wilhjelm, M. S. Jensen, S. K. Jaspersen, B. Sahl, E. Falk, *Visual and Quantitative Evaluation of Selected Image Combination Schemes in Ultrasound Spatial Compound Scanning*, IEEE Trans. on Med. Imaging, **23**(2) (2004) 182 – 190.
5. K. J. Opielinski, *Full angle ultrasound spatial compound imaging*, Proc. of 7<sup>th</sup> Forum Acusticum 2014 joined with 61<sup>st</sup> Open Seminar on Acoustics, Polish Acoustical Society - Acoustical Society of Japan, European Acoustics Association, Krakow 2014.
6. J. Camacho, L. Medina, J. F. Cruza, J. M. Moreno, C. Fritsch, *Multimodal Ultrasonic Imaging for Breast Cancer Detection*, Arch Acoust, **37**(3) (2012) 253 – 260.
7. S. Forte, S. Dellas, B. Stieltjes, B. Bongartz, *Multimodal ultrasound tomography for breast imaging: a prospective study of clinical feasibility*, Eur Radiol Exp, **1**(1):27 (2017) p. 6.

## **Performance of Coherent Modulation Scheme Used in Acoustic Underwater Communication System**

Jan H. SCHMIDT

*Gdansk University of Technology, Faculty of Electronics,  
Telecomm. and Informatics, ul. Narutowicza 11/12, 80-233 Gdańsk,  
jan.schmidt@pg.edu.pl*

Aleksander M. SCHMIDT

*Gdansk University of Technology, Faculty of Electronics,  
Telecomm. and Informatics, ul. Narutowicza 11/12, 80-233 Gdańsk,  
aleksander.schmidt@pg.edu.pl*

Iwona KOCHAŃSKA

*Gdansk University of Technology, Faculty of Electronics,  
Telecomm. and Informatics, ul. Narutowicza 11/12, 80-233 Gdańsk  
iwona.kochanska@pg.edu.pl*

### **Abstract**

The development of an acoustic underwater communication system for shallow waters is still a big scientific and construction challenge. Currently, non-coherent modulations in combination with strong channel coding are used to achieve reliable communication with low rate in such a channel. To obtain transmission with a higher transmission rate, it is required to use coherent modulation. This paper presents the assumptions of such a transmission system and the results of data transmission carried out by this system in the channel with the Rician fading, which reflects the short range shallow water channel. A digital version of the carrier phase modulation known as Phase-Shift Keying was selected for simulation.

**Keywords:** underwater acoustic communications, shallow underwater channel, coherent modulation

### **1. Introduction**

The task of the underwater acoustic communication system is error-free data transfer of large amounts of data. In the shallow waters channel there are unfriendly conditions for its implementation. The obtained transmission rates are usually low, and the variability of propagation conditions in the channel contributes to transmission errors. In such a channel, the multipath propagation has an influence on transmitted signal due to its reflections from the boundary surfaces of the channel and objects present in the water. On the receiving side, the signal from a direct path and the paths obtained during reflections is received. The transmitted signal suffers the refraction, which is caused by a significant changes of sound velocity as a function of depth. Multipath propagation and refraction produce time dispersion in the transmitted signal. The movement of the communication system's transmitter and receiver causes the Doppler effect, resulting in the time-domain scaling of an original broadband communication signal. This phenomenon also has a significant impact on the communication system's performance. The development of an acoustic

underwater communication system for shallow waters is still a big scientific and construction challenge.

In order to implement reliable acoustic underwater communication systems, incoherent modulations are usually used in combination with strong channel coding. They enable data transmission at a rate of hundreds of bits per second [1-3]. Obtaining higher transmission rates for this modulation method is possible as a result of a significant widening of the system operating band, which is obtained by using multiple hydroacoustic transducers or antennas [4]. Avoiding these inconvenient hardware system complications requires the use of coherent modulation. This is a significant complication of the receiver requiring considerable computing power to implement the channel's equalizer algorithm [5-7]. However, with the computing power offered by modern signal processors, this does not constitute any significant limitation in the implementation of the system.

This article presents the performance of an underwater communication system using BPSK coherent modulation, in a configuration with one transmit and one receiving antenna (SISO). For the development of the communication system concept, simulation tests were carried out in channels with Rician fading, which reflect the short range shallow water channels.

## 2. Underwater communication system

The considered acoustic underwater communication system assumes the use of carrier phase modulation. This modulation changes the phase of the carrier wave signal at a fixed frequency in accordance with a predetermined scheme assigned to the given order. The digital variant of this modulation is called Phase-Shift Keying (PSK) [5]. The phase value from the range  $< 0, 2\pi >$  is used to transmit data and carrier phases of transmitted signals are:

$$\theta_m = \frac{2\pi m}{M}, \quad m = 0, 1, \dots, M - 1. \quad (1)$$

For binary phase modulation BPSK ( $M = 2$ ) there are two carrier phases  $\theta_0 = 0$  and  $\theta_1 = \pi$ . In turn, for QPSK ( $M = 4$ ) there are four carrier phases  $\theta_0 = 0$ ,  $\theta_1 = \pi/4$ ,  $\theta_2 = \pi/2$  and  $\theta_3 = 3\pi/4$ . Symbol  $k$  is the number of bits per transmitted symbol where  $M = 2^k$ . The transmitted signal waveforms in the symbol interval  $< 0, T >$  can be expressed as:

$$s_m(t) = \sqrt{\frac{2E_s}{T}} \cos\left(2\pi f_c t + \frac{2\pi m}{M}\right), \quad m = 0, 1, \dots, M - 1, \quad (2)$$

where  $E_s$  is the energy transmitted per symbol. A digital phase-modulated signals are represented geometrically as two-dimensional vectors:

$$\mathbf{s}_m = \left( \sqrt{E_s} \cos\left(\frac{2\pi m}{M}\right) \quad \sqrt{E_s} \sin\left(\frac{2\pi m}{M}\right) \right). \quad (3)$$

Space constellation diagrams for  $M = 2$  (BPSK) and  $M = 4$  (QPSK) are presented in Figure 1.

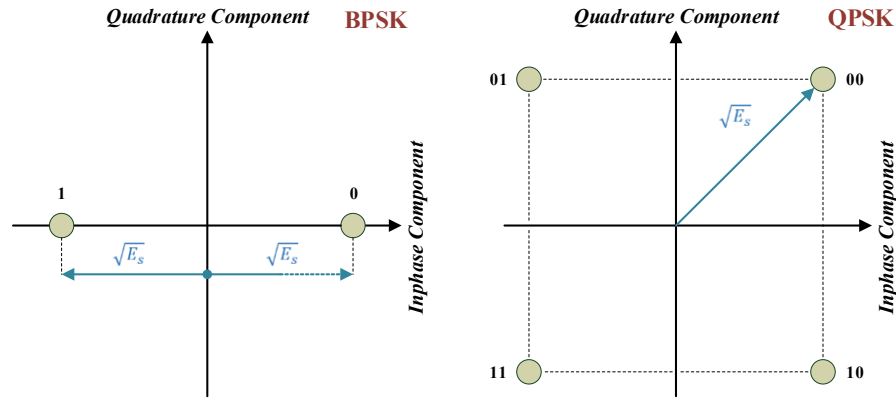


Figure 1. Space constellation diagrams for BPSK and QPSK

Figure 2 shows the determined error bit rate (BER) for the case of transmission data through channels with Additive White Gaussian Noise (AWGN) for BPSK and QPSK modulation. The BER is simulated in function of  $E_b/N_0$ , which is the ratio of the symbol energy to noise spectral density in dB.

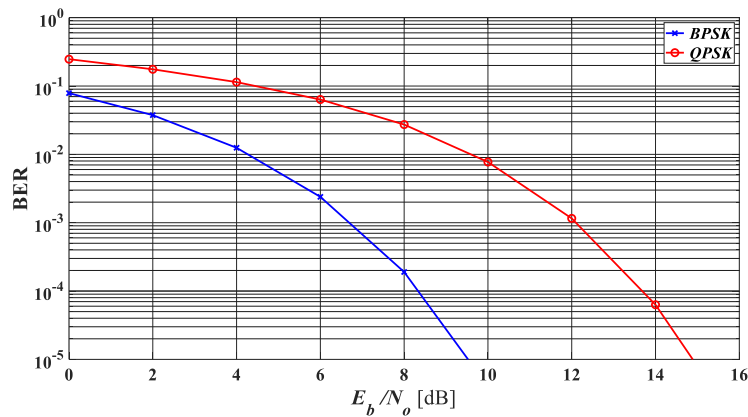


Figure 2. Bit error rate of BPSK and QPSK modulation in AWGN channel

To achieve  $\text{BER} = 10^{-3}$ , BPSK modulation required  $E_b/N_0 \approx 7$  dB and QPSK modulation  $E_b/N_0 \approx 12$  dB. For BPSK modulation, BER takes lower values than for QPSK modulation and higher order modulation. The obtained simulation results encourage the use of BPSK modulated signals in further studies.

Figure 3 shows the space constellation diagram for a series of 1000 symbols received, which were subjected to BPSK modulation and were transmitted in the AWGN channel with  $E_b/N_0 = 5$  dB.

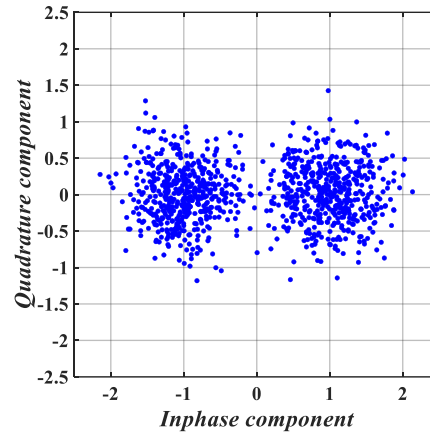


Figure 3. Space constellation diagram for BPSK modulation used in AWGN channel ( $E_b/N_0 = 5$  dB)

### 3. Simulations - transmission in channel with fading

As a result of multipath propagation phenomenon the hydroacoustic signal sent by the transmitter reaches the receiver in the form of its many components shifted relative to each other in time. This causes intersymbol interference (ISI) and fading of the received signal. Such an acoustic underwater channel can be presented as a linear filter whose low-pass impulse response  $h(\tau, t)$  is the sum of the impulse responses corresponding to all  $n$  propagation paths, where each of these paths is described by the corresponding amplitude  $a_n(t)$ , phase  $\varphi_n(t)$  and delay  $\tau_n(t)$ . Delay  $\tau$  applies only to paths obtained during reflections and the value of these paths depends on their geometry. The number  $N(t)$  is the maximum number of propagation paths during transmission, depending on  $t$ . It is assumed that the individual components of the incoming signal are mutually uncorrelated.

It is assumed that the transmitted signal  $s(t)$  has a form:

$$s(t) = \text{Re}[s_{LP}(t) e^{j2\pi f_c t}], \quad (4)$$

where  $s_{LP}(t)$  is the low-pass complex envelope signal  $s(t)$  and  $f_c$  is the nominal carrier frequency. Omitting the presence of noise, the received signal can be represented as:

$$x(t) = \sum_{n=1}^{N(t)} a_n(t) s[t - \tau_n(t)]. \quad (5)$$

By substituting equation (4) for equation (5) the expression (6) is obtained:

$$x(t) = \text{Re} \left[ \sum_{n=1}^{N(t)} a_n(t) e^{-j2\pi f_c \tau_n(t)} s_{LP}(t - \tau_n(t)) \right] e^{j2\pi f_c t} \quad (6)$$



and the equivalent low-pass reception signal has the form:

$$x_{LP}(t) = \sum_{n=1}^{N(t)} a_n(t) e^{-j2\pi f_c \tau_n(t)} s_{LP}(t - \tau_n(t)). \quad (7)$$

The complex low-pass impulse channel impulse response of a multipath underwater channel can be presented in the form (8) [4]:

$$h(\tau, t) = \sum_{n=1}^{N(t)} a_n(t) e^{-j\varphi_n(t)} \delta(\tau - \tau_n(t)), \quad (8)$$

where  $\varphi_n(t) = 2\pi f_c \tau_n(t)$ . The above form of low-pass complex impulse response  $h(\tau, t)$  relates to a non-stationary underwater channel, in which the received signal parameters (amplitude, phase and delay) are functions of time  $t$ .

However, for the shallow water environment there is a lack of distinguished, even generalized, classes of acoustic underwater channels. This fact results mainly from a large variety of existing channels, where almost every water reservoir is additionally characterized by strong changes in channel parameters over time.

Modeling using a deterministic underwater channel model a limited degree of suitability in practical solutions due to the absence of expected universality. It is usually used to analyze channels with known parameters such as, for example, a measuring pool. Therefore, a stochastic model of the underwater channel is commonly used to carry out research of data transmission in underwater channel with fading. Although this model does not take into account all factors influencing the properties of the communication channel, it largely describes the statistical conditions of the considered environment and has a large practical aspect. In order to model the communication channel, it is considered as a stochastic process for which the impulse response values of the channel are random variables with given probability distributions [5].

In this article, the considered configuration of the communication channel assumes the occurrence of a direct propagation path (LOS - *line-of-sight*). Therefore, the Rician channel model has been used, which is used when both, the direct and the reflected signals are received. The dominant direct component is usually associated with the direct propagation path with the highest power. The channel model with Rician fading is suitable for the simulation of a small range transmission channel [5].

Assuming that the envelope of the impulse response  $R = |h(\tau, t)|$  has a real component  $R_r$  and imaginary component  $R_i$ , which are independent random variables with Gaussian distributions  $N(\mu_r, \sigma^2)$  and  $N(\mu_i, \sigma^2)$  with respective mean values  $\mu_r$  and  $\mu_i$ , and variance  $\sigma^2$ , where  $R = \sqrt{R_r^2 + R_i^2}$ , the envelope of the received signal has the Rice distribution with the probability density function described as (9):

$$p_{Rice}(r) = \frac{r}{\sigma^2} e^{\frac{-(r^2 + s^2)}{2\sigma^2}} I_0\left(\frac{rs}{\sigma^2}\right) \quad \text{for } r \geq 0, \quad (9)$$

hence, the amplitude of the direct path signal has the form (10):

$$s = \sqrt{\mu_r^2 + \mu_i^2}. \quad (10)$$

$I_0$  is the modified Bessel function of the first kind and zero order. An important parameter of the Rician model is  $K$  factor, which determines the ratio of the direct path component (dominant component) to the power in the all the other components. It can be expressed as (11):

$$K = \frac{s^2}{2\sigma^2}. \quad (11)$$

The phase  $\varphi_n(t)$  of received signal is uniformly distributed in the interval  $[0, 2\pi]$ .

In order to estimate the performance of BPSK modulation, the first simulation tests were carried out in a channel with flat fading (non-selective channel), for which the influence of multipath effect is neglected. In these tests, BER was determined depending on the value of Rice  $K$  factor. In flat fading channel, all frequency components of transmitted signal are subject to the same attenuation and phase shift. For slow fading case, the attenuation and phase can be considered as a constant during at least one symbol interval. The received signal reaches the receiver through single fading path and this signal can be expressed as (12):

$$x(t) = a_n(t)e^{-j\varphi_n(t)}s_{LP}(t) + n(t), \quad (12)$$

where  $n(t)$  is the random variable of white Gaussian process. The obtained tests results for different Rice  $K$  factor values are presented in Figure 4. With the increase of the  $K$  value, smaller BER values are obtained for specific  $E_b/N_0$  values. To obtain  $\text{BER}=10^{-3}$ ,  $E_b/N_0 \approx 12.5$  dB ( $K = 3$  dB),  $E_b/N_0 \approx 9.5$  dB ( $K = 6$  dB) and  $E_b/N_0 \approx 7$  dB ( $K = 9$  dB), is required.

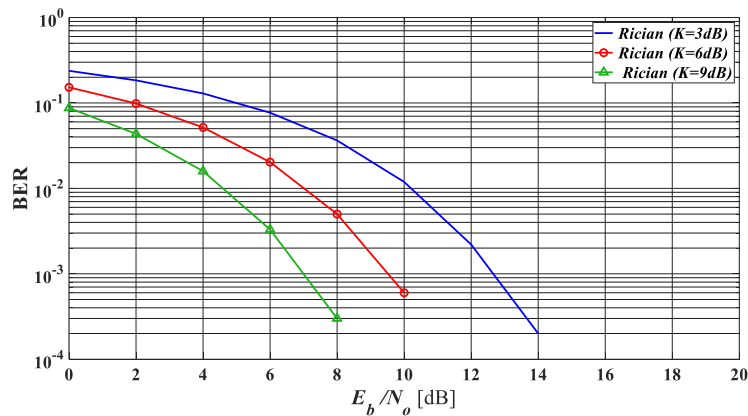


Figure 4. Bit error rate of BPSK modulation in Rician channel with flat fading

The second tested channel is a channel with frequency-selective fading, characteristic of the influence of multipath effects [8]. In frequency-selective fading case, the received

signal consists of multiple copies of transmitted signal, attenuated and delayed in time. The path parameters for a specific transmitter-receiver configuration have been pre-determined, i.e. delays of multipath components and their average path gain, where average path gain is normalized to 0 dB. The parameters are supposed to reflect the transmission between the transmitter and the receiver, which are 300 m away, in a 20 m deep water reservoir, with the same transmitter and receiver depth of 10 m.

Simulation tests included the execution of numerous simulations, and in the article are present results for two sets of parameters. Figure 5 presents the determined BER values for different Rice  $K$  values and the following path parameters: delay  $\tau_1 = 4.75$  ms and average path gain  $g_1 = -3$  dB. To get  $\text{BER} = 10^{-3}$ ,  $E_b/N_o \approx 17$  dB (for  $K = 3$  dB),  $E_b/N_o \approx 10$  dB ( $K = 6$  dB) and  $E_b/N_o \approx 6$  dB ( $K = 9$  dB), is required.

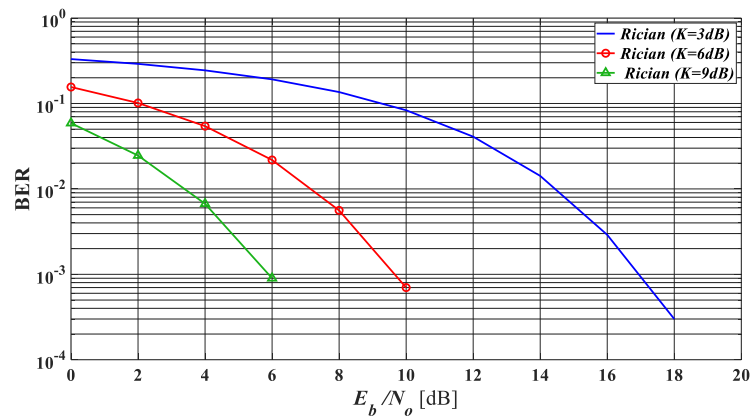


Figure 5. Bit error rate of BPSK modulation in Rician channel with frequency-selective fading ( *direct path + one delayed path* )

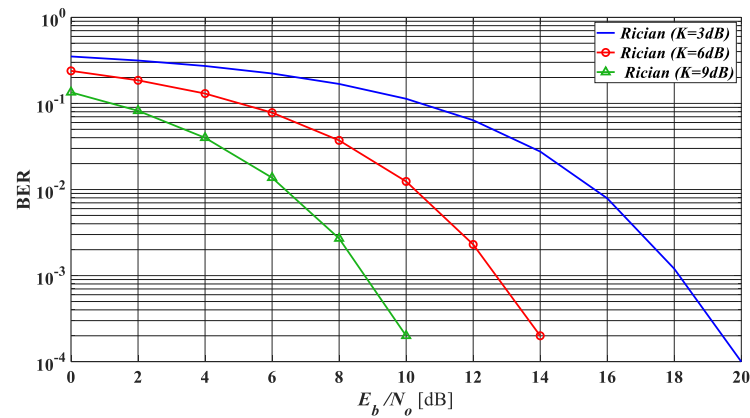


Figure 6. Bit error rate of BPSK modulation in Rician channel with frequency-selective fading ( *direct path + two delayed paths* )

Figure 6 contains the determined BER values for different  $K$  values and the following parameters: (first path) delay  $\tau_1 = 4.75$  ms and average path gain  $g_1 = -3$  dB, (second path) delay  $\tau_2 = 7.45$  ms and average path gain  $g_2 = -4.6$  dB. It follows that to ensure the transmission with the value of  $\text{BER} = 10^{-3}$ ,  $E_b/N_0 \approx 18$  dB (for  $K = 3$  dB),  $E_b/N_0 \approx 12.5$  dB ( $K = 6$  dB) and  $E_b/N_0 \approx 8.5$  dB ( $K = 9$  dB), is required.

#### 4. Conclusions

Simulation tests of the acoustic underwater data transmission system using coherent modulation have been carried out. This modulation allows to achieve high transmission rates in comparison with non-coherent modulations. Obtaining a reliable transmission with  $\text{BER} = 10^{-3}$  for the AWGN channel requires  $E_b/N_0 \approx 7$  dB, while for the channel with flat fading  $E_b/N_0 \approx 12.5$  (for  $K = 3$  dB) is required. In turn, for two tested channels with frequency-selective fading of specified parameters to obtain the same BER can be achieved using  $E_b/N_0 \approx 17$  (for  $K = 3$  dB) and  $E_b/N_0 \approx 18$  (for  $K = 3$  dB). For the channels with fading, when the increase of Rice  $K$  factor, a reduction in the required  $E_b/N_0$  is observed.

However, these are still significant values of  $E_b/N_0$ , when the efficient communication system is considered. The conducted tests, in particular those concerning the channel with frequency-selective fading with assumed simple channel parameters, showed the necessity of applying countermeasures to transmitted of signal fading, eg channel coding and channel correction techniques.

#### References

1. D. Garrood, *Applications of the MFSK Acoustic Communications System*, OCEANS 81, DOI: 10.1109/OCEANS.1981.1151697.
2. J. Schmidt, K. Zachariasz, R. Salamon, *Underwater communication system for shallow water using modified MFSK modulation*, *Hydroacoustics*, **8** (2005) 179–184.
3. K. Zachariasz, J. Schmidt, R. Salamon, *Code signals transmission using MFSK modulation in shallow waters*, *Hydroacoustics*, **4** (2001) 261–264.
4. J. H. Schmidt, A. M. Schmidt, I. Kochańska, *Multiple-Input Multiple-Output Technique for Underwater Acoustic Communication System*, *Proceedings of 2018 Joint Conference - Acoustics*, DOI: 10.1109/ACOUSTICS.2018.8502439.
5. J. G. Proakis, M. Salehi, G. Bauch, *Contemporary Communication Systems using Matlab (Third Ed.)*, Prentice Hall, Cengage Learning 2013.
6. M. Stojanovic, J. A. Catipovic, J. Proakis, *Phase-Coherent Digital Communications for Underwater Acoustic Channels*, *IEEE Journal of Oceanic Engineering*, **19**(1) (1994).
7. H. S. Dol, P. Casari, T. van der Zwan, R. Otnes, *Software-Defined Underwater Acoustic Modems: Historical Review and the NILUS Approach*, *IEEE Journal of Oceanic Engineering*, **42** (2017) 722–737.
8. I. Kochańska, J. H. Schmidt, *Estimation of Coherence Bandwidth for Underwater Acoustic Communication Channel*, *Proceedings of 2018 Joint Conference- Acoustics*, DOI: 10.1109/ACOUSTICS.2018.8502331.

## **Development of Predictive Model for Vibro-Acoustic Protections in Industrial Hall**

Janusz PIECHOWICZ

*AGH-University of Science and Technology,  
Krakow Al. Mickiewicza 30 30-059 Krakow,  
piechowi@agh.edu.pl*

Jakub PIECHOWICZ

*AGH-University of Science and Technology,  
Krakow Al. Mickiewicza 30 30-059 Krakow,  
qba.piechowicz@gmail.com*

### **Abstract**

The paper presents stages of developing a prognostic model for implementing noise protection in an industrial hall. The main source of vibroacoustic problems is the glass hardening furnace, which was installed in a new production hall. It generates vibroacoustic problems at the furnace operators' work stations and at work stations at other devices placed in this hall and in the adjacent hall. These problems are mainly result from mistakes made during the preparation of the furnace placement, as well as the assembly of the machine and accessories. The conducted vibroacoustic tests were aimed at diagnosing problems and developing appropriate ways to minimize the risk of noise and vibration hazards. The basic research tool was to create the correct hall acoustic model, determine the sound power of partial sources and calibrate the model. Modelling the sound protection efficiency of each noise reducing element has allowed the selection of appropriate solutions in terms of acoustic efficiency and approximate application costs. Measurements were carried out according to standard procedures. The article presents only some of the results of vibroacoustic tests directly related to the developed acoustic model. Analysis of vibration paths and noise propagation, indication of assembly defects is a recommendation to formulate vibroacoustic requirements for newly installed devices.

**Keywords:** noise, workstation, predicting software, work environment

### **1. Introduction**

The principles of combustion noise in industrial and power systems using gas and liquid fuels apply generally to all types of combustion systems [1]. The new glass hardening lines are distinguished by their efficiency, energy efficiency, high performance and the quality of tempered glass. Most functions are automated, fully controlled and managed by the PLC system from the computer. The glass panes, after being loaded into the furnace, are tempered. The glass formats are heated to the required temperature with the assumed speed of movement and the temperature curve, then the glass is tempered and cooled.

The authors present a case study of many vibroacoustic problems, which appeared after installing a glass hardening furnace in the industrial hall. Problems occur not only at workplaces, both in cases of excessive vibration and noise.

To analyze the possibilities of efficiency different types of noise protection solutions, Odeon, specialized software of room acoustics, was used. The purpose of computer simulations was to achieve the required noise levels by using effective noise abatement with the lowest possible investment cost.

For the design of low-noise indoor environments it is crucial that both the noise sources and the surrounding room surfaces are taken into account [2, 3]. Using a room acoustic computer model the first step is to model the room surfaces in terms of absorption and scattering properties. Next the real sources are modelled, and while some sources may be sufficiently modelled by a simple point source, this is not sufficient for large noise sources [5, 7].

In the design process the computer model offers a handful of tools like grid maps of noise distribution, reverberation time, spatial decay rate, identification of the most important sources, and possible effects of various noise control measures [4, 5, 8]. The distribution of early reflected sound over the particularly areas in halls could be investigated, especially with respect to the shape of halls [6, 9].

The usage of predictive acoustic software in industrial hall environment is frequent. In this paper, an acoustic noise analysis of an operating industrial plant is presented.

## 2. Problem statement

The plant has two big halls separated by a shared wall. In one of these halls, a furnace for glass hardening was placed. This furnace has been subjected to vibroacoustic tests as a source of excessive noise and vibration in the hall environment, causing noise pollution at the furnace work stations and neighboring machines. Overall dimensions of furnace are 22 m x 3.8 m x 3.8 m, and its weight is 32 tons. The heating power of the furnace is 624 kW, and its cooling capacity is 1130 kW. During operation, in the furnace are tempered glass plates with a thickness of 3.2 mm to 19 mm. The furnace is operated by two two-person teams of operators whose workplaces are located at the beginning and the end of the glass hardening processing line.

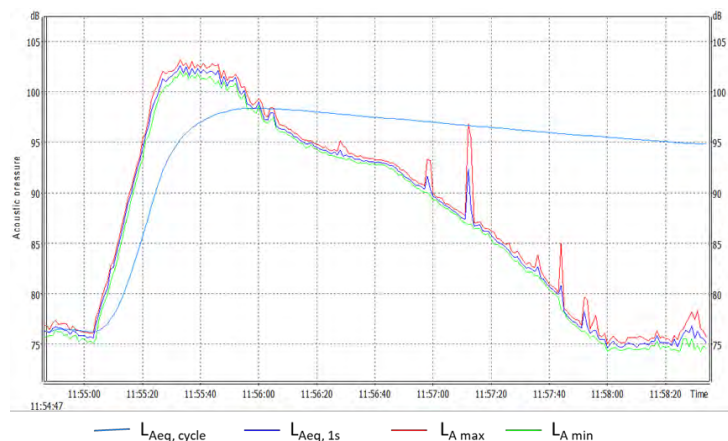


Figure 1. One cycle of hardening glass process

The technical process involves loading two panes of glass on feed rollers, that guide the glass panes through the furnace chambers, where the process of hardening occurs, and then operators pick up the glass pane from conveyors and placement them in special racks.

The cycle of two glass panes passing through the furnace is on average about 4 minutes (Fig. 1). The glass hardening process is accompanied by high level noise related to the blowing air into the furnace. The furnace is supported by four blowers equipped with 220 kW electric motors operating at a speed of 1450 rpm.

### 3. Noise measurements and hall modelling

#### 3.1. Measurements

To assess the impact of noise from furnace devices for glass hardening, sound measurements were made at operator work stations and at selected points (Fig. 2) in both halls, using the measurement methods described in the standards [11 - 13]:

- noise measurements around 3 fan motors that blow air into the furnace at a distance of 1 m from the machine contour [11];
- measurements of noise levels around the furnace at a distance of 1 m from the furnace contour [12];
- measurements at operators' work stations [13];
- measurements at neighboring work places for the glass hardening furnace [13],
- measurements in selected locations in the neighboring hall.

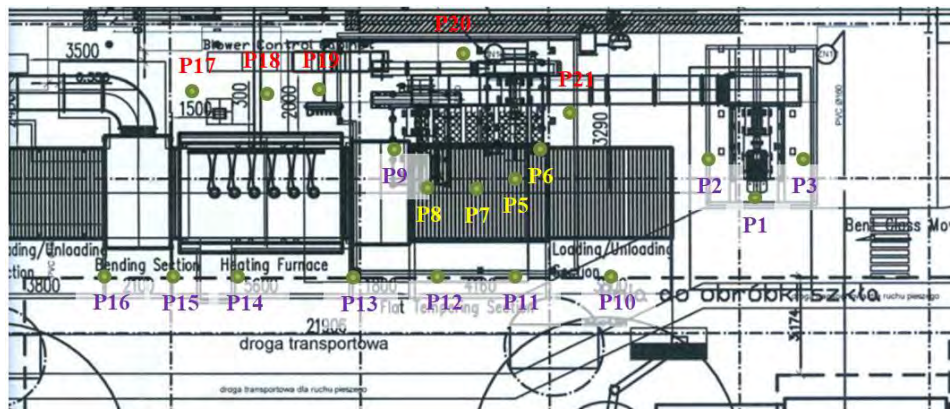


Figure 2. Layout of noise measurement points

All noise measurements were made using the SVAN 971 sound level meter. The measurement results were used to assess the noise hazard at all work stations in the hall, as well as to calibrate the acoustic model of the hall.

Table 1. Noise measurement results (measuring points shown at Fig. 2)

No	No measurement point	A-weighted sound pressure level $L_{Aeq}$ [dB]	A-weighted sound pressure level MIN $L_{Amin}$ [dB]	A-weighted sound pressure level MAX $L_{Amax}$ [dB]	Noise background level $L_{Abg}$ [dB]
<b>FAN MOTOR W1</b>					
1	P1	<b>93.2</b>	91.0	94.8	64.5
2	P2	<b>95.4</b>	94.0	96.5	
3	P4	<b>93.1</b>	91.5	94.3	
<b>FAN MOTOR W2</b>					
4	P5	<b>94.6</b>	91.1	96.2	64.5
5	P6	<b>96.1</b>	91.7	98.0	
6	P7	<b>93.2</b>	81.3	98.9	
<b>FAN MOTOR W3</b>					
7	P7	<b>93.2</b>	81.3	98.9	64.5
8	P8	<b>95.0</b>	89.4	97.5	
9	P9	<b>95.7</b>	90.9	97.4	
Emission A-weighted sound pressure levels (PN EN ISO 11202)					
10	P10	<b>92.1</b>	74.3	97.3	64.5
11	P11	<b>94.0</b>	74.0	99.0	
12	P12	<b>95.7</b>	78.6	100.3	
13	P13	<b>93.0</b>	76.6	98.1	
14	P14	<b>89.7</b>	75.3	95.0	
15	P15	<b>88.3</b>	75.8	93.3	
16	P16	<b>87.6</b>	74.1	92.8	
17	P17	<b>91.6</b>	72.7	91.6	
18	P18	<b>92.9</b>	72.6	99.8	
19	P19	<b>94.7</b>	73.3	101.2	
20	P20	<b>96.2</b>	74.3	103.2	
21	P21	<b>96.2</b>	76.0	101.4	

### Noise at the workplaces

Table 2 contains the noise risk assessment at operator workplaces SP1 and SP2 of the glass hardening furnace:

Table 2. Noise hazard assessment for 8 hours of working time shift in the operator workplace at the glass hardening furnace

No	No work station	Equivalent sound level for 8 h of work $L_{Aeq8h}$ [dB]	Maximum value sound level $L_{Amax}$ [dB]	Peak sound level $L_{Cpeak}$ [dB]	Background sound level $L_{Abg}$ [dB]
		Limit value $L_{Aeq8h} = 85$ dB	Limit value $L_{Amax} = 115$ dB	Limit value $L_{Cpeak} = 135$ dB	
1	SP1	90.6	100.3	116.3	64.5
2	SP2	93.9	103.2	120.4	64.5



The timetable of 8 hours work day of the furnace operators consist of:

- 105 cycles (one cycle 4 minutes) = 420 minutes,
- 30-minutes – breakfast break (background sound level  $L_{Abg} = 64.5$  dB),
- 30 minutes-preparing for the shift beginning and end of the shift (background sound level  $L_{Abg} = 64.5$  dB).

### 3.2. Acoustical model of the industrial hall and noise distribution computer simulation

Geometrical acoustics are used as a standard model for room acoustic design and predicting. Research on room acoustic simulation focuses on a more accurate modelling of propagation effects such as diffraction and other wave effects in rooms [10]. Odeon Combined 15 was used as a software for acoustic analysis of this hall. This software allows to provide sound pressure calculations for selected sources both in single points and grid of points using hybrid method based on Image Source Method and Ray-Radiosity.

ODEON is an energy-based room acoustic modeller, but sound waves are represented by rays, straight lines connecting sources and receivers. Reflections are represented by image and secondary sources, where the sound is still considered as rays. These simplifications make it possible to calculate the acoustic response in large spaces, while maintaining a short computation time with a high level of accuracy [5]. All calculations in ODEON are performed in the energy domain - not in the pressure domain. We use not only point sources for proper simulation of noise sources. Point, line, surface and multi surface sources are available in Odeon. For each source in the model could be assigned a certain radiated sound power in octave bands. The geometry of the noise sources should be be approximately modelled, and the surfaces should be used as sources radiating sound [7].

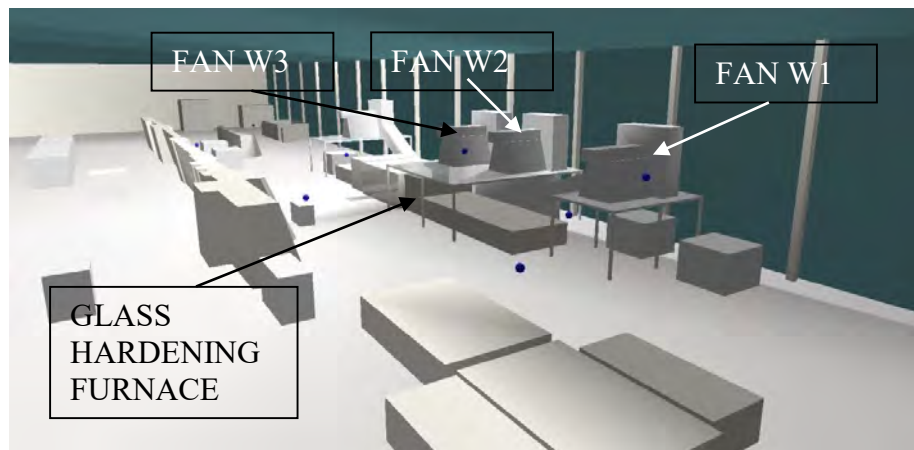


Figure 3. View of 3D model of hall with the glass hardening furnace

The geometrical model of hall was created in SketchUp software and then exported to ODEON (Fig. 3) using SU2ODEON plug-in recommended by ODEON developer,

what allowed to simplify geometrical preparation of hall as well as later added models of noise protection solutions [5].

The receivers are placed exactly at the workplace and represented by 3D Models. As input data to the model, the designated sound power levels of individual devices and the measured the Emission Sound Pressure Levels. Calibration of the sound power level of noise sources is done using the measurement results. For selected workplaces, full noise analysis is provided to take into account the impact of each hypothetical operating noise source (including reflections from walls).

#### 4. Discussion

Using the Odeon software, a series of computer simulations and calculations were made that gave the possibility to choose the best anti-noise solution for use in in situ conditions. Among the proposed solutions, calculations of the semi-open sound-absorbing and sound insulating enclosure of the ventilators were made, acoustic panels hang from the ceiling above the furnace, acoustic adaptation of the back wall behind the furnace, increased acoustic insulation of the air supply ducts to the furnace, soundproofing curtain separating the back wall of the furnace.

Different configurations of the proposed anti-noise measures were also analyzed by plotting maps of the A-weighted sound level distribution above the floor surface of the hall. The distribution of A-weighted sound level on 1.5 m over floor of the hall in the present state, without noise protection means has shown on Figure 4 and after the simulation of the application of soundproofing enclosures to the fan motors has shown on Figure 5.

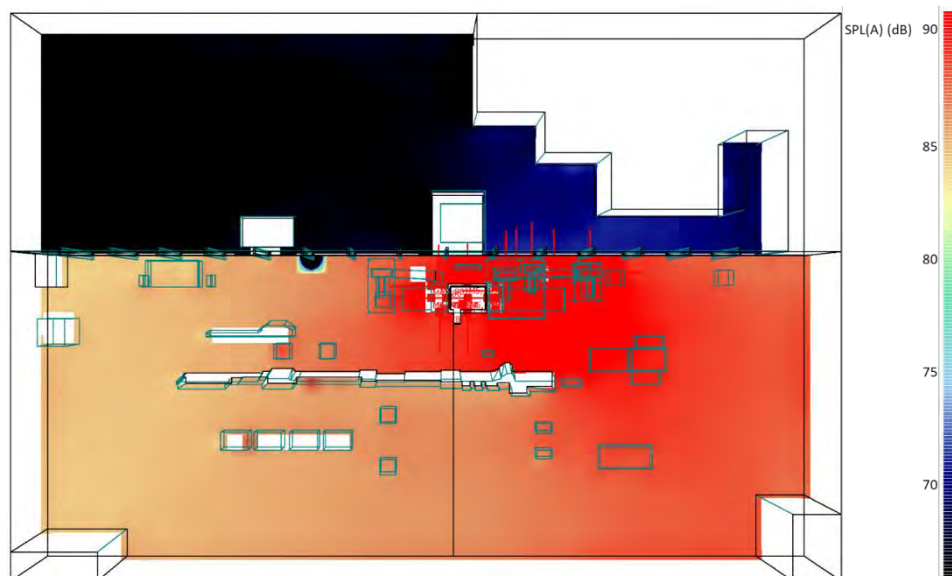


Figure 4. A-weighted sound level distribution – current status

Noise maps obtained from the simulation of sound level distributions in the hall for individual configurations of the proposed noise abatement protections point at the possibilities of noise reduction at workplaces at the furnace as well as at sites adjacent to the furnace. Analysis of the expected effects allows for the selection of measures that are acoustically effective and estimating the costs of their application. In the present case of the hall, taking into account the amount of exceeding the admissible values, it is necessary to apply measures with high acoustic efficiency, such as soundproofing enclosures for fan motors, but it isn't insufficient. The reduction of the A-weighted sound level at workplace SP2 is over 3 dB (A). In addition, adaptation of the back wall should be used (acoustic efficiency is about 1.5 dB (A) and sound insulation enhancement of air duct (about 0.8 dB (A)). The same applied noise protections, in the work place on the neighboring machine, give more than 6 dB (A) acoustic effectiveness. Discussion with the plant owner will allow you to choose the most effective from the point of view of acoustics and the most economical solution.

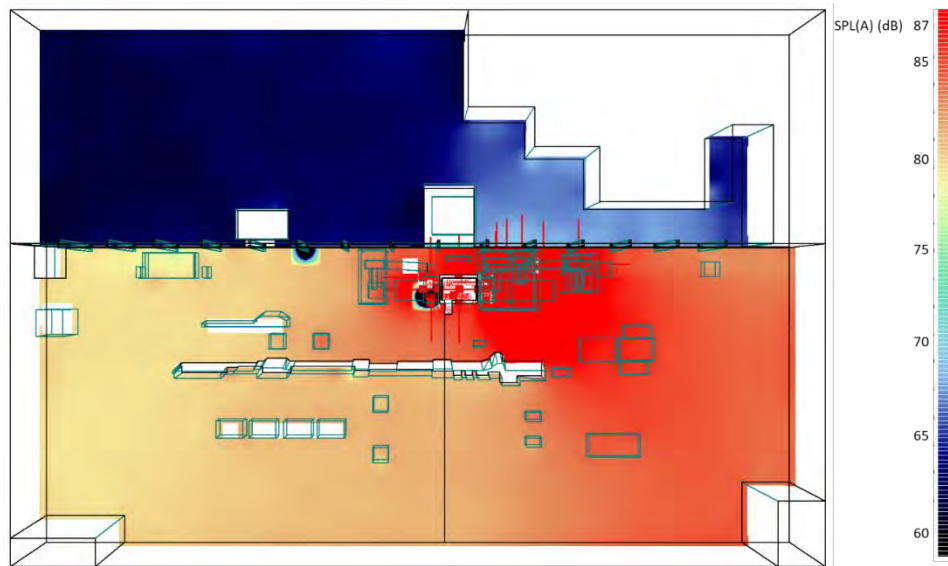


Figure 5. A-weighted sound level distribution after applied enclosures for fan motors No1, No2, No3

## 5. Conclusions

The reliability of computer modeling is confirmed by comparing the calculated and measured sound pressure levels at reference points. We created the 3D acoustic model of the hall, then calibrated it with the results of measurements and finally received a tool to solve the noise problems that we set at the beginning. Our project was created in order to analyze various scenarios of many possible solutions for noise reduction, calculating their acoustic efficiency and choosing the most effective variant.

One of the most valuable simulation results is the comparison of the share of individual noise sources in the values of A-weighted sound level at the reference points. These values allow estimating the acoustic efficiency of each proposed anti-noise solution. Noise maps showing the distributions of A-weighted sound levels above the floor surface of the hall enable the estimation of noise hazards at workstations located in this hall.

### Acknowledgments

Work has done as part of the statutory work 16.16.130.942 of the Department of Mechanic and Vibroacoustic of the AGH-UST in Krakow

The authors would like to thank J. Wiciak, B. Borkowski and R. Trojanowski for their support in carrying out research tests on the furnace for glass hardening.

### References

1. M. Crocker, *Encyclopedia of Acoustics*, J. Wiley & Sons, Inc., New York, Ch. 77, 1997.
2. Z. Engel, J. Piechowicz, L. Stryczniewicz, *Foundamentals of industrial vibro-acoustics*, AGH, Kraków, 2003.
3. L. L. Beranek, I. L. Vér, *Noise and Vibration Control Engineering – Principles and Applications*, John Wiley & Sons, New York (2006).
4. C. L. Christensen, J. H. Rindel, *A new scattering method that combines roughness and diffraction effects*, Forum Acousticum, Budapest. (2005).
5. C. L. Christensen, G. Koutsouris, *Odeon Room Acoustics Software*, User Manual, Ver. 12, 2nd Ed., Ch. 6.3. (2013).
6. A. Krokstad A., S. Strom, S.Sorsdal, *Calculating the acoustical room response by the use of a ray tracing technique*, J. of Sound and Vibration, **8**(1)(1968) 118 – 125.
7. J. H. Rindel, *Computer simulation techniques for acoustical design of rooms*, Acoustics, Australia, **23** (1995) 81 – 86.
8. J. H. Rindel, *Odeon and the scattering coefficient*, Power Point Presentation from Odeon Workshop at Baltic-Nordic Acoustical meeting, Mariehamn, (2004).
9. L. Savioja, U. P.Svensson, *Overview of geometrical room acoustic modeling techniques*, JASA, **138**(2) (2015) 708 – 730.
10. M. Vorländer, *Computer simulations in room acoustics. Concepts and uncertainties*, JASA, **133**(3) (2013) 1203 – 1213.
11. PN-EN ISO 3746:2010, *Acoustics - Determination of sound power levels and sound energy levels of noise sources using sound pressure - Survey method using an enveloping measurement surface over a reflecting plane*.
12. PN-EN ISO 11202:2010, *Acoustics - Noise emitted by machinery and equipment - Determination of emission sound pressure levels at a work station and at other specified positions applying approximate environmental corrections*.
13. PN-ISO 9612:2011, *Acoustics - Determination of occupational noise exposure - Engineering method*.

## **Binaural Speech Segregation System on Single Board Computer**

Tsuyoshi USAGAWA

*Kumamoto University, 2-29-1 Kurokami, Chuo, Kumamoto, 860-8555, Japan,  
tuie@cs.kumamoto-u.ac.jp*

IRWANSYAH

*Kumamoto University, 2-29-1 Kurokami, Chuo, Kumamoto, 860-8555, Japan,  
irwan@hicc.cs.kumamoto-u.ac.jp, iru.one.syah@gmail.com*

### **Abstract**

A pocket-size binaural speech segregation system has been developed and assembled with available consumer hardware. It can enhance a target speech in a certain direction while attenuating interfering sounds from other directions. This system is based on the frequency domain binaural model (FDBM) and it segregates multiple speeches based on the directivity. This real-time system is implemented on a low-cost single board computer which might be used as a hearing assistance device. Performance of the system is evaluated in a normal laboratory room as well as anechoic chamber. Even in a room with reverberation, the system works well and show the almost same performance obtained in anechoic chamber.

**Keywords:** Frequency Domain Binaural Model, Single Board Computer, Hearing Assistance, Open-Source Software

### **1. Introduction**

Very rapid aging of society in Japan, hearing assistant systems and hearing aids attract the attention of senior members as well as ones who have hearing disables. There are several very sophisticated hearing aids, namely digital ones, on a market, however, comparing the senior glasses, the market price has a significant differences between hearing aids and senior glasses. Also monaural hearing assistance does not always work well as monocular glass does not, thus it is recommended to use a pair of hearing aids for binaural assistance, which costs almost the double. Beside other reasons, this price gap makes a hearing aid difficult to be popular.

In this paper, the development of an “open-source” hearing assistance device with consumer hardware in affordable price range. The proposed hearing device is intended to assist a listener to focus her/his attention on a speaker that she/he is talking with. This binaural system can segregate and enhance sounds which come from specified direction using binaural cues by a frequency domain binaural model (FDBM) proposed by Nakashima *et al.* in 2003. This model uses the binaural cues, namely interaural level difference (ILD) and interaural phase difference (IPD) to estimate directions of sources in multiple sound source condition and segregates sounds based on estimated directions. Also, since FDBM works in the frequency domain, it has a relatively low computational

complexity compared to time domain based models. Using this advantage, FDBM is implemented on a cost-effective single board computers (SBCs) and it is controlled by Android-based smart phone by an application named “OpenFDBM” which allows a user to control the hearing assistance device via graphical interface. The performance of implemented FDBM as a binaural hearing assistance system is measured in an anechoic chamber as well as ordinary laboratory room with reverberation. Source code of FDBM written in Python which has rich audio signal processing libraries, run as an real-time application on SBC. Source codes including instructions to build the hearing device are freely available on Github and demonstration video is also available online.

## 2. Frequency Domain Binaural Model

This section shortly describes the overall overview of the original frequency domain binaural model [1]. However, here the original algorithm is slightly modified to use zero padding of fast Fourier transform (FFT) in order to maintain a high-frequency resolution when using a shorter frame length. Figure 1 shows the overall FDBM-based speech segregation scheme. Let  $x_R(i)$  and  $x_L(i)$  be the observed signals received at left and right microphones, defined here as

$$x_R(i) = s(i) * h_{r,\phi_s}(i) + v(i) * h_{r,\phi_v}(i) = s_R(i) + v_R(i) \quad (1)$$

$$x_L(i) = s(i) * h_{l,\phi_s}(i) + v(i) * h_{l,\phi_v}(i) = s_L(i) + v_L(i) \quad (2)$$

where  $s(i)$  and  $v(i)$  are the target and interfering speakers, respectively, and this notation “\*” indicates the convolution operator. When someone is talking in front of us, the speech signal received by left and right ears would be different, depending on the direction of the sound source. These acoustical differences perceived by left and right ears can be summarized by head related transfer functions (HRTFs).  $h_{l,\phi}(i)$  and  $h_{r,\phi}(i)$  are then defined as the left and right HRTFs as a function of  $\phi$  azimuth. We divide the FDBM into six stages: Framing Stage, ILD-IPD Calculation, Database Comparison, Weight Combination, Segregation Filter and Separation Stage.

### Stage 1. Framing Stage:

For conversion from time domain signals into frequency domain one, a short-time Fourier Transformation (STFT) was used in the original FDBM [2], however a perfect reconstruction method with a half-cycle sine window is introduced [3]. STFTs of observed signals are obtained by FFT for each windowed block:

$$X_R(\lambda, k) = FFT\{x_R^{zp}(i)\}; X_L(\lambda, k) = FFT\{x_L^{zp}(i)\} \quad (3)$$

where  $\lambda$  is the frame index and  $k$  is the frequency index. The zero-padded windowed block,  $x_{\{\cdot\}}^{zp}(i)$  is defined as follows

$$x_{\{\cdot\}}^{zp}(i) = \begin{cases} x_{\{\cdot\}}(i) \cdot w(i - \lambda \frac{N}{2}), & 0 \leq i \leq N - 1 \\ 0, & N \leq i \leq M - 1 \end{cases} \quad (4)$$

where  $N$  is the window length we want to shorten, and  $M$  is the frame length. As discussed in [4], the quality of the segregated sounds depends on the frequency resolution of FFT, which is required at least a 32-ms frame length. For that reason,  $M$  is set to 32 ms.

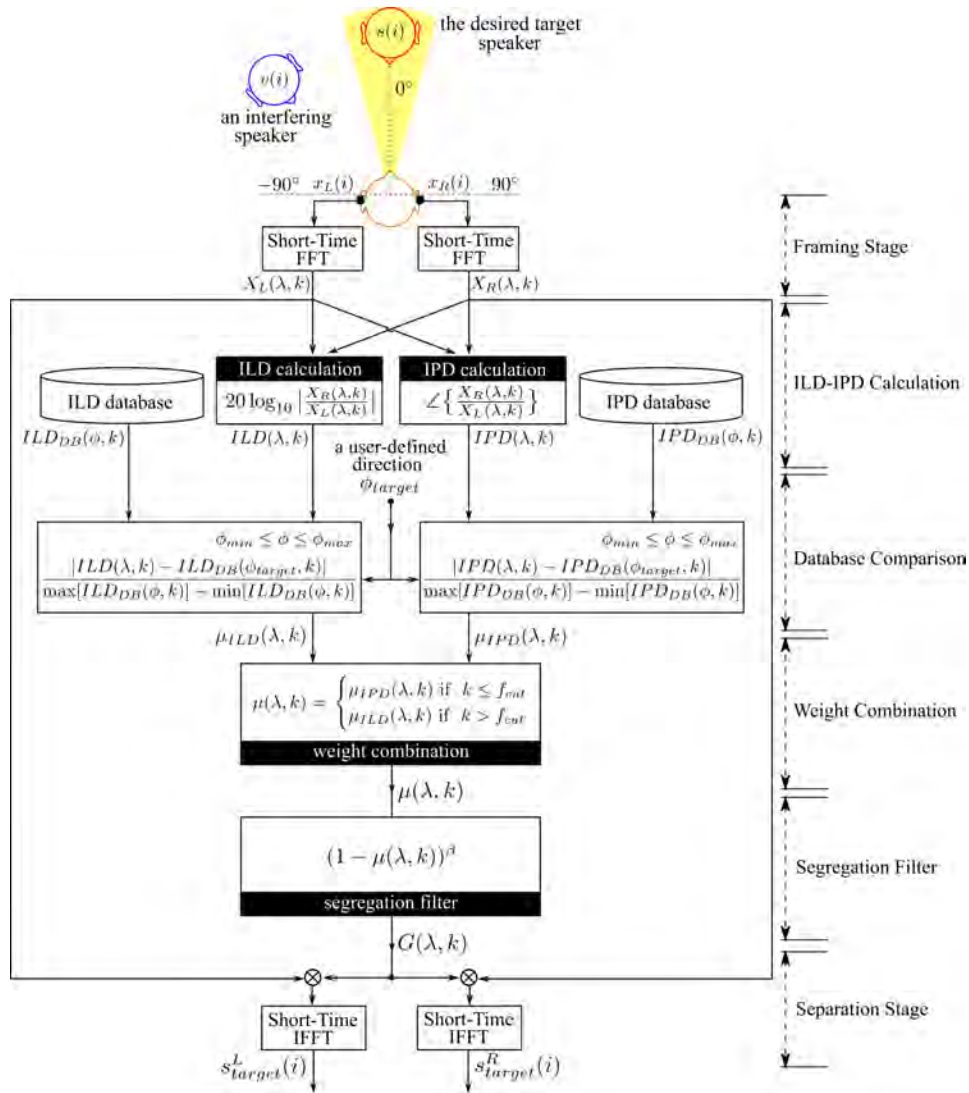


Figure 1. Overall structure of the FDBM-based speech segregation scheme

Stage 2. ILD-IPD Calculation:

It is well known that humans are able to localize a sound source in space by exploring the differences in time and level between the sound reaching the left and right ears. These relative differences are known IPD and ILD, respectively. If HRTFs are available in advance, a target speech signal in a particular direction is possible to be extracted. Using ILD and IPD information, the FDBM is developed to segregate the desired target signal. The ratio of the STFTs of the observed signals, known as the interaural spectrogram  $X_{IS}(\lambda, k)$ , can be obtained using Eq.(6) and ILD and IPD can be calculated as follows;

$$X_{IS}(\lambda, k) = \frac{X_R(\lambda, k)}{X_L(\lambda, k)} \quad (5)$$

$$ILD(\lambda, k) = 20 \log_{10} |X_{IS}(\lambda, k)|; \quad IPD(\lambda, k) = \angle X_{IS}(\lambda, k) \quad (6)$$

where  $\angle$  denotes the phase angle.

Stage 3. Database Comparison:

To extract a target speech signal in a particular direction  $\phi_{target}$ , HRTFs are required by the FDBM. Fortunately, MIT Media Lab has provided HRTFs of KEMAR Dummy-Head microphone [5]. Let  $H_{DB,R}(\phi, k)$  and  $H_{DB,L}(\phi, k)$  are right and left HRTF database as a function of azimuth in the frequency domain. Similar to the previous equations, ILD and IPD database can be formulated as

$$ILD_{DB}(\phi, k) = 20 \log_{10} \left| \frac{H_{DB,R}(\phi, k)}{H_{DB,L}(\phi, k)} \right|; \quad IPD_{DB}(\phi, k) = \angle \left\{ \frac{H_{DB,R}(\phi, k)}{H_{DB,L}(\phi, k)} \right\} \quad (7)$$

HRTF database provided by MIT covers the full azimuth from 0 to 360 with 5-degree resolution. By comparing ILD and IPD of the observed signals with the ILD and IPD database, weight factors used for extracting the desired target signal in a specific direction target can be obtained. The calculation of weight factors can be defined as

$$\mu_{ILD}(\lambda, k) = \frac{|ILD(\lambda, k) - ILD_{DB}(\phi_{target}, k)|}{\max[ILD_{DB}(\phi, k)] - \min[ILD_{DB}(\phi, k)]} \quad (8)$$

$$\mu_{IPD}(\lambda, k) = \frac{|IPD(\lambda, k) - IPD_{DB}(\phi_{target}, k)|}{\max[IPD_{DB}(\phi, k)] - \min[IPD_{DB}(\phi, k)]} \quad (9)$$

$\max[f(\phi, k)]$  indicates the maximum value of  $f$  at a particular frequency index  $k$  when  $\phi$  is varying from  $\phi_{min}$  to  $\phi_{max}$ , depending on the available HRTF database.

Stage 4. Weight Combination:

The weight factors,  $\mu_{ILD}(\lambda, k)$  and  $\mu_{IPD}(\lambda, k)$ , from the previous stage could not work optimally in all frequency region.  $\mu_{IPD}(\lambda, k)$  works well only in the low-frequency region. On the other hand,  $\mu_{ILD}(\lambda, k)$  is good for the high-frequency region



[1]. Therefore, both need to be combined. The combination of both weight factors is defined as follows:

$$\mu(\lambda, k) = \begin{cases} \mu_{IPD}(\lambda, k) & \text{if } k \leq f_{cut} \\ \mu_{ILD}(\lambda, k) & \text{if } k > f_{cut} \end{cases} \quad (10)$$

where  $f_{cut}$  is assumed to be the cutoff frequency which we consider to be 1250 Hz. To avoid amplification, the upper limit of the weight factor is set to one.

#### Stage 5. Segregation Filter:

The segregation filter is then defined as

$$G(\lambda, k) = [1 - \mu(\lambda, k)]^\beta \quad (11)$$

where  $\beta$  is the gain control parameter which is set to 16 obtained from preliminary tests.

#### Stage 6. Segregated Signal Reconstruction

To segregate a target speech signal, the segregation filter  $G(\lambda, k)$  is applied to both the left and right observed signals as follows

$$S_{target}^R(\lambda, k) = G(\lambda, k) \cdot X_R(\lambda, k); \quad S_{target}^L(\lambda, k) = G(\lambda, k) \cdot X_L(\lambda, k) \quad (12)$$

The target speech signals are then obtained by overlap-adding the windowed outputs

$$s_{target}^R(i) = \sum_{\lambda} IFFT\{S_{target}^R(\lambda, k)\}_{trunc} \cdot w(i - \lambda \cdot \frac{N}{2}) \quad (13)$$

$$s_{target}^L(i) = \sum_{\lambda} IFFT\{S_{target}^L(\lambda, k)\}_{trunc} \cdot w(i - \lambda \cdot \frac{N}{2}) \quad (14)$$

where  $IFFT\{\cdot\}_{trunc}$  means that the output signal of IFFT must be truncated (due to the zero-padding) to restore the original size  $N$  of the signal.

### **3. FDBM Implementation on SBCs**

For synchronization of audio and visual information, a latency between input and output signal must be negligibly short in order to reduce the chance to be detected by a user. In many real-time audio applications, low latency is very important. Especially in the case of hearing assistance system, latency may cause an echo in the perceived sound when users is listening to their own voice [6]. Thus, reducing latency is one of major issue on FDBM implementation.

Zero padding of FFT is used on SBC to implement FDBM to maintain a high-frequency resolution using a short frame length. As well known that, in a real-time application, a frame length refers to a buffer size where if it is too short we can suffer

crackles and audio dropouts. Thus, performing real-time processing on recent devices, such as ARM-based single board computer (SBC), needs to be discussed.

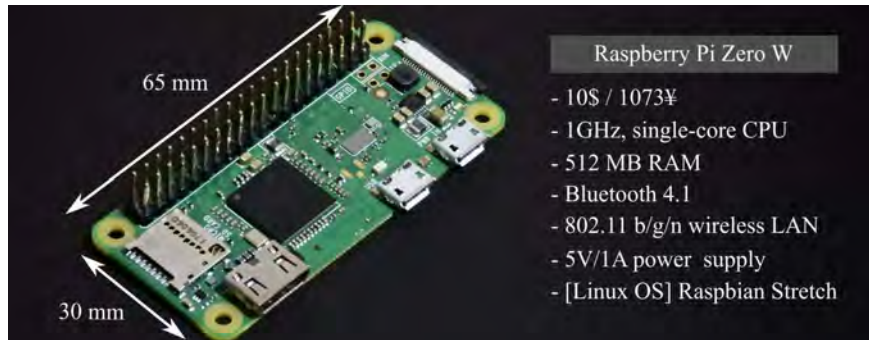


Figure 2. ARM-based single board computer (SBC)

Figure 3(a) shows the simulation condition where the target speaker is located at  $0^\circ$ , and the interferer at  $-45^\circ$ . The sampling frequency is 16 kHz, and SNR is set to 0 dB. Fifty male and fifty female speech signals were used as target speaker and interferer, respectively. The aim of this simulation is to investigate whether or not using zero-padding allows the use of a shorter buffer size ( $< 32$  ms) and can still maintain the quality of segregated sounds, evaluated by Perceptual Evaluation of Speech Quality (PESQ). Figure 3(b) shows the mean of PESQ scores as a function of frame length based on the simulation condition shown in Fig. 3(a). It can be seen that a shorter frame length could still be used where a similar PESQ score to the original FDBM was obtained when  $16\text{ms} < N < 32\text{ms}$ .

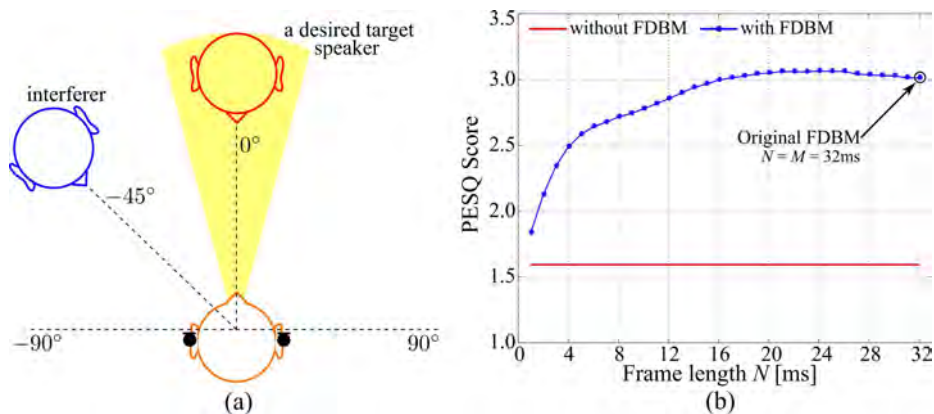


Figure 3. Simulation condition (SNR = 0 dB) and obtained PESQ against  $N$

**Hardware:**

Here the hearing assistance device was assembled with available consumer hardware for about US\$73 in total as shown in Fig. 4. The signal processing unit to which interfaces (i) an Andrea SB-205B – binaural microphone equipped earphones – is (ii) SBC: Raspberry Pi Zero W. This SBC runs the Raspbian, Debian-based OS, stored on (iii) a microSD card. It requires a power supply of 5V/1 A so that a lithium polymer (LiPo) battery: (iv) 3.7V 1300 mAH is equipped with (v) a step-up DC-DC converter. (vi) A little LiPoly charger is used to charge the battery via USB mini-B connector. (vii) A case from LEGO® was also made to cover all components. (viii) A case from LEGO® was also made to cover all components.

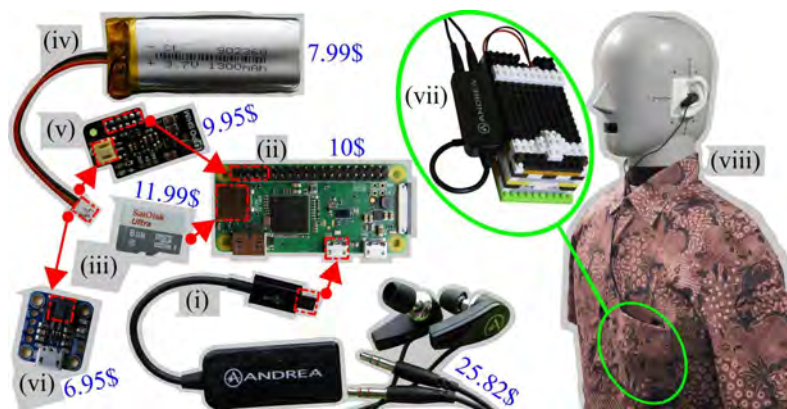


Figure 4. A hearing assistance system built with consumer hardware

**Software:**

The Python is the main language we used to run the system in real-time. Dummy head on right hand side of Fig. 4 shows an example of how to use the prototype on (viii) a dummy head. The list of hardware and software is available on our Github (<https://github.com/shiinoandra/OpenFDBM>).

**4. Evaluation of FDBM on SBC****Directivity Pattern Measurement:**

Figure 5 illustrates the procedure of the directivity pattern measurement as a function of direction. A dummy head wearing ear microphones/earphones was placed on the turntable and rotated every 5 degrees while five female and five male speech signals were produced by a BOSE 101VM loudspeaker located 1.4 m from the center of the dummy head. The prototype hearing device is running at sampling frequency of 16 kHz which we considered to be suitable for the Raspberry Pi Zero W to run the FDBM smoothly. When it was running, the computer used about 0.23A which can last over 3 hours when using a 1300-mAH battery. To control the device, an Android application so-called “OpenFDBM” has been developed using available open-source software.

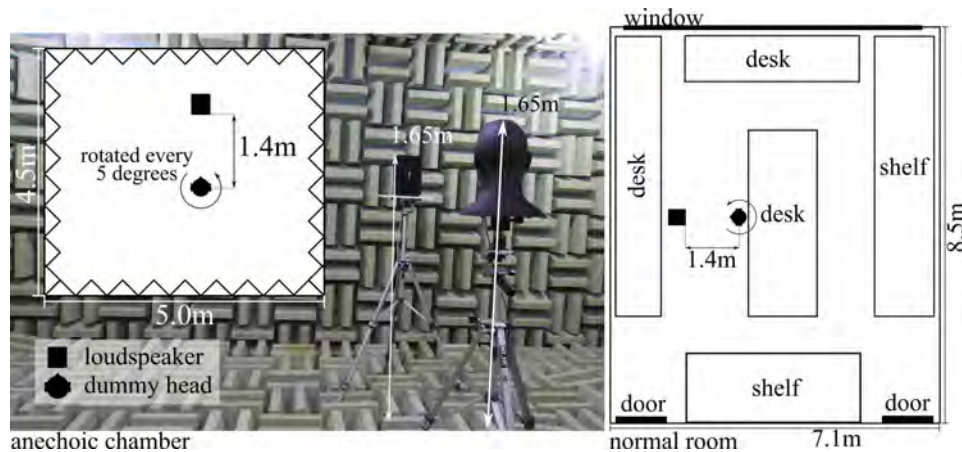


Figure 5. Schematic diagram to measure directivity pattern in an anechoic chamber and a normal laboratory room with a reverberation time (RT60) of 300 ms

Figure 6 shows the directional characteristics of the hearing device. It can be seen that the device running the FDBM enables a user to focus on a speaker in a certain direction, which is now only available for three directions: front, left and right sides (see a demo video <https://youtu.be/FxcLsXWcGg0>). In addition, “Front” mode does not attenuate the speech signals from the back (see Fig. 6). This phenomenon occurs because of a “front–back” confusion problem when using two microphones. A video example demonstrating how the “OpenFDBM” software is controlling the device in real-time is available at <https://youtu.be/avdnHMR2AR4>.

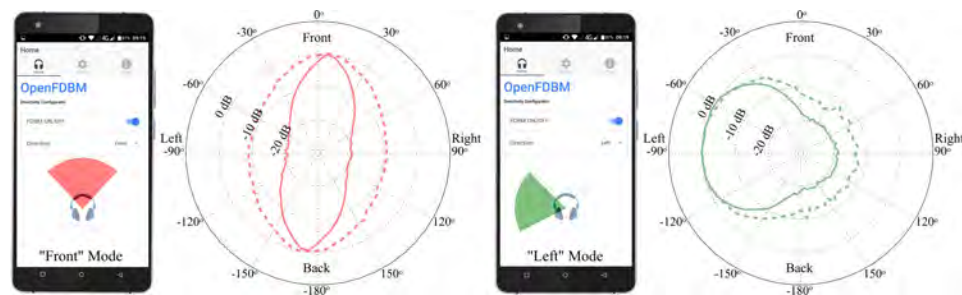


Figure 6. Directional characteristics of the Raspberry Pi hearing assistance device evaluated in an anechoic chamber (solid line) and a normal room (dashed line)

## 5. Conclusion

Real-time binaural speech segregation system using single board computer (SBC) has been successfully implemented. An “open-source” Raspberry Pi hearing assistance

device is assembled with available consumer hardware for about US\$73 in total. Implemented FDBM allows to focus user's attention on a target speaker in a particular direction. An android application so-called "OpenFDBM" has also been developed to help users to choose the direction of a speech they want to focus on. Source codes including instructions to build the system are available on GitHub and demonstration video of this system is also available online on YouTube.

## References

1. H. Nakashima, Y. Chisaki, Y. Usagawa, M. Ebata, *Frequency domain binaural model based on interaural phase and level differences*, Acoustical Science and Technology, **24** (2003) 172 – 174.
2. Y. Chisaki, K. Matsuo, K. Hagiwara, H. Nakashima, T. Usagawa, *Real-time processing using the frequency domain binaural model*, Applied Acoustics, **68** (2007) 923 – 938.
3. I. W. Selesnick, *Short-time fourier transform and its inverse*, 2009. (In: url [http://eeweb.poly.edu/iselesni/EL713/STFT/stft\\_inverse.pdf](http://eeweb.poly.edu/iselesni/EL713/STFT/stft_inverse.pdf) visited on 06/24/2019).
4. N. Hiruma, R. Kouyama, H. Nakashima, Y. Fujisaka, *Low delay wind noise cancellation for binaural hearing aids*. INTER-NOISE, (2016) 4844 – 4854.
5. MIT-Media-Lab, *HRTF Measurements of a KEMAR dummy-head microphone*, 2019. (<https://sound.media.mit.edu/resources/KEMAR.html> visited on 06/24/2019).
6. J. Agnew, J. M. Thornton, *Just noticeable and objectionable group delays in digital hearing aids*, J. the American Academy of Audiology, **11** (2000) 330 – 336.



## **Intonation Accuracy and Pitch Stability During Crescendo as the Voice Quality and Singer's Experience Indicator Among Choral Singers**

Mateusz GAWLIK

*AGH University of Science and Technology,  
30–059 Cracow, Mickiewicza 30,  
matgawel@agh.edu.pl*

Wiesław WSZOLEK

*AGH University of Science and Technology,  
30–059 Cracow, Mickiewicza 30,  
wwszolek@agh.edu.pl*

### **Abstract**

The ability to correctly reproduce notes by the voice is one of the essential features of the singing task and called intonation. In combination with other parameters like timbre, formants, and sound attack, it affects the reception of listening impressions. In this paper, we present results of the examination concerning the automatic evaluation of intonation among the nonsingers, untrained and trained choral singers. We performed both pitch error during vocalization and pitch stability in crescendo task analysis among studied groups. We used Zero Band Filtering method to determine fundamental frequency from the singing signal. We noticed significant differences between singers with different skills and experience, and the possibility to classify the level of advancement of the singer by using intonation characteristic.

**Keywords:** intonation, singing voice, pitch accuracy, jitter, signal processing

### **1. Introduction**

Nonprofessional singers tend to get common disorders due to too high laryngeal muscles tension or other reasons connected with unskillful singing. Lack of singing technique leads to vocal disorders that can cause the loss of the ability to sing or even a complete loss of voice. Correct determination and elimination of incorrect techniques is therefore essential task during singing training. The researchers of the singing voice assessment analyzed it in various ways. It is possible to assess advancement level by sound pressure level and phonation threshold pressure [1–3], subglottal pressure [4], glottal flow parameterization [5, 6], formants, and timbre evaluation [7, 8]. One of the factors useful for assessing the technical skills of a singer is also intonation [9]. Music experts can judge the sound as purely based only on pitch accuracy. Intonation is also an important factor when determining the singer's tessitura.

In [10] the researchers used intonation and vibrato parameter to classify singing technique into two groups (poor or good) with accuracy to the semitone. Authors used low-pass filtering to remove  $f_0$  fluctuations (connected with vibrato). They observed

that pitch interval accuracy was constant and close to zero for good singers and irregular, with a positive trend to 100 cents for poor ones.

In [11] authors defined the intonation as the ability to reflect the pattern tone accurately. The singers were classified into four classes: unsatisfactory, satisfactory, good and very good, based on the correctly performed number of phrases according to pitch and slope values. They set up an acceptable range of tone within a deviation of 20 cents and the slope absolute value  $< 0.01$ . They noticed that poor singers tend to overstate or understate the pitch in certain sound intervals.

Researchers of the singing voice often use the standard deviation of the  $f_0$  as an indicator of pitch accuracy [12]. In [13] the average of the standard deviation was no greater than 15 cents for the different tones among good singers of an amateur choir. It took about 150 ms for a singer to adjust to a heard change [14], where the ability of skilled choir members to adapt to sudden tones shifting by 50 or 100 cents. Those results are similar to [15], where rapid pitch changes of 120 ms duration at the beginning and end of a tone were observed.

The study [16] analysed pitch accuracy among professional singers during ascending and descending intervals. Their results correspond to [12] and proved that, on average, the singer's intonation was comparable to the patterns with the standard deviation of 20 cents.

Table 1. Singers' experience level and middle tones for different voice types.  
NS – nonsingers, US – untrained singers, TS – trained singers

NS	US	TS	Pattern tones	
p3	p2	p1	Bass	E2
p4	p6	p5	Baritone	G3
p7	p8	p13	Tenor	B3
p15	p9	p17	Alto	E4
p21	p10	p20	Mezzo	G4
	p11		Soprano	B4
	p12			
	p14			
	p16			
	p17			
	p18			
	p19			

There are a lot of studies related to intonation evaluation, here we reviewed only some of them due to article volume limitations. A detailed description of the current state of the art is included in [12] and [15].

Previous papers described intonation analyzes among singers at a similar level of skills or years of experience. We noticed that authors focus on a wider research group rarely. This observation encouraged us to answer the question: Is it possible to distinguish singers' advancement or experience level among nonsingers, untrained singers and trained singers by  $f_0$  analyzing, based on pitch accuracy? Additionally, based on  $f_0$  and vocal intensity relationships observed in [17] we hypothesized that pitch stability during crescendo is connected with experience and vocal training.



## 2. Material and Methods

### 2.1. Data acquisition

Sixteen choir members with different singing skills, two students of theatre academy and three ordinary students were recorded in the anechoic chamber at the University of Science and Technology in Cracow. We used G.R.A.S 40 AF high-precision condenser microphone, always positioned at the 50 cm distance from the singer's mouth. The receiver was connected to G.R.A.S 12AA 2-channel power module and M-Audio PROFIRE 610 audio interface with Octanes preamplifiers. We posted information about the studied group in Table 1.

### 2.2. Recording session

We asked singers to perform crescendo task at vowel *a* at the certain pitch. Table 1 contains detailed information about tones, used for that exercise. We divided the studied group into three subgroups, according to the level of experience:

- Nonsingers (NS) – people without any singing experience
- Untrained singers(US) – members of Cracow choirs with different singing skills and experience
- Trained singers (TS) – graduates of the vocal music academy.

### 2.3. Signal parametrization

We looked for useful parameters to find differences in intonation in nonsingers and singers, as well as if the differences occur for singers group on various advancement levels. First, we calculated fundamental frequency ( $f_0$ ) of the signal, as the primary intonation indicator. We also appointed additional parameters such as jitter and linear regression coefficient. We used frame-based analysis of the signal with 25 ms frames length and 50% overlapping. Fundamental frequency was obtained by Zero Band Filtering method [18]. We used the Kruskal–Wallis test to check whether differences between groups are statistically significant. It was used in because of non-normality of the distribution of the parameters.

## 3. Results

At first, we calculated pitch error among all examined groups (Fig. 1 left) as the absolute value of difference from the pattern pitch. The results are presented in Tab. 2. The results are not surprising, and most of the tested singers (except p3, p9, and p16) intoned pattern pitch with high accuracy (low pitch error). However, we found a statistically significant difference between the median values of each group for all vowels (Kruskal–Wallis test,  $p < 0.01$ ). Additionally, trained singers obtained the lowest values of medians of pitch error for almost all vowels that ensures frequency resolution range 10–25 cents, which is similar to [11].

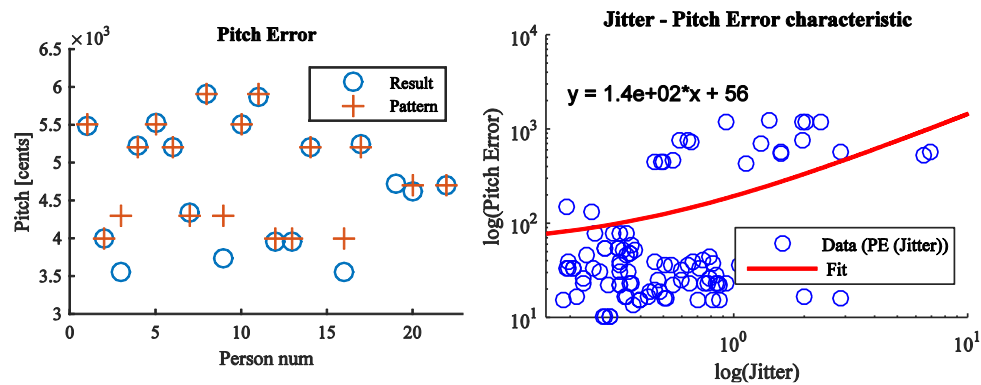


Figure 1. Pitch errors between patterns and sang vowels (left), boxplots median of linear regression coefficients from *a* vowel during (right)

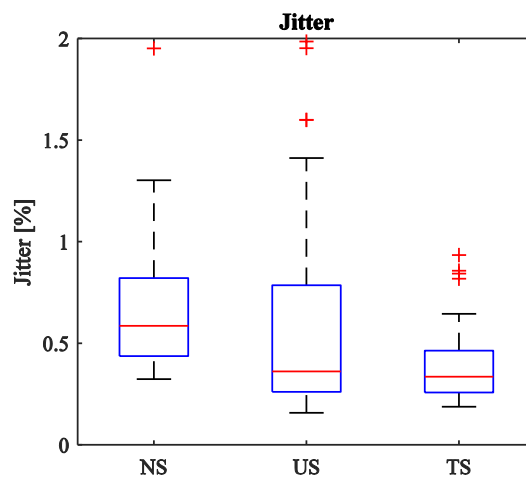


Figure 2. Boxplots median of linear regression coefficients from a vowel during crescendo task. NS – nonsingers, US – untrained singers, TS – trained singers

Next, the pitch error was expressed as a function of jitter parameter (Fig. 1 right). We observed that an increase of pitch error is connected with higher jitter values. Therefore untrained singers obtained higher values of this parameter than other groups (Fig. 3), but there was no statistically significant difference between these groups (Kruskal–Wallis test,  $p > 0.1$ ). Finally, we carried out a trend line examination during crescendo sounds. In all studied cases, untrained and trained singers obtained lower values of linear regression coefficients than nonsingers (Figs. 4 – 7). The median values were presented in Tab. 3. Additionally, we observed oscillation in  $f_0$  characteristics in singers groups, which was not present in nonsingers (Fig. 6). Moreover, the oscillating trend is more regularly for trained, then untrained singers (Fig. 5 and Fig. 6). The linear

regression coefficient values are the lowest for trained singers group. It is possible to differ this group from others at a 5% significance level ( $p = 0.04$  for Kruskal–Wallis test).

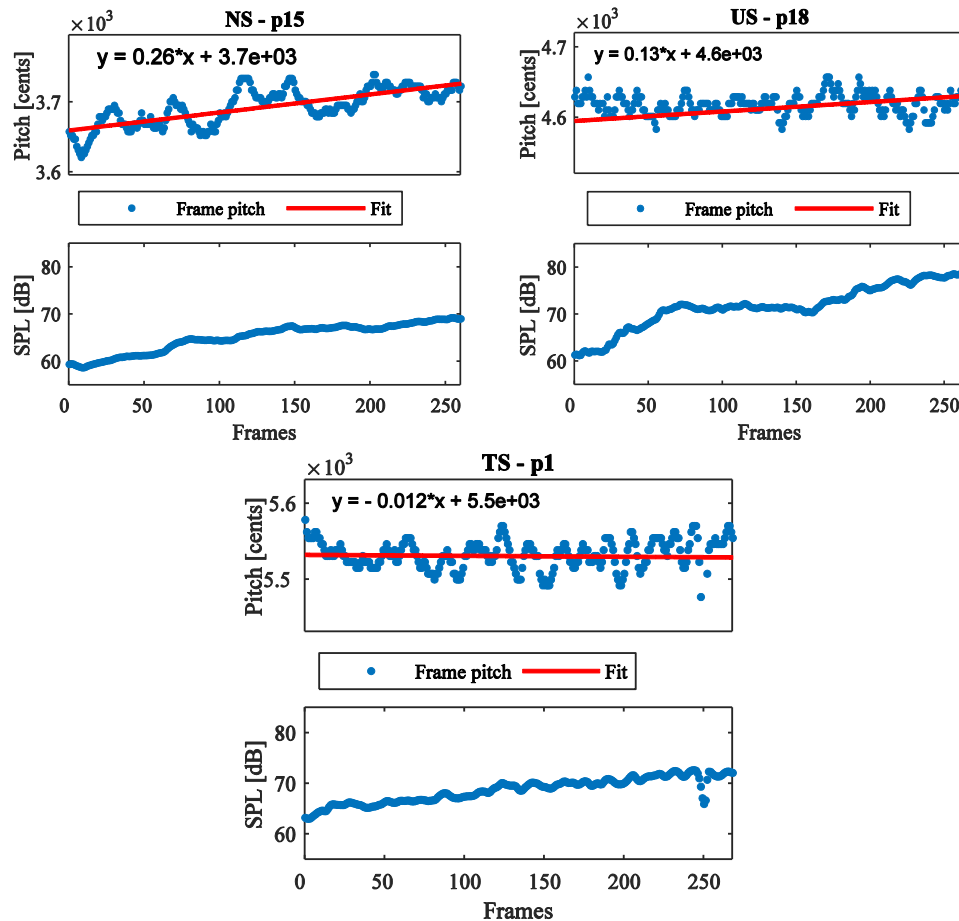


Figure 4. Example of the estimated trend line for nonsinger (NS), untrained (US) and trained (TS) singer subject. The bottom plots present Sound Pressure Level values

#### 4. Discussion

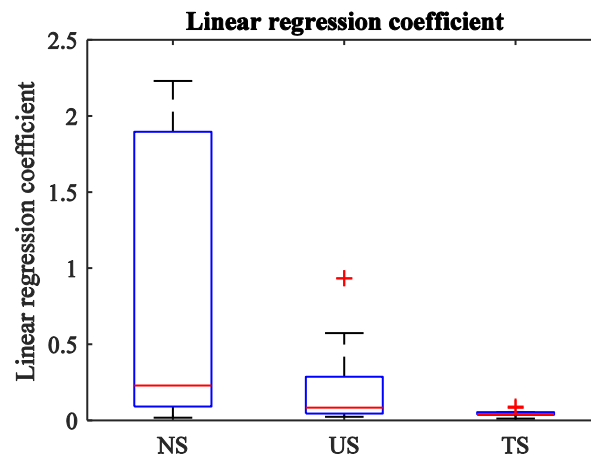


Figure 7. Boxplots median of linear regression coefficients from a vowel during crescendo task. NS – nonsingers, US – untrained singers, TS – trained singers

The wrong intonation for p3, and p9 is connected with the level of training: both singers are at the beginning of the voice lessons process. Additionally, p9 is an untrained singer with a couple of years of experience, whereas p3 is nonsinger. It caused differences in errors (170 cents).

The frequency resolution in all studied groups is lower than one tone (mean < 150 cents) and can be used as the voice advancement indicator (mean NS: 120 cents, US: 55 cents, TS: 20 cents). Mean pitch error at the level of 20 cents for trained singers group is similar to [11] results. The positive correlation between pitch error and jitter has not been noticed in the literature before. This trend results from the fact that pitch error is calculated as the absolute value of difference from the pattern tone. It means that greater frequency fluctuations cause the parameter to increase. Lower jitter values for trained subjects (median NS: 0.59%, US: 0.58%, TS: 0.21%) confirms pitch interval accuracy estimation for poor and good singers from [10]. Low values of Jitter indicate the ability to maintain a fixed frequency throughout the singing when it is needed.

The highest, positive regression coefficients for nonsingers during crescendo exercise proved the tendency to increase tone together with Sound Pressure Level as [3]. Additionally, this parameter can be used with others to distinguish trained from untrained singers (median US: 0.06, TS: 0.001), which corresponds to standard values from the literature [11]. Vibrato was not observed in the nonsingers group, whereas it was more regularly for trained than untrained singers. This behavior is well known in the literature [19], but it is worth to sign that it can be useful for the determination level of voice training and advancement by the regularity factor.

Table 2. Statistical parametrization for pitch error.  
NS – nonsingers, US – untrained singers, TS – trained singers

Vowel	NS [cents]			US [cents]			TS [cents]		
	Median	Min	Max	Median	Min	Max	Median	Min	Max
a	22.93	144.99	27.66	40.42	64.75	26.76	10.36	17.23	21.73
e	18.89	50.35	28.01	40.93	64.67	30.04	15.78	21.48	21.93
i	28.33	154.86	27.79	23.42	47.75	22.86	16.31	23.81	31.39
o	28.25	144.25	24.79	32.85	56.51	26.76	15.47	20.13	24.03
u	28.22	109.44	167.29	21.87	46.23	53.92	23.42	30.04	33.81

Table 3. Linear regression analysis results from crescendo sounds.  
NS – nonsingers, US – untrained singers, TS – trained singers

NS			US			TS		
Median	Min	Max	Median	Min	Max	Median	Min	Max
0.160	0.123	0.615	0.064	0.137	0.279	0.001	0.011	0.019

## 5. Conclusions

We confirmed that there are parameters like pitch error, jitter, linear regression coefficients, and vibrato that differ significantly among nonsingers, untrained and trained singers. Low values of linear regression coefficients for trained and untrained voices indicates the ability to keep the pitch constant during a crescendo, which is the pitch stability and the voice quality indicator.

Choosing the promising parameters is the first step to create singers voice quality system. In the future, we will be looking for other ones to perform the classification of the advanced level of singing by, e.g., HMM, ANN, SVM.

## References

1. L. Enflo, J. Sundberg, A. McAllister, *Collision and Phonation Threshold Pressures Before and After Loud, Prolonged Vocalization in Trained and Untrained Voices*, *Journal of Voice*, **27** (September 2013) 527 – 530.
2. I. R. Titze, *Phonation threshold pressure: a missing link in glottal aerodynamics*, *The Journal of the Acoustical Society of America*, **91** (May 1992) 2926 – 35.
3. I. R. Titze, J. Sundberg, *Vocal intensity in speakers and singers*, *The Journal of the Acoustical Society of America*, **91** (May 1992) 2936 – 2946.
4. J. Sundberg, *Flow Glottogram and Subglottal Pressure Relationship in Singers and Untrained Voices*, *Journal of Voice*, **32** (January 2018) 23 – 31.
5. N. Henrich, B. Roubeau, M. Castellengo, *On the use of electroglottography for characterisation of the laryngeal mechanisms*, *Proceedings of the Stockholm Music Acoustics Conference*, August 6–9, 2003.
6. N. Henrich, C. D'Alessandro, B. Doval, M. Castellengo, *Glottal open quotient in singing: measurements and correlation with laryngeal mechanisms, vocal intensity, and fundamental frequency*, *The Journal of the Acoustical Society of America*, **117** (March 2005) 1417 – 30.

7. E. Pórolniczak, M. Kramarczyk, *Formant analysis in the assessment of the quality of choral singers*, Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), (September 2013) 200 – 204.
8. J. Sundberg, *Articulatory interpretation of the “singing formant”*, The Journal of the Acoustical Society of America, **55** (April 1974) 838 – 844.
9. T. Murry, *Pitch-matching accuracy in singers and nonsingers*, Journal of Voice, **4** (January 1990) 317 – 321.
10. T. Nakano, M. Goto, Y. Hiraga, *An automatic singing skill evaluation method for unknown melodies using pitch interval accuracy and vibrato features*, INTERSPEECH 2006, April 2006.
11. E. Pórolniczak, M. Łazoryszczak, *Quality assessment of intonation of choir singers using  $f_0$  and trend lines for singing sequence*, Metody Informatyki Stosowanej, **4** (2011) 259 – 268.
12. J. Sundberg, *Intonation in singing*, The Oxford Handbook of Singing, chapter 13, April 2019.
13. S. Ternstrom, J. Sundberg, *Intonation precision of choir singers*, Journal of the Acoustical Society of America, **84** 59 – 69.
14. A. Grell, J. Sundberg, S. Ternstrom, M. Ptak, E. Altenmuller, *Rapid pitch correction in choir singers*, Journal of the Acoustical Society of America, **126**(1) (2009) 407 – 413.
15. J. Dai, S. Dixon, *Intonation trajectories within tones in unaccompanied soprano, alto, tenor, bass quartet singing*, Journal of the Acoustical Society of America, **146** (August 2019).
16. A. Vurma, J. Ross, *Production and perception of musical intervals*, Music Perception, **23** 331 – 344.
17. P. Gramming, J. Sundberg, S. Ternstrom, R. Leanderon, W. Perkins, *Relationship between changes in voice pitch and loudness*, Journal of Voice, (1988) 118 – 126.
18. K. Deepak, *Epoch extraction using zero band filtering from speech signal*, Circuits Systems and Signal Processing, **34** (2014).
19. E. Prame, *Measurements of the vibrato rate of ten singers*, The Journal of the Acoustical Society of America, **96** (October 1994) 1979 – 1984.

## **Determination of Sound Power Level by Using a Microphone Array and Conventional Methods**

Grzegorz BOGUSŁAWSKI

*Lodz University of Lodz, 266 Piotrkowska Street, 90-924 Lodz, Poland,  
grzegorz.boguslawski@p.lodz.pl*

Joanna Maria KOPANIA

*Lodz University of Lodz, 266 Piotrkowska Street, 90-924 Lodz, Poland,  
joanna.kopania@p.lodz.pl*

Patryk GAJ

*Institute of Power Engineering - Thermal Technology Branch "ITC" in Lodz,  
113 Dabrowskiego Street, 93-208 Lodz, Poland, patryk.gaj@itc.edu.pl*

Kamil WÓJCIAK

*Institute of Power Engineering - Thermal Technology Branch "ITC" in Lodz,  
113 Dabrowskiego Street, 93-208 Lodz, Poland, kamil.wojciak@itc.edu.pl*

### **Abstract**

Sound power is measured to make objective comparisons between the same type of products, but also because legislation requires it. To release a new product, it is often compulsory to certify it according to International Organization for Standardization (ISO) standards, and also with national and local regulations. Determining of sound power is not a straightforward process. Sound power can be determined through the measurement of sound pressure (series 3740 methods) or sound intensity (series 9614 methods). Selecting one of the above methods depends on the purpose of the test, as well as the available equipment, desired grade of accuracy, background noise level or the test environment. Nowadays the additional methods, such as microphone arrays are used to located of the noise source and determined of pressure sound level. But the results obtained with acoustic cameras cannot be, for now, used for legislative purposes (are not ISO compliant). In this work the differences in the determination of sound power level by using conventional ISO methods and microphone arrays are determined. The system composed of a loudspeaker and a fan were used as a sound source of the noise. Sound power levels according to ISO 3746 and ISO 9614-1 were determined and were compared with the developed method by using microphone arrays techniques.

**Keywords:** sound power level, ISO methods, microphone array, intensity probe

### **1. Introduction**

Sound power quantity used in order to describe the noise emitted by sources is independent of the acoustic surroundings. The value of sound power is reproducible for any test conditions and is, therefore, an excellent indicator for comparing noise sources. Sound power measurement is employed in order to facilitate machinery noise reduction and to determine whether the operation of a machine is consistent with noise legislation and standards. Machines and equipment must be designed and manufactured in such a way that the risk of noise hazard is as low as possible and, additionally, in European Union the "Noise and machinery directive 2006/42/EC" requires that noise levels must

be determined by measurement, and be declared in the product documentations [1]. The aim of the directive is that users and buyers must be able to assess and compare machines on all aspects of health and safety, including noise [2-6].

The three different qualities describe the sound: sound pressure level, sound intensity and sound power level. Sound power level can be calculated through sound pressure or sound intensity measurements. Sound power level is consistent, comparable, and more practical for noise control measures. There are various methods, having different accuracy levels, in different acoustic fields. Generally, the preferred method is the method in which sound pressure level is measured over an enveloping surface, surrounding the noise source hypothetically, and then sound power level is calculated by using surface averaged sound pressure level and measurement surface area. The selection between pressure and intensity measurements is based on numerous factors such as the acoustic nature of the environment, ease of application, access to measurement instruments, and speed of the experiment.

To determine the sound power level of machinery or equipment, based on the measurements of sound pressure, the two most important series of standards are ISO 3740 and ISO 11200 are used [6, 7]. The ISO 3740 series describes the methods of measuring noise emissions from machinery in terms of sound power level, both A-weighted and in frequency bands. The ISO 11200 series describes methods of measuring emission sound pressure levels in a controlled acoustical environment. The sound intensity methods for determining of sound power level are standardized in ISO 9614 methods [8]. These methods can be used in-situ in almost any acoustic environment. The sound intensity method is best suited for research and development, for engineering testing and is used for engineering measurements in the development of a new product, and also in-situ measurements.

Nowadays, because of the development of Industry 4.0, where automated processes are used, the system of data collection and transfer from any location should be used. Thanks to this many data analysis, procedures can be used to create shop floor schedules or prevention system in industry process [9, 10]. But many manufacturing situations have got dynamic nature, due to the movement of material. Dynamic processes are harder to analyze due to the change in mass and shape of the systems. This makes real-time analysis difficult. The development of new measured techniques like microphone array or vibrations detectors makes monitoring of these process is much easier and faster.

In this work the differences in the determination of sound power level by using conventional methods according to ISO 3746 and ISO 9614-1 and microphone arrays (acoustic camera) are determined. The system composed of a loudspeaker and a fan were used as a sound source of the noise. The calculated values of sound power level and standard uncertainty of the mean are compared and discussed.

## 2. Experimental method

The measurements were performed on the specially constructed test stand at the Institute of Power Energy in Lodz – Figure 1. The anechoic test chamber is cubic, approximately 350 m<sup>3</sup> in size and has walls that are acoustically treated with foam wedges providing



a reflection-free environment. The system composed of a loudspeaker and a fan were used as a sound source of noise – Figure 1.



Figure 1. System of noise source composed from loudspeaker and fan, and measurement with sound analyser (on the left), intensity probe (in the centre) and acoustic camera (on the right) are being made

Three types of equipments were used in these studies:

1. sound analyzer as a sound meter (SVAN 958A with preamplifier and measurement microphone MK 255). The microphone was calibrated before commencing the acoustic test;
2. intensity probe with the configuration of microphones face to face and 12 mm solid cylindrical spacer separate, connected with analyzer and post-processing system (type 3545 with analyzer 2241). The probe was calibrated in sound intensity calibrator type 3541;
3. acoustic camera (Noise Inspector), with 40 MEMS microphones and HD camera and software which allow real-time sound imaging for quick results. The microphones were not calibrated before the measure because of due to the construction they do not require frequent calibration.

The first two measuring devices were used for testing the standard ISO methods (conventional). These methods, PN EN ISO 3746 and PN EN ISO 9614-1, specify a measuring the overall weighted sound pressure levels at prescribed microphone positions around the noise source. The A-weighted sound power level is calculated from these measured values. The minimum number of the key microphone positions over "box measurement surface" depends on dimensions of the noise source, measurement distance and also on the type of partial grids, dividing the measurement surface [11]. In this study, eight key microphone positions at 0.6 m distance were used in accordance, over the parallelepiped measuring surface (but for the acoustic camera because of its technical limitations, the distance of 1m was used). In each eight key microphone positions, the three measurements were done by using each measurement equipment. The measurement surface, enclosing the source under test was defined, as is seen in Figure 2.

The third equipment used in this study was an acoustic camera. Microphone array (beamforming) seem to be the most promising techniques to separate different sources. The beamforming technique starts from the microphone signals, perceived by some transducers fixed to a frame, whose geometry has to be properly designed, and reconstructs the sound field in each location within a pre-defined region of space.

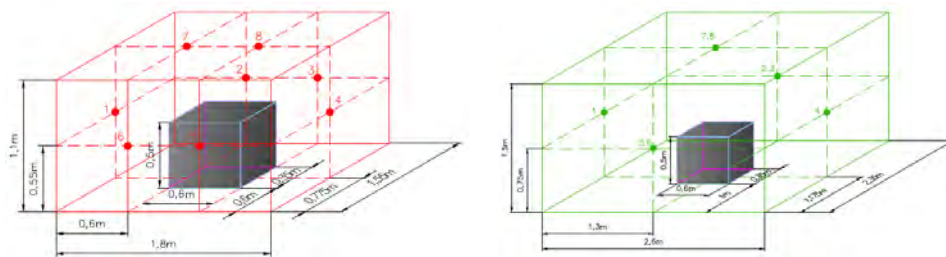


Figure 2. Scheme of measurement points for sound analyser and intensity probe (on the left) and for acoustic camera (on the right)

Literature gives many methods to process array measurements in order to identify sound sources. The oldest, but very simple, is the Delay & Sum Technique [12]. The “delay and sum” beamformer takes a set of the time difference of arrival which determine where the beamformer is steered and computes the outputs  $s_B(t)$  as (1):

$$s_B(t) = \frac{1}{K} \sum_{m=1}^K s_m(t + \tau_{ml}) \quad (1)$$

where  $l$  is a reference microphone which can be chosen to be the closest microphone to the sound source so that all  $\tau_{ml}$  are negative and the beamformer is causal. To steer the beamformer, one selects TDOAs (time difference of arrival) corresponding to a known source location. Noise from other directions will add incoherently, and decrease by a factor of  $K^{-1}$  relative to the source signal which adds up coherently, and the beamformed signal is clear [13].

The images from the acoustic camera were done in the same planes as in conventional methods. From each plane, in the strongest signal (red area – in maximum values) the 1/3 octave spectrum was calculated by the chosen algorithm, and thanks to that, the single global (for all plane) and local (the same points like used in conventional methods) values of sound pressure were calculated. Signal not processed by algorithm Delay & Sum image and processed by this algorithm from the acoustic camera is presented in Figure 3 and Figure 4.

### 3. Determination of sound power

The purpose of the tests was determining sound power level the source of noise, composed of two individual sound sources. The sound sources were placed on the floor. The setting the loudspeaker and fan simulated a machine as an object with a clear directional sound character, which in industry area is common. The average values

sound pressure level in specific points determined according to ISO methods and for the acoustic camera in the measurement planes are presented in Table 2.

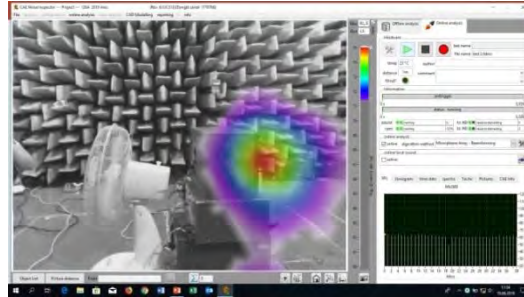


Figure 3. Non-processed image from acoustic camera for point 2 and 3

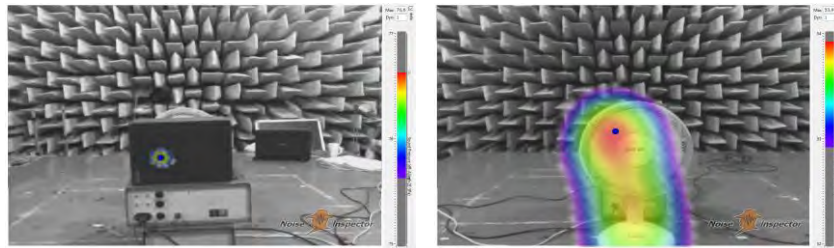


Figure 4. Calculated image by Delay and Sum algorithm from acoustic camera for point 1 and 4

Sound analyzer SVAN 958 used in these studies, is treated as a standard method, because of this equipment is under metrological supervision. There are differences between the values of sound pressure level (SPL) obtained by different methods. It is seen from Table 1, that in point 1 the highest value of SPL is from SVAN 958 (83.9 dB), and the lowest from local signals of the acoustic camera (66.8 dB). Similar values of SPL are between obtained from SVAN958 (83.9 dB), intensity probe (80.7 dB) and global signal of acoustic camera (79.1 dB). Values obtained from local points of the acoustic camera (corresponding to measure points of conventional methods) are too lower and should not to be taken into account.

The main goal of this work was the determination of the sound power level. Sound power is the rate per unit time at which airborne sound energy is radiated by a source, and its unit is Watt [W]. The sound power levels  $L_W$  is defined as (2):

$$L_W = 10 \log_{10} \left( \frac{W}{W_0} \right), \text{dB} \quad (2)$$

where the reference sound power  $W_0$  is 1 pW.

Table 1. Average sound pressure in selected points got by using different methods

Measurement point	Intensity probe $L_{pA}$	Sound analyser SVAN 958 $L_{pA}$	Acoustic camera - $L_{pA}$ in maximum (red area)	Acoustic camera - $L_{pA}$ global	Acoustic camera - $L_{pA}$ local
1	80.7	83.9	76.7	79.1	66.8
2	68.7	67.4	58.8	61.4	53.5
3	66.1	60.2			48.7
4	66.3	60.0	53.9	57.3	52.8
5	66.9	63.2	61.2	63.7	50.5
6	70.6	69.5			58.3
7	70.9	71.0	63.8	67.2	63.8
8	67.0	64.3			55.0

Sound power can only be calculated or determined based upon either sound pressure measurement or sound intensity measurement. So, given the measurement surface according to Figure 2 for sound analyser and intensity probe and for the acoustic camera, also the environmental correction and measurement conditions the sound power were calculated for each method and additionally type A uncertainties were estimated. The type A uncertainty was calculated as the standard uncertainty of the mean due to sampling, according to point D.4.2.9 of standard ISO 3746:

$$u_{mean} = \frac{s}{\sqrt{N_M}} = \frac{1}{\sqrt{N_M}} \sqrt{\frac{1}{(N_M-1)} \sum_{j=1}^{N_M} (L'_{pi(ST)} - L'_{pav})^2} \quad (3)$$

where:

$N_M$  – number of microphone positions,  $L'_{pav}$  – means of measured values  $L'_{pi(ST)}$ .

The result of these calculations is represented in Table 2.

Table 2. Sound power level and standard uncertainty of the mean got by use different methods

Parametr	Intensity probe	Sound analyser SVAN 958	Acoustic camera in maximum (red area)	Acoustic camera - global	Acoustic camera - local
Sound power $L_{WA}$ , dB	83.3	84.1	83.8	85.8	72.7
Standard uncertainty of the mean, $u_{mean}$ , dB	1.7	2.7	3.8	3.7	2.3

When we consider the values of sound power obtained from different methods it is clean that four results can be taken into account: from SVAN 958, intensity probe, the maximum signal of acoustic camera and global signal of acoustic camera, where differences around 2.5 dB seem acceptable. The sound power from local signals (corresponding to measure points of conventional methods) of the acoustic camera is too

lower, so in the case of the acoustic camera do not use the same reference points (measuring) as in the conventional methods (ISO).

The second aspect of tested methods was determining of standard uncertainty of the mean. In our study, the standard uncertainty of the mean for sound power level is the lowest for intensity probe but the highest for the maximum signal of the acoustic camera. That may be related to the precision of place the selected point on the image from the acoustic camera, in which the 1/3 octave spectrum is calculated necessary to calculate sound pressure level in this point and finally to determine sound power level. Relationship between standard uncertainty of the mean on the values of sound power and apply measurement methods is presented in Figure 4. The value obtained by the acoustic camera seems too high and requires further study and analysis.

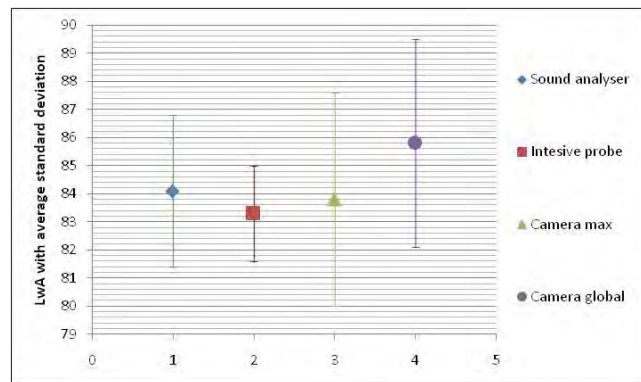


Figure 5. Relationship between standard uncertainty of the mean on the values of sound power and apply measurement methods

#### 4. Conclusions

The three methods were used in these studies: sound analyzer as a sound meter, intensity probe with the analyzer and acoustic camera. Obtained results showed a quite good agreement between using methods. Difference between using the sound analyzer or intensity probe or acoustic camera depends on the characteristics of the source, the characteristics of the environment and distance of the measurement positions and the source. For the sound pressure methods, many restrictions have to be made for source characteristic and measurement environment in order to make the determination of uncertainty the sound power level within acceptable limits. Measurements of sound intensity are now fairly routinely made to determine the sound power of sources, machinery noise source identification and the transmission loss (sound reduction index) of structures. But acoustic camera seen nowadays as possibly the most progressive method might be also used to determine the sound power level. The benefits of this non-contact method can be used at a variety of machinery and equipment, even at greater distances. But for faster identification of faults or its prediction, it is necessary to create a database of samples with frequency characteristics of individual faults, which in the future may reduce costs and speed up this fault detection technology.

## References

1. Directive 2006/42/EC of the European Parliament and of the Council of 17 May 2006 on machinery, and amending Directive 95/16/EC (recast).
2. R. F. Barron, *Industrial noise control and acoustics*, Marcel Dekker Inc, New York, 2001.
3. A. D. Bies, C. H. Hansen, *Engineering Noise Control. Theory and Practice*, Fourth Edition, Spon Press of Taylor & Francis Group, New York, 2009.
4. M. Aliabadi, R. Golmohammadi, A. Ohadi, *Empirical comparison of the in situ methods for determining sound power of typical embroidery machine located in industrial workroom*, International Journal of Occupational Hygiene, **5**(3) (2013) 89 – 95.
5. C. Tomozei, A. Astofli, V. Nedeff, G. Lazar, *Noise sources characterization inside and outside factory*, Environmental Engineering and Management Journal, **3** (2012) 701 – 708.
6. International Organization for Standardization, *Engineering method in essentially free-field over a reflective plane* (ISO 3744:2011); *Precision method in anechoic or semi-anechoic room* (ISO 3745:2012); *Survey method with no special test environment* (ISO 3746:2010); *Engineering or survey method in situ in a reverberant environment* (ISO 3747:2010).
7. International Organization for Standardization, *Acoustics -- Noise emitted by machinery and equipment -- Guidelines for the use of basic standards for the determination of emission sound pressure levels at a work station and at other specified positions* (ISO 11200:2014).
8. International Organization for Standardization, *Precision, engineering or survey discrete measurement points method* (ISO 9614-1:2010); *Engineering or survey scanning measurement method* (ISO 9614-2:1996); *Precision scanning measurement method* (ISO 9614-3:2002).
9. W. Ji, L. Wang, *Big data analytics based fault prediction for shop floor scheduling*. J Manuf. Syst., **43** (2017) 187 – 94.
10. D. Pleban, *Method of acoustic assessment of machinery based on global acoustic quality index*, Arch Acoust, **35** (2010) 223 – 235.
11. E. Bilgic, H. Mutaf, C. Kirbas, E. Sadikoglu, *Evaluation of uncertainty contributions of measurement surface and number of microphone positions in determination of sound power levels*, Special issue of 3<sup>rd</sup> International Conference on Computational and Experimental Science and Engineering (ICCESEN 2016), Acta Physica Polonica, **132** (2017) 642 – 649.
12. Don H. Johnson, Dan E. Dudgeon, *Array Signal Processing: Concepts and Techniques*, Prentice Hall Signal Processing Series, 1993.
13. M. S. Brandstein, D. B. Ward (editors), *Micro-phone Arrays: Signal Processing Techniques and Applications*, Springer-Verlag, Berlin, Germany, 2001.

## Tuning the Selected Acoustic Helicoidal Resonator with a Short Flat Bar – Numerical Analysis

Wojciech ŁAPKA

*Poznan University of Technology, Faculty of Mechanical Engineering and Management,  
Institute of Applied Mechanics, Jan Pawel II 24 Street, 61-139 Poznan,  
wojciech.lapka@put.poznan.pl*

### Abstract

This paper describes the possible way of tuning the selected acoustical helicoidal resonator placed in straight cylindrical duct by the use of a short flat bar. The acoustic attenuation performance (transmission loss) of helicoidal resonator has emphatically changed with the change of the length and the degree of rotation of a short flat bar placed close to the resonator. The finite element numerical calculations of the acoustical systems were made in COMSOL Multiphysics computational environment. The results show that the change of length and rotation of short flat bar can widely change the resonance frequencies of helicoidal resonator. So in this work were presented the possible simple tuning options for the acoustic helicoidal resonator applied in ducted systems.

**Keywords:** acoustic helicoidal resonator, tuning, sound propagation, flat bar, ducts and mufflers

### 1. Introduction

The helicoidal acoustic resonator belongs to the family of mufflers and resonators [1]. Its work is similar to the Helmholtz or a quarter-wave resonator. The acoustic resonance is achieved here by placing the helicoidal profile with proper dimensions inside a straight cylindrical duct [2, 4, 8-10]. There are a lot of different researches of this resonator. The most important works describe the experiments of using this device. The interactions between the helicoidal acoustical resonator and the dissipative silencer are presented in work [3]. The article [5] describes the 90-degree duct elbow with the presented resonator. In the article [7], the authors wrote about a simple comparison to Helmholtz resonator. The behavior of resonator when the helicoidal profile made from elastic material is presented in paper [6]. The last researches about interaction with 90-degree duct elbow [5] showed the possible way of tuning the helicoidal resonator by its rotation in the duct axis. The researches presented in this work show the new way of tuning the resonator. The mentioned method is basing on placing the short flat bar inside the resonator.

### 2. Investigated acoustical system and numerical model

Similarly to the previous researches [2, 5, 8] investigated acoustical system consist of a cylindrical duct with the diameter  $d = 2a = 125$  mm, where  $a$  is a radius of the duct, with helicoidal resonator inside, which has the constant relationship  $s/d = 1.976$ , between helicoidal pitch  $s$  to the cylindrical duct diameter  $d$ , and the number of helicoidal turns  $n$  equals 0.671. The helicoidal profile thickness

$g$  to the cylindrical duct diameter  $d$ , ratio  $g/d$  equals 0.04, and the value of relationship between mandrel diameter  $d_i$  and cylindrical duct diameter  $d$ , ratio  $d_i/d$  equals 0.24. The short flat bar thickness  $g_b$  has the same value as the thickness of helicoidal profile  $g$ , ratio  $g_b/d = 0.04$ , and its width  $w_b$  had the same value as radius  $a$  of the duct. The four radial positions of a short flat bar were considered as presented in Figure 1. The change in length of a short flat bar  $l_b$  was considered in the range from  $0.16a$  to  $1.6a$  with the step of  $0.16a$ . The 0 degree position consisted in placing the beginning of the helicoidal profile directly at the beginning of the short flat bar.

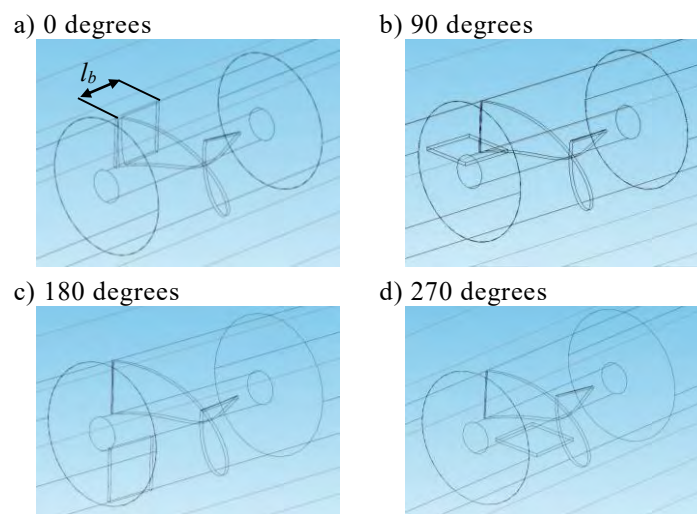


Figure 1. Visualisation of four radial positions of a short flat bar as a tuning element inside cylindrical duct of inner diameter  $d = 2a = 125$  mm ( $a$ —radius) with an acoustic helicoidal resonator inside with ratio  $s/d = 1.976$  and number of turns  $n = 0.671$ ,  $l_b$  is the length of a short flat bar

The COMSOL Multiphysics software was used to solve the Helmholtz equation [11] for the three dimensional models by the use of finite element method (FEM). The following parameters were used:

- type of medium - air with temperature 20°C (speed of sound in air  $c = 343$  m/s) without mean flow,
- computations in the frequency domain with the use of time-harmonic pressure acoustics physics,
- defined boundary conditions: hard walls with perfect reflection defined for all surface elements of helicoidal resonator, cylindrical duct walls and short flat bar walls; plane wave radiation defined for inlet (incident pressure  $p = 1$  Pa) and outlet circular surfaces (anechoic termination) of the cylindrical duct,
- modeled finite element mesh was automatically generated with the rule of minimum five finite elements per wavelength [11] for the highest analyzed frequency.



The wave number  $k = 2\pi f/c$  and Helmholtz number  $ka$  was used to present the non-dimensional frequency [12, 13]. The Transmission Loss ( $TL$ ) [2, 11, 12] was calculated for the most interesting frequency range  $f$  [Hz] from 800 Hz ( $ka = 0.916$ ) to 1.6 kHz ( $ka = 1.83$ ) with the calculation step of 1 Hz.

### 3. Acoustical results – transmission loss

In Figure 2 to Figure 5 are presented the transmission losses of helicoidal resonator with respectively 0, 90, 180 and 270 degrees of radial position of a short flat bar with the length  $l_b$  range from  $0.16a$  to  $1.6a$  with the step of  $0.16a$ .

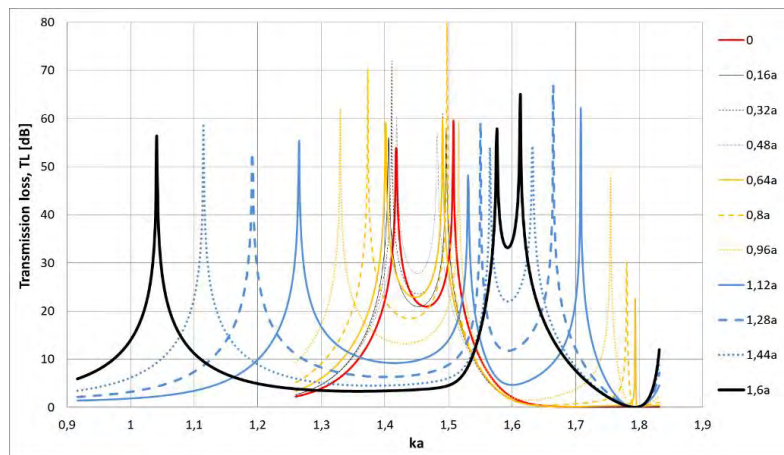


Figure 2. Transmission loss of helicoidal resonator for 0 degree of radial position of a short flat bar with the length  $l_b$  from  $0.16a$  to  $1.6a$  and the step of  $0.16a$

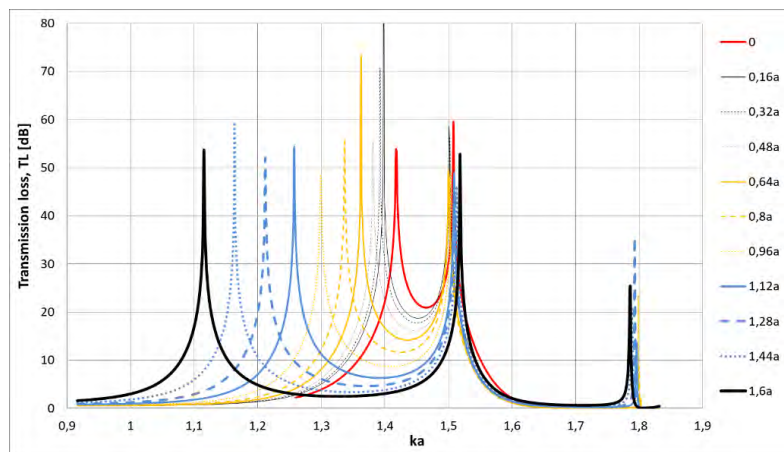


Figure 3. Transmission loss  $TL$  of helicoidal resonator for 90 degree of radial position of a short flat bar with the length  $l_b$  range from  $0.16a$  to  $1.6a$  and the step of  $0.16a$

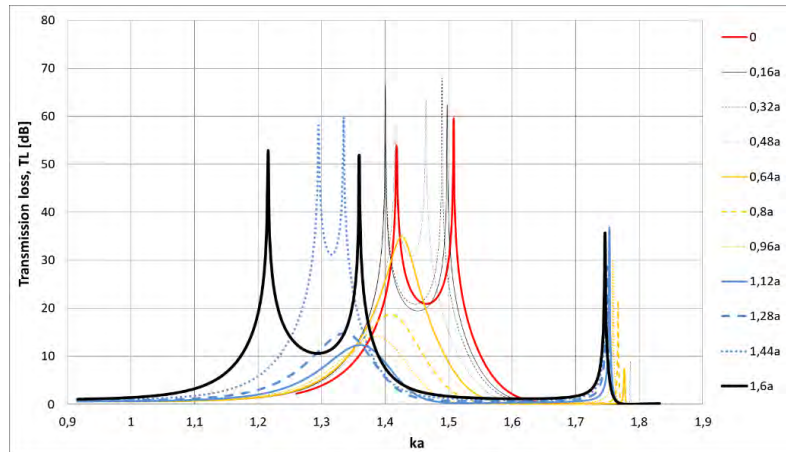


Figure 4. Transmission loss of helicoidal resonator for 180 degree of radial position of a short flat bar with the length  $l_b$  range from  $0.16a$  to  $1.6a$  and the step of  $0.16a$

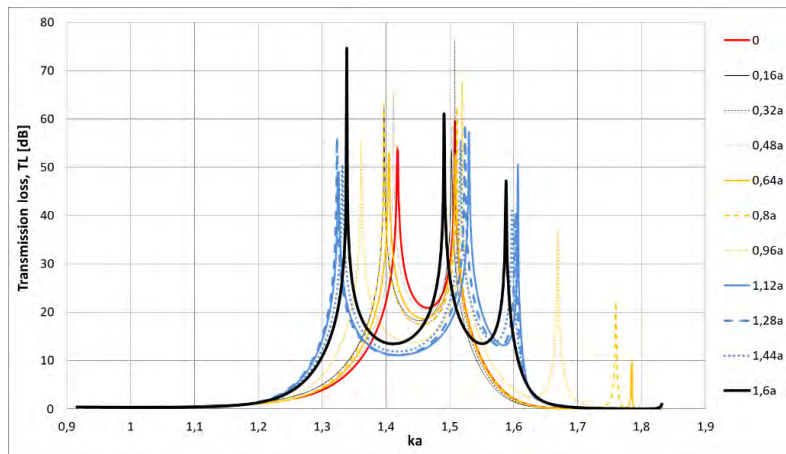


Figure 5. Transmission loss of helicoidal resonator for 270 degree of radial position of a short flat bar with the length  $l_b$  range from  $0.16a$  to  $1.6a$  and the step of  $0.16a$

Considered acoustical systems with a four radial positions of a short flat bar with different length placed at the beginning of the acoustic helicoidal resonator inside cylindrical duct, show that the widest range of Helmholtz number  $ka$  and so on the tuning possibilities for investigated helicoidal resonator was obtained for 0 degree of radial position of a short flat bar. In this case the highest values of narrowband sound attenuation can be obtained for Helmholtz number from  $ka = 1.04$  to  $ka = 1.76$ . The smallest change of sound attenuation in  $ka$  range gives the 270 degree position of a short flat bar, from  $ka = 1.32$  to  $ka = 1.61$ . But nearly every kind of obtained transmission loss characteristics are very interesting and tends to be implementable.

For the 0 degrees of radial position the higher change in the length  $l_b$  of a short flat bar slightly changes the frequency of two resonances of investigated helicoidal resonator for  $l_b$  range from  $0.16a$  to  $0.48a$ , and then transforms into three resonance  $TL$  characteristics for the length  $l_b$  from  $0.64a$  to  $1.6a$ . All received transmission loss characteristics are saddled with the highest saddle level  $TL = 33$  dB between second and third resonance for the longest flat bar with  $l_b = 1.6a$ , and the lowest saddle level  $TL = 3.3$  dB between first and second resonance for the same flat bar with  $l_b = 1.6a$ .

For the 90 degrees of radial position the higher change in length  $l_b$  of a short flat bar mainly changes the frequency of a first resonance of investigated helicoidal resonator, and the second resonance stays almost in the same value of  $ka$  – the inconsiderable change in the range from  $ka = 1.5$  to  $ka = 1.52$ . All received transmission loss characteristics have two resonances and all are saddled with the highest saddle level  $TL = 18$  dB for the shortest flat bar ( $l_b = 0.16a$ ), and the lowest saddle level  $TL = 2.5$  dB for the longest flat bar ( $l_b = 1.6a$ ).

Very different and interesting results were obtained for the 180 degrees of radial position of a short flat bar. For the length  $l_b$  from  $0.16a$  to  $0.48a$  the  $TL$  characteristics are saddled and consists of two resonances, which are getting closer to each other, and then for  $l_b = 0.64a$  the highly resonant acoustical systems attenuation decreases and transforms into one resonance characteristics with lower  $TL$  levels for  $l_b$  range from  $0.8a$  to  $1.28a$ . For the last analysed two lengths  $l_b$  of the short flat bar, which equal  $1.44a$  and  $1.6a$ , the  $TL$  characteristics have three resonances, where the first and second resonance have high  $TL$  level between  $ka = 1.21$  and  $ka = 1.36$  with the  $TL$  saddle between them, and the third resonance appears for the Helmholtz number  $ka$  that equals about 1.75.

The 270 degrees of radial position of a short flat bar with different length  $l_b$  gives the smallest change in sound attenuation in  $ka$  domain. For the range of lengths from  $0.16a$  to  $0.64a$  the  $TL$  characteristics doesn't change very much, first resonance changes in the range of  $ka$  from 1.39 to 1.42, and the second resonance changes in the range of  $ka$  from 1.5 to 1.51. For the range of lengths from  $0.8a$  to  $1.6a$  the  $TL$  characteristics have clearly visible three resonances in the range of Helmholtz numbers  $ka$  from 1.32 to 1.76. The interesting  $TL$  characteristics were obtained for short bar lengths  $l_b$  from  $1.12a$  to  $1.6a$ , where the lowest  $TL$  levels in the saddle obtain higher than 10 dB values on almost same level for both saddles between first and second resonance and a second and third resonance.

#### 4. Conclusions

Presented paper described the numerical analysis of the acoustic system with helicoidal resonator inside a cylindrical duct with a short flat bar, as a tuning element. Undertaken research work showed that the widest tuning possibilities for investigated helicoidal resonator could be obtained for the 0 degree of radial position of a short flat bar. For this case the highest values of narrowband sound attenuation were obtained for Helmholtz number  $ka$  from 1.04 to 1.76. The smallest change of  $ka$  gives the 270 degree position of

a short flat bar, from  $ka = 1.32$  to  $ka = 1.61$ . But nearly every kind of obtained transmission loss characteristics are very interesting and tends to be implementable.

### Acknowledgments

Author gratefully acknowledges that the study was financed by the Ministry of Science and Higher Education as a project 02/21/DSPB/3546 and 02/21/SBAD/3558.

### References

1. M. H. Khairuddin, M. F. M. Said, A. A. Dahlan, K. A. Kadir, *Review on Resonator and Muffler Configuration Acoustics*, Archives of Acoustics, **43**(3) (2018) 369 – 384.
2. W. Łapka, *Transmission loss and pressure drop of selected range of helicoidal resonators*, Vibrations in Physical Systems, **26** (2014) 121 – 128.
3. W. Łapka, *Acoustic attenuation performance of a round silencer with the spiral duct at the inlet*, Archives of Acoustics, **32**(4) (Supplement) (2007) 247 – 252.
4. W. Łapka, *Insertion loss of spiral ducts - measurements and computations*, Archives of Acoustics, **34**(4) (2009) 537 – 545.
5. W. Łapka, *Numerical Analysis of Sound Propagation in Selected Acoustical System with Helicoidal Resonator Placed in Cylindrical Duct with 90 Degree Elbow*, Vibrations in Physical Systems, **29** (2018) 2018018-1 – 2018018-6.
6. W. Łapka, *Numerical Study of Acoustic-Structure Interaction of Selected Helicoidal Resonator with Flexible Helicoidal Profile*, Archives of Acoustics, **43**(1) (2018) 83 – 92.
7. W. Łapka, C. Cempel, *Acoustic attenuation performance of Helmholtz resonator and spiral duct*, Vibrations in Physical Systems, **23** (2008) 247 – 252.
8. W. Łapka, C. Cempel, *Computational and experimental investigations of a sound pressure level distribution at the outlet of the spiral duct*, Archives of Acoustics, **33**(4) (Supplement) (2008) 65 – 70.
9. W. Łapka, C. Cempel, *Noise reduction of spiral ducts*, International Journal of Occupational Safety and Ergonomics (JOSE), **13**(4) (2007) 419 – 426.
10. W. Łapka, C. Cempel, *Acoustic filter for sound attenuation in ducted systems*, Polish Patent, PAT.216176, Date of publication 31.03.2014.
11. S. Marburg, B. Nolte, *Computational Acoustics of Noise Propagation in Fluids – Finite and Boundary Element Methods*, 578, Springer-Verlag, Berlin, Germany, 2008.
12. M. L. Munjal, *Acoustics of Ducts and Mufflers with Application to Exhaust and Ventilation System Design*, Inc., Calgary, Canada, John Wiley & Sons, 1987.
13. A. Snakowska, *Teoria pola akustycznego zastosowana do badania układów o symetrii cylindrycznej*, ISBN 978-83-66016-08-8, Wydawnictwo AGH, Kraków, 2018.

## **Effectiveness of Acoustic Banners Depending on their Arrangement in the Concert Hall – Case Study**

Piotr Z. KOZŁOWSKI

*Wrocław University of Science and Technology, Department of Acoustics  
and Multimedia, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław,  
piotr.kozlowski@pwr.edu.pl*

### **Abstract**

One of the most frequently used methods of adjusting the room acoustics of concert halls is the temporary introduction of additional acoustic absorption in the form of acoustic banners. Banners are implemented in the form of fabric surfaces placed vertically or horizontally along the walls of concert halls. In practice, one or two layers of heavyweight fabrics are used, characterized by a high value of the sound absorption coefficient. Simplified methods of designing acoustic banners assume estimating the effectiveness of adjusting the reverberation time on the basis of the static theory. The paper presents the results of measuring the effectiveness of acoustic banners carried out in the concert hall. The possibility of tuning the interior acoustics was verified depending on the position of the banner in the room as well as the degree of its opening.

**Keywords:** room acoustics, adjusting room acoustics, acoustic banners, multipurpose venues

### **1. Introduction**

The necessity to hold varied events in one room leads to the need for adjusting to very different acoustic requirements [1-3] within a single room, depending on each event's particular characteristics. Experiences gathered when designing [7-9] and executing adjustable acoustics for venues located in Poland were used as well.

In the case of adjusting the acoustics of concert halls with relatively small cubic volume, there are problems related to the installation of acoustic banners with active surface sufficient for obtaining the expected level of the reverberation time adjustment capacity. Therefore, the decision was made to research the issue of the impact of the arrangement of banners on their effectiveness. This topic will be further explored in subsequent papers on the subject of methods of acoustic adjustment in multi-purpose venues.

The article is a case study based on results of measurements made in the recently designed concert hall at Mieczysław Karłowicz Primary & Secondary Music School in Katowice [6].

### **2. Methods of Adjusting Acoustics**

Some discussion about methods of adjusting acoustics was presented in the previous papers [4, 5]. Classic methods of acoustic adjustment are based, among others, on the well-known Sabine's equation (1).

$$RT = \frac{0.161 \times V}{A} = \frac{0.161 \times V}{S_1 \cdot \alpha_1 + S_2 \cdot \alpha_2 + \dots + S_n \cdot \alpha_n} \text{ [s]} \quad (1)$$

The reverberation time  $RT$  is directly proportional to the cubic volume of the room  $V$  and inversely proportional to the acoustic absorption  $A$  of the entire room. The acoustic absorption  $A$  of each of  $n$  surfaces/elements in the room is the product of the surface area  $S$  and the acoustic absorption coefficient  $\alpha$  of a given element. We can, therefore, directly influence the reverberation of the room by changing its cubic capacity (using reverberation chambers or lowering the acoustic ceiling), or introducing elements (rotating acoustic panels, acoustic roller banners) with a different acoustic absorption coefficient  $\alpha$ .

The most popular method of mechanical adjustment of the reverberation time is the introduction of textile curtains – acoustic roller banners – to the room [5]. The basic principle is to place in front of a surface with a small coefficient  $\alpha$  value a piece of fabric with a significantly larger  $\alpha$  value. Thus, the sound energy reflected from this surface is effectively decreased, which leads to decreasing reverberation energy in the room, and thus, directly decreasing the  $RT$  value.

### 3. Tested arrangement of acoustic banners

The study was conducted in the concert hall at Mieczysław Karłowicz Primary & Secondary Music School in Katowice, a typical shoebox-shaped space. In 2016, design documentation for a complete renovation was prepared [6] and used in 2017-2018 to perform a comprehensive overhaul. Dimensions and shape of the hall's external walls result from the space initially intended for the concert hall. They could not be modified during the renovation due to volume and construction limitations of the school building.

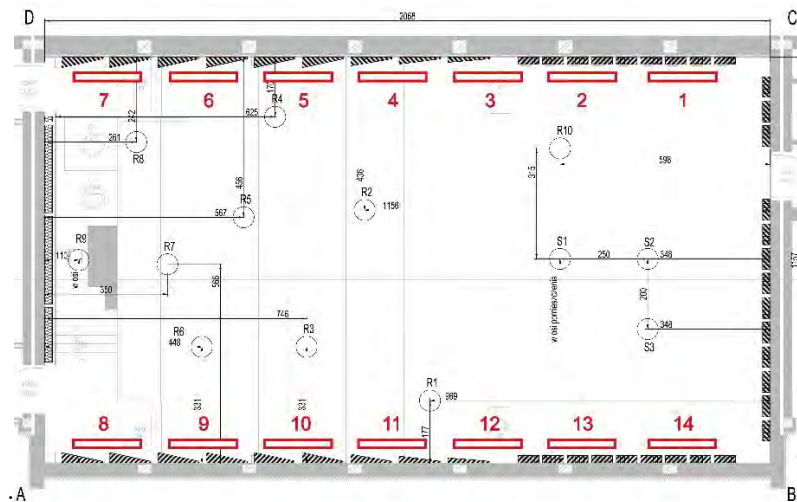


Figure 1. Plan of the measured concert hall. Placement of the acoustic treatment, acoustic banners (numbered in red) and measurement (source and receiver) points

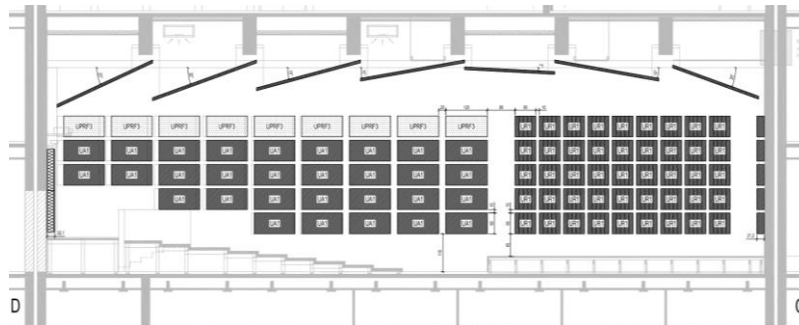


Figure 2. Section view of the left wall of the measured concert hall

Figure 1 features the plan of the room with the placement of the wall acoustic treatment, acoustic banners (numbered) and measurement points used in the acoustic study. The room is symmetric with respect to its main axis, except for the location of the entrance door to the stage and stairs leading to the sound engineer's station.

Above the stage and auditorium, reflective acoustic ceiling panels are placed (Figure 2). On the front (Figure 4) and side walls (Figures 1, 2) in the stage area, dispersing panels are located. Reflective and absorbing panels for the controlled scope of frequencies are located on side walls (Figure 2) in the auditorium area. The back wall (Figure 4) is entirely covered with broad-range absorption panels. The stage has the classic construction of a philharmonic floor with planks laid on wooden joists. The auditorium floor is slightly ascending. There is a sound engineer's *Front of House* station in front of the back wall, separated from the audience with a reflective barrier.

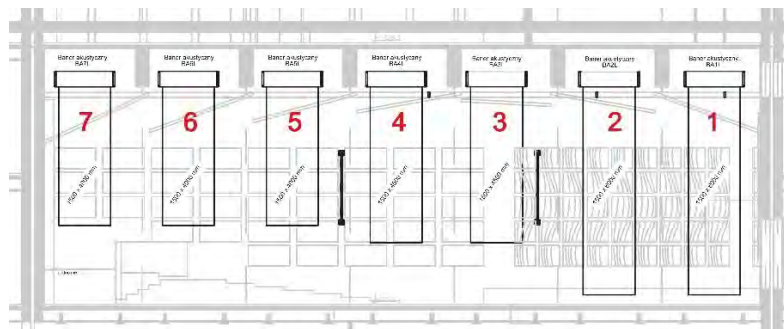


Figure 3. Acoustic banners' placement on the left wall of the measured concert hall

Acoustic banners are placed symmetrically along the side walls of the room (Figures 1, 3). Banner cases are placed above acoustic ceiling panels, and thus when completely closed, banners are not visible from the room, nor do they affect the room acoustics. The banners used are composed of a single layer of fabric, 150 cm wide, with the acoustic absorption class A. Figure 3 features banners opened to the degree established in the design documentation as the maximum opening which precludes collision with audience members walking the stairs. For situations described below as banners

opened at 100%, the lower edge of banners almost reached the floor of the stage or auditorium.

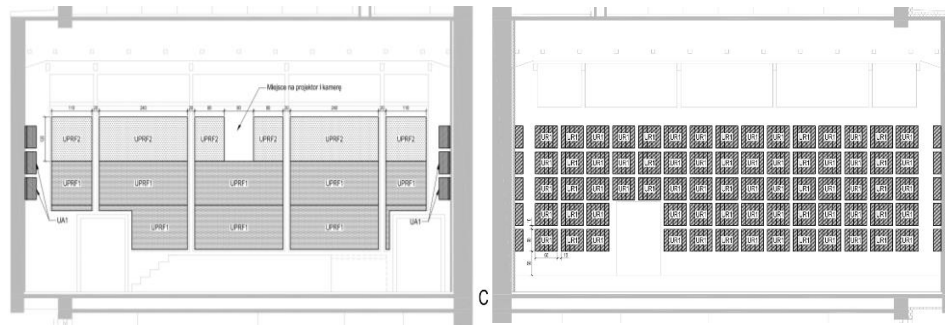


Figure 4. Section view of the back wall (left) and front wall (right) of the measured concert hall

During the study, 3 locations of measurement sources (S1–S3) were used, as well as 10 locations of measurement microphones (R1–R10), which amounts to 30 measurements in total for each studied situation (banner arrangement). The study used a multi-channel measurement system with a 12-sided sound source as well as omnidirectional measurement microphones compliant with the current standards of interior acoustics measurements, including the PN-EN ISO 3382-1:2009 norm. Measurement points were selected at the stage of designing the hall acoustics mainly due to their representative character in displaying acoustic properties of the room during modelling, and then during measurements.

#### 4. Measurement results for various arrangement of acoustic banners

The banner control system enables rolling out both individual banners and their groups independently, and to different degrees. A fully opened banner was marked as 100%, while the absorbent fabric completely hidden in its case was marked as 0%.

Figure 6 presents the results of the first series of measurements in which all banners were opened to the same degree. T20 reverberation time values were measured with:

- banners fully hidden (0%),
- 6 situations when banners were rolled out by additional 15% (15%, 30%, 45%, 60%, 75%, 90%),
- banners fully opened to 100%.

The analysis of results presented in Fig. 5 indicates that the smallest changes in the T20 reverberation time were obtained for 60%-100% situations. The greatest changes were obtained in the first phase of opening, i.e. 0%-45%. This allows for the conclusion that in the case of the studied hall, the first portion of the opened acoustic banners has a stronger impact on the resulting outcome of frequency characteristics of the T20 reverberation time than the last portion of these banners. It is probably due to the fact that the increase of acoustic absorption  $A$  is higher at the beginning of banner's opening phase than at the end of opening phase.



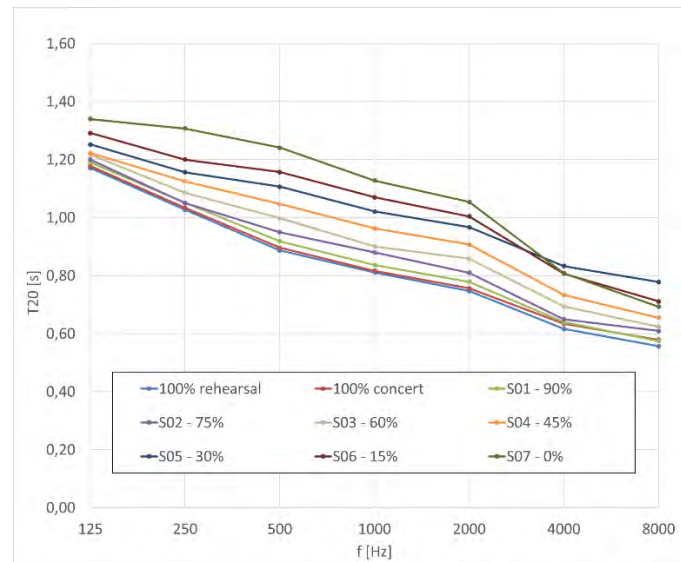


Figure 5. T20 in the function of banners' opening degree

This situation can be explained alternatively by the fact that lower portions of side walls, due to the close proximity of very absorbent audience seats, are less effective in creating the reverberation effect resulting from reflecting sound waves. This conclusion should be, of course, verified with research on a larger sample of concert halls, which has been included in the further scheduled studies on the topic.

Subsequent tests concerned the impact of the arrangement of acoustic banners in the room on the obtained reverberation time values. Within this test, selected groups of banners were being completely opened (100%), while others were fully closed (0%).

In the second series of measurements (Figure 6), banners were rolled out in groups of 2 on each side of the room. Therefore, in this series, there were always 4 banners (2 on each side) fully opened (100%) at any given time. The series of measurements started from the situation T11, when banners (1, 2, 14, 13) located on the stage were opened. Subsequent measurements were made for groups of opened banners increasingly distant from the stage (T12, T13, T14, T15). In the last tested situation T16, the last two banners (6, 7, 9, 8) on each side were opened.

The third series of measurements (Figure 7) was made similarly to the second one, only in this one groups of 3 banners were opened on each side, starting with the ones closest to the stage (situation T21 – banners 1, 2, 3, 14, 13, 12), ending in the back of the auditorium (situation T25 – banners 5, 6, 7, 10, 9, 8).

The fourth series of measurements (Figure 8) was made similarly to previous ones, only in this one groups of 4 banners were opened on each side, starting closest to the stage (situation T31 – banners 1, 2, 3, 4, 14, 13, 12, 11), and ending in the back of the auditorium (situation T34 – banners 4, 5, 6, 7, 11, 10, 9, 8).

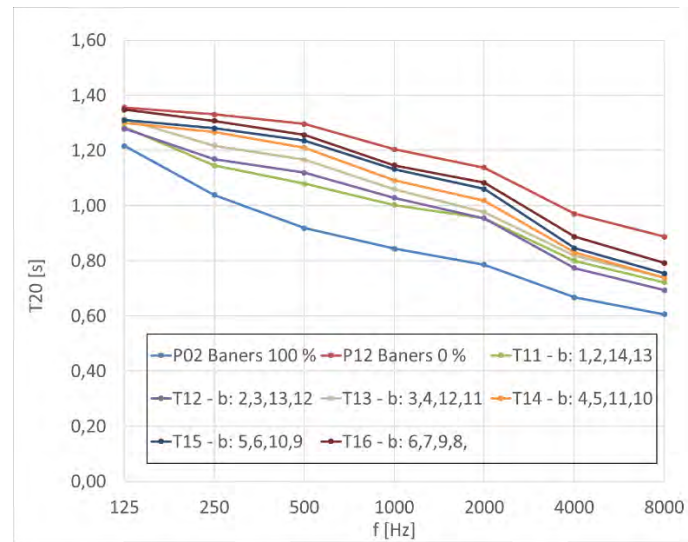


Figure 6. T20 in the function of banner positions.  
Banners 100% opened in groups of 2 per side wall

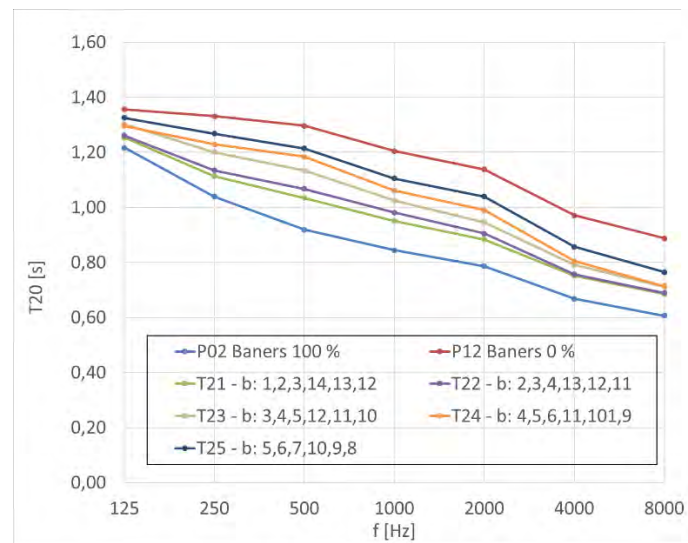


Figure 7. T20 in the function of banner positions.  
Banners 100% opened in groups of 3 per side wall

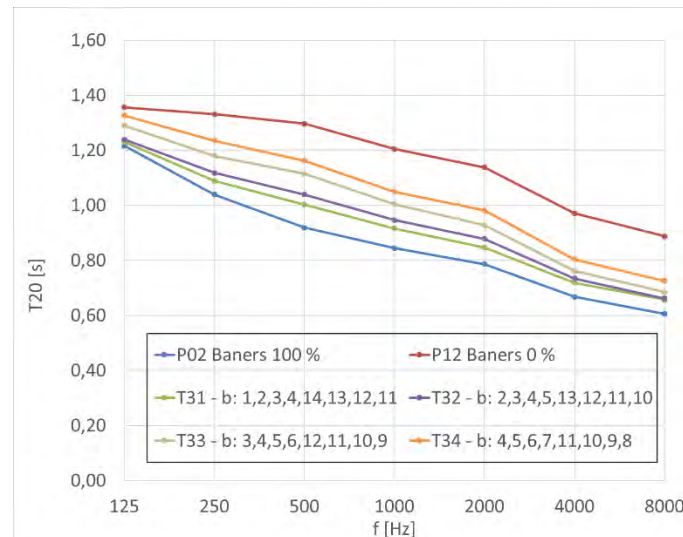


Figure 8. T20 in the function of banner positions.  
Banners 100% opened in groups of 4 per side wall

The analysis of results featured in Figures 6–8 indicates that the impact of acoustic banners on the obtained value of T20 reverberation time drops as the distance from the stage grows. This relationship is more pronounced when 4 banners per side were opened (Figure 8) than when groups of 2 banners per side were opened (Figure 6).

This situation can be explained in two ways. The first factor influencing this fact is a slightly decreasing surface of banners as the distance from the stage increases, which results from the ascending auditorium floor. However, differences between the surface of banners are not as large as differences between RT values obtained as a result of their opening.

The second possible explanation is that when banners closer to the stage are opened, suppressed are side surfaces that create mostly first order reflections. In the case of surfaces further from the stage, suppressed surfaces are those responsible to a larger degree for creating reflections of higher orders, and to a smaller extent of first order reflections. Thus, surfaces located further from the stage are less effective in creating the reverberation effect, which makes it less effective to cover them with acoustic banners in order to readjust reverberation time, than covering surfaces closer to the stage.

## 5. Conclusions

The conducted tests indicate that acoustic banners located closer to the stage are more effective in adjusting the reverberation time value than banners located in the back portion of auditorium side walls in the presented case.

Moreover, a stronger impact on the change of the reverberation time was noted in the case of banner surfaces opened at the initial phase of their opening, in comparison to surfaces activated at the end of the banner's opening.

The presented conclusions were derived from measurements and studies conducted in just one concert hall. To verify the accuracy of the results featured in this paper, further broader research including other concert halls is planned.

## References

1. L. Beranek, *Concert Halls and Opera Houses*, Springer Science + Business Media (2004).
2. M. Mehta, J. Johnson, J. Rocafort, *Architectural Acoustics Principles and Design*, Prentice Hall (1998).
3. M. Barron, *Auditorium Acoustics and Architectural Design*, Spon Press, London (2010).
4. P. Z. Kozłowski, *How to Prepare Typical Cinema Theatre to Become Multipurpose Music Venue*, Proc. of 146<sup>th</sup> Audio Engineering Society Convention, paper 10188, Dublin 2019.
5. P. Z. Kozłowski, *How to Adjust Room Acoustics to Multifunctional Use at Music Venues*, Proc. of 2018 Joint Conference ACOUSTICS, pp. 152-157, Polish Acoustical Society, IEEE Ustka 2018.
6. P. Z. Kozłowski, M. Pawelec, *Design Documentation for the Reconstruction of the Concert Hall at the M. Karłowicz Primary & Secondary Music School in Katowice*, Pracownia Akustyczna, Wrocław 2016 (in Polish).
7. P. Z. Kozłowski, et al., *Design Documentation for the Expansion of the Concert Hall of the Municipal Artistic School in Mińsk Mazowiecki. Electroacoustics, Reverberation Enhancement System, video preview*, Pracownia Akustyczna, Wrocław (2016) (in Polish).
8. P. Z. Kozłowski, R. Bolejko et al., *Design Documentation for the new headquarters of the Polish National Radio Symphony Orchestra in Katowice. Detailed design. Room acoustics, noise control*, Konior Studio / Pracownia Akustyczna, 2011 (in Polish).
9. P. Z. Kozłowski, K. Rudno-Rudziński, et al., *Design Documentation for the Castle Opera in Szczecin. Detailed design. Room acoustics, noise control*, Pracownia Akustyczna 2014, (in Polish).

## **Numerical Analyses of the Effectiveness of an Integrated Disc Based Piezoelectric Sensor-Actuator**

Roman TROJANOWSKI

*AGH University of Science and Technology, Al. Mickiewicza 30, 30-059 Krakow,  
roman.cz.trojanowski@agh.edu.pl*

Jerzy WICIAK

*AGH University of Science and Technology, Al. Mickiewicza 30, 30-059 Krakow,  
wiciak@agh.edu.pl*

### **Abstract**

This paper deals with numerical analyses of plates vibration reduction effectiveness of an integrated disc based piezoelectric sensor-actuator compared to standard type disc based piezoelectric actuator. For that purpose 400 x 400 x 2 mm plate clamped on all sides was modelled with 2 piezo elements attached to it. One was a standard square based piezoelectric actuator used to excite the plate. The second one is disc based and can be either a standard element or an integrated sensor-actuator and is used for vibration reduction.

The harmonic analyses were performed for the 1st, 2nd, 4th and 5th mode. Voltage used for plates excitation was always set to 100 V. The amplitude of voltage applied to the actuator was selected using internal ANSYS optimization procedures. The goal function for this was the minimum of the displacement vector sum of  $n$  nodes of the plate, with  $n$  having 3 possible values.

**Keywords:** AVC, FEM, ANSYS

### **1. Introduction**

The concept of an active vibration reduction was introduced at the end of XIX century. Sometime later - around the 1930s - a concepts for an active noise reduction were formulated. One of the first work dealing with the problem of reduction of structural sounds transmitted to acoustic surrounding was the article of C. R. Fuller and J. D. Jones published in 1987 [1]. Authors used a single electrodynamic actuator and were able to reduce the acoustics pressure levels emitted by an external monopole source inside a cylindrical shell by about 10-20 dB. Still the field of the active vibration control [2, 3] is constantly moving forward. Developments in material engineering, increase of technological potential and computing power allowed carrying advanced computer simulations [4] as well as to control processes leading to vibration and noise reduction.

Analytical field is also a subject of a continuous development. New theories and mathematical models are widely used for problems of objects vibrations [5] and sound radiation [6].

Other works deal with control type and algorithms [7, 8], combining active and passive methods [9].

In their previous works authors concentrated on piezo actuators with a step change in materials properties. The idea behind them was that perhaps they could serve as a simpler replacements for functionally graded actuators. The idea was first introduced

with fairly simple models [10]. After some refining the improved models and certain analytical analyses were presented [11]. Another improvement was introducing the acoustic surrounding of analysed plate [12]. These works generally concluded that introduced change in material properties of piezo actuators didn't produce any substantial change in the reduction levels of plates vibration. Another work presented the attempts to verify numerical results with physical experiments [13]. Unfortunately small number of samples (they had to be custom made) and the differences between these samples made the results unsatisfactory.

However another idea was born from this. If the changes of the inner part of a piezo element have a negligible influence on its effectiveness in obtained reduction levels perhaps it could be used as a sensor. The numerical results of this idea are presented in the hereby work.

## 2. Numerical Models

To test if a concept of a sensor-actuator could actually be a feasible construct numerical models were created using ANSYS software. Models consisted of a steel plate clamped on all sides with 2 piezoelectric elements attached. One of them was square based and used to excite the plate. The other one was a disc based actuator which was either used to full capacity as a standard actuator or had a inner part “turned off” as to somewhat simulate the behaviour of a sensor-actuator. This piezoelectric element was used for vibration reduction. The modelled plate with piezo elements placement can be seen on Figure 1.

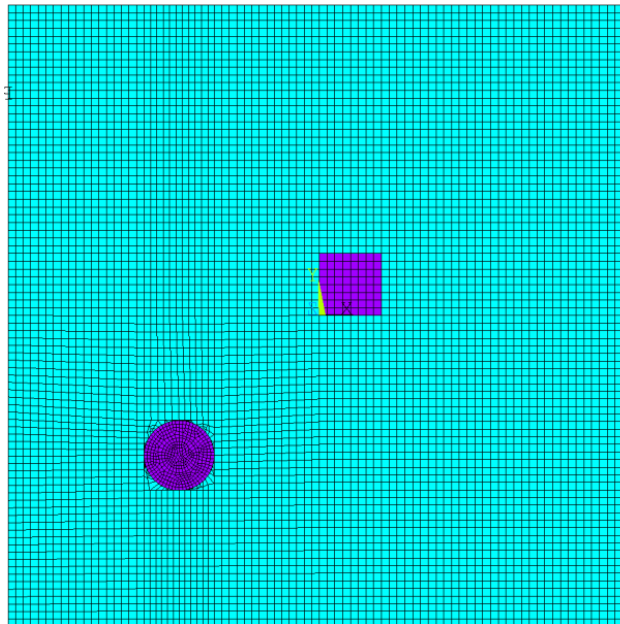


Figure 1. Modelled plate with piezoelectric actuators

Both piezo elements had a reference area of 400 mm<sup>2</sup> and thickness of 1 mm. Both were modelled using material properties of PZ 28. For the disc based actuator when used as an sensor-actuator the area of the part that was “turned off” was ¼ of the area of the whole actuator (Figure 2).

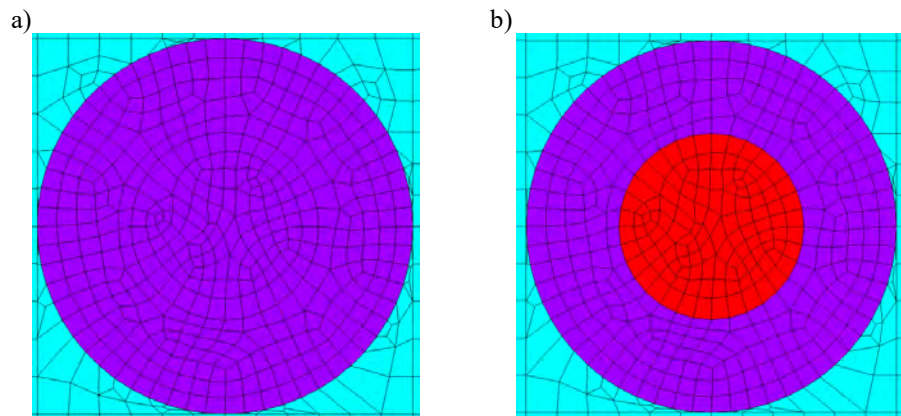


Figure 2. Disc based actuator a) when used as a standard actuator, b) when simulating a sensor-actuator

Model parameters can be found in Table 1.

Table 1. Models parameters

Structural element	Element used for modeling	Properties
Plate	SOLSH190	$E = 1.93 \times 10^{11}$ Pa, $\nu = 0.29$ , $\rho = 7800$ kg/m <sup>3</sup>
Piezo elements	SOLID226	Properties of PZ 28
Air	FLUID30	$\rho = 1.2$ kg/m <sup>3</sup> $c = 343$ m/s

Harmonic analyses were performed for the 1<sup>st</sup>, 2<sup>nd</sup>, 4<sup>th</sup> and 5<sup>th</sup> mode. During each harmonic analysis voltage of 100 V was applied to the square based piezo actuator to excite the plate. Then an optimization procedure was performed to find the disc based actuator to find the amplitude of the voltage and reduce the excited vibrations. This was done using an internal ANSYS procedures. One cycle of optimization was set up to have no more than 30 steps. After the completion of a cycle the final value of voltage amplitude was taken as a starting value of the next cycle but with a narrower voltage range. The range of the first cycle was 0–500 V, the last was  $\pm 2.5$  V of the last starting value.

The phase of the voltage applied to the disc based actuator was not a part of the optimization procedure as previous works shown that it would be  $0^\circ$  for the 1<sup>st</sup> and 4<sup>th</sup> and  $180^\circ$  for the 2<sup>nd</sup> and 5<sup>th</sup> mode.

The goal function used in the optimization procedure is given by equation 1:

$$J_1 = \min \sum_{i=1}^n |\mathbf{X}_{\text{sum}}(i)| \quad (1)$$

where: min is the smallest value of sum;  $\mathbf{X}_{\text{sum}}(i)$  is the displacement vector sum of the  $i$ -th node;  $n$  is the number of nodes used for calculations. There are 3 possible values for  $n$  depending on the case:

- $n$  is equal to every node making the back of the plate (here the back of the plate is the side to which the piezoelectric elements are not attached), that amounts to 7296 nodes and is considered a best case scenario;
- $n$  is equal to 81 nodes forming a “virtual” sensor the size of the disc based actuator placed on the same diagonal as piezo actuators but in the upper level side of the plate (near 1/4th of its length);
- $n$  is equal to 402 nodes forming a “virtual” sensor the size of the “turned off” part of sensor-actuator placed directly under the centre of the disc based actuator.

### 3. Results

The reduction of vibration was calculated as:

$$L_{\text{red}} = 20 \log \frac{\sum_{i=1}^n |\mathbf{X}_{1\text{sum}}(i)|}{\sum_{i=1}^n |\mathbf{X}_{2\text{sum}}(i)|} \quad (2)$$

where  $\mathbf{X}_{1\text{sum}}(i)$  is the displacement vector sum in  $i$ -th node before the reduction and  $\mathbf{X}_{2\text{sum}}(i)$  is the displacement vector in the  $i$ -th node after the reduction and  $n$  is the number of nodes used (as per 3 cases mentioned before).

Table 2 presents the results of vibration reduction obtained for the case where the whole back area of the plate is used as a sensor. This will be treated as a base as it is a best case scenario for the global approach to the vibration reduction. It can be seen that there are no significant differences between the results for the standard actuator and the one simulating the sensor-actuator (the biggest difference being about 0.4 dB). These differences can be attributed to the optimization algorithm finding the optimal value, but the actual best value might differ a little (about 0.1-0.2 V).

It should be noted that although when using sensor-actuator we were able to obtain almost identical vibration reduction levels the voltage applied to it had to be significantly higher. This is of course the result of the working area of the actuator being smaller. It should be possible to somewhat mitigate this by making the sensor area of sensor-actuator smaller, but further test are needed to determine how would that affect it.



Table 2. Results obtained when  $n = 7296$  nodes (whole back area of the plate); mode - number of mode; type - full actuator, actuator-sensor;  $U_a$  - amplitude of voltage applied to actuator;  $\varphi_a$  - phase of the voltage applied to the actuator;  $L_{red}$  - vibration reduction

mode	type	$U_a$ [V]	$\varphi_a$ [°]	$L_{red}$ [dB]
1	actuator	371.94	180.00	39.9
2		58.71	360.00	43.3
4		12.12	180.00	25.8
5		161.98	360.00	35.1
1	actuator-sensor	499.57	180.00	39.9
2		79.02	360.00	43.3
4		16.48	180.00	26.0
5		219.78	360.00	35.5

Table 3 presents the results of the vibration reduction when introducing a virtual sensor placed on the same diagonal as actuators, but in the upper right quarter of the plate. Apart from the results of vibration reduction from the virtual sensor the Table also presents what would be the reduction level from these voltage amplitudes calculated for every node making the back of the plate.

It can be seen that there are no significant differences between a standard actuator and simulated sensor-actuator (no more than 0.2 dB).

Table 3. Results obtained when  $n = 81$  nodes (“virtual” sensor on the diagonal in the upper side of the plate); mode - number of mode; type - full actuator, actuator-sensor;  $U_a$  - amplitude of voltage applied to actuator;  $\varphi_a$  - phase of the voltage applied to the actuator;  $L_{red}$  - vibration reduction;  $L_{redf}$  - vibration reduction calculated for all nodes making the back of the plate

mode	type	$U_a$ [V]	$\varphi_a$ [°]	$L_{red}$ [dB]	$L_{redf}$ [dB]
1	actuator	371.99	180.00	44.6	39.9
2		58.75	360.00	53.3	43.1
4		12.26	180.00	23.8	26.0
5		162.68	360.00	31.5	35.4
1	actuator-sensor	499.88	180.00	44.6	40.0
2		78.94	360.00	53.5	43.2
4		16.53	180.00	23.8	26.0
5		220.01	360.00	31.5	35.5

Table 4 presents the results of the vibration reduction when using the virtual sensor placed under the centre of the actuator used for reduction. It can be seen that for the sensor-actuator the obtained vibration reduction levels are slightly higher (up to 1.2 dB). This is actually a somewhat “false” reading. When we compare the reduction calculated by using all the nodes of the plate ( $L_{ref}$ ) it can be seen that again there are almost no

changes between different types of actuators used (less than 0.6 db). Possible reason for the difference in readings from the “virtual” sensor is the fact that for the standard actuator this area is working, whereas for sensor-actuator- it is simply “turned off”, hence the lower reading when the voltage is applied to the actuator.

Table 4. Results obtained when  $n = 402$  nodes (“virtual” sensor placed under the centre of disc based actuator); mode - number of mode; type - full actuator, actuator-sensor;  $U_a$  - amplitude of voltage applied to actuator;  $\varphi_a$  - phase of the voltage applied to the actuator;  $L_{red}$  - vibration reduction;  $L_{redf}$  - vibration reduction calculated for all nodes making the back of the plate

mode	type	$U_a$ [V]	$\varphi_a$ [°]	$L_{red}$ [dB]	$L_{redf}$ [dB]
1	actuator	370.41	180.00	25.2	38.6
2		58.67	360.00	41.6	43.2
4		12.24	180.00	31.8	26.0
5		162.36	360.00	33.0	34.9
1	actuator-sensor	497.95	180.00	25.7	38.8
2		79.01	360.00	42.3	43.3
4		16.51	180.00	32.0	26.0
5		219.60	360.00	34.2	35.4

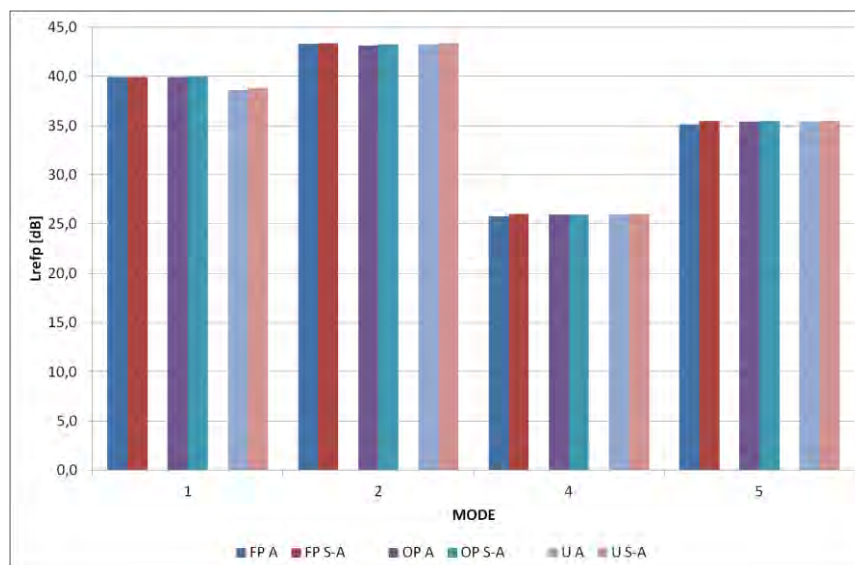


Figure 3. Comparison of vibration level reduction when recalculated to the  $L_{redf}$ ; the FP stands for when using the whole plate as a sensor; OP then using sensor placed on the diagonal in the upper part of the plate; U when using sensor placed under the actuator; A is the “standard” actuator and S-A is the sensor-actuator

Figure 3 shows the comparison of vibration levels obtained when using different sensors which are recalculated for every node making the back area of the plate. This was done to show how a sensor-actuator will behave when compared to other cases. It can be seen that for the 1<sup>st</sup> mode using the sensor placed directly under the actuator (simulating the sensor part of the sensor-actuator) results in slightly lower level of vibration reduction (1.1-1.3 dB). For the other modes there are basically no notable differences. The question remains whether this difference will become more significant as the sensor part of the sensor-actuator becomes smaller (which should bring down the amplitude of voltage required to obtain the results). If so it could limit the possible uses of such a hybrid.

### 3. Conclusions

This paper presents the results of preliminary simulations of a piezoelectric sensor-actuator hybrid. There were 2 questions to be answered. First, how will said hybrid behave when compared to a “standard” piezoelectric actuator in terms of vibration reduction levels. And the second question was how useful will the sensor part of such system be.

The results show that there are no significant differences between the sensor-actuator and the standard actuator in terms of vibration reduction levels. It did however required a significantly higher amplitude of voltage to obtain similar results as a “standard” actuator (this should be somewhat mitigated by reducing the size of the sensor part of the sensor-actuator).

As for the usefulness of the sensor part of sensor-actuator. It can be seen that depending on the mode of the plate using the sensor placed inside an actuator did result in slightly worse vibration reduction levels compared to other sensors.

### Acknowledgments

The work described in this paper has been executed within statutory activities of the Faculty of Mechanical Engineering and Robotics of AGH - University of Science and Technology

### References

1. C. R. Fuller, J. D. Jones, *Experiments on Reduction of Propeller Induced Interior Noise by Active Control of Cylinder Vibration*, Journal of Sound and Vibration, **112** (1987) 389 – 395.
2. C. R. Fuller, S. J. Elliott, P. A. Nelson, *Active Control of Vibration*, Academic Press, London 1996.
3. C. H. Hansen, S. D. Snyder, *Active Control of Noise and Vibration*, E&FN Spon, London 1997.
4. E. Augustyn, M. S. Kozień, *A study on possibility to apply piezoelectric actuators for active reduction of torsional beams vibrations*, Acta Physica Polonica A, **125**(4A) (2015) A-164-A-168.

5. E. Zołopa, A. Brański, *Analytical determination of optimal actuators position for single mode active reduction of fixed-free beam vibration using the linear quadratic problem idea*, Acta Physica Polonica A, **125**(4) (2014) A-155-A-158.
6. K. Szemela, W. P. Rdzanek, W. Żyłka, *The radiation efficiency measurements of real system of a thin circular plate embedded into a thick square baffle*, Archives of Acoustics, **43**(3) (2018) 413 – 423.
7. P. Gardonio, S. J. Elliott, J. Acoust. Soc. Am., **117**, 2046 (2005).
8. P. Gardonio, *Sensor-Actuator Transducers for SmartPanels*, ACTIVE, Sydney 2006.
9. A. Kundu, A. Berry, J. Acoust. Soc. Am., **129**, 726 (2011).
10. R. Trojanowski, J. Wiciak, *Comparison of Efficiency of Different Shapes of Homogeneous and Two-Part Piezo Elements on Vibration Reduction*, Acta Physica Polonica A, **122**(5) (2012) 905-907.
11. M. Wiciak, R. Trojanowski, *Numerical Analysis of the Effectiveness of Two-part Piezoactuators in Vibration Reduction of Plates*, Acta Physica Polonica A, **125**(4A) (2014) A-183-A-189.
12. J. Wiciak, R. Trojanowski, *Comparison of Vibration and Acoustic Pressure Reduction Using Different Types of Piezo actuators*, Acta Physica Polonica A, **128**(1A) (2015) A-62-A-66.
13. J. Wiciak, R. Trojanowski, *Evaluation of the Effect of a Step Change in Piezo Actuator Structure on Vibration Reduction Level in Plates*, Archives of Acoustics, **40**(1) (2015) 71 – 79.

## Research of Acoustical Impedance of Human Skin

Przemysław PLASKOTA

*Wrocław University of Science and Technology, 50-370 Wrocław,  
Wybrzeże Wyspiańskiego 27, przemyslaw.plaskota@pwr.edu.pl*

### Abstract

This paper presents the methodology of measuring acoustic impedance of human skin and discusses the results of the tests. Measurement of the acoustic impedance of human skin is a complex issue. The most difficult thing is related to the nature of the object of research. Cooperation of volunteer participating in the experiment is crucial in collecting accurate data. Appropriate measuring conditions must be provided because of the length of time required to conduct measurement. The Kundt tube method has been used as a research technique applied for evaluation of the material impedance.

It was assumed that the results would differ with regard to different points chosen for measurement, located on the body of volunteer as well as with regard to the features of the skin structure of the volunteers participating in this experiment. Results have revealed that for different people similar parameters have been obtained. Moreover, results are alike for various measured points localized on person's body.

**Keywords:** acoustical impedance, human skin, HRTF

### 1. Introduction

The measurements of the acoustic impedance of the human skin were conducted in order to obtain the parameter required for numerical human head model. The model can be used for numerical calculation of Head-Related Transfer Function (HRTF) [1]. To calculate HRTF using numeric methods it is necessary to establish certain border conditions. The acoustic impedance appears to be appropriate border condition [2].

The subject of the acoustic impedance of human skin in the literature is related to medical ultrasound diagnostic methods [3]. Typical methodology is to provide one average value for ultrasound frequencies. It cannot be assumed that the impedance is constant in wider range of frequencies. For practical use in ultrasound diagnosis, resulting in one average value describing the impedance is sufficient.

This paper presents the methodology of measurement of the acoustic impedance of human skin in the audible frequencies range. The results of measurements are presented and a short description of difficulties related to conducting the experiment is given.

### 2. Methodology

The measuring of the acoustic impedance can be done using standing wave ratio method [4] or using transfer-function method [5]. For the acoustic impedance of human skin standing wave ratio method was applied for the range of frequencies between 800 Hz and 6300 Hz and transfer function method was used for the range between 160 Hz and 1600 Hz.

The test was performed with the participation of two people whose skin structures were different. Volunteer A had more extensive fat tissue than volunteer O. Measuring was done on forehead, cheek, abdomen, hand and lower back (loins). The end of the impedance tube was placed on chosen part of the body and the volunteer was asked to press against the tube. The difficulty faced at this stage was that the person was asked to maintain possibly the same degree of pressure on the tube. It should be noted that there is no method to assure constant pressure of the tube's rim on person's skin. No special means improving tightness in the place of junction between tube's rim and person's skin were used.

Uneven pressure on the tube's rim in the various spots is a significant obstacle in conducting the test. While the pressure is changing, the tested spot is changing its location in relation to the tube's rim to some extent. Increasing the pressure results in a shortening of the part of the tube in operation, whereas decreasing the pressure results in an extension of the part of the tube in operation.

For invariable measuring conditions equally invariable pressure is required. Taking into consideration the nature of the object in the test, achieving constant pressure while conducting the experiment is impossible. In the same time when smaller, 30 mm diameter tube is used the uneven pressure on skin results in insignificant variations in the part of the tube in operation. The variations are in the range of  $\pm 1$  mm. However, when the bigger 100 mm diameter tube is used, the range of variations also increases up to  $\pm 5$  mm.

To validate the obtained data the transfer function method has been applied next. Minimal amount of time needed to obtain measurement assures better accuracy in measuring the depth of tester. In addition, the accuracy of measurement is increased by repeating the operation many times and averaging the obtained results.

### 3. Results

The results of measuring the human skin acoustic impedance are presented. Figure 1 illustrates module of acoustic impedance of human skin for different localizations. Figure 2 shows phase of the acoustic impedance of human skin. It should be noted that even though the test was done on skin in various parts, the results are consistent and oscillate around certain average value. It could be expected that different results would be obtained for the skin on forehead and for the part of abdomen. This expectation is justified with the fact that the thickness and the structure of forehead and abdominal skin vary to big extend. Furthermore, there is a hard structure – skull bone – under the skin on forehead and soft structures – body organs – beneath abdominal skin. The convergence of results is also found in the phase of acoustic impedance of human skin. In this case the range of variations between the obtained values is minimal and does not exceed the value of  $\pi/2$ .

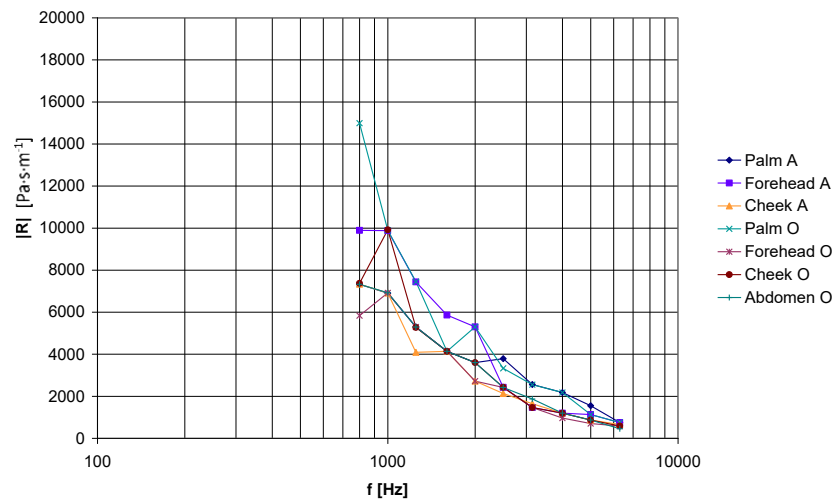


Figure 1. Module of acoustic impedance of human skin measured with standing wave ratio method

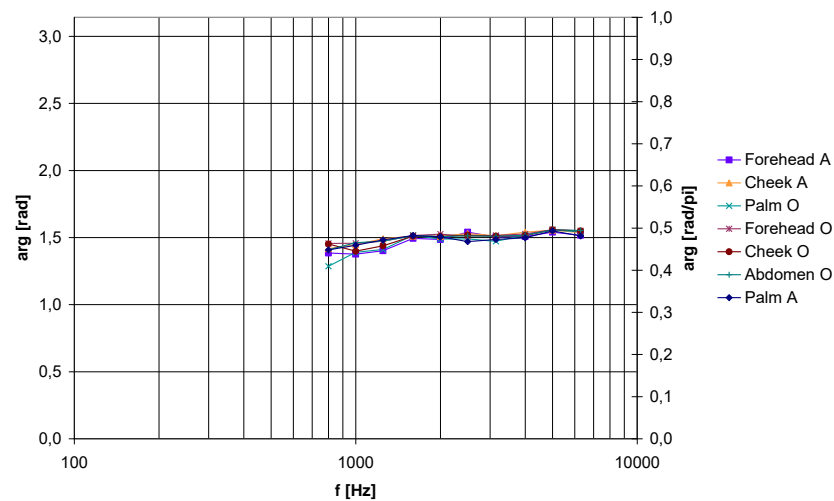


Figure 2. Phase of acoustic impedance of human skin measured with standing wave ratio method

Figures 3 and 4 present the same results averaged for all measured points. What is distinctive is the fact that while the frequency is increasing the variation in obtained values decreases.

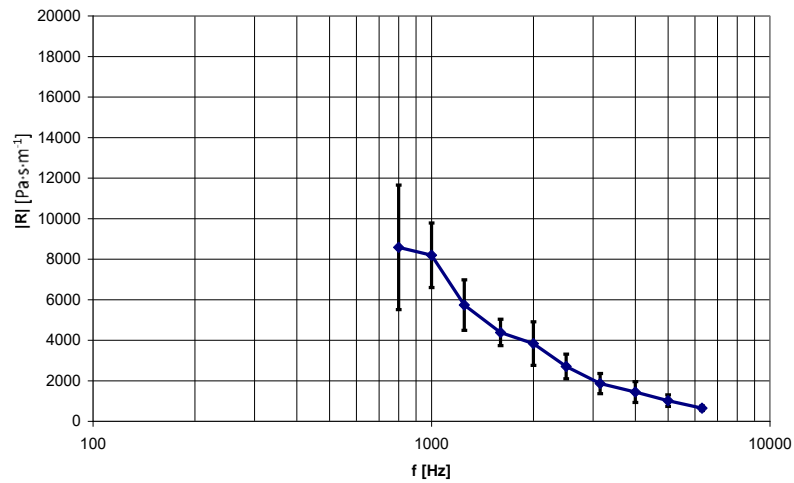


Figure 3. Average value of module of human skin acoustic impedance measured with standing wave ratio method

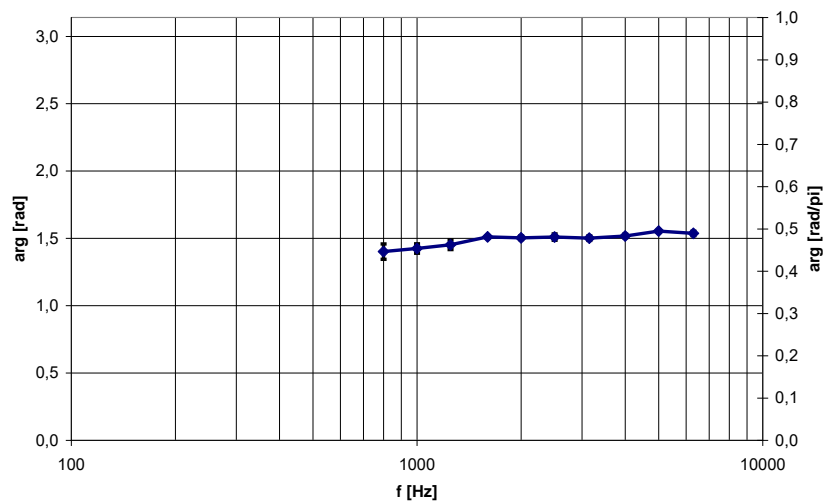


Figure 4. Average value of phase of human skin acoustic impedance measured with standing wave ratio method

Figures 5 and 6 illustrate the outcome of measuring of human skin acoustic impedance in the ranges of frequencies between 160 Hz and 6300 Hz. The chart presents average value of impedance in all measured points. This type of results presentation is justified in the light of the convergence of results for all measured points. It should be noted that the skin impedance on abdomen and on back were measured using lower frequency



range. Due to the big diameter of the tube it could not be applied, for instance, to test skin on hand because the size of hand is smaller than that of the tube.

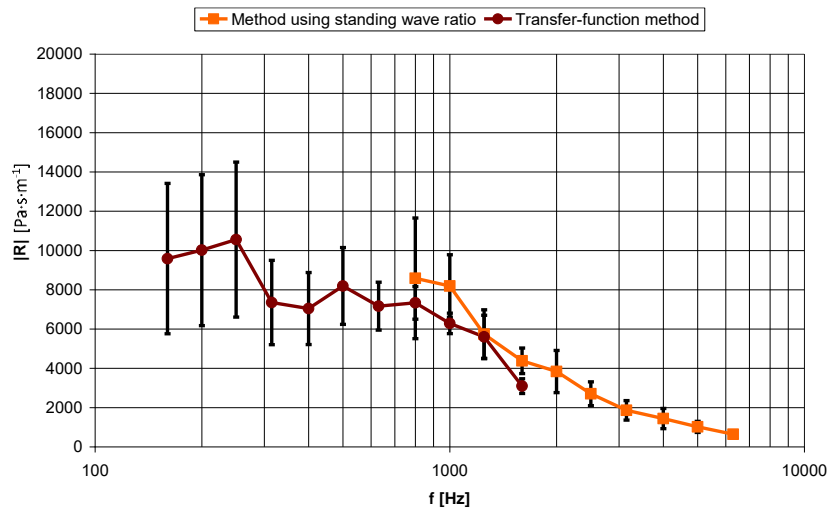


Figure 5. Average value of module of human skin acoustic impedance measured with standing wave ratio method and transfer-function method

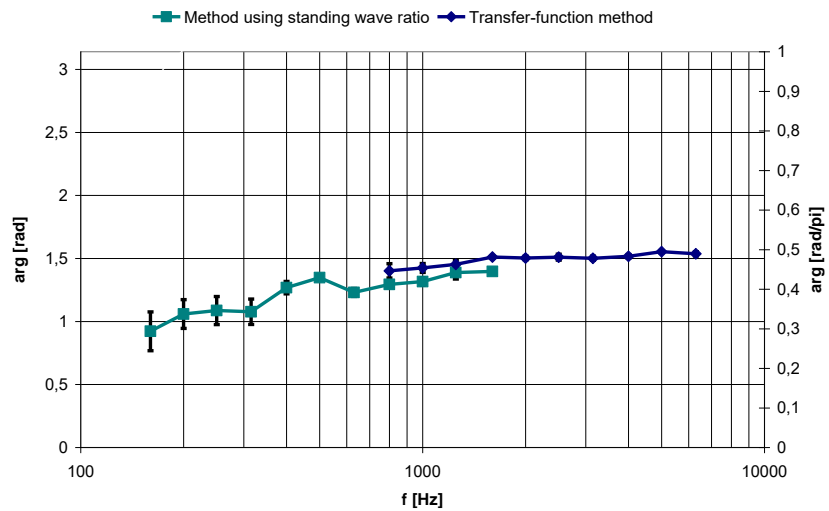


Figure 6. Average value of phase of human skin acoustic impedance measured with standing wave ratio method and transfer-function method

On the base of obtained results it could be concluded that the value of the module of human skin acoustic impedance is decreasing as the frequency increases. Conversely, the value of the phase is increasing, but  $\pi/2$  seems to be the maximum value.

In the common measuring range (frequencies between 800 Hz and 1600 Hz) the results are within the range of standard deviation for both methods. This fact exist with regard to both – module and the phase of human skin acoustic impedance.

### 3. Conclusions

The paper presents the methodology of human skin impedance measurement. On the basis of the conducted tests it can be assumed that the human skin impedance depends on the frequency of the sound and not on the structure of the skin (i.e. the thickness of fat tissue) neither on the structure of the tissues located directly beneath the skin. It can also be assumed that the impedance value is only correlated with the outer layer of tissue.

Although only two person participated in the experiment, author supposed that the result may differ for larger population. Nevertheless, this preliminary results can be used as a boundary condition for numerical calculation.

### References

1. A. B. Dobrucki, P. Plaskota, *Computational modelling of head-related transfer function*, Archives of Acoustics, **32**(3) (2007) 659 – 682.
2. P. A. Nelson, Y. Kahana, *Spherical harmonics, singular-value decomposition and the head-related transfer function*, Journal of Sound and Vibration, **239**(5) (2000)607 – 637.
3. K. Opieliński, *Application of transmission of ultrasonic waves for characterization and imaging of biological media structure*, Oficyna Wydawnicza Politechniki Wrocławskiej, Wrocław 2011 (in Polish).
4. ISO 10534-1:2004, *Acoustics - Determination of sound absorption coefficient and impedance in impedance tubes - Part 1: Method using standing wave ratio*, ISO, 2004.
5. ISO 10534-2:2003, *Acoustics - Determination of sound absorption coefficient and impedance in impedance tubes- Part 2: Transfer-function method*, ISO, 2003.

## Analysis of Doppler Tomography in Circular Geometry as a Novel Method of Imaging Tissue Cross-Sections *in vivo*

Tomasz ŚWIETLIK

*Department of Acoustics and Multimedia, Faculty of Electronics, Wrocław University of Science and Technology, Wyb. Wyspiańskiego 27, 50-370 Wrocław, Poland, tomasz.swietlik@pwr.edu.pl*

Krzysztof J. OPIELIŃSKI

*Department of Acoustics and Multimedia, Faculty of Electronics, Wrocław University of Science and Technology, Wyb. Wyspiańskiego 27, 50-370 Wrocław, Poland, krzysztof.opielinski@pwr.edu.pl*

### Abstract

Currently, methods such as conventional ultrasound B-mode scanning (US), computerized X-ray tomography (CT), magnetic resonance imaging (MRI), standard X-ray diagnostics, radioisotope imaging and thermography are used to visualize the internal structure of tissue *in vivo* and to diagnose the patient. Doppler tomography (DT) is an innovative method of reconstructing the image of the tissue section using ultrasonic waves and Doppler effect. In contrast to the currently applied solutions (US), this method uses a continuous wave, which, in theory, allows one to operate with higher energy and to detect smaller inclusions within the examined tissue. This study focuses on the analysis of DT simulation in circular geometry, where a two-transducer ultrasonic probe circulating around the tested object is used to measure the useful signal. In this paper, the influence on the tested object's cross-section imaging quality of both the simulated Doppler signal's registration parameters, and the calculation algorithm's parameters, were analyzed.

**Keywords:** Doppler tomography, Doppler signal, continuous wave ultrasonic tomography, image reconstruction

### 1. Introduction

Currently, the most popular use in medicine of both ultrasound wave and Doppler effect is the measurement of blood flow velocity in blood vessels. The Doppler Tomography method also uses these two elements but for a completely different kind of measurement. The main task of this method is the reconstruction of the image of the tissue section. Just like in the classical measurement of blood velocity, also in this case, an ultrasonic two-transducer probe generating a continuous wave is used. The difference is that in DT this probe is put in motion to trigger the Doppler effect, while the imaged tissue section is stationary.

Due to the way the probe moves, we distinguish two data archiving geometries: linear geometry and circular geometry [1, 2]. In this paper, we focused on the second geometry in which the probe moves around the object being examined.

## 2. The principle of data acquisition in the DT method for circular geometry

As already mentioned in the DT method, we use a two-transducer probe that transmits and receives an ultrasonic wave. This probe generates a continuous wave with the frequency  $f_T$  and moves around the tissue being examined as shown in Figure 1a. At a given rotation angle  $\theta$ , the wave is reflected from the stationary inclusions and returns to the probe with the changed frequency  $f_R$ . For a single inclusion, we register the appropriate Doppler frequency. In the DT method, the image is reconstructed based on the measurement of this frequency during the rotation of the probe. This frequency can be calculated using formula:

$$f_d = \frac{2 \cdot f_T \cdot v \cdot \cos(\theta)}{c} \quad (1)$$

where:  $f_T$  – ultrasonic wave frequency generated by the probe,  $v$  – linear component of the scatterer instantaneous velocity moving toward the ultrasonic wave propagation,  $\theta$  – rotation angle,  $c$  – ultrasonic wave speed in the tissue (the cross-section of the examined object). This formula can also be converted to the following form:

$$f_d = 2 \cdot f_T \cdot \omega_{turn} \cdot r \cdot \frac{\cos(\theta)}{c} \quad (2)$$

where  $\omega_{turn}$  is the angular speed of the object under examination (the cross section of the tissue),  $r$  is the distance between the inclusion and the centre of rotation.

It should also be noted that in the DT measurement system, the imaging tissue and the probe should be immersed in water.

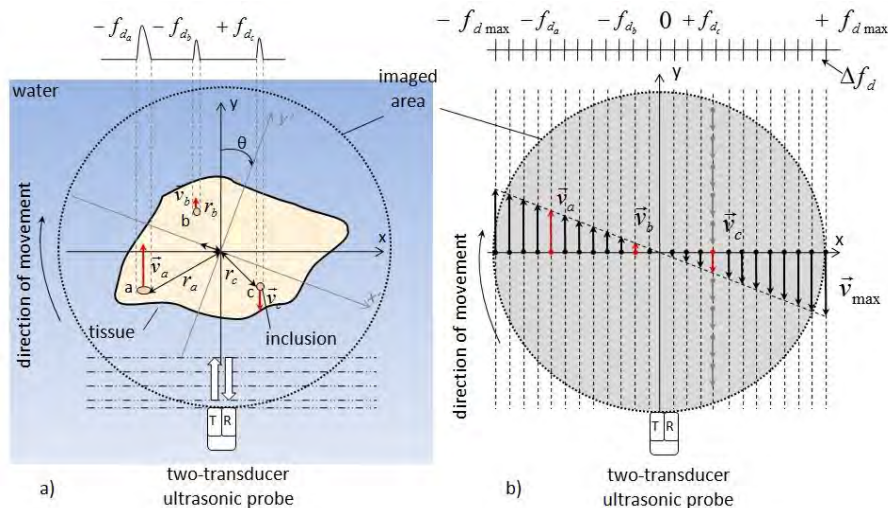


Figure 1. The method of data acquisition for DT in the circular geometry (a) and the nature of velocity distribution  $v$  and Doppler frequencies  $f_d$  in the imaging area (b)

Data acquisition in the DT method requires the registration of data from the rotating probe. However, to understand the essence of image reconstruction it will be easier when we assume that the object being imaged rotates around its own axis, while the probe remains stationary. From the point of view of the Doppler effect, the Doppler frequency is the same when the source moves, and the object remains stationary or the object moves, and the source is immovable.

For example, let's assume that as shown in Fig. 1a, we have three inclusions marked with the letters **a**, **b**, **c**, in the section of the imaged tissue. It should be noted that for the rotation angle equal to  $0^\circ$  the component linear velocity in the direction of wave propagation for these three inclusions is  $v_a$ ,  $v_b$  and  $v_c$  (Fig. 1a). Each of these velocities will be related to the generated Doppler frequency in accordance with the formula (2). In this case, Doppler frequencies will be  $f_{da}$ ,  $f_{db}$ ,  $f_{dc}$ , respectively. Suppose that at a given rotation angle, the incidence moves in the direction from the probe. For such a case, the Doppler frequency is taken into account with the plus sign (+). If the inclusions move towards the probe, we take the frequency with the minus sign (−).

To understand the DT method thoroughly one should notice two facts. First, as we move away from the center of rotation, the linear velocity component in the direction of wave propagation increases to the maximum value of  $v_{max}$  (Fig. 1b). This velocity determines the maximum Doppler frequency. Assuming the given diameter of the imaging area, this frequency can be calculated using the formula (2). The second fact is that on lines perpendicular to the direction of wave propagation (vertical lines in Fig. 1b), the velocity components  $v$  have the same values. In Fig. 1b this can be seen on the example for  $v_c$  velocity. In this case, by using the formula (1) it can be concluded that the inclusions on the vertical lines (Fig. 1b) will generate the same Doppler frequency at a given rotation angle  $\theta$ .

To reconstruct the image in the DT method, the maximum Doppler frequency should be calculated and the range from  $-f_{dmax}$  to  $+f_{dmax}$  should be created. Next, the range should be divided into equal bands with the width  $\Delta f_d$ . These bands are called Doppler ranges. At a given rotation angle, each of these bands represents a portion of the imaged zone as shown in Fig. 1b. If one or more inclusions appear in a given zone, they will generate Doppler frequencies of a given value. In the DT method, at each rotation angle to a given band, we write the sum of the amplitudes of the Doppler frequencies. As a result, we get a matrix called a sinogram, where in the following lines the angles of rotation change, while in the columns we have the sum of Doppler frequency amplitudes in the given Doppler bands. An example of such a sinogram is shown in Fig. 2a. Doppler frequencies have been calculated from the equation (2). The calculations were carried out for the case in which we simulate one inclusions in the position of 4 cm from the center of rotation and set at an angle of  $0^\circ$  in relation to the ultrasound probe. The inclusion is infinitely small and scatters the ultrasonic wave in the same way in every direction. The remaining parameters of the experiment are:  $f_T = 4.7$  MHz,  $\omega_{turn} = 2$  turns per second,  $c = 1482$  m/s, the diameter of the imaging area is equal to 10 cm, the number of read angles for half rotation is equal to 500, the number of Doppler bands is 99. After the creation of sinogram, the image should be reconstructed by means of the algorithm used in ultrasound tomography. In our case, it is a fast algorithm under the name of Filtered Back Projection (FBP) [4]. An additional advantage of

the algorithm is the ability to reconstruct the entire image from data recorded only from the half of the rotation of the tested object (or ultrasound probe). An example of the image of an inclusion reconstructed by FBP algorithm from a previously calculated sinogram (Fig. 2a) is shown in Fig 2b.

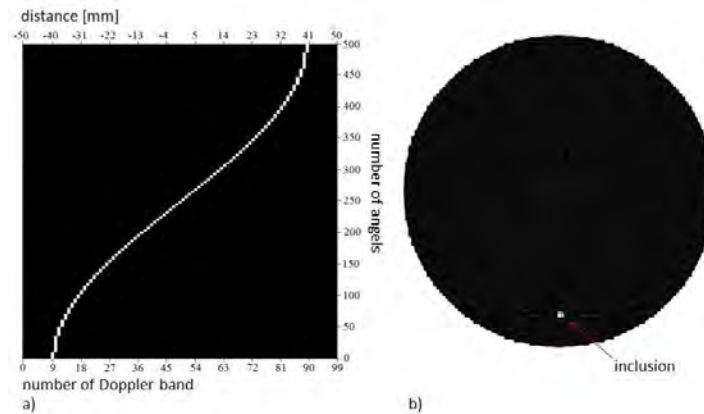


Figure 2. Simulation of the sinogram (a) and the reconstructed image of a single inclusion (b) placed at a distance of 4 cm from the center of rotation and at an angle of  $0^\circ$  in relation to the probe

### 3. Doppler signal simulation

Doppler signal contains only Doppler frequencies. Because this signal is used to reconstruct the image, its simulation and thorough examination is necessary for a better understanding of the DT method.

It is worth noting that this is a *chirp* type signal. Starting from the definition of this signal, it is possible to determine the formula:

$$s(t) = A \sin \left( \frac{4 \cdot \pi \cdot f_T \cdot r_0}{c} \cdot (\sin(2\pi \cdot f_{rot} \cdot t + \alpha_0) - \sin \alpha_0) \right) \quad (3)$$

which allows to simulate a Doppler signal from the incident being at a distance  $r_0$  from the center of rotation, at an angle  $\alpha_0$  in relation to the ultrasound probe [4]. In this equation,  $f_{rot}$  is the rotation frequency of the object,  $t$  is time, and  $A$  is the amplitude of the signal. To be able to simulate more than one inclusions, we calculate the Doppler signal for each incident separately, and then add their values for the appropriate time samples.

### 4. Modifications of the DT method due to the improvement of the imaging quality

In the DT method, one should record the Doppler signal section for each probe rotation at angle  $\theta$  and calculate its spectrum. The next step is to divide the frequency range of the determined spectra into individual Doppler bands in which we have the sum of

amplitudes of Doppler frequencies derived from inclusions. However, we come to a disturbing conclusion. In order to register Doppler signals for particular angles of rotation  $\theta$ , the recording of this signal should be switched on in a proper measuring time even hundreds of times. This leads to a relatively complicated measurement system. The second important problem is the limited time period and the number of Doppler signal samples for particular angles of rotation. In this approach, the signal spectrums have a low resolution, which translates into low accuracy in determining Doppler frequencies. This is due to the fact that the spectral resolution  $\Delta f$  is equal to the sampling frequency  $f_s$  divided by the number of samples  $N$ . It seems natural that if we increase the number of signal recording samples, the spectrum resolution should increase. However, it turns out that by increasing the number of  $N$  samples we also increase the sampling frequency  $f_s$  and consequently their ratio, i.e. resolution, remains the same. This fact means that we cannot determine the right amount of Doppler bands. Please note that the higher the number of bands, the better the image resolution [3].

The solution to both the first and the second problem can be registering the Doppler signal from the entire turnover and applying the signal overlap procedure for individual angles of rotation  $\theta$ . This procedure is illustrated in Fig. 3.

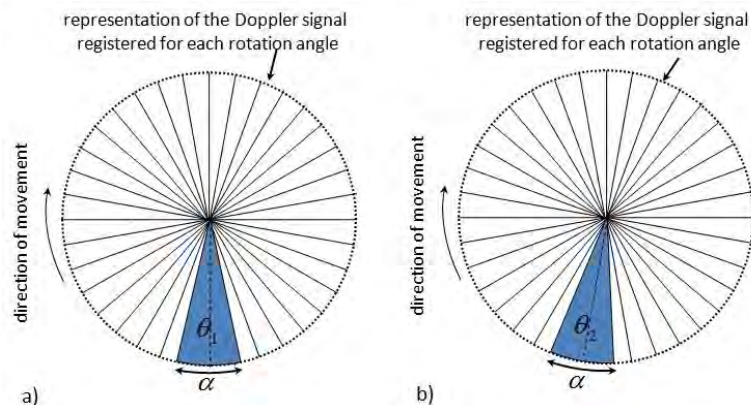


Figure 3. Example of Doppler signal determination for acquisition angle: (a)  $\theta_1$  and (b)  $\theta_2$  with  $\alpha$  length in the overlay procedure

The step forward for this procedure is the division of the Doppler signal recorded from the total turnover into the individual acquisition angles  $\theta$ . For example, for 36 reading angles, these angles will be:  $\theta_1 = 0^\circ$ ,  $\theta_2 = 10^\circ$ , ...,  $\theta_{100} = 350^\circ$ . In Fig. 3, the angles  $\theta_1$  and  $\theta_2$  are presented respectively. Then, each Doppler signal sample should be given a rotation angle at which it was recorded instead of time. In the next step, for determined angles  $\theta$  we are calculating the middle sample of the Doppler section of the  $\alpha$  length which is taken for analysis. For each angle  $\theta$  is designated Doppler range of  $\alpha$ . Then we calculate the spectrum of this signal and Doppler frequencies, as a result we create the sinogram and reconstruct the image.

The influence of the  $\alpha$  parameter on the imaging quality was examined on the example of rotation of an infinitely small inclusions placed on the rotating platform

4 cm from the center of rotation and at an angle of  $0^\circ$  in relation to the ultrasound probe. The value of the  $\alpha$  parameter is given in relation to the length of the Doppler signal recorded from full rotation, i.e.  $360^\circ$ . During the experiment, it took values  $\alpha = 3.6^\circ$ ,  $7.2^\circ$ , ...,  $36^\circ$ . The other simulation parameters are:  $f_T = 4.7$  MHz,  $\omega_{turn} = 2$  turns per second,  $c = 1482$  m/s, the diameter of the imaging area is 10 cm, the number of acquisition angles for half rotation is 500, the number of Doppler samples per half of the rotation is 25000.

Figure 4 shows images of inclusions reconstructed on the basis of accurately calculated Doppler frequencies, i.e. without determining these frequencies from the Doppler signal. Each image has a dimension of 15 mm x 15 mm. This simulation later in the article will be called an perfect case. It should be noted that this method cannot be applied in practice. This simulation gives the opportunity to examine the changes in the imaging that introduces the inaccuracy of Doppler frequency determination. For each of the values of the  $\alpha$  parameter, the corresponding number of Doppler bands was determined. These are values of 21, 39, 59, ..., 199, respectively. This information allows to create appropriate sinograms and, as a result, to determine the image of the inclusion. Figure 4 clearly shows that the higher the value of  $\alpha$  corresponding to the larger number of Doppler bands, the better the resolution.

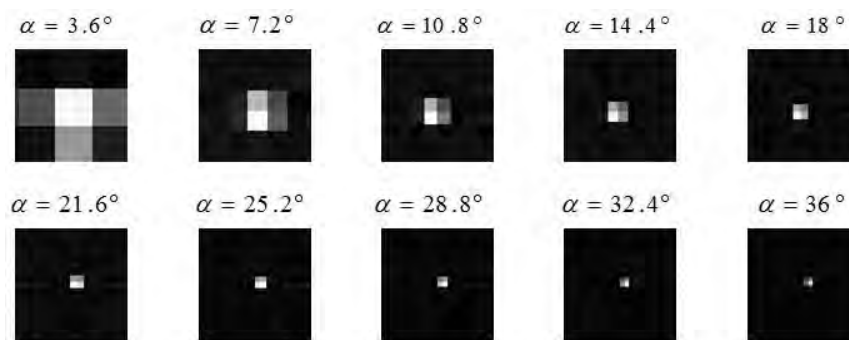


Figure 4. Image reconstruction for a single inclusion in the perfect case simulation for different values of  $\alpha$  parameter

Figure 5 shows the results of image reconstruction for calculating the frequency from the Doppler signal and applying the overlapping procedure in similar way as in the case of real measurement method. This simulation will be referred to as the real case later in the article. It can be seen that as for the previous case, as the  $\alpha$  parameter increases, the resolution of the image improves. However, it is clearly visible that for  $\alpha = 21.6^\circ$  the resolution in the direction of the  $OX$  axis deteriorates and the image of the object is extended. To explain this phenomenon, the values of the Doppler signal spectrum resolution for the read angle  $\Delta f$  were calculated depending on the given  $\alpha$ . The values of  $\Delta f$  are equal to: 398.39 Hz, 199.59 Hz, 133.15 Hz, 99.90 Hz, 79.93 Hz, 66.62 Hz, 57.14 Hz, 49.97 Hz, 44.42 Hz, 39.98 Hz. When calculating the maximum Doppler frequency for this simulation equal to 3984.27 Hz, both the previous reconstruction of images can be used to determine the number of Doppler bands. Considering only the  $\Delta f$



spectrum resolution, one should expect that the accuracy of Doppler frequency determination and the image quality should improve with the increase of  $\alpha$ .

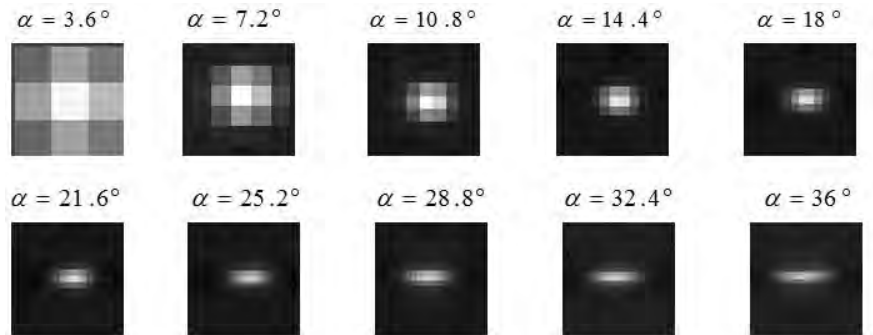


Figure 5. Image reconstruction for a single inclusion in real case simulation for different values of  $\alpha$  parameter

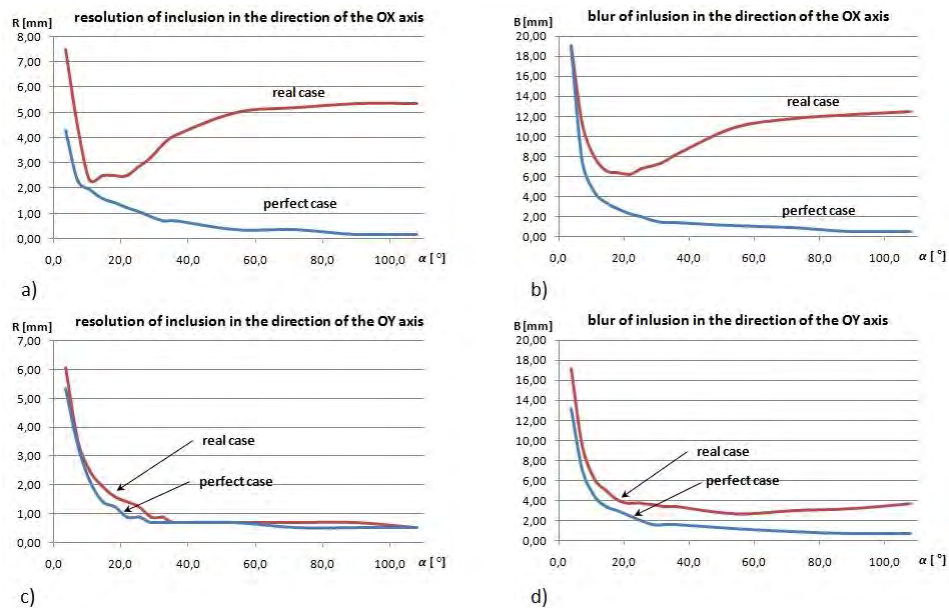


Figure 6. Inclusion image (a) resolutions and (b) blurs in direction of the  $OX$  axis and inclusion image (c) resolutions and (d) blurs in direction of the  $OY$  axis for the real case and perfect case simulation for different  $\alpha$  parameter

It should be noted that by increasing the width of overlapping  $\alpha$ , we also increase the path on which we observe the moving inclusions for one angle  $\theta$ . In this causes we register a greater number of velocities for inclusions (in the direction of wave

propagation), and this leads to a greater number of Doppler frequencies for a single acquisition angle  $\theta$ . Therefore, an additional elongation and blurring of the object appears on the image caused by averaging. In order to investigate this phenomenon carefully and determine which value  $\alpha$  gives the best resolution and the smallest blur in the direction of the  $OX$  and  $OY$  axes, the values for individual pixels of the images were examined. Point resolution was determined at 3 dB decrease in pixel values and blur at 90 % of maximum value. The results are shown above, in Figure 6. Four additional values of the  $\alpha$  parameter were considered:  $54^\circ$ ,  $72^\circ$ ,  $90^\circ$  and  $108^\circ$ . The charts in Fig. 6c and Fig. 6d show that both the resolution and image blur in the direction of the  $OY$  axis for both simulations for the perfect and real case remain practically the same. It can therefore be concluded that the overlap procedure has no significant effect on imaging in the referenced direction. In contrast, charts in Fig. 6a and Fig. 6b clearly show a deterioration in both resolution and blur for  $\alpha$  above  $21.6^\circ$ .

## 5. Conclusions

Doppler tomography is a promising method of reconstructing the image of the tissue *in vivo*. The advantage of this method is the use of ultrasonic waves that can be safe and the use of a single head with only two transducers, which makes it possible to design a relatively cheap device. But in order for it to be used in practice, modifications are necessary. It should be remembered that the Doppler signal should be registering for the full rotation of the probe (or imaging object). It is not necessary to record this signal for each rotation angle  $\theta$  separately during measurements. Secondly, a method should be used to accurately determine the Doppler frequencies from the recorded Doppler signal. The solution may be the overlapping procedure, where the length of the signal for a particular rotation angle  $\theta$  is  $\alpha$ . It is possible to set the  $\alpha$  parameter value so that the image resolution of a single point will be approximately 2 mm, and the image blur will not exceed 7 mm. These results suggest that one of the practical applications of this method in the future may be the diagnosis of female breasts in detection of cancerous changes.

## References

1. H-D. Liang, M. Halliwell, P. N. T. Wells, *Continuous wave ultrasonic tomography*, IEEE Trans Ultrason Ferroelectr Freq Control, **48** (2001) 285 – 292.
2. H-D. Liang, Ch. S. L. Tsui, M. Halliwell, P. N. T. Wells, *Continuous wave ultrasonic Doppler tomography*, Interface Focus, (2011) 665 – 672.
3. T. Świetlik, K. J. Opieliński, *The use of Doppler Effect for Tomographic Tissue Imaging with Omnidirectional Acoustic Data Acquisition*, In: Piętka, E. et al. (eds.), Information Technologies in Medicine, Advances in Intelligent Systems and Computing **471**, Springer International Publishing, Switzerland (2016) 219 – 230.
4. T. Świetlik, K. J. Opieliński, *Analysis of the possibility of doppler tomography imaging in circular geometry*, In: Piętka, E. et al. (eds.), Information Technologies in Medicine, Advances in Intelligent Systems and Computing **762**, Springer International Publishing, Switzerland (2019) 52 – 63.

## Fluid-Fluid Phononic Crystal with Elastic Coat Working in Audible Frequencies

Aleksandra KLIMEK

Wrocław University of Science and Technology, Wyspiańskiego 27, 50-370 Wrocław,  
aleksandra.klimek@pwr.edu.pl

### Abstract

Phononic Crystals are receiving rising attention in the field of modern acoustic materials. PCs are artificial structures of periodically arranged scatterers. Such a structure enables creating a band gap in which, due to the Bragg diffraction phenomenon, vibrations are restrained or even forbidden. In this paperwork, the fluid-fluid PC is tested and simulated - the scatterers are constructed of water cylinders with an ethylene propylene diene monomer coat (EPDM – a hyperelastic rubber) and are embedded in air. The band gap is calculated to emerge in the audible range of frequency. Every simulation is performed with the use of the finite element method.

**Keywords:** Phononic Crystal, band gap, Finite Element Method, viscoelasticity

### 1. Introduction

Phononic Crystals (PCs) are periodic structures of matrix-embedded scatterers in a host medium. The crystal mode of action is based on the difference between the inclusions and the medium acoustic impedance. Such composition reveals an intriguing behaviour. Its band structure may hold band gaps in which the acoustic wave propagation is prohibited.

Band gaps can be formed in the crystal band structure in two ways. First one is by the Bragg Diffraction and second is correlated with the localized modes inside the unit cell. Bragg diffraction occurs when scatterers repetitiously reflect acoustic wave repetitiously. The equidistant between every inclusion causes creation of a standing wave in the medium, thus confinement of the acoustic energy inside the PC.

The band structure depends on multiple parameters: the elastic properties of components, inclusions shape, lattice type (most common are: square, hexagonal and honeycomb) and spatial distribution of inclusions. The last one is described by lattice parameter  $a$  and filling factor  $\beta$ , which is the ratio of the volumes of the inclusion and the whole unit cell.

The periodicity of Phononic Crystals enables to derive the dispersion curves basing on single unit cell, i.e. with the application of Bloch Theorem. Primarily, for a given matrix, one must define the corresponding Brillouin Zone, which is a primitive, unit cell of a reciprocal lattice (which is, in turn, Fourier transform of regular lattice). In this work, the scatterers are arranged in a square lattice (Figure 1.a). The corresponding Brillouin Zone is presented in Figure 1.b. The dispersion curves may be obtained afterwards by calculating the eigenfrequencies for every wavevector.

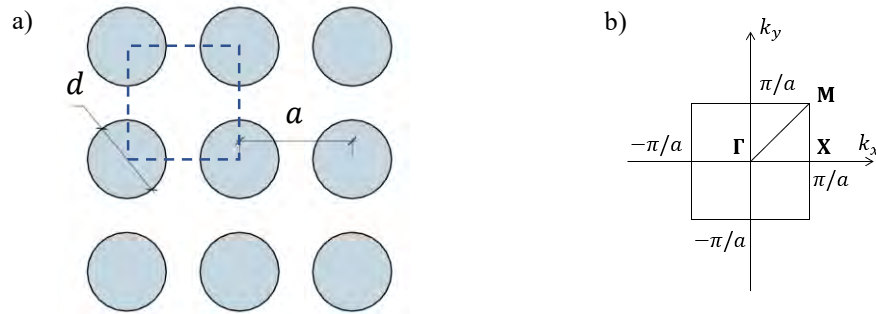


Figure 1. a) The square lattice. The dashed line represents unit cell; b) the corresponding first irreducible Brillouin Zone.  $\Gamma, X, M$  are highly symmetric points

In the following test, the host medium is air and the scatterers are constructed of water cylinders with an ethylene propylene diene monomer (EPDM rubber) coat. Due to the highly nonlinear nature of coat material and thus highly nonlinear behaviour of the water-coat inclusion pair, the prediction of dispersion curves is complicated. The Table 1. contains basic acoustic parameters of EPDM rubber.

Table 1. Basic properties of EPDM rubber [1]

Density [ $10^3 \frac{\text{kg}}{\text{m}^3}$ ]	Elastic Modulus [MPa]	Poisson Ratio	Sound Speed [ $\frac{\text{m}}{\text{s}}$ ]	Acoustic Impedance [ $10^3 \frac{\text{Pa s}}{\text{m}}$ ]
0.87	50.7	0.3	280	243.6

## 2. The FEM simulation results

The simulation was performed in COMSOL Multiphysics software. To simplify the calculations, it was assumed that the boundary between the medium (air) and cylinder (EPDM rubber) is hard. The Bloch-Floquet periodic boundary conditions were:

$$p_n = p_{n-1}e^{-ika} \quad (1)$$

where:

- $p_{n-1}, p_n$  – the acoustic pressure on the left and right edge of unit cell respectively, in the  $x$  direction and on the bottom and top edge in the  $y$  direction;
- $k$  – the wavevector component in the  $x$  or  $y$  direction respectively ( $k_x$  or  $k_y$ ).

Every geometric parameter of the simulation and real model are presented in the Table 2.

Table 2. Lattice geometric parameters

Quantity	Symbol	Unit	Value
Lattice parameter	$a$	[mm]	50
Water cylinder radius	$r_w$	[mm]	17.5
EPDM coat thickness	$d_E$	[mm]	4
Cylinder radius	$r_E$	[mm]	21.5
Filling factor	$\beta$	-	0.58
Rows number (among $y$ -axis)	-	-	6
Columns number (among $x$ -axis)	-	-	5

Resultant band structure for the first 5 eigenfrequencies is as shown on the Figure 2.

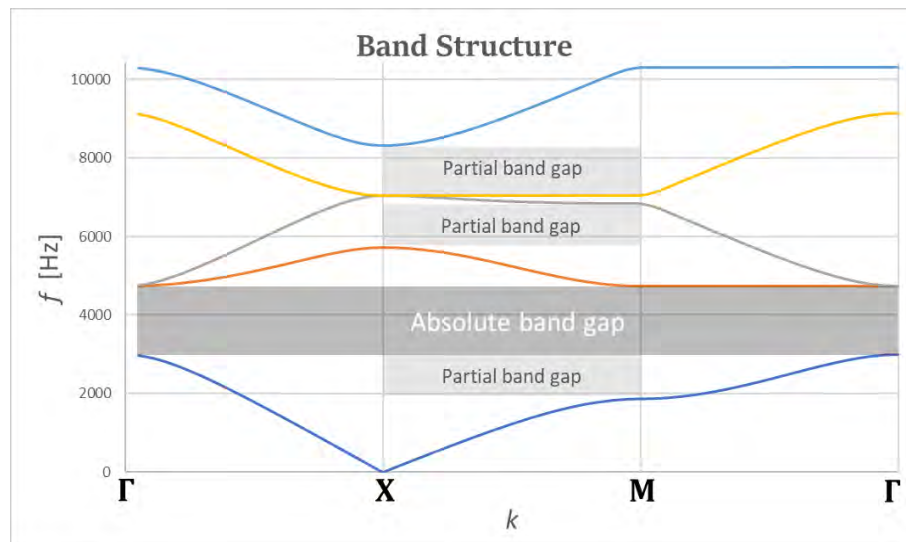


Figure 2. Band structure

In the given frequency range (0÷10 kHz) there is one absolute band gap (in every direction) and 3 partial band gaps, only for direction **X-M**. The Table 3 contains upper and lower frequency of every band gap.

Table 3. Band gaps frequencies

Band gap type	Partial (1)	Absolute	Partial (2)	Partial (3)
Lower frequency [Hz]	1872	2998	5726	7033
Upper frequency [Hz]	2998	4740	6831	8325

The sound pressure level distribution among the matrix acutely demonstrates the effects of absolute and partial bandgaps (Figure 3).

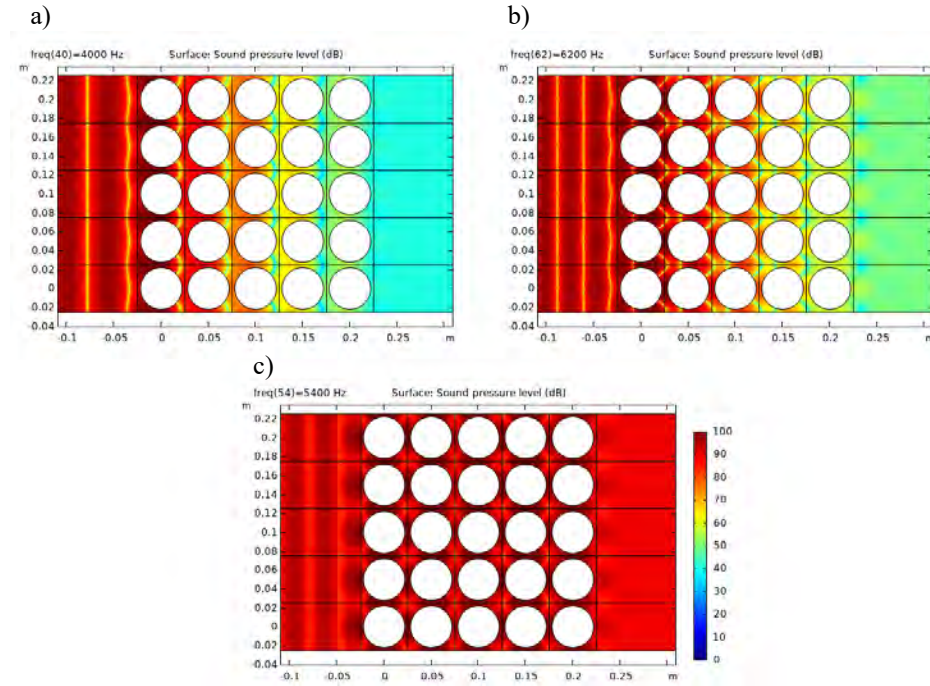


Figure 3. Sound pressure level distribution among the 5 columns of inclusions in the host medium domain (air).

- a) The frequency of 4000 Hz in the absolute band gap,
- b) The frequency of 6200 Hz in the partial band gap,
- c) The frequency of 5400 Hz (between the first two) in the conductive band.

The difference in SPL on the right edge of the crystal between conductive band and bandgap reaches 50 dB for the absolute and 40 dB for the partial band gap.

### 3. Test results

For the aim of this work, sonic crystal with every geometry parameter as given in Table 2. was constructed. Cylinders were fixed rigidly between two MDF boards to ensure its precise equidistant. There were *no* absorbing materials on the sides of the crystal. The research consisted of measuring the transmission loss between the no-PC and with the PC system (Figure 4.). In addition, a transmission loss with an empty coat (i.e. filled with air) was checked. The single broadband speaker was used to generate the sine sweep between 100 Hz and 10 kHz.

Due to rubber acoustic parameters, the measurements were expected to show shifting the band gap towards the lower frequencies. The frequency of local resonances of water cylinders *or* strings formed as a result of mounting the cylinder rigidly on both sides are located outside the frequency range in the study. However, according to the author's

opinion, the appearance of other resonances resulting from the superposition of rubber and water vibrations in the scatterer is believable. In such a case, in the resonance locations in the graph of transmission loss, one will be able to observe large differences between the curves representing water-filled and air-filled cylinders.

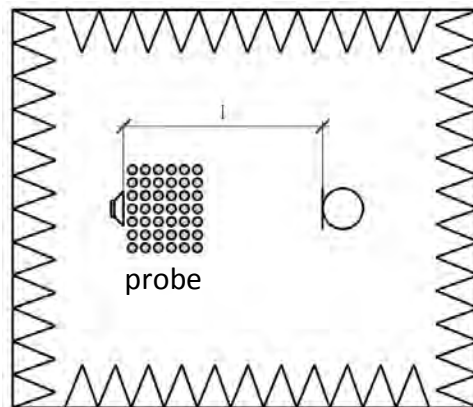


Figure 4. Measurement system diagram (top view).  $l$  is the distance between the broadband speaker and the microphone

The distance between microphone and speaker was variable. In total, measurements were taken for 16 positions of the microphone, with  $l$  between 50 cm and 2 m. Every transmission loss curve is 1/48 oct. smoothed and is a result of 3 measurements average. Selected representative results for  $l = 1$  m and  $l = 1.5$  m are in the Figure no 5 and 6 respectively.

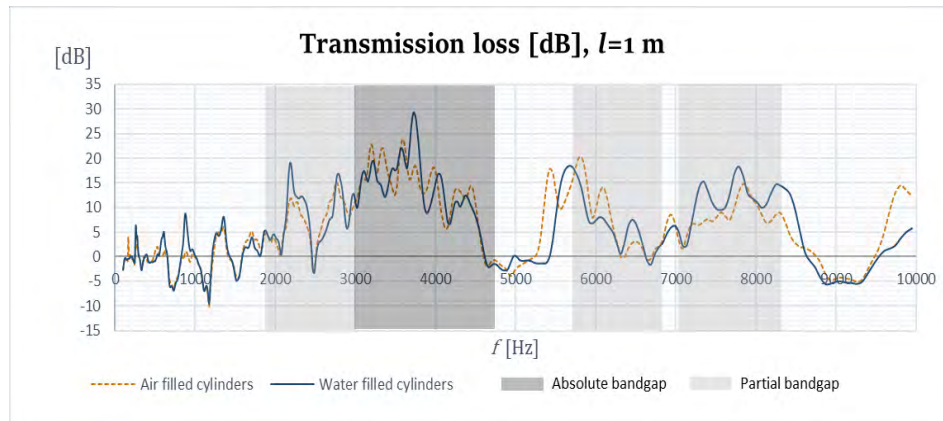


Figure 5. Transmission loss for  $l$  equal to 1 m



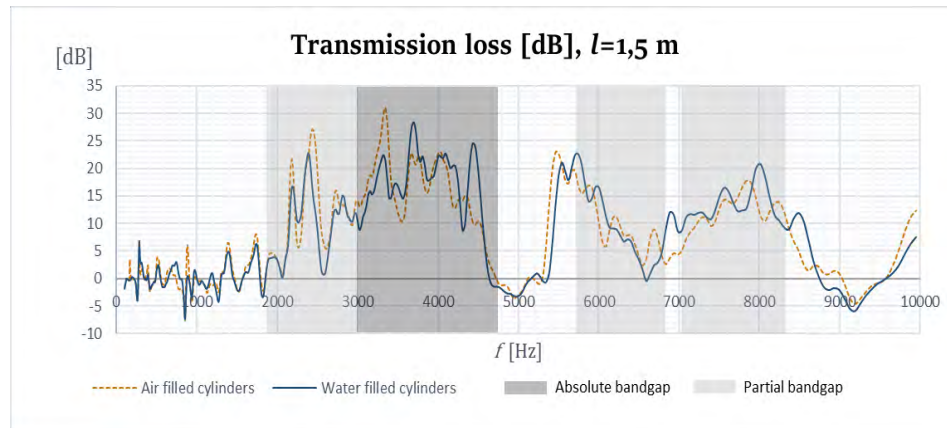


Figure 6. Transmission loss for  $l$  equal to 1.5 m

#### 4. Conclusions

The measurement results demonstrate a bandgap shift towards lower frequencies. This phenomenon is slightly more visible for the air-filled cylinders and becomes particularly evident for the absolute bandgap upper frequency (around 4.6 kHz), 2nd partial bandgap lower frequency (around 5.4 kHz) and narrow conductive band (around 6.6 kHz). The difference between air-filled cylinders and water-filled cylinders (around 5.4 kHz) allows supposing that this effect may occur due to the changed mass of the entire cylinder, and thus the creation of local resonance giving the impression of frequency shift. Other possible hypothesis explaining this effect is an impact of rubber impedance, which is not infinitely high. To confirm first or second hypothesis, in the future it would be necessary to compare the results with band structure of fluid-solid phononic crystal constructed of full EPDM cylinders.

The transmission loss curve is oscillating around 0 dB in the conductive band and reaches 20 dB in the presence of the bandgap (for single frequencies even 30 dB). Once more, this is an easily predictable effect, because the boundary between the scatterer and host medium was not completely hard. The lower difference in impedances must result in a weaker reflection of the wave between the scatterers.

Apart from the effect appearing around 5.4 kHz and described above, no evidence for localized modes was observed. Admittedly, there are single difference peaks (around 4 kHz for  $l = 1$  m graph and around 4.4 kHz for  $l = 1.5$  m), however, there is no constant difference for each microphone position. This allows to draw the conclusion that the observed differences do not result from the properties of the photonic crystal, but rather are the effect of the irregularity of the acoustic field in the measurement zone.

This work is an entry for further testing of viscoelastic materials in the application in photonic crystals and in acoustic metamaterials.



## References

1. Y. Fei, W. Fang, M. Zhong, H. Kadou et al., *Morphological Structure, Rheological Behavior, Mechanical Properties and Sound Insulation Performance of Thermoplastic Rubber Composites Reinforced by Different Inorganic Fillers*, *Polymers*, **10**(3) (2018) 276.
2. A. Khelif, A. Adibi, *Phononic Crystals, Fundamentals and Applications*, Springer, New York 2016.
3. P. A. Deymier, *Acoustic Metamaterials and Phononic Crystals*, Springer-Verlag Berlin Heidelberg, New York 2013.

