ARCHIVES OF ACOUSTICS
23, 1, 3–31 (1998)

HEARING DAMAGE FROM EXPOSURE TO MUSIC

A. JAROSZEWSKI, T. FIDECKI, P. ROGOWSKI

Fr. Chopin Academy of Music
(00-368 Warsaw, Okólnik 2, Poland)

Sound pressure levels and exposures in discotheques and youth clubs and during training sessions of music students were measured and analysed. Effects of exposure in the form of permanent and temporary threshold shift were determined in the samples of young discotheque attendants and in music students. The consequences of the threshold shift in the perception of pitch, loudness, and time are discussed.

1. Introduction

Music has been recognised as a source of acoustic trauma, and a danger of hearing loss from loud music has aggravated substantially over the past quarter of a century. Although live music can also be potentially dangerous to the hearing system, most traumatic effects observed in the samples of young population in Poland (and also western countries) were caused by music from portable cassette players, high power home sonic equipment and very high power electroacoustic systems in discotheques or pop and rock concert halls. Significant effects of overexposure were found in some music students due to very high sound pressure levels in training sessions.

Traumatic effects of acoustic overstimulation that have long been considered as a decrease in sensitivity only, manifest themselves in several psychophysical spaces. The decrease in sensitivity or hearing loss is thus usually associated with poorer frequency discrimination and frequency resolution. In the perception of intensity a distortion of loudness function, known as a recruitment is often observed. Severe distortions of signals caused by acoustic trauma result from the poorer perception of auditory events with time. In the proximity of an (elevated) hearing threshold and up to moderate sensation levels acoustic trauma often causes a lack of tonality and/or non-linear distortions. Cumulative effects of a pronounced acoustic trauma change the characteristics of signals incoming auditory pathways to such an extent that spoken messages cannot be understood.

The question of the danger of hearing loss in young people seems to be open to various interpretations and has recently brought conflicting answers. However, the findings from the audiometric laboratory examinations of statistically significant samples of young musicians in Poland show that 68% of them bear marks of acoustic overexposure while

50% of them show selective hearing loss of 20 dB or more in at least one ear. The observations seem to indicate that the danger of hearing loss from long term socially administered overexposures to music might have been underestimated yet.

2. SPL's and exposures

The deterioration of hearing from exposure to loud and to amplified music has been studied and described by large number of investigators over the past quarter of a century. Their efforts resulted in large supply of experimental data that has been critically reviewed by HUGHES *et al.* [40], WEST and EVANS [100], CLARK [13], DIBBLE [17] and in some measure by BRADLEY *et al.* [8], HELLSTRÖM and AXELSSON [39], AXELSSON *et al.* [1] and many others.

A number of investigators have measured and analysed data referring to music (noise) exposures, i.e. corresponding sound pressure levels and their frequency and time distributions and also to the duration and character of exposures.

Analysed were the data pertaining both to live and amplified pop/rock music, e.g. KOWALCZUK (1967), RINTELMAN and BORUS [83], RINTELMAN *et al.* [84], CABOT *et al.* [10], CLARK and BOHNE [14], CLARK [13], BORSCHGREVINK [7], ISING *et al.* [41], JAROSZEWSKI and RAKOWSKI [44], JAROSZEWSKI [56], and to live and recorded and/or amplified symphonic music e.g. LEBO and OLIFANT [65], AXELSSON and LINDGREN [2], WESTMORE and EVERSDEN [101], RABINOVITZ *et al.* [79], JANSSON and KARLSSON [42], FRY [29], SCHACKE [89], WOOLFORD *et al.* (1988), and CLARK [13], JAROSZEWSKI *et al.* [60].

Regrettably, the abundant published data from various authors show very large variance with reference to the sound pressure levels and their distribution in the frequency scale and in time and, consequently, also with reference to exposures.

However, almost all researchers report sound pressure levels that are potentially dangerous to hearing as for example: BIKERDIKE and GREGORY [5], $L_{\mathrm{Aeq}} = 88-113$ dB, DIBBLE (1988), $L_{\mathrm{Aeq}} = 94 - 99$ dB, while MAWHINNEY and MC CULLAGH [69] report SPL's in excess of 95–115 dB with peaks from approximately 105 dB to over 125 dB.

In investigations of sound pressure levels and their distribution in 10 Warsaw discotheques present authors found even larger values of SPL (JAROSZEWSKI *et al.* [57]). L_{Aeq} was found to be in the range from 90 to 116 dB. Long time average spectra in 1/3 octave bands in the tested discotheques are given in Fig. 1. SPL cumulative distribution functions for these discotheques are given in Fig. 2.

Impulsiveness of music noise tested, determined as a difference between the peak and equivalent level is given in Fig. 3 showing presence of peaks reaching 30 dB with median value of approximately 22 dB.

In investigation of sound exposures in Swedish symphonic orchestras JANSSON and KARLSSON [42] reported mean L_{Aeq} values from 89 dB up to 93 dB with maximum values reaching 98.6 dB. They have not observed peak values exceeding 125 dB. Similar data, L_{Aeq} from 85 dB up to 90 dB with weekly equivalent of 85 dB were reported by AXELSSON and LINDGREN [2]. SCHACKE [89] recorded sound pressure levels in orchestra

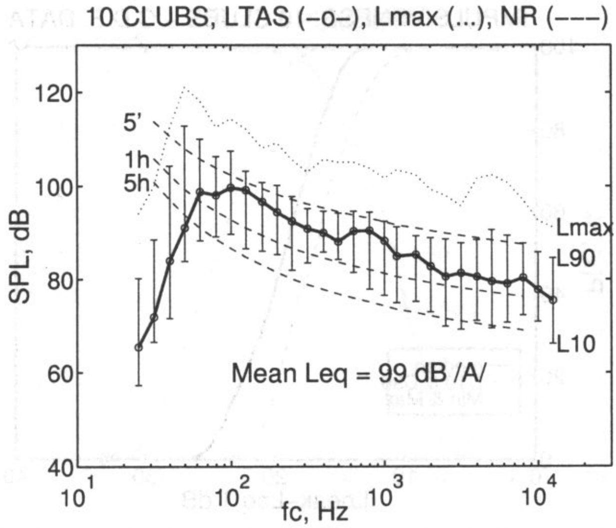


Fig. 1. Long time average spectra in 1/3-octave bands and noise rating curves for pop/rock music in 10 Warsaw discotheques.

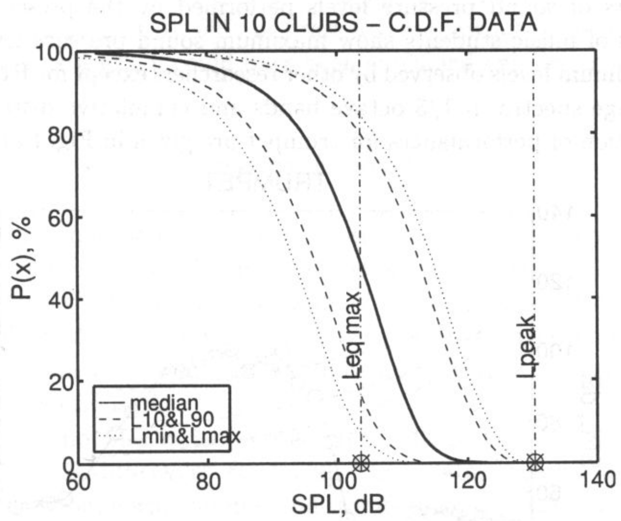


Fig. 2. SPL cumulative distribution functions for pop/rock music in 10 Warsaw discotheques.

of the Deutsche Oper Berlin and found average A levels for brass ranging from 87 dB to 96 dB with peaks reaching 122 dB, and average levels for woodwinds varying between 88 dB and 97 dB with peaks reaching 117 dB. In the report by SCHACKE [89] L_{Aeq8h} for wind instruments was determined at 87.7 dB which is about twice as much as maximum permissible exposure according to German regulations. Alarming data on L_{Aeq} values were reported by ROYSTER *et al.* [87]. They found L_{Aeq} values ranging from 74.7 to 94.7 dB with peaks in the range from 112 to 143 dB.

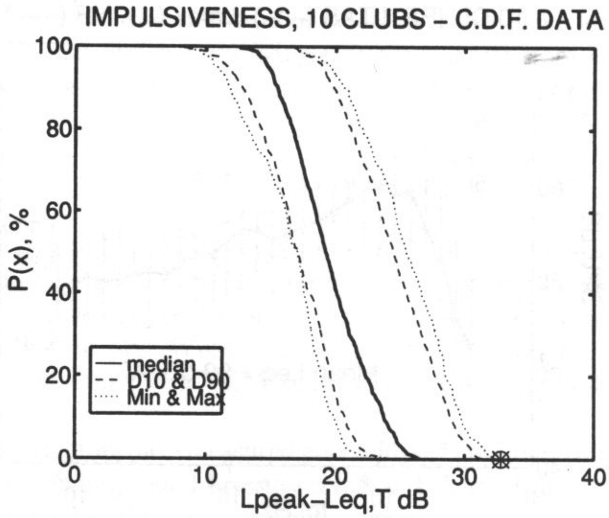


Fig. 3. Impulsiveness as a difference between peaks and equivalent SPLs of the tested discotheque music.

Measurements of sound pressure levels performed by the present authors during training sessions of music students show maximum sound pressure levels substantially higher than maximum levels observed by other researchers except for ROYSTER *et al.* [87]. Long time average spectra in 1/3 octave bands and cumulative distribution functions for a short selection of performances for trumpet are given in Fig. 4 and Fig. 5.

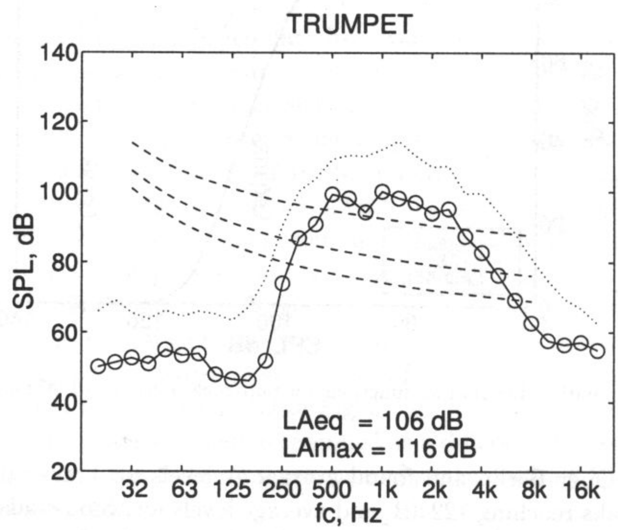


Fig. 4. LTAS in 1/3-octave bands of trumpet sound during daily practice of music student. Noise rating curves for 5 min, 1 hr and 5 hrs are also shown on the diagram.

Present findings indicate also that the exposures experienced by music students during their training sessions are far in excess relative to the permissible doses. Predictably,

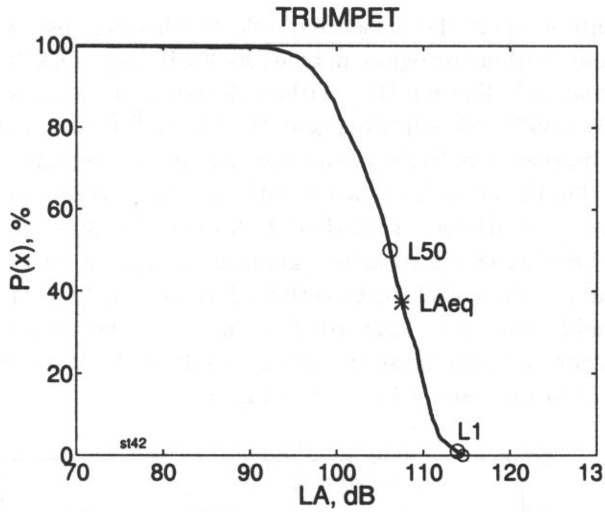


Fig. 5. SPL cumulative distribution functions of trumpet sound during practice hours of music student.

such exposures lead to aggregation of NIPTS acquired over a long period of time and as such should not be ignored. In 50% of brass and woodwind student players PTS of 10 dB to 25 dB was found which could have been explained only by acoustic trauma.

3. Hearing thresholds

In audiometric examination of classical musicians AXELSSON and LINDGREN [2], WESTMORE and EVERSDEN [101], RABINOWITZ *et al.* [79], KARLSSON *et al.* [63], WOOLFORD [105], JOHNSON *et al.* [61, 62], JANSSON *et al.* [43], OSTRI *et al.* [75], ROYSTER *et al.* [87], and others, all have found audiometric patterns corresponding to the noise induced hearing loss. The audiograms showed notches, mostly at frequency 6 kHz in 30% to over 50% of the tested sample.

The depth of these notches varied between HTL 10 dB (OSTRI *et al.* [75], ROYSTER [87]) and 20 to 25 dB (RABINOWITZ *et al.* [79]. The greatest hearing loss was found in musicians playing bassoon, horn, trumpet and trombone (e.g. AXELSSON and LINDGREN [23]). However, some investigators e.g. WESTMORE and EVERSDEN [101], AXELSSON and LINDGREN [2], have declared that only "slight degree of hearing loss was found in the average hearing thresholds", even that they also found notch shaped audiograms with the dip at 6 kHz in the tested samples.

To relate the exposure data from loud and amplified pop/rock music to the damage of hearing many authors measured and analysed the permanent and temporary threshold shift in music performers, personnel and attendants to discotheques and youth clubs e.g. LIPSCOMB [66], SKRAJNAR [92], CATALANO and LEVIN [11], FEARN [19], FEARN and HANSSON [20, 21], CLARK [13], JAROSZEWSKI and RAKOWSKI [44], JAROSZEWSKI *et al.* [57].

Many investigators report the presence of selective hearing loss of various depths at 6 kHz. However, some authors recognise notches 20-30 dB deep at 6 kHz as "not exceeding limits of normal hearing" (DIBBLE [17]). Others declare in such cases that "the hearing losses are relatively small" (MAWHINNEY and MC CULLAGH [69]) or show little concern that the notches observed in individual data were lost in the averaging procedure (WEST and EVANS [100]). On the other hand, some authors express their serious concern about the notches that are 7-10 dB deep at 6 kHz (FEARN and HANSON [20, 21], FEARN [19]).

In audiometric testing of statistically significant sample of music students, present authors arrived at the data that seem rather alarming. In 68% of a sample of 214, permanent threshold shift of at least 10 dB or more in one ear at 6 kHz was found which clearly indicates a music noise trauma. A result of the statistical analysis of the audiometric data from this sample is given in Fig. 6.

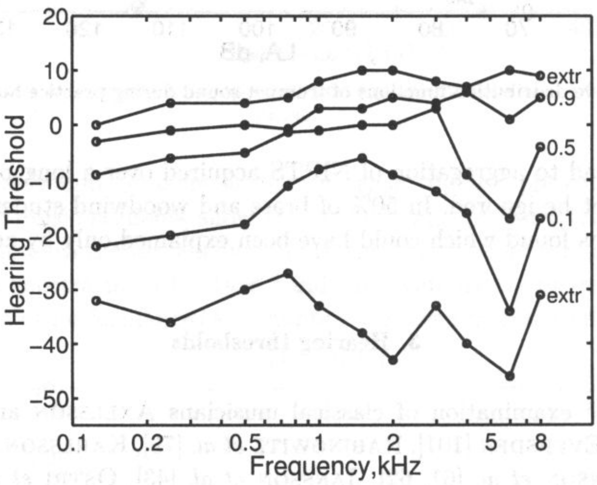


Fig. 6. Hearing threshold data. Sample of 214 subjects aged 16 to 27.

It has been learned from interviews that in all cases of notch shaped PTS the subjects were attendants to discotheques and/or used portable cassette or CD players and/or used home audio equipment at very high level.

Audiometric data from a group of 14 attendants to discotheques show similar results of overexposure to music in the form of the notches at 6 kHz, 20 dB deep, and a considerable amount of temporary threshold shift TTS2 measured immediately after cessation of over 6 hrs exposure in the discotheque, Fig. 7.

Quite dramatic audiometric data were obtained from 4 music performers in one of youth clubs. For example these data show hearing loss of 20 to 50 dB in both ears of one of the subjects tested while the temporary threshold shift measured 2 min. after 5 hrs exposure reaches from 50 to 81 dB, Fig. 8.

Also, both hearing loss and temporary threshold shift occupy wide range of frequencies. Maximum sound pressure levels in this night club were reaching 125 dB (JAROSZEWSKI and RAKOWSKI [44]).

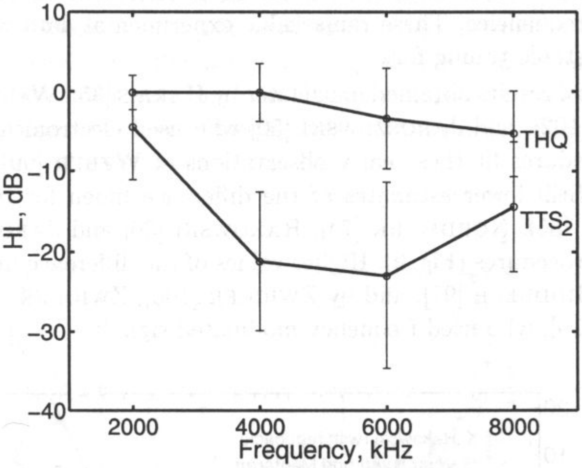


Fig. 7. Hearing threshold in quiet (THQ) and temporary threshold shift (TTS2) in 14 discotheque attendants after 6 hrs of exposure to music.

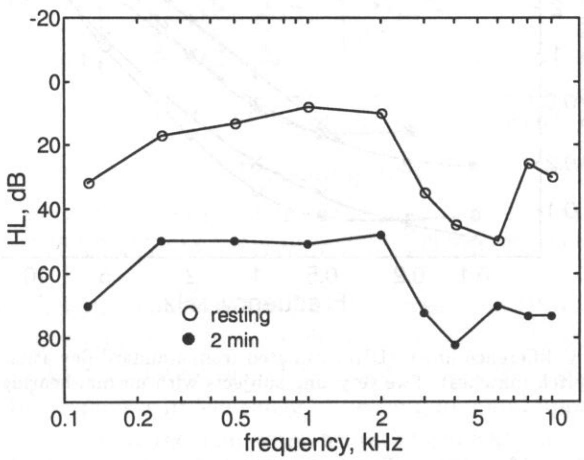


Fig. 8. Hearing threshold in quiet (open circles) and temporary threshold shift (closed circles) in one pop/rock musician after 5 hrs of exposure.

The data obtained seem to indicate that the hazard of listening to very loud music from high power electronic equipment in night clubs or discotheques or from low-power portable players delivering very high sound pressure levels may be substantially larger then it is generally assumed.

4. Frequency discrimination

It has long been known that the frequency discrimination in the hearing system is astonishing. Early observations by WEBER [99] and PREYER [78] showed that very experienced musicians can discriminate 64 to 83 pitches in a semitone in the middle range of audible frequencies. These remarkable experimental data were obtained with the use of an adjustable tuning fork.

The experimental results obtained much later by HARRIS [35], WALLISER [98], MOORE [70], WIER *et al.* [102], and JAROSZEWSKI [50] who used electronically controlled constant stimuli procedures fit these early observations of WEBER [99] and PREYER [78] surprisingly well. Still lower estimates of the difference limen for frequency were obtained by RITSMA [85], NORDMARK [74], RAKOWSKI [80] and JAROSZEWSKI [50] who used adjustment procedures (Fig. 9). Higher values of the difference limen were obtained by SHOWER and BIDDULPH [91], and by ZWICKER [106], ZWICKER and FELDTKELLER [107], and FASTL [19], who used frequency modulated signals.

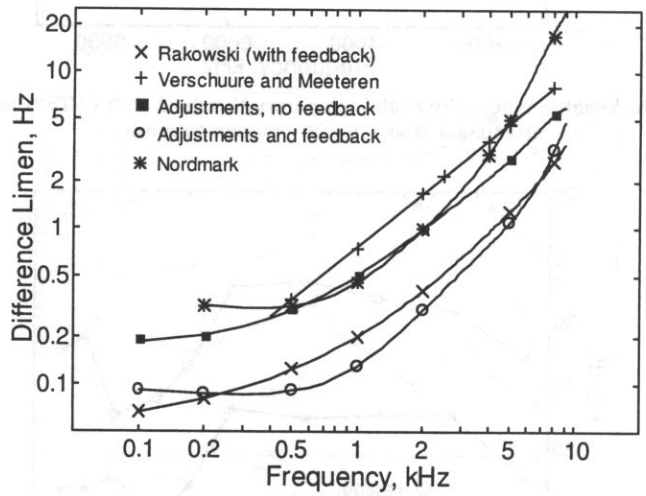


Fig. 9. Frequency difference limen (DL) estimated from standard deviation of adjustments (pitch matches). Twelve young subjects with normal hearing.

Numerous studies of frequency discrimination in hearing impaired subjects show that in majority of cases with sensorineural hearing loss frequency difference limens for pure tones are larger than in normal hearing subjects. In the data from e.g. TYLER, WOOD and FERNANDES [96], HALL and WOOD [32], FREYMAN and NELSON [27], which were obtained from subjects with comparatively large hearing loss of 30 to 60 dB and relatively flat over the range of audiometric frequencies, the DL estimates are on average several times larger than in normal hearing subjects. However, DL values obtained by these authors are rather large both for normal and for hearing impaired subjects. More recent data from FREYMAN and NELSON [28], show also several times worse DL's in hearing impaired than in normal hearing subjects, but the DL estimates are also large (e.g. 4.5 Hz/1200 Hz and 15 Hz/1200 Hz for normal and hearing impaired correspondingly) while the group of subjects is less consistent relative to the amount of hearing loss over the frequency scale.

Much lower DL values which were found in normal hearing young musicians by RAKOWSKI [80] and JAROSZEWSKI [50], expressed by approximately 0.1 Hz for frequencies below 1 kHz indicate the difference in hearing ability in musicians and in non musicians. Poorer DL estimates were found in elder but musically educated and musically trained subjects outside the range of frequencies at which hearing loss was measured in them, JAROSZEWSKI and FIDECKI [55], Fig. 10.

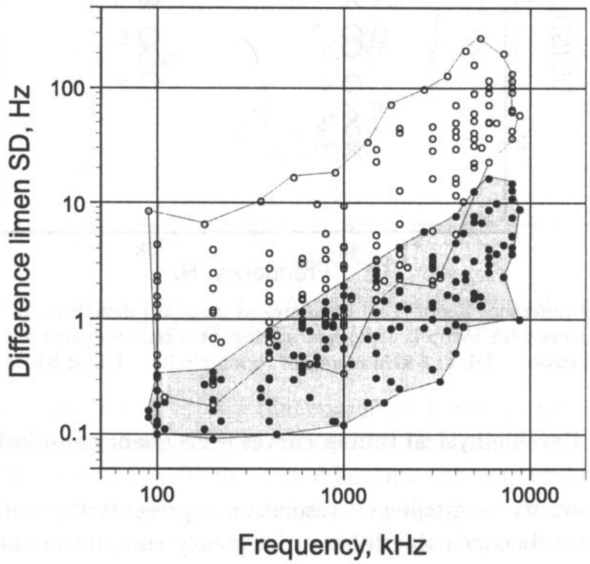


Fig. 10. Scatter diagram of difference limen (DL) for 12 subjects with normal hearing (closed circles) and for 12 subjects with large hearing loss of 60-70 dB in high frequency range (open circles).

On the other hand, larger contrast was found between the DL estimates in normal hearing and in hearing impaired subjects reaching 1 : 30 to even 1 : 90 in extreme cases, while it amounted to approximately only 1 : 10 in the data from other investigations. Also, the DL was always poorer in the range of hearing loss and approximately normal outside this range, which is in agreement with the data from FREYMAN and NELSON [28].

Severe loss in pitch discrimination ability pertains to the cases of large hearing loss which was observed in musicians active in pop and rock music performances (as shown in section 3) and in those involved in operation of discotheques. Less severe effects of hearing loss on pitch discrimination were found in musicians with selective hearing loss at 6 kHz. (so called "notch") of moderate and larger depth of 20 to 40 dB, Fig. 11.

While the difference limen is within normal limits outside the frequency of hearing loss, it is from 2 to 6 times worse at the frequency of impairment. In majority of cases the worse DL was associated with larger hearing loss. However, in some cases of notch shaped audiograms the DL estimates were approximately normal at the frequency of loss.

The present observations show that even moderate selective hearing loss acquired from overexposure to music often affects also other important characteristics of the hearing system and this effect on the DL is also selective.

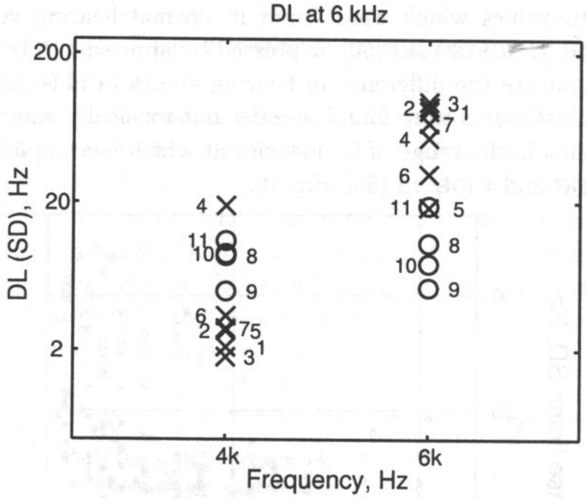


Fig. 11. Frequency difference limen (DL) estimates as standard deviations of adjustments (pitch matches) for subjects with V-dip (notch) hearing loss at 6 kHz. Subjects are depicted by their numbers. Crosses – DL at 6 kHz increased, open circles – DL at 6 kHz unaffected.

5. Psychophysical tuning curves – Frequency resolution

Frequency selectivity or frequency resolution represents the ability of the hearing system to distinguish between the different frequency components of a complex sound. Although different measures and procedures were used to determine the frequency resolution as described by FLORENTINE *et al.* [25], TYLER *et al.* [95], MOORE and GLASBERG (1987), LUTMAN and WOOD [68], LUTMAN *et al.* [67], COX and ALEXANDER [15], the psychophysical procedure described by CHISTOVICH [12] and SMALL [93] seems to be most widely used. In this procedure a faint test tone signal at fixed frequency and level is masked by variable frequency masker. The level of masker necessary to mask a test tone as a function of frequency results in the so called psychophysical tuning curve.

Psychophysical tuning curves measured in normal hearing subjects have extremely steep flanks, reaching 3×10^3 dB/oct or 3 dB/Hz in the upper flank when measured in forward masking and with off-frequency listening. The measure of frequency selectivity or frequency resolution is usually represented by the slope of the upper flank of the tuning curve or by a Q value defined as the center frequency divided by the bandwidth of the tuning curve at certain arbitrary level above the level of the test tone.

The first demonstrations of extremely steep flanks of psychophysical tuning curves in normal hearing subjects in forward masking (Fig. 12) were given by JAROSZEWSKI and RAKOWSKI [59] and JAROSZEWSKI [45, 46] and by MOORE [71].

Contemporarily, psychophysical tuning curves were measured in normal-hearing and hearing-impaired subjects by e.g. FLORENTINE *et al.* [25], TYLER *et al.* [95, 96] showing, regardless of the procedure, pronounced differences in frequency selectivity between these groups. However, these authors, similarly as later LUTMAN and WOOD [68] and LUTMAN *et al.* [67] used but continuous maskers of various characters. Since, their data do not

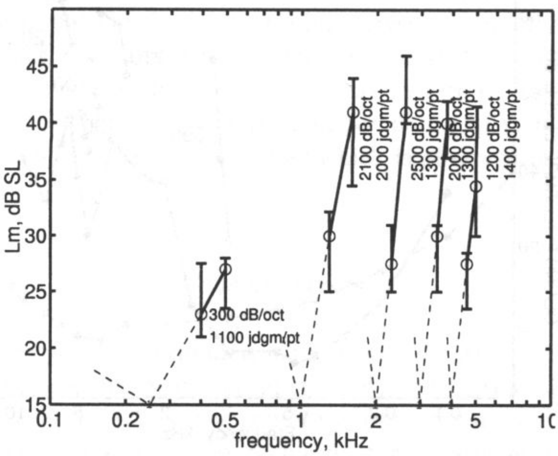


Fig. 12. Psychophysical tuning-curves in forward masking in normal hearing subjects. (2-IFC procedure. Lm = 15 dB SL).

reflect the utmost frequency selectivity that manifests better in forward masking. Using pulsed test tone and continuous masker, FLORENTINE *et al.* [25] for example, reported Q values at 4000 Hz of 6.06 to 6.92 in normal hearing subjects and only 3.0 in subjects with noise induced hearing loss and corresponding steepness of the upper flank of approximately 180 dB/oct. for normal hearing subjects, and approximately 100 dB/oct for hearing impaired. The Q value for young normal hearing musicians measured in forward masking reached 16.7 to 29.0, JAROSZEWSKI ET AL. [57], which is in agreement with the data from WIGHTMAN *et al.* [104], and the slope of upper flank 1.2×10^3 dB/oct, JAROSZEWSKI and RAKOWSKI [59], JAROSZEWSKI [45, 46].

In musicians with large sensorineural noise induced hearing loss a substantial decrease of both the steepness of tuning curves and of the Q values was observed, JAROSZEWSKI [47]. Contrary to the frequency discrimination estimates which were decreased in the range of hearing loss and preserved at normal level where threshold hearing level was normal, the decrease of Q values and of the steepness of upper flank occured in the frequency range of large hearing loss and also outside these frequencies, Fig. 13. These data are consistent with the data reported by FLORENTINE *et al.* [25], and early data from WIGHTMAN *et al.* [104]. However, conflicting data have been reported in studies, which demonstrated frequency discrimination performance at normal level in subjects, who were performing abnormally in frequency resolution, e.g. WIGHTMAN [103], TYLER *et al.* [96].

The psychophysical tuning curves measured in young musicians with only relatively small and/or moderate selective sensorineural hearing loss at 6 kHz, show somewhat lower values of Q and less steep upper flank at the frequency of loss (6 kHz), than above and below this frequency, JAROSZEWSKI *et al.* [52]. This observation indicates that even small amount of selective hearing loss affects also frequency resolution similarly as frequency discrimination. However, the performance of subjects in these tasks does not

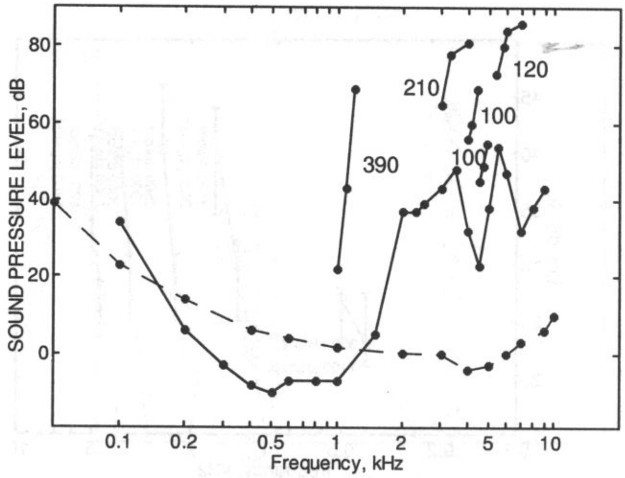


Fig. 13. Slopes of the upper flank of the psychophysical tuning-curves in hearing impaired with noise-induced sloping high frequency hearing loss as indicated by heavy line. Dashed line – normal threshold in quiet.

conform to a rigid pattern as follows from the evidence reported by e.g. WIGHTMAN *et al.* [107], FLORENTINE *et al.* [25], JAROSZEWSKI [47], JAROSZEWSKI *et al.* [57].

As indicated by TYLER *et al.* [96] frequency resolution and frequency discrimination may operate on the basis of different mechanism: frequency discrimination may depend on temporal coding while frequency resolution on place coding. FLORENTINE *et al.* [25] observed that, cit.: "the most sensitive measures of reduced frequency selectivity are the Q values of the tuning curves", even that they used pulsed test tone and continuous

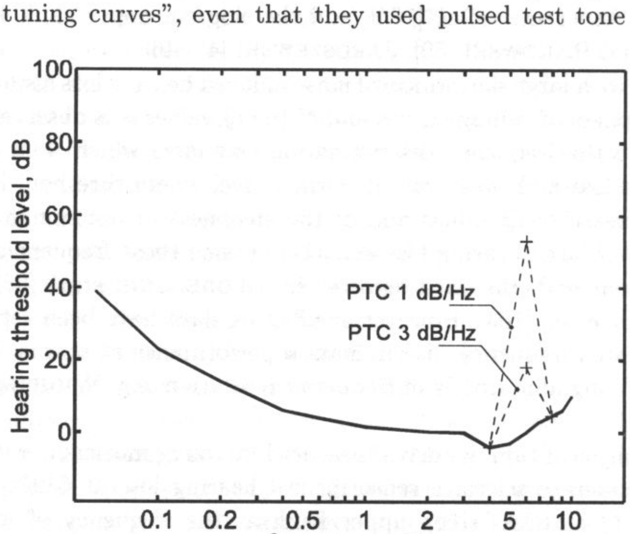


Fig. 14. Hearing threshold in quiet for subjects with V-dip (notch) hearing loss at 6 kHz and the slopes of the upper flank of psychophysical tuning-curve for V-dip 50 dB and 15 dB.

masker. The data from WIGHTMAN *et al.* [104] and JAROSZEWSKI [47], JAROSZEWSKI *et al.* [54], seem to indicate that Q value of the tuning curves measured in forward masking may be still more sensitive measure which "not only reveals even small amounts of hearing impairment, it also provides a measure of the degree of cochlear impairment" (after FLORENTINE *et al.* [25]). The slopes of upper flanks of psychophysical tuning curves are also affected by selective hearing loss which is observed in many subjects at 6 kHz, see Fig. 14.

6. Constant errors

Constant errors in frequency discrimination received little attention of the researchers exploring operation of the auditory system in spite of the fact that their existence was discussed already by FECHNER [22] at the end of the past century. Later constant errors were revisited by KELLOG [64] in a remarkable doctoral dissertation cited up to the present day.

Constant errors measured in normal hearing subjects, were present in the results of pitch adjustment to unison in the experiments of RAKOWSKI and HIRSH [81, 82], and more recently by the senior author of the present report, JAROSZEWSKI [49, 50]. It is interesting to note that small errors were observed in the procedures with feedback and large ones in the procedures without it.

One of the better illustration of the constant error manifestation are response density functions in the adjustment procedure. These functions, represent the ensembles of the values of frequency in which in the experimental run the subject decided that the pitch of the adjusted (matched) signal is too large or that it is too small, JAROSZEWSKI [52], Fig. 15 a, b, c.

The widths of these functions, which are usually normal in their nature, determine the dispersion of decisions or standard deviation in the final result. The ratio of one of the density functions over the sum of both is by definition the psychometric function, Fig. 15 c, while their situation along the frequency scale is the constant error or systematic time error.

As demonstrated earlier, JAROSZEWSKI [53, 54] constant errors are small in normal hearing subjects and comparatively large in hearing impaired with noise induced sensorineural hearing loss. Also constant errors reflect the amount of cochlear impairment being relatively small at frequencies where normal hearing is preserved and increasing dramatically at frequencies of hearing loss. The CE often increases by a factor of 10 to 100, see Fig. 16.

In the case of noise induced high frequency sloping hearing loss e.g. such as in Fig. 16 a, it means that while 1 kHz is perceived by the subject at its normal pitch, the pitch of 5.5 kHz is shifted by 100 Hz (i.e. 1/5 of a semitone). It seems to be interesting to note that in some cases, in the range of hearing loss, constant error changes abruptly its sign as demonstrated in the examples in Fig. 16 a, b. The intersubject change of sign in constant errors was observed in the earlier experiments by RAKOWSKI and HIRSH [81, 82]. The intersubject abrupt change of sign in constant errors from frequency to fre-

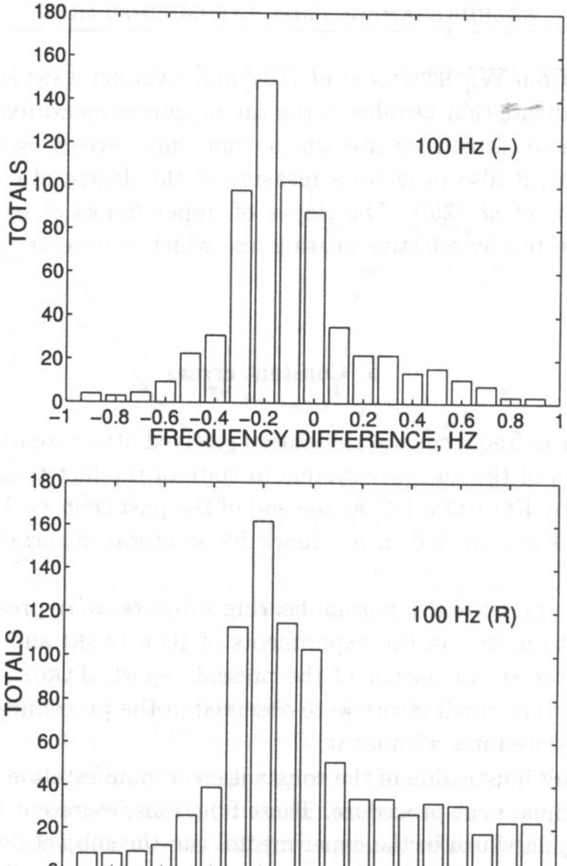
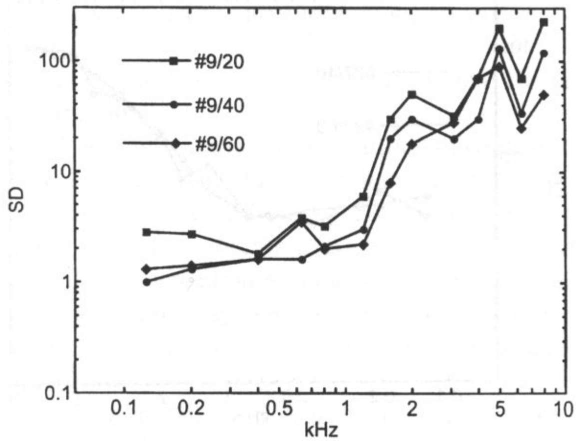
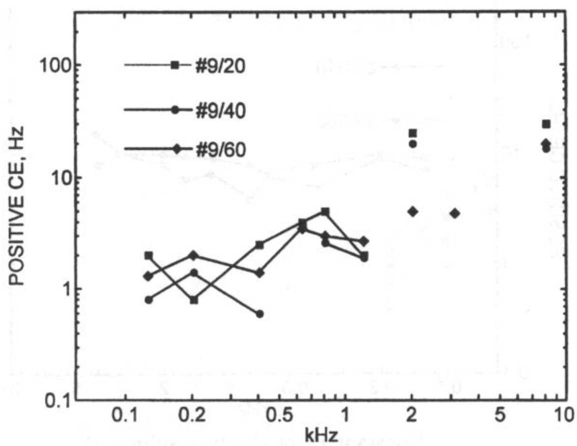


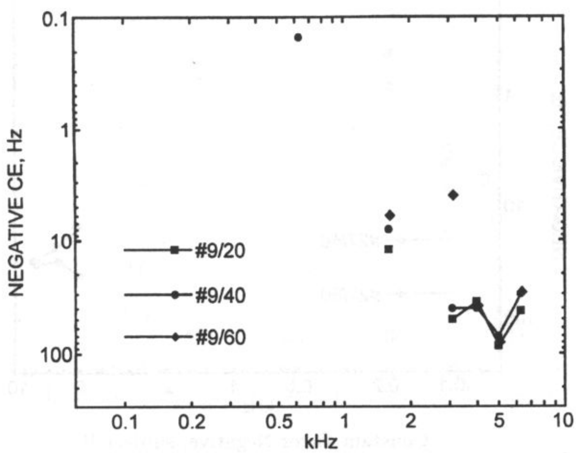
Fig. 15. a) Response density function for 100 Hz averaged over 12 subjects. Total number of "moves" made from the situation in which the variable signal was estimated as "too low" in pitch. b) Response density function for 100 Hz averaged over 12 subjects. Total number of "moves" made from situation in which the variable signal was estimated as "too low" and "too high" in pitch. c) Psychometric function derived from the response density functions given in Fig. 15 a and 15 b.



Frequency difference limen (DL), subject A

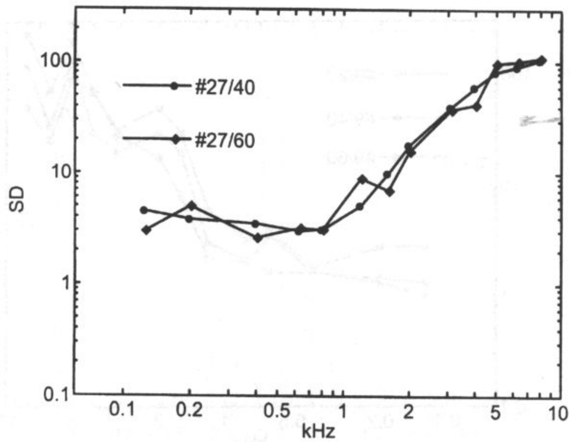


Constant Error Positive, subject A

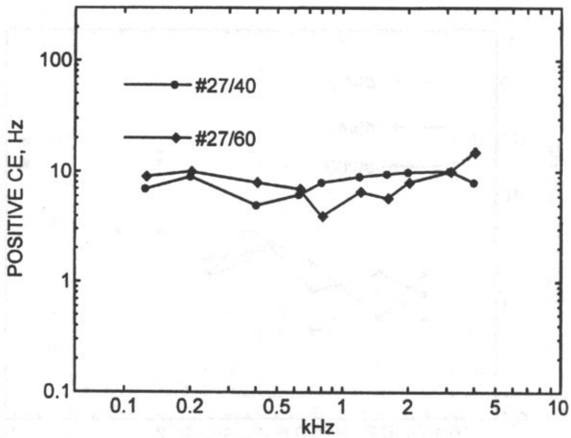


Constant Error Negative, subject A

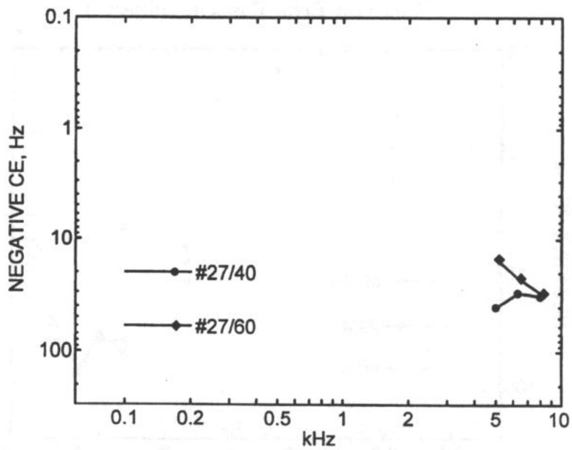
Fig. 16.



Frequency difference limen (DL), subject B



Constant Error Positive, subject B



Constant Error Negative, subject B

Fig. 16. Frequency dependence of the frequency difference limen DL and of the constant error CE for two subjects A and B, with large high frequency sloping hearing loss and with abrupt change in CE sign. The data for sensation levels 20, 40 and 60 dB for subject A and for 40 and 60 for subject B.

quency was reported only recently, JAROSZEWSKI [53, 54], and the origin of this strange phenomenon is totally unclear.

In all cases of normal hearing an absolute magnitude of constant errors was strongly positively correlated with the difference limen estimator for frequency. The same dependency was observed in cases of impaired hearing of sensorineural nature both in cases of deep impairment and in relatively moderate hearing loss, JAROSZEWSKI [53, 54]. Therefore, constant errors should be recognised as indicators of cochlear impairment reflecting one more deficiency of the auditory system resulting from, or accompanying the noise induced decrease of sensitivity. This deficiency shifts the pitches of perceived signals and it may add to their distortion.

In cases of presently often measured in young musicians and non-musicians selective noise induced hearing loss at 6 kHz ("notch"), the constant errors confirm the principle observed in large hearing loss, being small outside the frequency of hearing loss and larger by a factor of 3 to 8 at the frequency of impairment, Fig. 17.

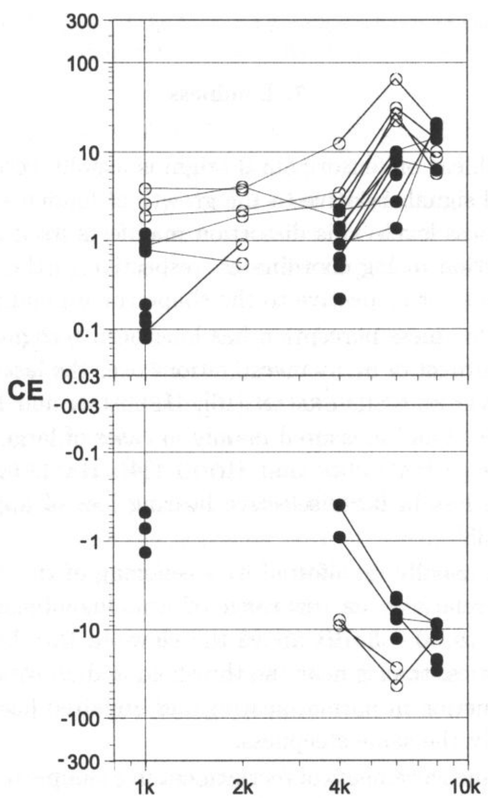


Fig. 17. Scatter diagram of constant errors CE data in subjects with V-dip ("notch") hearing loss at 6 kHz, at the frequency of loss and outside of this region.

However, this is true only for the "notches" or selective hearing loss exceeding HL 20 dB. For smaller selective hearing loss of approximately 10 dB in 60% of cases the CE was found to be the same or comparable to the CE outside hearing loss.

Whilst it is doubtful if relatively small or moderate constant error can affect the perception of pitch in normal human communication i.e. speech, music or warning signals in a detectable degree, it doubtless reflects the state and the operation of the auditory system which is abnormal. Even small deviations from what should be normal must be regarded as an evidence of the initiated process of degradation.

Statistical analysis of the large data base for frequency discrimination and constant errors indicates that both are independent from the sensation level within the whole range of frequencies investigated i.e. from 100 Hz to 9 kHz. This result is statistically significant at level of 0.01 for sensation levels 20 and 40 dB SL and at level 0.03 for sensation level 60 dB, which is in agreement with the earlier observations. However, significant intrasubject and intersubject variability was found with reference to the magnitude of constant error and its dependence on frequency. Analysis of the data base indicated that the dependence of constant error on frequency reflects individual characteristics of the auditory system, JAROSZEWSKI [53, 54].

7. Loudness

Hearing loss of cochlear or sensorineural origin is usually accompanied by distortion of perception of sound signals relative to the growth of loudness as a function of sound pressure level or loudness level. This distortion manifests itself often in increase of the slope of loudness function in log coordinates, respective to the slope observed outside the region of hearing loss or respective to the shape corresponding to normal hearing.

Such distortion of loudness perception has long been recognised as the so called recruitment and was a subject of many investigations over the last 70 years (e.g. FOWLER [26], BÈKÈSY [4], DAVIS and SILVERMAN [16], HELLMAN and MEISELMAN [37]). The recruitment was observed and measured mainly in cases of large, over 40 dB, and widespread hearing loss (e.g. HALLPIKE and HOOD [34], HALLPIKE [33], HELLMAN and MEISELMAN [37, 38]), less in large selective hearing loss of approx. 70 dB, HELLMAN and MEISELMAN [37, 38].

The recruitment is usually manifested by steepening of the loudness function in log coordinates only over relatively narrow range of stimulus intensity, which spreads from approx. 5 – 15 dB up to 30 – 35 dB above the elevated threshold. Below this limited range of sound intensities, that is near the threshold and above 30 dB over the elevated threshold, loudness function in normal-hearing and impaired-hearing with sensorineural hearing loss has usually the same steepness.

The data from the measurements of recruitment in a sample of 149 hearing – impaired subjects with noise induced hearing loss show that in the range of recruitment the slope of loudness function is always in excess of approx. 0.5 and reaches 6.0 in some cases, showing substantial intersubject variability, Fig. 18.

In normal hearing, or in hearing impaired subjects in the frequencies where normal hearing was preserved, the slope of loudness function is usually lower than 1.5 and sometimes is as low as 0.5, as reported by JAROSZEWSKI *et al.* [58]. These data are rather

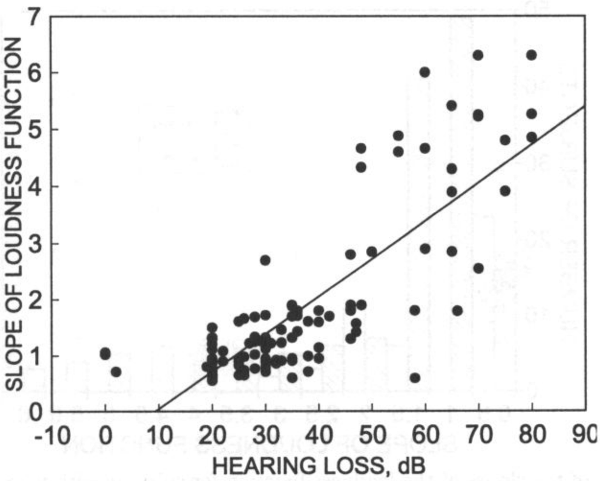


Fig. 18. Scatter diagram of the slope of the loudness function relative to the degree of hearing loss in 149 subjects with noise-induced hearing loss.

in agreement with the recent data from HELLMAN [36] and HELLMAN and MEISELMAN [37, 38], and also with the early data from HALLPIKE [33].

In normal hearing subjects and in hearing impaired outside of hearing loss, inter-subject variability of the shape of loudness function is much smaller than in hearing impaired with wide-spread hearing loss or in hearing impaired with selective hearing loss, at frequencies of loss.

Statistical analysis of the data from the sample of 76 subjects with high frequency sloping noise induced hearing loss clearly indicates that steeper slopes of loudness function are strongly positively correlated with the amount of hearing loss. The same holds for the slopes of loudness function measured in subjects with selective V-dip or notch shaped hearing loss. While the slope is correlated with the amount of hearing loss, the variance of slope values observed reflect the variability of the depth of hearing loss. It is obvious thus, why the variance of the slopes for normal hearing that is also demonstrated in histogram in Fig. 19 is much smaller than in hearing impaired.

An example of typical behaviour of loudness function in musician with large sloping high frequency hearing loss is given in Fig. 20, and similar behaviour was observed in 50 subjects in a sample of 76 with high frequency sloping hearing loss.

Typical slopes of the loudness function in normal hearing subject are given for comparison in Fig. 21.

Individual data on recruitment in selective V-dip music noise induced hearing loss in young musician are given in Fig. 22, showing significant change of the slope of loudness function at the frequency of hearing loss and at sensation level of 15 dB. At sensation level of 50 dB and at the same frequency of hearing loss (4 kHz) the slope of loudness function is less than at adjacent frequencies at both sides of the impairment.

This kind of recruitment was observed in 30% of the sample of 76 examined ears with noise induced selective hearing loss at 6 kHz or at 4 kHz.

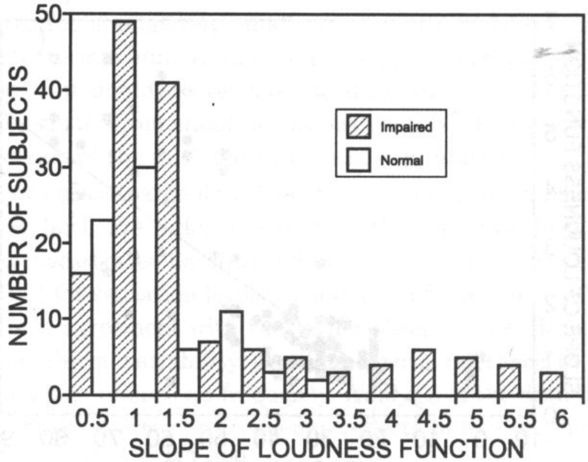


Fig. 19. Distribution of the slopes of the loudness function for subjects with noise induced hearing loss and with normal hearing.

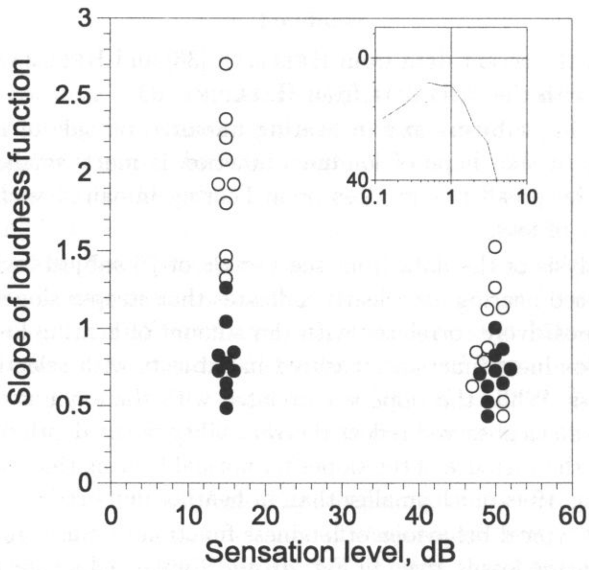


Fig. 20. The slope of the loudness function in musician with sloping high frequency hearing loss.

In some cases of selective noise induced hearing loss measured in young musicians a distortion of loudness function of the nature reversed respective to the recruitment was observed. This type of distortion is manifested, similarly as the recruitment, in the range of approximately 5 dB up to 35 dB SL above the elevated threshold. This type of distortion of the loudness function was also observed in other laboratories and is known as a "decruitment". Typical example of "decruitment" accompanying music noise induced hearing loss of V-dip "notch" type, 35 dB deep at 4 kHz is represented in Fig. 23.

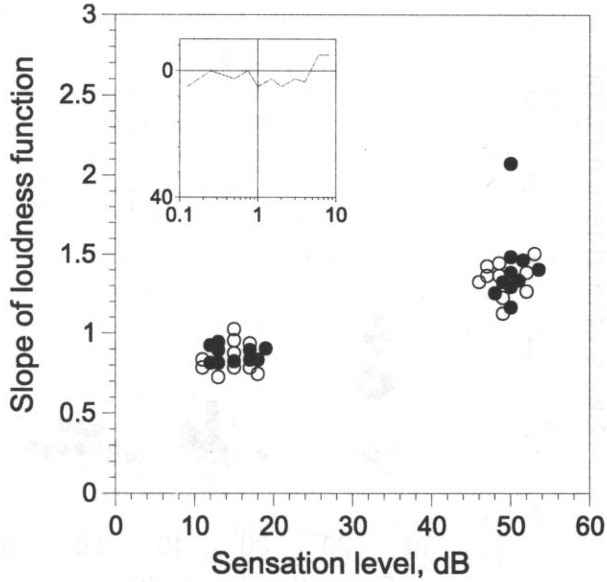


Fig. 21. Typical slope of the loudness function in normal hearing subject.

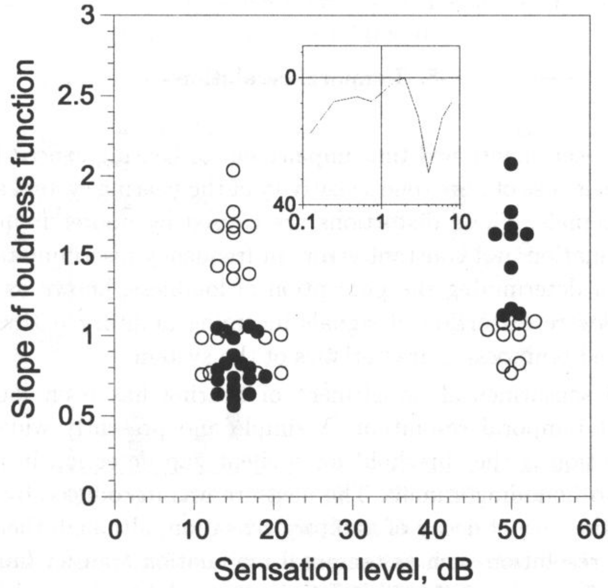


Fig. 22. Recruitment in the selective V-dip ("notch") music-induced hearing loss.

While flattening of the loudness function is largest at 15 dB, in a smaller degree flattening is also present at 35 dB and at 50 dB. Unfortunately, no data are available for sensation levels larger than 50 dB. It should also be observed that in some cases of recruitment at 15 dB SL a decruitment was measured at 50 dB SL.

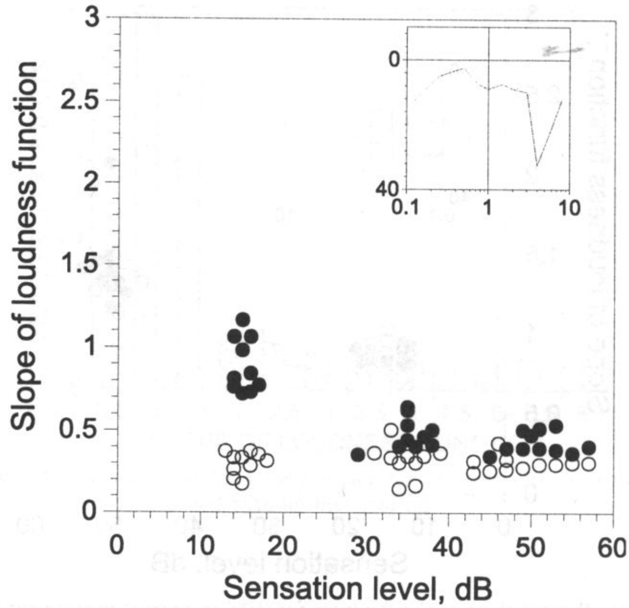


Fig. 23. "Decruitment" in the selective V-dip ("notch") music-induced hearing loss.

8. Temporal resolution

It has already been mentioned that impairment of hearing caused by acoustic overstimulation results in loss of pure tone sensitivity of the hearing system and in distortions of the incoming signals. These distortions are caused by poorer frequency resolution, frequency discrimination and constant errors in frequency discrimination and by deformation of function determining the perception of loudness, known as recruitment and as decruitment. Severe distortion of signals incoming auditory pathways may also be produced by altered temporal characteristics of the system.

Noise induced sensorineural impairment of hearing has been found by some researchers to affect temporal resolution. A simple and presently widely used measure of temporal resolution is the threshold for a silent gap detection in a continuous signal or between two bounding stimuli. The measure was introduced by PENNER [76] for determination of the rate of decay of auditory sensation, although there are other measures of temporal resolution such as temporal modulation transfer function introduced by VIEMEISTER [97], or discrimination of time-reversed signals, used by RONKEN [86], or gap difference limen, RUHM *et al.* [88], TYLER *et al.* [95].

Many researchers e.g. BOOTHROYD [6], TRINDER [94], FITZGIBBONS and WIGHTMAN [23, 24], BUUS and FLORENTINE [9], GLASBERG *et al.* [30] have found that gap detection thresholds are larger in subjects with cochlear impairment than in normal hearing. On the other hand MOORE and GLASBERG [72], MOORE *et al.* [73] observed that, for sinusoids at least, gap detection thresholds in normal and in impaired hearing are about the same, while TYLER *et al.* [95] obtained worse scores in gap detection for some hearing impaired listeners.

With reference to these data GLASBERG *et al.* [30], MOORE and GLASBERG [72], and PLACK and MOORE [77] declared that poorer gap detection in hearing impaired could result from distortion of loudness perception. This interpretation was supported by experiments with simulated effect of loudness recruitment, GLASBERG and MOORE [31]. On the other hand, PLACK and MOORE [77] found reduced temporal resolution in impaired hearing in one of the three subjects tested.

Gap detection in narrow band noise but without notched noise masking at center frequency performed with normal hearing music students shows a little lower gap detection thresholds than those reported by SHAILER and MOORE [90]. The experiment was performed at only 4, 6 and 8 kHz to avoid the influence of fluctuations of noise and at sensation level of 20 dB SL. Thresholds in 6 normal listeners were almost all lower than 3 ms.

The same experiment was performed with 9 hearing impaired musicians with severe noise induced high frequency sloping hearing loss of 50 to 65 dB for frequencies above 1 kHz. They had normal hearing at frequencies lower than 1 kHz except for one who had normal hearing below 750 Hz only. In seven cases the measured gap detection thresholds were substantially worse than in normal hearing young subjects, reaching from 5.0 to 25.0 ms, see Fig. 24. In two cases gap detection was close to the threshold measured in young normal hearing musicians, amounting to 3.0 and 4.0 ms.

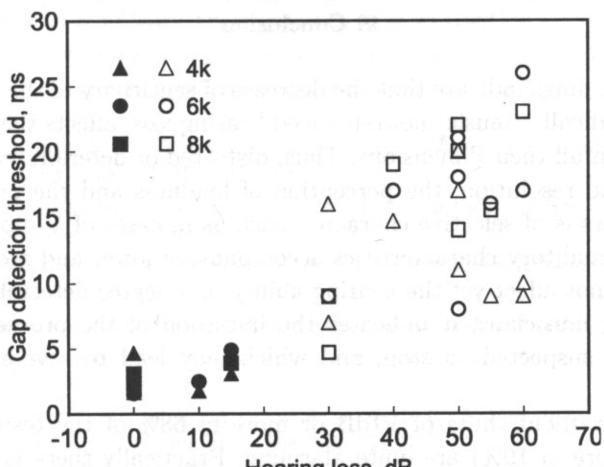


Fig. 24. Gap detection threshold in normal and hearing impaired subjects related to the degree of hearing loss at the test frequency.

Measurements of the gap detection thresholds in 2 music students with "notch-shaped" selective noise induced hearing loss at 6 kHz performed at the frequency of loss and at two adjacent frequencies above and below 6 kHz gave the results comparable to those for normal-hearing subjects. No difference was obtained between the data at the frequency of loss and outside of it.

It seems to be not easy to interprete these results with existing major controversy, Moore (1993), as to whether noise induced sensorineural impairment of hearing affects temporal resolution. Present data clearly suggest that such dependence exists at least in cases of relatively deep and wide-spread cochlear hearing loss. The present data, JAROSZEWSKI *et al.* [57] are in some conflict with the data from GLASBERG *et al.* [30] and MOORE *et al.* [73] and with suggestions by PLACK and MOORE [77], who recognised "gap detection as not adequate way of measuring temporal resolution". However, the data are consistent with findings by TYLER *et al.* [95, Fig. 1, p. 749] who also found large proportion of their impaired subjects performing worse than normal hearing.

Much better gap detection thresholds at high frequencies result, as it was pointed out by FITZGIBBONS and WIGHTMAN [23] and by TYLER *et al.* [95], from the small time constants and large bandwidths of high frequency auditory filters.

Referring to the present consideration whether gap detection experiment is a good or worse measure of temporal resolution, it is evident that in a good number of reports (and the present) the data for the gap detection threshold show marked differences between normal and hearing impaired subjects. No indication of worse temporal resolution from the gap detection experiment with noise induced hearing loss in music students is ambiguous. It may be that temporal resolution in cases of selective noise induced hearing loss is preserved at normal level. However, the gap detection threshold procedure applied may be not sensitive enough.

9. Conclusion

The present findings indicate that the decrease of sensitivity of the hearing system or hearing loss, specifically (music) noise-induced hearing loss, affects the perception of the auditory signals in all their dimensions. Thus, distorted or deteriorated is the frequency discrimination and resolution, the perception of loudness and the perception of time, even if hearing loss is of selective character, such as in cases of V-dips at 6 kHz. While the distortion of auditory characteristics accompanying lower and moderate degrees of hearing loss may not affect yet the hearing ability in a degree detectable in professional activity of young musicians, it indicates the initiation of the process of impairment, which cannot be suspected to stop, and which may lead to severe disability of the hearing system.

Permanent threshold shifts of 10 dB or more in 68% of the tested sample of 214 (and 30 dB or more in 10%) are quite alarming. Practically there is less than 10% of the population of young musicians in which no marks of overexposure were found. Main sources of acoustic trauma are discotheques, portable cassette or CD players, high power home sonic equipment and very loud training sessions of young musicians.

Acknowledgement

This work was supported by Grant 4 S404 041 07 from The State Committee for Scientific Research.

References

[1] AXELSSON A., ELIASSON A., ISRAELSSON B., *Hearing in pop/rock musicans: A follow-up study*, Ear and Hearing, **16**, 245–253 (1995).

[2] AXELSSON A., LINDGREN F., *Pop music and hearing*, Ear and Hearing, **2**, 64–69 (1981).

[3] AXELSSON A., LINDGREN F., *Early noise induced hearing loss in teenage boys*, Scand. Audiol., **10**, 91–96 (1981).

[4] BÈKÈSY G. VON, *Loudness recruitment*, Trans. Am. Otol. Soc., **55**, 1–9 (1965).

[5] BIKERDIKE J., GREGORY A., *An evaluation of hearing damage risk to attenders at discotheques*, Leeds Polytechnic School of Constructional Studies, Dept. of Environment, Noise Advisory Council, DoE Project Report, 1–158 (1980).

[6] BOOTHROYD A., *Detection of temporal gaps by deaf and hearing children*, S.A.R.P. 12, Clarke School for the Deaf., Northampton, M. A. 1973.

[7] BORSCHGREVINK H.M., *One third of 18 years old male conscripts show noise induced hearing loss 20 dB before the start of military service. The incidence being doubled since 1981 reflecting in- creased leisure noise?* Noise as a Public Health Problem, 2, Council of Building Research, 27–32, 1988.

[8] BRADLEY R., FORTNUM H., COLES R., *Research note: patterns of exposure of schoolchildren to amplified music*, British Journal of Audiology, **21**, pp. 119–125 (1987).

[9] BUUS S., FLORENTINE M., *Gap detection in normal and impaired listeners: The effect of level and frequency* [in:] Time Resolution in Auditory Systems, A. NICKELSON [Ed.], Springer, Berlin 1985.

[10] CABOT R.C., GENTER C.R. II, LUCKE T., *Sound levels and spectra of rock music*, J. Audio Eng. Soc., **27**, 267–284 (1979).

[11] CATALANO P.J., LEVIN S.M., *Noise-induced hearing loss and portable radios with head-phones*, International Journal of Pediatric Otolaryngology, **9**, 59–67 (1985).

[12] CHISTOVICH L.A., *Frequency characteristics of the masking effect*, Biofiz., **2**, 743–755 (1975).

[13] CLARK W.W., *Noise exposure from leisure activities: A review*, J. Acoust. Soc. Amer., **90**, 175–181 (1991).

[14] CLARK W.W., BOHNE B.A., *Temporary hearing losses following attendance to a rock concert*, J. Acoust. Soc. Amer. Suppl. 1, **79**, s. 48 (1986).

[15] COX R.M., ALEXANDER G.C., *Indices of frequency and temporal resolution as a function of level*, J. Acoust. Soc. Amer., **93**, 1068–1075 (1993).

[16] DAVIS H. and SILVERMAN S.R., *Hearing and deafness*, Holt, Rinehard and Winston, New York, 2-nd edition, 1978.

[17] DIBBLE K., *Hearing loss and music*, J. Audio Eng. Soc., **43**, 251–266 (1995).

[18] FASTL H., *Frequency discrimination for pulsed versus modulated tones*, J. Acoust. Soc. Amer., **63**, 275–277 (1978).

[19] FEARN R.W., *Hearing level of young people as a function of attendance rate at amplified pop music performances*, J. Sound and Vibr., **104**, 351–354 (1986).

[20] FEARN R.W., HANSON D.R., *Hearing thresholds of young people aged 11-25 years: a review and overall assessment*, J. Sound and Vibr., **138**, 155–161 (1990).

[21] FEARN R.W., HANSON D.R., *Hearing level of young subjects exposed to amplified music*, J. Sound and Vibr., **128**, 509–512 (1989).

[22] FECHNER G.T., *Revision der Hauptpunkte der Psychophysik*, BREITKOPF und HARTEL [Eds.], Leipzig 1882.

[23] FITZGIBBONS P.J., WIGHTMAN F.L., *Temporal resolution in normal and hearing impaired listeners*, J. Acoust. Soc. Amer. Suppl. 1, **65**, S133 (1979).

[24] FITZGIBBONS P.J., WIGHTMAN F.L., *Gap detection in normal and hearing impaired listeners*, J. Acoust. Soc. Amer., **72**, 761–765 (1982).

[25] FLORENTINE M., BUUS S., SCHARF B. and ZWICKER E., *Frequency selectivity in normal hearing and hearing-impaired observers*, J. Speech Hear. Res., **23**, 646–669 (1980).

[26] FOWLER E.P., *Marked deafened areas in normal ears*, Arch. Oto-laryngol., **8**, 151–155 (1928).

[27] FREYMAN R.L., NELSON D.A., *Frequency discrimination as a function of signal frequency and level in normal-hearing and hearing-impaired listeners*, J. Speech Hear. Res., **34**, 1371–1386 (1986).

[28] FREYMAN R.L., NELSON D.A., *Frequency discrimination as a function of tonal duration and excitation pattern slopes in normal and hearing-impaired listeners*, J. Acoust. Soc. Amer., **79**, 1034–1044 (1991).

[29] FRY H.J.H., *Overuse syndrome in musicians – 100 years ago: a historical review*, The Medical Journal of Australia, **145**, 620–625 (1986).

[30] GLASBERG B.R., MOORE B.C.J., BACON S.P., *Gap detection and masking in hearing-impaired and normal-hearing subjects*, J. Acoust. Soc. Amer., **81**, 1546–1556 (1987).

[31] GLASBERG B.R., MOORE B.C.J., *Effects of envelope fluctuation on gap detection*, Hear. Res., **64**, 81–92 (1992).

[32] HALL J., WOOD E.J., *Stimulus duration and frequency discrimination for normal hearing and hearing-impaired subjects*, J. Speech Hear. Res., **27**, 252–256 (1984).

[33] HALLPIKE C.S., *The loudness recruitment phenomenon: A clinical contribution to the neurology of hearing*, [in:] Sensoneural Hearing Processes and Disorders, A.B. GRAHAM [Ed.], Little, Brown, Boston 1967.

[34] HALLPIKE C.S. and HOOD J.D., *Observations upon the neurological mechanism of the loudness recruitment phenomenon*, Acta Otolaryngol., **50**, 472–480 (1959).

[35] HARRIS J.D., *Pitch discrimination*, J. Acoust. Soc. Amer., **24**, 750–755 (1952).

[36] HELLMAN R.P., *Loudness functions in noise induced and noise simulated hearing losses*, [in:] Proc. 5th Int. Congress on Noise as a Public Health Problem, 1988.

[37] HELLMAN R.P. and MEISELMAN C.H., *Loudness relations for individuals and groups in normal and impaired hearing*, J. Acoust. Soc. Amer., **88**, 2596–2606 (1990).

[38] HELLMAN R.P. and MEISELMAN C.H., *Rate of loudness growth for pure tones in normal and impaired hearing*, J. Acoust. Soc. Amer., **93**, 966–975 (1993).

[39] HELLSTRÖM P.A. and AXELSSON A., *Sound levels hearing habits and hazards of using portable cassette players*, J. Sound and Vibr., **127**, 521–528 (1988).

[40] HUGHES E., FORTNUM H.M., DAVIS A.C., HAGGARD M.P., COLES R.R.A., LUTMAN M.E., *Damage to hearing arising from leisure noise*, British Journal of Audiology, No. 20, 157–164 (1986).

[41] ISING H., HANEL J., PILGRAMM M., BABISCH W., LINDTHAMMER A., *Gehörschadensrisiko durch Musikhören mit Kopfhörern*, Phoniatrie und Pdaudiologie HNO, **42**, 764–768 (1994).

[42] JANSSON E., KARLSSON K., *Sound levels recorded in the symphony orchestra and risk criteria for hearing loss*, Scandinavian Audiology, **12**, 215–221 (1983).

[43] JANSSON E., AXELSSON A., LINDGREN F., KARLSSON J., OLAUSSEN T., *Do musicians of the symphony orchestra become deaf?: A review of the investigations of musicians in Gothenburg and Stockholm*, [in:] Acoustics for choir and orchestra, S. TERNSTROEM [Ed.], Publication No 52, Royal Swedish Academy of Music, 62–74 (1986).

[44] JAROSZEWSKI A., RAKOWSKI A., *Loud music induced threshold shifts and damage risk prediction*, Archives of Acoustics, **19**, 3, pp. 311–321 (1994).

[45] JAROSZEWSKI A., *A new method for determination of frequency selectivity in post-stimulatory masking*, Proc. 9th ICA, Madrid 1977, H40.

[46] JAROSZEWSKI A., *A new method for determination of frequency selectivity in post-stimulatory masking*, Acustica, **41**, 327–333 (1979).

[47] JAROSZEWSKI A., *Psychophysical diagnosis of the hearing system* [in Polish], [in:] Akustyka w Technice, Medycynie i Kulturze, MALECKI I. [Ed.], IPPT-PAN, Warszawa, pp. 196–304 (1993).

[48] JAROSZEWSKI A., FIDECKI T., ROGOWSKI P., MONIAK R., *Loudness perception in selective and frequency limited hearing loss* [in Polish], Proc. Open Seminar on Acoustics, Gliwice-Ustroń 96, Vol. 1, 305–312 (1996).

[49] Jaroszewski A., *Pitch perception near physical unison and expansion of frequency intervals*, Acustica, **72**, 269–274 (1990).

[50] Jaroszewski A., *Effect of feedback on pitch discrimination*, Acustica, **76**, 137–141 (1992).

[51] Jaroszewski A., *A study of constant (systematic) errors in pitch discrimination of short tone pulses*, Acustica, **77**, 106–110 (1992).

[52] Jaroszewski A., *The case for the adjustment procedure*, Acustica, **82**, 852–862 (1996).

[53] Jaroszewski A., *Constant errors in the perception of pitch in tone pulses in severe hearing loss* [in Polish], Proc. XLII Open Seminar on Acoustics, Gliwice-Ustroń 17–20.09.1996, Vol. 1, 305–312 (1996).

[54] Jaroszewski A., *Constant errors in pitch discrimination in the range of wide band moderate hearing loss* [in Polish], Proc. IV Symp. On Electroacoustics of Hearing Aids, Czerniejewo 8–9.10.1996, 63–66 (1996).

[55] Jaroszewski A., Fidecki T., *Pitch discrimination scatter in normal hearing and in large hearing loss* [in Polish], IV Symposium in Electroacoustics: Electroacoustical transducers and acoustics of hearing aids, Inst. of Acoustics, A. Mickiewicz University, Poznań, 1996, 63–66.

[56] Jaroszewski A., Fidecki T., Rogowski P., Moniak R., *Exposures and hearing thresholds in loud pop/rock music*, [in:] Akustyka w Technice i Medycynie, IPPT-PAN [Ed.], Kraków, pp. 291–300 (1995).

[57] Jaroszewski A., Fidecki T., Rogowski P., Moniak R., *Sound pressure level distributions in 10 Warsaw discotheques and TTS2 in attendants*, Proc. of Internoise'97, Budapest, Vol. II, 1163–1166 (1996).

[58] Jaroszewski A., Fidecki T., Rogowski P., Moniak R., *Noise induced hearing loss and its consequences in pitch and loudness perception*, Proc. of XLIV Open Seminar on Acoustics, Jastrzębia Góra, Vol. I, 337–342 (1996).

[59] Jaroszewski A., Rakowski A., *Psychoacoustical equivalents of tuning curves as determined by post-stimulatory masking technique*, Archives of Acoustics, **1**, 127–135 (1976).

[60] Jaroszewski A., Fidecki T., Rogowski P., *Exposures and hearing thresholds in music students due to training sessions*, Proc. of XLIV Open Seminar on Acoustics, Jastrzębia Góra, Vol. I, 343–348 (1997).

[61] Johnson D.W., Scherman R.E., Aldridge J., Lorraine A., *Effects of instrument type and orchestral position on hearing sensivity for 0,25 to 20 kHz in orchestral musician*, Scandinavian Audiology, **14**, 215–221 (1985).

[62] Johnson D.W., Sherman R.E., Aldridge V., Lorraine A., *Extended high frequency hearing sensivity: a normative threshold study in musicians*, Ann. Otol. Rhinol. Otolaryngol., **95**, 196–202 (1986).

[63] Karlsson K., Lundquist P., Olaussen T., *The hearing of symphony orchestra musicians*, Scandinavian Audiology, **12**, 257–264 (1983).

[64] Kellog W.N., *An experimental comparison of psychophysical methods*, Ph.D. Dissert. Faculty of Philosophy, Columbia University, 1929.

[65] Lebo P.Ch., Oliphant K.P., *Music as a source of acoustic trauma*, Laryngoscope, **78**, 1211–1218 (1968).

[66] Lipscomb D.M., *Ear damage from exposure to rock-and-roll music*, Arch. Otolaryngol., **90**, 29–39 (1969).

[67] Lutman M.E., Gatehouse S., Worthington A.G., *Frequency resolution as a function of hearing threshold level and age*, J. Acoust. Soc. Amer., **89**, 320–328 (1991).

[68] Lutman M.E., Wood E.J., *A simple clinical measure of frequency resolution*, Brit. J. Audiol. 19, 1–8 (1985).

[69] Mawhinney C.H., McCullagh G.C., *Music induced hearing loss in musicians playing different musical instruments*, IoA Reproduced Sound & Conference, Windermere, Proc IoA, Vol. 14, part 5, 95–105 (1992).

[70] Moore B.C., *Frequency difference limens for short-duration tones*, J. Acoust. Soc. Amer., **54**, 610–619 (1973).

[71] Moore B.C.J., *Psychophysical tuning curves measured in simultaneous and forward masking*, J. Acoust. Soc. Amer., **63**, 524–532 (1978).

[72] Moore B.C.J., Glasberg B.R., *Gap detection with sinusoids and noise in normal, impaired and electrically stimulated ears*, J. Acoust. Soc. Amer., **83**, 1093–1101 (1988).

[73] Moore B.C.J., Glasberg B.R., Donaldson E., McPherson T., Plack C.J., *Detection of temporal gaps in sinusoids by normally hearing and hearing-impaired subjects*, J. Acoust. Soc. Amer., **85**, 1266–1275 (1989).

[74] Nordmark J.O., *Mechanism of frequency discrimination*, J. Acoust. Soc. Amer., **44**, 1533–1540 (1968).

[75] Ostri B., Eller N., Dahlin E., Skylv G., *Hearing impairment in orchestral musicians*, Scand. Audiol., **18**, 243–249 (1989).

[76] Penner M.J., *Detection of temporal gaps in noise as a measure of the decay of auditory sensation*, J. Acoust. Soc. Amer., **61**, 552–557 (1977).

[77] Plack C.J., Moore B.C.J., *Decrement detection in normal and impaired ears*, J. Acoust. Soc. Amer., **90**, 3069–3076 (1991).

[78] Preyer W., Engl. transl. On the limits of the perception of tone. Proc. of the Musical Assoc. p. 1–32, (Über die Grenzen der Tonwahrnehmung) W. Preyer, Physiol. Inst. Jens Ed., pp. 1–32 (1876).

[79] Rabinowitz J., Hausler R., Bristown G., Ray P., *Study of the effects of loud music on musicians of The Suisse Romande Orchestra*, Journal de Médicine et Hygiéne, **19**, 1909–1921 (1982).

[80] Rakowski A., *Pitch discrimination at the threshold of hearing*, Proc. 7-th ICA, Budapest 1971, 20H6.

[81] Rakowski A., Hirsh I., *Post-stimulatory pitch shift of pure tones*, J. Acoust. Soc. Amer., **68**, 467–474 (1980).

[82] Rakowski A., Hirsh I., *Post-stimulatory pitch shift and its dependence on signal parameters*, J. Acoust. Soc. Amer., **63**, Suppl. 1, 50 (1979).

[83] Rintelman W.F., Borus Q.F., *Noise induced hearing loss and rock and roll music*, Arch. Otolaryngol., **88**, 377–385 (1968).

[84] Rintelman W.F., Lindberg R.F., Simitely E.K., *Temporary threshold shift and recovery patterns from two types of rock-and-roll music presentation*, J. Acoust. Soc. Amer., **51**, 1249–1255 (1972).

[85] Ritsma R.J., *Pitch disrimination and frequency discrimination*, [in:] Reports 5-e Congrés International d'Acoustique, Liége 1965, Vol. Ia, B22 (1965).

[86] Ronken D., *Monaural detection of a phase difference between clicks*, J. Acoust. Soc. Amer., **47**, 1091–1099 (1970).

[87] Royster J.D., Royster L.H., Killion M.C., *Sound exposures and hearing thresholds of symphony orchestra musicians*, J. Acoust. Soc. Amer., **89**, 6, 2793–2803 (1991).

[88] Ruhm H.B., Mencke E.O., Milburn R., Cooper W.A., Rose D., *Differential sensitivity to duration of acoustic signals*, J. Speech Hear. Res., **9**, 371–384 (1966).

[89] Schacke G., *Sound pressure levels within an opera orchestra and its meaning to hearing*, Abstract of paper presented to the 22nd International Congress on Occupational Health, Sydney, Australia, 1987.

[90] Shailer M.J., Moore B.C.J., *Gap detection as a function of frequency, bandwidth and level*, J. Acoust. Soc. Amer., **74**, 467–473 (1983).

[91] Shower E.G., Biddulph R., *Differential pitch sensitivity of the ear*, J. Acoust. Soc. Amer., **3**, 275–287 (1931).

[92] Skrajnar S.F., Royster L.H., Berger E.H., Pearson R.G., *Do personal radio headsets provide hearing protection?*, Sound and Vibration, **5**, 16–19 (1985).

[93] Small A.M., *Pure-tone masking*, J. Acoust. Soc. Amer., **31**, 1619–1625 (1959).

[94] TRINDER E., *Auditory fusion: a critical interval test with implications in differential diagnosis*, Brit. J. Audiol., **13**, 143–147 (1979).

[95] TYLER R.S., SUMMERFIELD Q., WOOD E.J., FERNANDES M., *Psychoacoustic and phonetic temporal processing in normal and hearing-impaired listeners*, J. Acoust. Soc. Amer., **72**, 740–752 (1982).

[96] TYLER R.S., WOOD E.J., FERNANDES M., *Frequency resolution and discrimination of constant and dynamic tones in normal and hearing-impaired listeners*, J. Acoust. Soc. Amer., **74**, 1190–1199 (1983).

[97] VIEMEISTER N.F., *Temporal modulation transfer functions based on modulation thresholds*, J. Acoust. Soc. Amer., **66**, 1364–1380 (1979).

[98] WALLISER K., *Zusammenwirken von Hüllkurvenperiode und Tonheit bei der Bildung Periodentonhöhe*, Ph.D. Dissertation, TU München 1986.

[99] WEBER E.H., [in:] Helmholtz H., Engl. transl. On the sensations of tone. Dover Publications Inc., New York 1954, p. 147, 1887.

[100] WEST P.D.B., EVANS E.F., *Early detection of hearing damage in young listeners resulting from exposure to amplified music*, British Journal of Audiology, **24**, pp. 89–103 (1990).

[101] WESTMORE G.L., EVERSDEN I.D., *Noise - induced hearing loss in orchestral musicians*, Arch. Otolaryngology, **107**, 761–764 (1981).

[102] WIER C.C., JESTEADT W., GREEN D.M., *Frequency discrimination as a function of frequency and sensation level*, J. Acoust. Soc. Amer., **61**, 178–184 (1977).

[103] WIGHTMAN F., *Psychoacoustic correlates of hearing loss*, [in:] New perspectives in noise - induced hearing loss, R.P. HAMMERNICK, D. HENDERSON, R. SALVI [Eds.], New York, Raven, pp. 375–394 (1982).

[104] WIGHTMAN F., MCGEE T., KRAMER M., *Factors influencing frequency selectivity in normal and hearing-impaired listeners*, [in:] Psychophysics and Physiology of Hearing, EVANS E.F., WILSON J.P. [Eds.], Academic Press, London, New York, pp. 295–310, 1977.

[105] WOOLFORD D.H., *Sound pressure levels in symphony orchestras and hearing*, Preprint 2104(B-1), Australian Regional Convention on The Audio Engineering Society, Melbourne Sept. 25–27 (1984).

[106] ZWICKER E., *Direct comparison between the sensation produced by frequency modulation*, J. Acoust. Soc. Amer., **34**, 1245–1250 (1962).

[107] ZWICKER E., FELDTKELLER R., *Das Ohr als Nachrichtenempfänger*, Hirzel-Verlag, Stuttgart, 2 erw. Auflage, 1967.

ARCHIVES OF ACOUSTICS
23, 1, 33–38 (1998)

THE COMPUTER PROCEDURES TO ANALYSE THE PARAMETERS OF THE AUDITORY BRAINSTEM RESPONSES IN ADVANCED RETROCOCHLEAR TUMORS OF THE AUDITORY PATHWAY

W. BOCHENEK, J. HATLIŃSKI,

ENT Clinic Stomatology Department, Medical Academy,
(00-739 Warszawa, Stępińska 19/25, Poland)

Z. RANACHOWSKI

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

This investigation was aimed at a computer- supported analysis of the Auditory Brainstem Responses (ABR) recorded in both ears of 8 patients with unilateral tumor of the retrocochlear extracanalicular auditory pathway. Applying the computer procedure designed for processing the averaged ABR waveforms, the latencies in the time domain of waves I, III, V were calculated and the ratio of the amplitudes of those waves in both ears of the same patient was evaluated and correlated with the surgically verified size of the tumor.

1. Introduction

The tumors of the VIII, cochleo-vestibular (stato-acoustical), cranial nerve in its intracanalicular part and of the pono-cerebellar angle (extracanalicular) do not infiltrate the environmental structures and do not make methastases. Nevertheless, while growing-up they affect not only the hearing and the sense of equilibrium but the facial nerve function as well and, in the later stage, they may compress the breathing and cardiac centers in the brainstem creating an immediate life threat.The most important diagnostic tool in these cases of tumors are now advanced radiological techniques. Since they are expensive and not completely harmless, some audiological tests should be made as soon as possible before sending the patient to the radiologist. Among the appropriate audiological tests, the registration of the Auditory Brainstem Responses (ABR) can provide the most valuable data. The ABR waveforms are registered as variations of the electrical potential of the electrodes placed on the head of the subject after application of an acoustic stimulus into the ear [1]. Positive and negative surface electrodes are placed on both mastoids and a ground one on the forehead. To supress the influence of the environmental noises, the stimulus is repeated ca. 1000 times and the registered waveforms are averaged. The stimulus usually used is a 100 μs click delivered *via* the headphone with a repetition rate of 31 per second at the sound level choosen in the range from 0 to 100 dB HL and with alternating polarity that prevents some artefacts to occure [2]. The time analysis of the recorded ABR applies to the first 10 ms after the stimulus onset [1]. In this period, the auditory nerve and the primary auditory brainstem centers are activated. The normal ABR waveform includes five-seven maxima (waves) among which the first, third and fifth are the dominant ones. It is generally accepted that the source of the first wave is the distal part of the cochlear nerve, the third wave is generated mostly in cochlear nuclei and the fifth one in the lateral lemniscus nuclei [3].

The presented data concern the ABR registered for patients operated in the ENT Clinic of the Institute of Stomatology of the Warsaw Medical Academy. It seems that the proposed computer analysis can to some extent facilitate additionally in the evaluation of the ABR recordings in these cases.

2. Material and method

The material comprises a series of 8 cases, 5 women and 3 men aged from 38 to 50, with unilateral eight-nerve neuroma. Five of the cases were rightsided and 3 leftsided, all expanded already into the cerebello-pontine angle but not compressing yet the brainstem centers. They were operated upon in our Clinic in the years 1996-97. The recordings were made from the affected side while those from the opposite side served for inter-ear comparing purposes. The ABR from the two normal hearing young subjects (No 1 and No 2 in Table 1) served as referrences and 20 other normal ABRs were analysed to test the software used. The ABR were recorded on the ABR Acquisition System EPTEST, made in Poland [4]. Each record included 1000 integer numbers, representing the current ABR potential sampled every 10 μs; in this way the ABR run within the first 10 ms after the stimulus onset were stored.

The authors have designed the software to visualize the ABR waveforms and to determine the diagnostic parameters listed in the Introduction. The time plots of the recored waveforms were completed by the lines reffering to generally accepted time limits for normal latencies measured in healthy ears in response to the stimulation by the 90 dB nHL click (see Figs. 1, 2 and 3). The latencies accepted as normal are as follows [4]:

- the latency of the I wave – 1.6 ± 0.4 ms,
- the latency of the III wave – 3.8 ± 0.4 ms,
- the latency of the V wave – 5.6 ± 0.4 ms,
- the inter-ear latency difference – ± 0.2 ms.

In spite of the latency data, the authors did not found any generally accepted values concerning the inter-ear differences of the wave amplitudes. The lack of such values is certainly due to the great intersubject shape variability of the ABR waves. A step to

Table 1. The results of the evaluation of the ABR waveform parameters referring to the investigated subjects.

Subject No.	Size of the recognized pathology (mm)	Latency of wave I (miliseconds)	Latency of wave III (miliseconds)	Latency of wave V (miliseconds)	Amplitude of the ABR waves ratio: reference/pathological ear		
		! denotes the value beyond the limit			I wave	III wave	V wave
1	0	1.59 1.66	3.74 3.79	5.64 5.55	2.75	1.05	1.35
2	0	1.42 1.46	3.49 3.50	5.37 5.48	1.17	1.13	2.0
3	11	1.44 1.99	3.64 4.86!	5.44 6.07!	0.24	2.40	0.91
4	13	– 2.28!	3.67 3.68	– 5.54	–	4.57	–
5	15	1.42 1.48	3.61 4.29!	5.37 5.12	0.76	1.42	1.25
6	15	1.69 –	3.88 –	5.75 6.05!	–	–	1.50
7	25	1.59 1.42	3.89 –	5.18 –	2	–	–
8	30	1.14 1.70	3.55 –	– 6.94!	0.21	–	–
9	34	1.53 1.38	3.88 –	5.87 5.85	1.40	12.0	1.09
10	40	1.48 –	3.75 4.50!	– 5.42	6	16.5	–

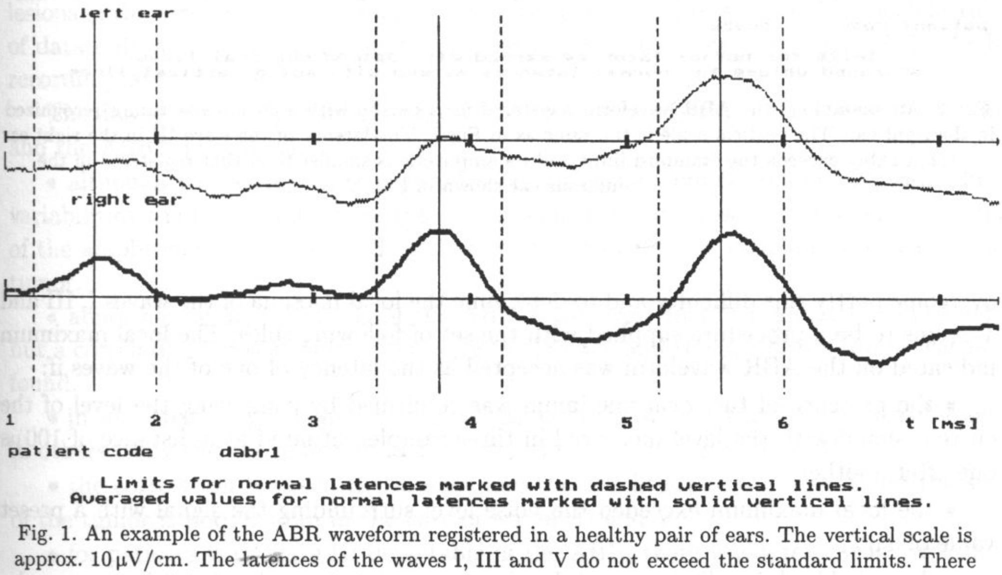


Fig. 1. An example of the ABR waveform registered in a healthy pair of ears. The vertical scale is approx. 10 μV/cm. The latences of the waves I, III and V do not exceed the standard limits. There are slight inter-ear differences in latences and amplitudes of the three waves.

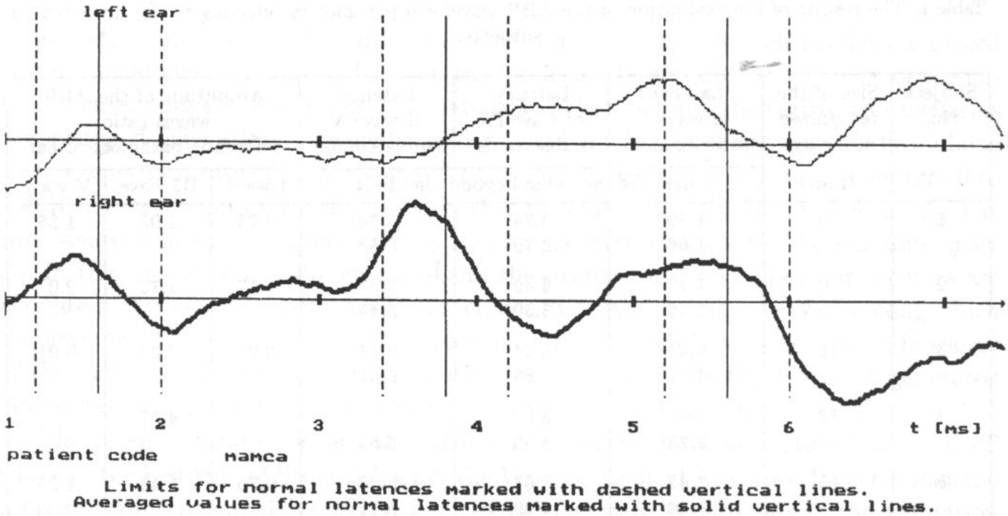


Fig. 2. An example of the ABR waveform registered for a person with a 15 mm size tumor recognized in the left ear. The vertical scale is the same as in Fig. 1. The latence of the wave III in the left ear remarkably exceeds the standard limit and its amplitude is smaller than that in the healthy ear.

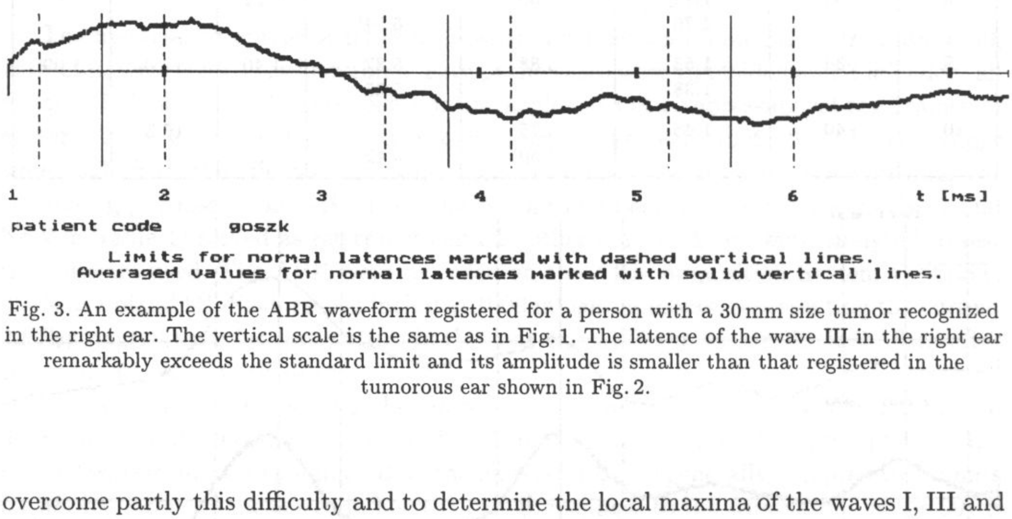


Fig. 3. An example of the ABR waveform registered for a person with a 30 mm size tumor recognized in the right ear. The vertical scale is the same as in Fig. 1. The latence of the wave III in the right ear remarkably exceeds the standard limit and its amplitude is smaller than that registered in the tumorous ear shown in Fig. 2.

overcome partly this difficulty and to determine the local maxima of the waves I, III and V seems to be a procedure supplied with the set of following rules: The local maximum indicated on the ABR waveform was accepted as the latency of one of the waves if:

• the presence of the local maximum was confirmed by comparing the level of the current signal with the level measured in three samples situated at a distance of 100 μs one after another,

• the local maximum exceeded the noise level surrounding the signal with a preset value of 1 μV,

• the individual time limits were taken into account to recognize a certain wave mode.

3. Discussion of the results

Table 1 shows the results of the evaluation of the ABR: the latencies of three waves and the ratio of the amplitudes of those waves registered in the healthy ear and in that with recognized pathology. The ABR parameters of the subjects starting from the No 3 up to No 10 are referred to the size of the tumor. The two first subjects, labelled No 1 and No 2, are of healthy, normal hearing persons.

An example of the ABR waveform recorded in a healthy ear is shown in Fig. 1. The wave latencies are within the above mentioned time limits, the amplitudes of all three maxima are distinct and high and there are no significant inter-ear differences.

An example of the ABR waveform from a patient with a relatively smaller (in this series) size of the tumor – 15 mm beyond the internal acoustic porus – is shown in Fig. 2. In this case the neuroma is recognized in the left ear however the right one is normal. The latencies of the waves III and V are beyond the normal time limits in the pathological ear but there are only slight differences in the wave amplitudes between a pathological and a normal ear.

An example of the ABR waveform registered for a patient with great size of the right ear tumor – 30 mm beyond the internal acoustic porus – is shown in Fig. 3. The amplitudes of the waves III and V are poor when compared with those recorded from the opposite ear. The latency of the wave I largely exceeds the normal limits. The inter-ear differences are substantial.

4. Conclusions

Because of the relatively small number of cases and the advanced size of tumoral lesions in the cerebello-pontine region, the material may be presented only as a collection of data only and some results of the proposed approach to the evaluation of the ABR recordings are preliminary.

The data presented in Table 1 suggest some correletion between the size of the tumor and the ABR's parameters:

• although the ABR wave amplitudes are subject to much greater interindividual variabilities than their latencies; there is a remarkable correlection between the ratio of the amplitudes of the wave III, registered in both subjects' ears, and the size of the tumor,

• abnormal latency patterns of the ABR waves were found in all pathological ears but a correlation between the extention of the latency and the size of the tumor was not found,

• in some pathological ears the identification of the wave III and V was feasible but in others using the above formulated criteria, no maxima could be found,

• the correlation between the latency and the amplitude of the wave I and the size of the tumor is poorer than in the case of waves III and V,

• for patients with wave III detectable in the ABR, there was a reversed correlation between the size of the tumor and the amplitude of wave III.

To summarize: Using the computer procedures described in the article to analyse the ABR waveforms in subjects with advanced unilateral retrocochlear tumors verified surgically, the following parameters seem to be most valuable:

- the presence/absence of the waves I, III and V,
- the normal/extended latencies of those waves,
- the inter-ear ratio of amplitudes of the wave III.

References

[1] H. Davis, *Principles of electric response audiometry*, The Annals of Otology, Rhinology and Laryngology, Suppl. 28, **85**, 3, Part 3 (1976).

[2] D.R. Watson, *The effects of cochlear hearing loss, age and sex on the auditory Brainstem response*, Audiology, **35**, 246–258 (1966).

[3] K.B. Campbell *et al.*, *Auditory potentials*, Proc. from the 1st Int. Workshop and Symposium on Evoked Potentials, Milan, Apr. 1980, Vol. 1, No. 1, Amplaid, pp. 21–31, Milano 1981.

[4] K. Kochanek, *The auditory Brainstem responses in audiological evaluation of children* [in Polish], Modern Medicine, **3**, 7, 41–44 (1996).

[5] A.R.D. Thornton, *Developments in Diagnostic Interpretation and Response Analysis*, Proc. from the 1st Int. Workshop and Symposium on Evoked Potentials, Milan, Apr. 1980, Vol. 1, 1, Amplaid, pp. 77–82, Milano 1981.

ARCHIVES OF ACOUSTICS
23, 1, 39–49 (1998)

SUBJECTIVE EVALUATION OF SOUND EMITTED BY A LOUDSPEAKER SYSTEM IN CORRELATION WITH ITS STEADY-STATE AND TRANSIENT RESPONSES

M. NIEWIAROWICZ, A. FURMANN, E. HOJAN, P. PERZ

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

Objective parameters of a loudspeaker system, which describe its transmission properties in a transient- and steady-state, have been determined. These parameters have been related to the attributes of the sound perception space (fullness, clearness, sharpness, spaciousness, loudness, lack of distortions) with respect to the overall subjective evaluation and the parametric one. A good correlation between sharpness, fullness, spaciousness and selected objective parameters have been obtained. An influence of the test signal on the results of the subjective evaluation was also discovered.

1. Introduction

The multi-dimensional character of the sensation space connected with the perception of acoustic signals creates a number of significant difficulties when attempts to correlate the subjective evaluations of such signals with their physical parameters are performed [10, 13, 16, 17]. Investigations of those problems indicated, among other things, the applicability of the multi-dimensional scaling technique to the analysis of the results of the subjective evaluation [1, 6] and to the modelling of physical parameters of signals on the basis of psychoacoustic data in the case of objective evaluations [4, 7]. The physical parameters of the signals can be directly connected to the physical parameters of their sources and, in this context, subjective evaluations of the sound can be related to the parameters of their sources [5]. This kind of investigations performed for electroacoustic transducers showed that it was necessary to include physical parameters, which characterise their work in both the steady and transient states [14]. Results presented in many papers showed that the subjective evaluation of the acoustic signals emitted by transducers depends on many factors, connected with the investigations procedure [3]. The procedure of investigations of the subjective evaluations must optimise all these factors whenever possible so that the final evaluation be recurrent, provided precisely defined conditions have been met.

The subjective evaluation of sounds reproduced by the loudspeaker systems depends on the parameters of the sound perceived by the subject. These parameters depend on the parameters of the test signal, on the physical parameters of the loudspeaker system and the acoustic properties of the room. When the acoustic parameters of the room are known, the same test signal reproduced by different loudspeaker systems result in different values of objective parameters and different subjective evaluations of the reproduced sounds. The present investigations aimed at the determination of the influence of changed amplitude and phase response of the loudspeaker systems on the subjective evaluation of sounds reproduced by these systems [9].

2. Procedure of the subjective investigations

2.1. Scenery

The investigations were performed using a pair of three way loudspeaker systems (with a passive radiator), available in the random production series (Factory of Loudspeakers TONSIL S.A. in Września, Poland). The geometry of the system and the structure of the crossover network is presented in Fig. 1.

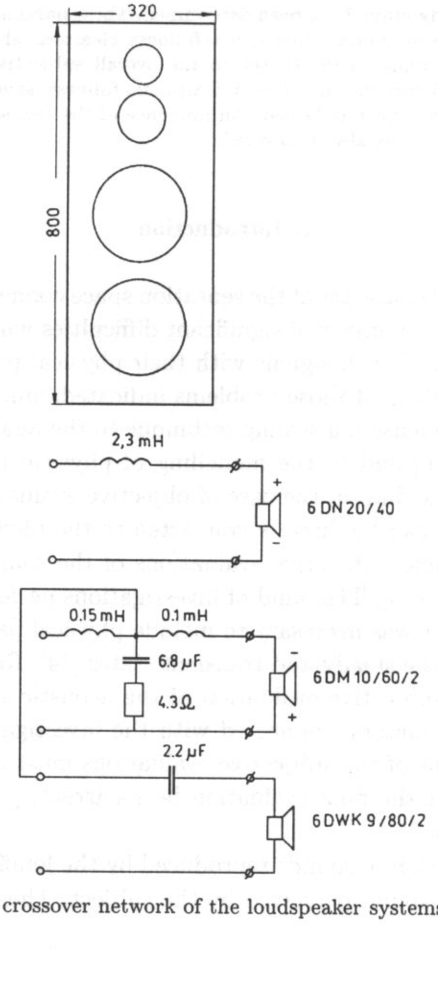


Fig. 1. Geometry and crossover network of the loudspeaker systems under investigation.

On the basis of those loudspeaker systems, measurement systems with different amplitude and phase characteristics have been artificially set up. For this purpose, special filters with a set of response curves, corresponding to the frequency responses of the loudspeaker systems, have been used. The filter characteristics are shown in Fig. 2.

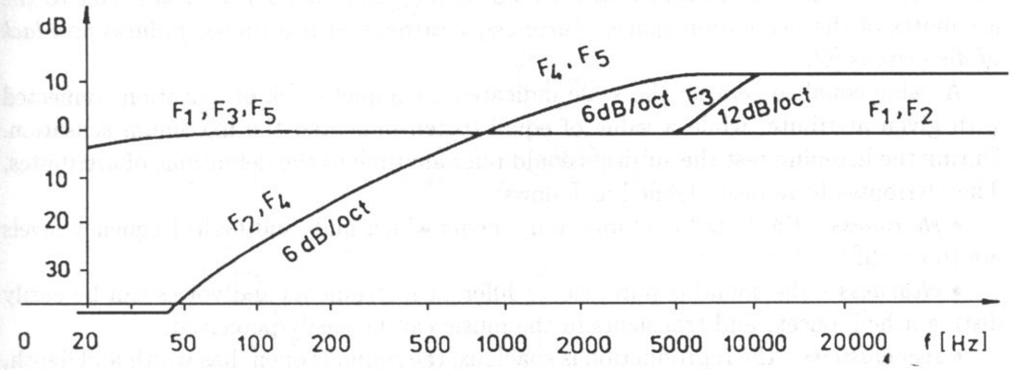


Fig. 2. Filter frequency characteristics.

In this way, the set-up imitating the five loudspeaker systems with different amplitude and phase characteristics was created.

Four one-minute test signals have been used in the investigations:

P1 – pink noise,

P2 – speech sound,

P3 – wide band music,

P4 – white noise.

The selection of the test signals was motivated by the desire to check whether the influence of the amplitude and phase characteristics of the loudspeaker system on the subjective evaluation of the reproduced sound depends on the test signal. Listening tests were presented on the stereophonic way and the level of the sound acoustic pressure was 80 dB (pink noise was the calibration signal). The measuring signals were tape-recorded and then reproduced using the same tape recorder. The white noise and pink noise signals were tape-recorded directly from the signal generator. The wide band music and speech signals were tape-recorded under natural conditions – during the musical performance in a concert hall and during the lecture in an auditorium, respectively.

Eight hi-fi experts were the subjects during the tests. The listening tests were carried out in a listening room, compatible with IEC standards.

2.2. Methodology

The subjective investigations were aimed at the determination of attributes of the perception space [8] which would help to explain the differences between the evaluation of sounds emitted by the loudspeaker systems. The subjective evaluation of sounds emitted by the loudspeaker systems consisted of two parts: an overall evaluation and a parametric evaluation.

The overall evaluation was conducted by means of the method of triadic comparisons [12]. The subject's task was to determine a pair of most similar sounds and another one of least similar sounds.

The parametric evaluation was carried out by means of the method of rating scale. The subject's task was to assess each sound on a 0 – 10 scale. Those scales refer to the attributes of the perception space: *sharpness, clearness, spaciousness, fullness* and *lack of distortions* [6].

A value equal to zero on the scale indicated a complete lack of sensation connected with given attribute, while a value of equal to ten indicated the maximum sensation. During the listening test the subjects could refer anytime to the definitions of attributes. The attributes have been defined as follows:

• *sharpness* – the sound contains components which mid- and high- frequency levels are too high,

• *clearness* – the sound is pure, clear; different instruments and voices can be easily distinguished, onsets and transients in the music can be easily perceived,

• *spaciousness* – the reproduction is spacious, the sound is open, has width and depth, fills the room, gives the impression of the subject's presence in the space surrounded by the sound,

• *fullness* – the sound contains the entire spectrum without any limitations, at least in the bass range

• *lack of distortions* – indicates a pure sound, without distortions, one which is not harsh, hiss or rumbling.

2.3. Results of the subjective evaluation

The results of the evaluation of dissimilarities between sounds have been analysed by the technique of multidimensional scaling, the method of individual differences (INDSCAL [1]) which showed that a two-dimensional space determines dissimilarities between the tested sounds. The calculations revealed that two dimensions are the optimal number of dimensions of the metric space, and enables to describe the configuration of the experimental data. The value of "stress", designating the difference between both the experimental and fitted data, was 5%. The investigations performed by KRUSKAL [11] showed that a value of stress less than 10% is an indicator of good fitness of configuration points to the experimental data. The dimensions of that space have been interpreted by the multiple regression analysis; the results of the evaluation of sound attributes have been used. It was found that one of the dimensions of the dissimilarity space is determined by the *sharpness* and the other one by the *fullness* and *spaciousness* (see Fig. 3).

On the basis of the analysis of the amplitude and phase characteristics (see Subsec. 3.1, Fig. 4) and the results of the subjective evaluation of sounds emitted by the loudspeaker systems under investigations, it has been found that:

1) the amplitude characteristic affects the *sharpness, fullness* and *spaciousness* of the sound,

2) the phase characteristic affects the *fullness* and *spaciousness* of sounds (these statements result from the qualitative analysis of the shape of the amplitude and phase characteristics),

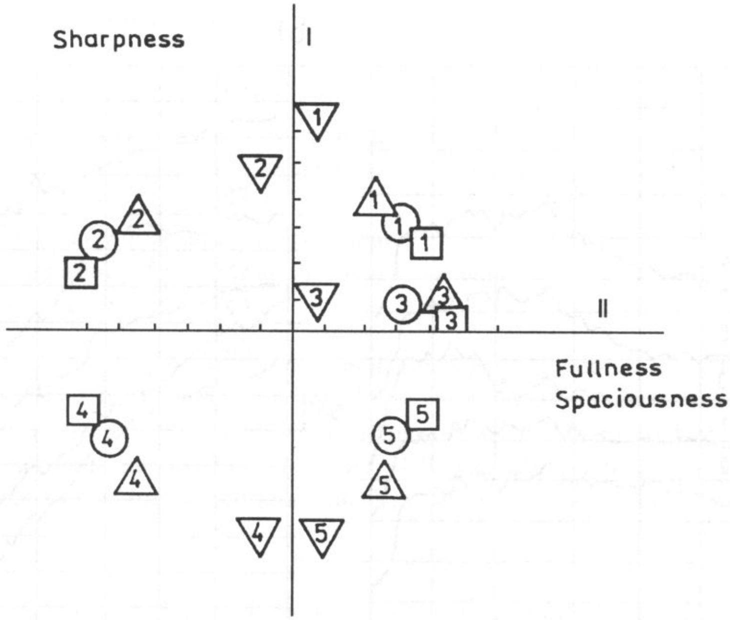


Fig. 3. Results of the subjective evaluation (symbols: △ – pink noise, ▽ – white noise, □ – speech, ○ – music; numbers inside symbols denote the number of filter).

3) the influence of the amplitude and phase characteristic depends on the test signal (in the case of white noise signal this influence is decisively different than for other test signals).

3. Procedure of the objective investigations

3.1. Steady state parameters

Amplitude and phase characteristics of loudspeaker systems have been determined in an anechoic chamber using the two-channel Bruel & Kjaer 20-32 analyser. White noise was the excitation signal and the measuring microphone was located at the distance of 1 m, on the reference axis of the loudspeaker system. Figure 4 shows the amplitude and phase characteristics of the loudspeaker systems under investigations (particular amplitude and phase characteristics were shifted one to the other of 10 dB and 50°, respectively).

Basing on the definition of the parameter, characterising the frequency response in a quantitative way [2], a new, similar parameter was introduced. Having defined limit frequencies, the value of its mean level was determined. Consequently, one can determine the mean value of ±2 dB deviations in the spectral extent by which the signal exceeded the mean level of the amplitude characteristic, which can be adopted as a measure defining the non-uniformity of this response.

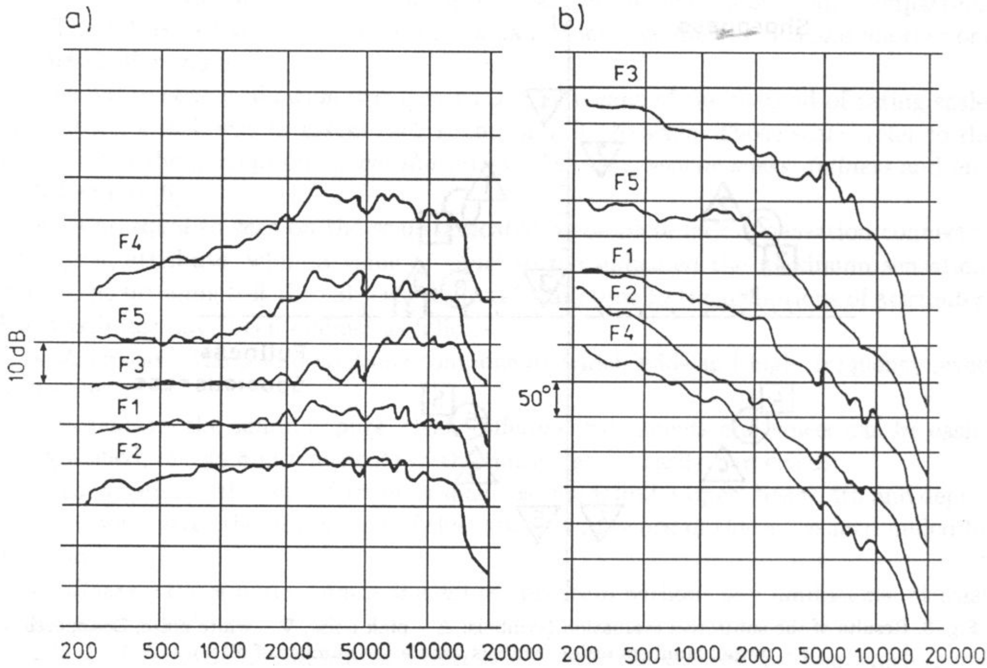


Fig. 4. Frequency (a) and phase (b) characteristics of loudspeaker systems.

Inversely, treating the amplitude characteristic as a spectrum of the acoustic signal emitted by the electroacoustic transducer, one can define the loudness in sones and the loudness level in phones, on the basis of Zwicker diagrams [15].

3.2. Transient parameters

The existence of a correlation between the subjective and objective evaluations, which we were looking for, can occur primarily when the objective evaluation is based on parameters which characterise the transient properties of a transducer [5]. The occurrence of transients in the transducers, resulting from the excitation of vibrations of mechanical elements that self-vibration frequencies, damping and reciprocal couplings are different, is related to the dynamic properties of the transducer and can make a basis for its quality evaluation. When selecting these parameters, we used the results of investigations of the perception of changes in the envelope of the transients, investigations of the transducers by impulse methods and of the perception of the deformation of impulses in the acoustic field of the loudspeaker systems; the modelling of objective parameters under the aspect of psychoacoustic phenomena has also been applied [7]. In order to obtain specific parameters, it was necessary to use both an appropriate kind of the forcing signal and a specific measurement set-up [14]. In the investigations under discussion the *tone burst* signal was used.

Its analytical form can be written as:

$$x(t) = A \sin \omega t \cdot g(t) \tag{1}$$

where $g(t)$ represents the rectangular gate function.

The objective evaluation of the loudspeaker systems was performed on the basis of the following transient parameters:

- duration of the initial and final transients,
- power coefficients of the initial and final transients,
- transient characteristic.

3.2.1. The duration of the initial and final transients. The duration of transient processes connected with the signal rise and decay is one of the fundamental measures of transient distortions introduced by the loudspeaker when it is excited by an impulse. Taking into account the response character of loudspeaker system excited with tone pulses, the duration of initial and final transients are defined as an intervals of time:

$$t_r = t_{01}, \qquad t_d = t_{23}. \tag{2}$$

Figure 5 shows the most characteristic forms of the transition between the steady state and transients. Hence, the determination of the mean value $\overline{x}$ of the signal at the steady state interval t_{12} helps specify, in accordance with the adopted definition, the duration of the transient from the moment when the envelope reaches 0.1 of the mean value in the steady state ($0.1\overline{x}$) until it reaches values lower than 0.9 or higher than 1.1 of the mean value in the steady state ($0.9\overline{x}$ or $1.1\overline{x}$).

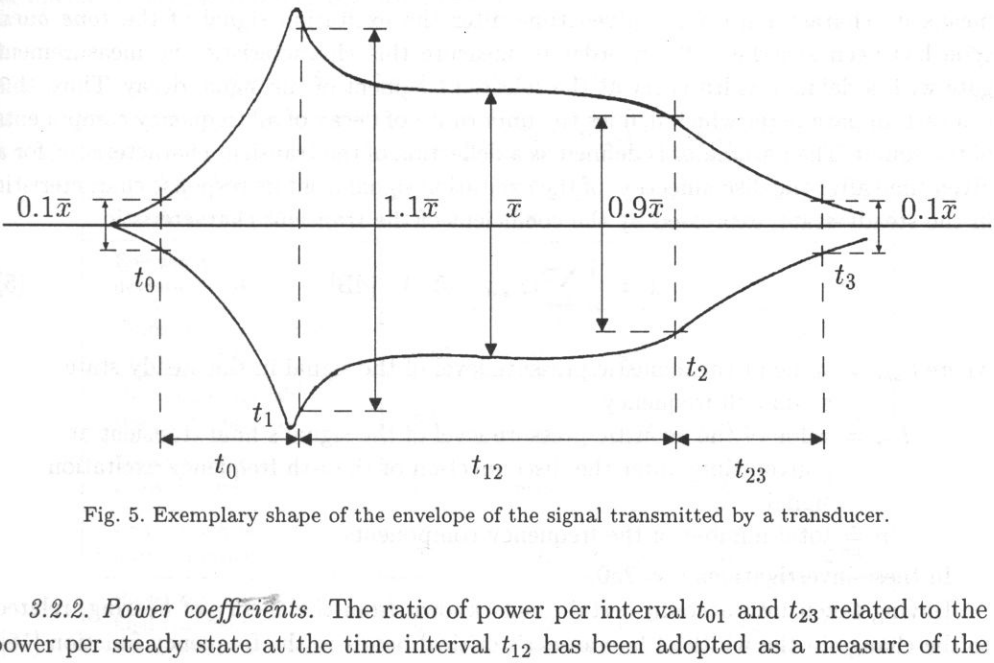


Fig. 5. Exemplary shape of the envelope of the signal transmitted by a transducer.

3.2.2. Power coefficients. The ratio of power per interval t_{01} and t_{23} related to the power per steady state at the time interval t_{12} has been adopted as a measure of the transient power at the signal rise and decay. They discrete forms which are the approximations of the integral formulas were defined as follows:

• for the initial transient

$$M_1 = \frac{\dfrac{\sum\limits_n |X_{01n}|^2}{n_{01}}}{\dfrac{\sum_k |X_{12k}|^2}{k_{12}}}, \tag{3}$$

where $X_{01n} \mathrel{\hat=} n^{\text{th}}$ value of the initial transient amplitude,
$n_{01} \mathrel{\hat=}$ number of samples for the initial transient,
$X_{12k} \mathrel{\hat=} k^{\text{th}}$ value of the steady state amplitude,
$k_{12} \mathrel{\hat=}$ number of samples for the steady state;

• for the final transient

$$M_2 = \frac{\dfrac{\sum\limits_n |X_{23n}|^2}{n_{23}}}{\dfrac{\sum\limits_k |X_{12k}|^2}{k_{12}}}, \tag{4}$$

where $X_{01n} \mathrel{\hat=}$ value of the final transient amplitude,
$n_{23} \mathrel{\hat=}$ number of samples for the final transient.

In the investigations under discussion the sample frequency was 50 kHz.

3.2.3. Transient characteristic. The loudspeaker system's transient characteristic means its characteristic for a given time after the excitation signal of the *tone burst* type has been switched off. In order to measure this characteristic the measurement gate with a defined width is set at the selected fragment of the signal decay. Thus, this is a certain parameter which defines the uniformity of decay of all frequency components of the sound. The parameter is defined as a deflection of the transient characteristic, for a given time after the disconnection of the excitation signal from its response characteristic in the steady state, expressed by the coefficient of the transient characteristic

$$D = \frac{1}{n} \sum_i (L_{ssi} - L_{\tau i}) \quad \text{[dB]}, \tag{5}$$

where $L_{ssi} \mathrel{\hat=}$ value of the acoustic pressure level of the signal in the steady state at its i-th frequency,
$L_{\tau i} \mathrel{\hat=}$ value of the acoustic pressure level of the signal's final transient at a given time, after the disconnection of the i-th frequency excitation signal,
$n \mathrel{\hat=}$ total number of the frequency components.

In these investigations $n = 250$.

Having taken into account certain facts known from the theory of hearing, related to the change in the width of human ear's critical bands in the frequency function [18], the manner of determining transient characteristics has been modified – a measurement gate with varying width has been used. It has been assumed that up to 500 Hz the gate width is 10 ms and above 500 Hz it is $1/0.17\,f$, where f denotes the frequency of the *tone burst* excitation signal. Since $f = 1/T$, the gate width can also be expressed as equal to about 6 T. If the transient characteristics are determined in the above manner, they are subject to a particular "weighting" and, consequently, while retaining the features of objective parameters, they comprise a certain aspect of subjective nature [7].

All objective parameters under investigation were determined in an anechoic chamber. In order to approximate diffraction conditions around the measurement microphone to real conditions around the subject's head, the measurement microphone (condenser microphone B & K 1/4" 4135) was mounted in an artificial head.

3.3. Results of the objective investigations

The results of the objective investigations are presented in Table 1. Upon analysis of these results the following can be stated:

1) the power coefficient of the initial transient M_1 and the coefficient D of the transient characteristic of the loudspeaker systems 2 and 4 differ from the values of the other,

2) the duration of the final transient of the loudspeaker systems 1, 2 and 4 differ from the values obtained for the loudspeaker systems 3 and 5,

3) the value of the transient characteristic deviation of the loudspeaker system 5 significantly exceeds those of the other systems.

Table 1. Objective parameters of loudspeaker systems.

PARAMETER	LOUDSPEAKER SYSTEM				
	1	2	3	4	5
Coefficient D (dB)	15.5	18.1	13.6	17.6	14.0
Frequency characteristic deviation (dB)	2.78	2.66	3.24	2.10	1.87
Transient characteristic deviation (dB)	4.82	3.03	4.91	3.94	7.23
Duration of initial transient (ms)	2.77	1.74	2.21	4.08	3.03
Duration of final transient (ms)	6.54	6.26	3.29	6.40	4.76
Power coefficient of initial transient	0.96	0.91	0.95	0.92	0.96
Power coefficient of final transient	0.11	0.09	0.08	0.12	0.10

4. Conclusion

The attempt to find significant causes (within the physical parameters of the loudspeaker systems) contributing to the occurrence of differences in the subjective evaluation of the sound emitted by a loudspeaker system seems to be successful. It has been found that the amplitude characteristic affects the *sharpness, fullness* and *spaciousness* of the sound and the phase characteristic affects the *fullness* and *spaciousness* of the sound. A simple correlation between the objective evaluation and the subjective parametric evaluation as well as between the objective evaluation and the subjective overall evaluation has been found to determine the relation between objective and subjective evaluations. The value of the coefficient D of transient characteristic and the power coefficient of the initial transient M_1 are the objective parameters which differentiate investigated loudspeaker systems (2, 4 and 1, 3, 5); as a result of the correlation analysis it can be proved that the *fullness* and *spaciousness* are the attributes responsible for the differentiation of the investigated loudspeaker systems. On the other side, a slight correlation between the duration of the initial transient and *sharpness* exists. It has been also shown that the influence of the frequency and phase characteristic depends on the test signal, especially for the white noise signal.

A detailed analysis of the results of the subjective evaluation of transducers showed that the subjective evaluation of sounds obtained was decisively influenced by the work of the transducers in transient states.

References

[1] J.D. Carroll and J.J. Chang, *Analysis of individual differences in multidimensional scaling via an N-way generalisation of "Eckart - Young" decomposition*, Psychometrika, **35**, 283–319 (1970).

[2] A. Dobrucki, C. Szmal and Z. Żyszkowski, *Dyspersja charakterystyk skuteczności głośników jako miara zniekształceń linearnych*, Prace Naukowe Inst. Telek. PWr, **21**, 19–27 (1975).

[3] A. Gabrielsson, *Dimension analysis of perceived sound quality of sound-reproducing systems*, Scand. J. Psychol., **20**, 159–169 (1979).

[4] A. Gabrielsson, H. Sjogren, *Perceived sound quality of sound reproducing systems*, J. Acoust. Soc. Am., **65**, 1919–1933 (1979).

[5] A. Furmann, E. Hojan, M. Niewiarowicz, P. Perz, *On the correlation between the subjective evaluation of sound and the objective evaluation of acoustic parameters for a selected source*, J. Audio Eng. Soc., **38**, 11, 837–844 (1990).

[6] A. Furmann, *Correlation between the subjective and objective evaluation of sounds emitted by a loudspeaker*, Ph. D. dissertation, Institute of Acoustics, UAM, Poznań 1991.

[7] E. Hojan, *Modelling of objective parameters of loudspeakers under the aspect of psychoacoustic data*, Proceedings of the XXXII Open Seminar on Acoustics, Kraków, 375–378 (1985).

[8] E. Hojan, A. Furmann, *Relation between selected attributes of the perception space and the evaluation of sound fidelity*, Archives of Acoustics, **16**, 95–105 (1991).

[9] E. Hojan, M. Niewiarowicz, P. Perz, A. Furmann, *Subjective evaluation of a sound emitted by a loudspeaker in correlation with its steady-state and transient responses*, Proceedings of the International Symposium on Subjective and Objective Evaluation of Sound, Poznań, 53–60 (1990).

[10] A. Illenyi, P. Korpasy, *Correlation between loudness and quality of stereophonic loudspeakers*, Acustica, **49**, 334–336 (1981).

[11] J.B. KRUSKAL, *Multidimentional scaling by optimizing goodness of fit to a nonmetric hypothesis*, Psychometrika, **29**, 1, 3–29 (1964).

[12] W.J.M. LEVELT, J.P. GEER, R. PLOMP, *Triadic comparisons of music intervals*, Brit. J. Math. Statist. Psychol, **19**, 163–179 (1966).

[13] H. MÖLLER, *Loudspeaker phase measurements, transient response and audible quality*, Brüel & Kjaer, 1-14 (1975).

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

[15] E. PAULUS, E. ZWICKER, *Programme zur automatischen Bestimmung der Lautheit aus Terzpegeln oder Frequenzgruppenpegeln*, Acustica, **27**, 253–266 (1972).

[16] H. STAFFELD, *Correlation between subjective and objective data for quality loudspeakers*, J. Audio Eng. Soc., **22**, 402–415 (1974).

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

[18] E. ZWICKER, H. FASTL, *Psychoacoustics – Facts and Models*, Springer Verlag, Berlin 1990.

ARCHIVES OF ACOUSTICS
23, 1, 51–66 (1998)

SYMMETRY AND ASYMMETRY AS A PHYSICAL AND PERCEPTUAL FEATURE OF THE COMPLEMENTARY PAIR OF BEATING SINUSOIDS PART I. AMPLITUDE AND FREQUENCY ENVELOPE RELATIONS

L. RUTKOWSKI

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

Beating sinusoids are an interesting case of a simultaneous change of intensity and frequency achieved without the need of a modulator. Studies of the perception of beats provide numerous data concerning also the sound pitch perception. Hitherto, the following conclusions have been made from those studies: i) if the amplitude of one tone is much larger than the amplitude of the other one, of the two-tone complex, the pitch shifts towards the frequency of the larger amplitude tone; ii) if the amplitudes of the two tones are the same, the pitch is localized precisely at the arithmetic average of the two tone frequencies. These statements imply therefore, that a symmetry with respect to the arithmetic average frequency of the two-tone beatings is present in the pitch localization on the frequency scale. Most recent studies show, however, that this symmetry is not always maintained. In the current study, divided into Part 1 and Part 2, an attempt is made, basing on the discussion and numerical analysis of the functions which describe the beatings, to determine the cause of this asymmetry. One of the arguments may come from the fact that the narrow-band condition for beating waveforms is only partially satisfied. This implies that the consequences of the relative rate of changes of the amplitude envelope to the resultant frequency envelope should be considered in the analysis of the beatings signal. The lack of symmetry is evidenced by the functions which reflect the influence of the magnitude of the ratio of the amplitudes of two signal components on the values of the normalised parameters EWAIF (Envelope Weighted Average of Instantaneous Frequency) and IWAIF (Intensity Weighted Average of Instantaneous Frequency) correlated with the sound pitch. In Part 2, two psychoacoustic experiments are described that aimed at the examination of the pitch of beatings in view of the symmetry arguments mentioned above. Main conclusions obtained in this part of the study are used throughout together with the literature available on this subject.

1. Introduction

In spite of its rudimentary character, the phenomenon of beating sinusoids continues to be the subject of interest of both the theory of modulation [9, 15, 16] as well as the studies of the perception of sounds with time-varying parameters.

The beating waveform, also referred to as TCC (Two-Component Complex [4]) as the superposition of two tones with very similar frequencies, may also, from the analytical standpoint, be regarded as an example of the elementary modulation process known as TTSS (Two-Tone Single-Sideband) [15]. The basic feature of the beating is the time-variation of the amplitude envelope which resembles that of the modulated waveforms. Usually, this variation does not arise from nonlinearities of the amplitude (AM) or the frequency modulation (FM) process, but follows from superposition of two signals which is evidently a linear process. In the signal resulting from superposition the features of the two components, in particular their individual time variations, cease to be important. The Fourier spectrum of the beating waveform contains two components of pre-determined frequencies and amplitudes. The width of the Fourier spectrum of beats is equal to the frequency difference of its two components. All these arguments out the elementary properties of the beats phenomenon as long as we confine ourselves to spectral analysis.

The beating waveform manifests itself as a periodic variation of the sound pressure level (SPL) with the repetition time equal to the inverse of the frequency difference of the two components. The SPL variations (envelope of the beating) can be monitored on an oscilloscope or directly perceived as an audible loudness change. It occurs however, that the beating cycle also features an instantaneous frequency variation [1, 2, 4, 8, 14, 15, 16], perfectly synchronised with the variations of the sound pressure level. But from the perceptual point of view and its physical detection, the former is not perceived as easily as the changes in the loudness. Both these features of the beating waveform, i.e. the change of the amplitude envelope and the instantaneous frequency variation, are not resolvable in the Fourier spectrum. Determination of the instantaneous frequency variations in the beating cycle requires frequency demodulation, an instrumental task which is not straightforward within the acoustic frequency range [10]. Hence, the procedure of the demodulation is usually performed employing Hilbert transform algorithms [1, 9] or time-frequency distributions TFD [9]. The perceptual complexity of the beatings results mainly from the simultaneous occurrence of changes of the amplitude envelope together with the instantaneous frequency variations, including the occurrences of extreme frequency changes at the moments which correspond to the minima of the sound amplitude envelope. Perception of the beatings is a part of the studies of signals of variable frequency and the amplitude envelope and provide an insight into problems of simultaneous perception of the loudness and the pitch of sounds.

Investigation of the physical properties and the perception of beatings is also important for the room acoustics, particularly with sounds of time-varying frequency propagating in a room. A beating-like effect occurs in a room [12] when two waves superpose, for instance, the direct wave of an advancing frequency and the reflected one having the same frequency changes but reproducing its previous history of the frequency change. In a specific point of the room there is some frequency difference which causes the beatings and which depends on the velocity, character of the frequency changes and the delay of the reflected wave.

While it is difficult to accept that sound propagation in our environment may involve nonlinear processes of amplitude and frequency modulation, it must be quite natural to expect, for the envelope changes and instantaneous frequency variations, that it results from acoustic waves superposition; there is little doubt that the two features should appear even in geometric structures of moderate complexity. Hence, it seems purposeful and justified to devote some attention to this kind of phenomena.

This study consists of two parts. Part I is mainly devoted to the physical aspect of variations of the beatings signal and its results establish some grounds for the subject of the perception of the beatings pitch that is dealt with in Part II. In its principal contents, the study deals with the so called complementary pair of signals usually referred to as SL and SH [3]. The signal abbreviated as SL (Stronger Low) consists of a pair of tones in which the lower frequency signal is of a higher sound pressure level whereas the reverse is implied for the Stronger High, abbrev. SH. The subject of particular interest of the authors is the intuitive possibility of a symmetry or its absence for the two complementary pairs of signals: SL-SH; those will be discussed on the physical grounds of the beatings (Part I) as well as the perception of the pitch of the two-tone complexes SL-SH. There is already a report by DAI [2] who notifies the problem of the asymmetry in the discrimination of the pitch. Thus far, the reported asymmetries in the perception of the sound pitch remain unexplained.

The analytical representation of the beatings waveform embraces its two cardinal attributes: the changes of amplitude envelope and the instantaneous frequency variations. There is a number of articles in which the changes of the amplitude envelope and the instantaneous frequency variations of the beats have been analysed [1, 3, 8, 14, 15]. In some of these articles the final formulae, describing the changes of amplitude envelope and instantaneous frequency variations of beatings, have been obtained in terms of the analytic signal concept. It looks however that in the majority of the contributions to this subject, the physical aspects of the beats have been treated only marginally. Consequently, the available description has been only superficial and has routinely ignored the essential elements of the problem pursued by the present investigation.

2. The beatings – its analytical description and properties

To derive a formula for the changes of the amplitude envelope and instantaneous frequency variations of a beatings signal, let us assume the real signal $r(t)$ to be a sum of two tones of different amplitudes x_L and x_H and the frequencies f_L and f_H (the subscripts L and H refer to the lower and the higher frequency tones, respectively).

$$r(t) = x_L \cos 2\pi f_L t + x_H \cos 2\pi f_H t . \tag{I.1}$$

Applying the Hilbert transform to the real signal $r(t)$ we obtain

$$\mathrm{Hi}\{r(t)\} = x_L \sin 2\pi f_L t + x_H \sin 2\pi f_H t . \tag{I.2}$$

The analytic signal associated with the real one (I.1) will be

$$r_a(t) = r(t) + j\,\mathrm{Hi}\{r(t)\}. \tag{I.3}$$

The envelope of the analytic signal is obtained from the formula:

$$|r_a(t)| = \sqrt{[r(t)]^2 + [\mathrm{Hi}\{r(t)\}]^2} . \tag{I.4}$$

This envelope is a real function equal to the real signal envelope. For the beatings (I.1), the envelope calculated with the formula (I.4) is

$$e(t) = |r_a(t)| = x_L \sqrt{1 + \delta^2 + 2\delta \cos 2\pi \Delta f t} , \tag{I.5}$$

where $\delta = x_H/x_L$ is the amplitude ratio, whereas $\Delta f = f_H - f_L$ is the frequency difference of the superposed signals which effect the beats.

The phase of the analytic signal equals

$$\varphi_a(t) = \operatorname{arctg}\left[\frac{\mathrm{Hi}\,\{r(t)\}}{r(t)}\right] \quad [\mathrm{rad}]. \tag{I.6}$$

The instantaneous frequency IF(t) is calculated as the time derivative of the phase (I.6) of the analytic signal

$$\mathrm{IF}\,(t) = \frac{1}{2\pi}\,\frac{d\varphi_a(t)}{dt} \quad [\mathrm{Hz}]. \tag{I.7}$$

In practice, for the complex signals covering a certain frequency range, the phase changes are defined with reference to a fixed, constant frequency, for instance f_L, its phase being a linear function of time $\varphi(f_L, t) = 2\pi f_L t$. Then, the instantaneous frequency variations can be written as

$$\mathrm{IF}\,(t) = \frac{1}{2\pi}\,\frac{d\widetilde{\varphi}_a(t)}{dt} + f_L\,, \tag{I.8}$$

where $\dfrac{1}{2\pi}\,\dfrac{d\widetilde{\varphi}_a(t)}{dt}$ is the time dependent part of the instantaneous frequency.

Hence, the resulting signal, associated to the real signal $r(t)$ (Eq. (I.1)) may be rewritten in the following form:

$$r(t) = |r_a(t)| \cos\left[2\pi \int \mathrm{IF}\,(t)\,dt\right] = \mathrm{Re}\left\{\exp\left[\ln|r_a(t)| + j2\pi \int \mathrm{IF}\,(t)\,dt\right]\right\}. \tag{I.9}$$

This equation points out the possibility of a generalisation of the signal phase (see also [13]) by regarding the changes of the amplitude envelope as a factor in the phase of the signal $r(t)$.

Therefore, the expression

$$\mathrm{CI\Phi}(t) = \ln|r_a(t)| + j2\pi \int \mathrm{IF}\,(t)\,dt = \ln|r(t)| \tag{I.10}$$

is the complex phase of the resultant signal. The complex instantaneous frequency of the signal becomes

$$\begin{aligned}\mathrm{CIF}\,(t) = \frac{1}{2\pi}\,\frac{d\,\mathrm{CI\,\Phi}(t)}{dt} &= \left[\frac{1}{2\pi}\,\frac{1}{|r_a(t)|}\,\frac{d|r_a(t)|}{dt} + j\,\mathrm{IF}\,(t)\right] \\ &= \frac{1}{2\pi}\,\frac{1}{|r(t)|}\,\frac{d|r(t)|}{dt} \quad [\mathrm{Hz}]. \end{aligned} \tag{I.11}$$

It follows from Eq. (I.11) that the complex instantaneous frequency is a function which most generally describes the changes of the real signal $r(t)$, disregarding them if they are associated with phase or the amplitude envelope changes. The magnitude of the complex instantaneous frequency (a real function describing the changes of frequency) of the resulting signal $r(t)$ can be calculated according to the equation

$$|\mathrm{CIF}\,(t)| = \sqrt{\left[\frac{1}{2\pi}\,\frac{1}{|r_a(t)|}\,\frac{d|r_a(t)|}{dt}\right]^2 + [\mathrm{IF}\,(t)]^2} \quad [\mathrm{Hz}]. \tag{I.12}$$

Note that only when the amplitude envelope is constant, the magnitude of the complex instantaneous frequency is equal to the instantaneous frequency IF(t).

Applying Eq. (I.8) and (I.11), the complex instantaneous frequency of the beatings will be

$$\mathrm{CIF}\,(t) = -\frac{\delta\Delta f\sin(2\pi\Delta ft)}{1+\delta^2+2\delta\cos(2\pi\Delta ft)} + j\frac{\delta\Delta f(\cos(2\pi\Delta ft)+\delta)}{1+\delta^2+2\delta\cos(2\pi\Delta ft)} + jf_L\,. \tag{I.13}$$

The first part of the complex instantaneous frequency (the real part) results exclusively from the amplitude envelope changes; the second and the third terms (imaginary part) involve instantaneous frequency variations and the constant tone frequency f_L. The real part of the instantaneous frequency has been usually omitted in the literature, however, for acoustic signals of a variable amplitude envelope or for signals which do not satisfy the narrow band condition, it ought to be taken into account.

The magnitude of the complex instantaneous frequency of the beats (a real function describing frequency variations) equals according to the equation (I.12)

$$|\mathrm{CIF}\,(t)| = \sqrt{\left[\frac{\delta\Delta f\sin(2\pi\Delta ft)}{1+\delta^2+2\delta\cos(2\pi\Delta ft)}\right]^2 + \left[\frac{\delta\Delta f(\cos(2\pi\Delta ft)+\delta)}{1+\delta^2+2\delta\cos(2\pi\Delta ft)} + f_L\right]^2}\,. \tag{I.14}$$

Generally, the variability of the beats is described with two basic real functions: the amplitude envelope (I.5),that is the magnitude of the analytic signal and the frequency envelope (I.14) which is the magnitude of the complex instantaneous frequency of the analytic signal.

We now differentiate $r(t)$ Eq. (I.9) against time

$$\frac{dr(t)}{dt} = r(t)\,\mathrm{Re}\left\{\left[\frac{1}{|r_a(t)|}\,\frac{d|r_a(t)|}{dt} + j\,\mathrm{IF}\,(t)\right]\right\}. \tag{I.15}$$

The expression in square brackets (I.15) is the complex instantaneous frequency. An essential physical interpretation follows from the Eq. (I.15), namely that the complex instantaneous frequency can be regarded as a generalised, relative rate of changes of the signal $\dfrac{1}{r(t)}\,\dfrac{dr(t)}{dt}$ involving both the envelope and/or the frequency changes. At the same time, it is a quantity which, to certain extent, describes the real time evolution of any signal because it contains only the first derivative of the signal with respect to time. Under certain conditions, however, namely for a narrow-band signal, the complex instantaneous frequency (Eq. (I.15)) may adequately account for the signal variability.

Generally, using polar coordinates, we may now write

$$\mathrm{CIF}\,(t) = |\mathrm{CIF}\,(t)|\exp\left[\varphi_{\mathrm{CIF}}(t)\right], \tag{I.16}$$

where

$$\varphi_{\mathrm{CIF}}(t) = \mathrm{arctg}\left[-\frac{\delta\Delta f(\cos(2\pi\Delta ft)+\delta) + f_L(1+\delta^2+2\delta\cos(2\pi\Delta ft))}{\delta\Delta f\sin(2\pi\Delta ft)}\right] \tag{I.17}$$

is the phase angle between the imaginary and real components of the complex instantaneous frequency variations of beatings. The nonlinear dependence of φ_{CIF} on time points out the existence of modulation of the phase angle of the complex instantaneous frequency.

3. Results of calculations of the beat changes

For the analysis of physical properties of the two-tone beatings, the formulae derived in section 2 were used in the numerical calculations. The outcome of these calculations, obviously limited to a few selected parameters of the beatings, present some cognitive value, and will be applied in Part II of the present study in the interpretation of results of the investigation of perception processes.

In Fig. 1, envelope changes $e(t)$ (normalised to unity) (a), changes of the real part of the complex instantaneous frequency Re CIF (t) (b), and the imaginary part of the

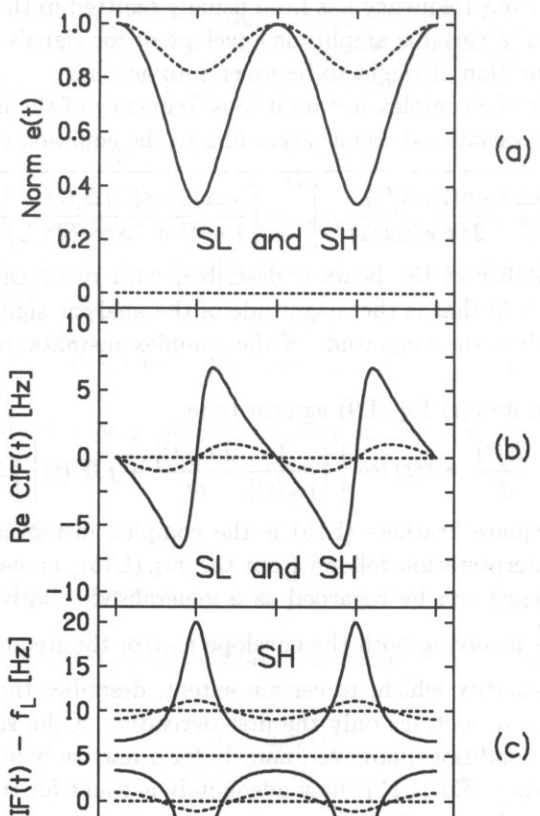


Fig. 1. Basic temporal relations for beating sinusoids of the lower frequency $f_L = 1000$ Hz and the frequency difference of 10 Hz for the complementary pairs SH – stronger high and SL – stronger low (see text); (a) the normalised envelopes, (b) the real parts of the Complex Instantaneous Frequency, (c) the imaginary parts of the Complex Instantaneous Frequency diminished by f_L. (SL – amplitude ratio 0.5 (solid line), amplitude ratio 0.1 (dashed line)), (SH – amplitude ratio 2 (solid line), amplitude ratio 10 (dashed line)).

complex instantaneous frequency changes $\operatorname{Im}\mathrm{CIF}(t)$ (c) diminished by $f_L = 1000\,\mathrm{Hz}$, the lower tone frequency; are shown. The frequency difference of the beats signal equals 10 Hz, the ratio of the amplitudes δ with SL marks is 0.5 (solid line), 0.1 (dashed line), whereas those marked with HL are $\delta = 2$ and 10 (solid and dashed lines, respectively). Thus, the situations discussed here, i.e. when $\delta < 1$ and $\delta > 1$ (or $\delta = 1/\delta$), refer to the two complementary pairs of signals which were labelled earlier SL ($\delta < 1$) and HL ($\delta > 1$).

The following conclusions can be drawn from the discussion of the plots in Fig. 1:

- variations of the imaginary part of the complex instantaneous frequency (Fig. 1c) exhibit a symmetry with respect to the arithmetic average of the two frequencies of the beating components SL and SH, both for small and large magnitudes of δ,
- at $\delta = 0.5$ (with the pair denoted SL) and at $\delta = 2$ (with the pair marked SH), the variations of the imaginary part of the complex instantaneous frequency are non-symmetrically displaced from the frequency f_L for $\delta < 1$ and from the frequency $f_H = f_L + 10\,\mathrm{Hz}$ for $\delta > 1$; respective (see also [3]) average values of these variations equal f_H and f_L,
- the changes of the amplitude envelope (Fig. 1a) are identical for the two complementary pairs of signals SL and SH; the minimum of the normalised amplitude envelope curve is 0.818 for $\delta = 0.1$ (both SL and SH) and 0.33 for $\delta = 0.5$ (SL and SH),
- identical changes of the amplitude envelope lead to identical variations of the real part of the complex instantaneous frequency (Fig. 1b) of zero mean value; it does not depend on which pair of signals, SL or SH, is discussed,
- at $\delta = 0.1$ with the SL pair and $\delta = 10$ with the SH pair, the graphs of all the functions presented in Figs. 1b and 1c look like sinusoidal curves, i.e. they oscillate more or less around f_L at $\delta < 1$ and $f_H = f_L + 10\,\mathrm{Hz}$ at $\delta > 1$,
- at the moments corresponding to the maximum value of the envelope amplitude, the magnitudes of the imaginary part of the complex instantaneous frequency (Fig. 1c) reach a maximum for the SL pair and a minimum for the SH pair.

4. Weighted changes of the frequency envelope

The concepts of amplitude-envelope weighted or squared envelope weighted signal frequency variations, is widely used, mainly in studies concerning the perception of the sounds in which changes of amplitude and frequency coexist. FETH *et al.* [4] proved a considerable correlation between the perceived pitch of the beatings and envelope-weighted average of instantaneous frequency (EWAIF). Next, ANANTHARAMANN *et al.* [1] proposed that the application of sound intensity (correctly: squared envelope) variations (IWAIF: *Intensity-Weighted Average of Instantaneous Frequency*) as a weighting function for evaluation of the pitch for beatings yields better accord with subjective investigations. IWAMIYA *et al.* [6, 7] used the signal envelope function to weight the instantaneous frequency variations in the determination of the so called principal pitch of sounds which were simultaneously amplitude and frequency modulated; a model for the vibrato achieved by musicians during instrumental performance was achieved in this way. The comparison of Figs. 1a and 1c shows also that the instantaneous frequency variations taking place in the vicinity of maximum of beatings amplitude has a more pronounced effect on the perception of the pitch than those occurring within the intensity minimum. From the perceptual point of view, these problems will later be discussed more thoroughly in Part II.

On the frequency axis, the envelope weighted average of instantaneous frequency [1, 2] defines the coordinate of a centre of gravity of the figure set out by the signal spectrum which, in the time domain, can be determined from the formula

$$\mathrm{EWAIF} = \frac{\int_0^t e(t) f(t)\, dt}{\int_0^t e(t)\, dt}, \tag{I.18}$$

where $e(t)$ is the amplitude envelope (weighting function) and $f(t)$ describes the frequency variations. For periodical changes of amplitude and frequency, the upper bound of integration ought to be made a multiple of the period of these changes. Similarly, the squared envelope (intensity) weighted average of the instantaneous frequency can be defined with [1, 3]

$$\mathrm{IWAIF} = \frac{\int_0^t e^2(t) f(t)\, dt}{\int_0^t e^2(t)\, dt}. \tag{I.19}$$

Some problems may arise, when attempting numerical calculations with the formulae (I.18) and (I.19), for instance, when the envelope amplitude approaches the zero value (two-tone beatings of amplitude ratio close to 1). For this reason ANANTHARAMAN [1] recommends a spectral method of the calculation of EWAIF and IWAIF. Given a preset resolution of analysis, the spectral method, however, is accurate only for large frequency separations of the two tones of the beatings. The error may be significant when these separations are small. The usage of the spectral method may be attributed rather to the wide availability of FFT algorithms. Investigation of the time evolution of the frequency $f(t)$ requires a demodulation of the signal frequency, involving the calculation of the Hilbert transform; the latter is not so easily available and so popular as the Fast Fourier Transform, FFT. From the perceptual point of view, not only the spectrum of the beatings, but also the pattern of the time variations of the instantaneous frequency is important.

The functions defined by Eq. (I.14) and illustrated as examples in Fig. 1 are necessary if the two weighting representations of the beatings, (I.18) and (I.19), have to be determined. For the appropriate complementary pairs of signals SH and SL the beatings envelopes are identical, whereas the frequency variations are described by distinctive functions. Also, it is of importance which function will be adopted for the description of the frequency dependence on time $f(t)$. For signals of constant amplitude and variable frequency, for example for FM signals, $f(t)$ describes directly the variations of frequency effected by the modulation process. However, if the signal, for which the weighted values of instantaneous frequency are to be evaluated, exhibits both envelope and frequency variations, like the beatings, the formulae (I.18) and (I.19) have to be modified by the frequency function expressed through the changes of the frequency envelope (I.14).

In order to demonstrate the importance of the frequency envelope variations in evaluating the weighted values defined by (I.18) and (I.19), the two quantities (normalised through dividing by $\Delta\omega$) were calculated as a function of the amplitude ratio of the two components SL and SH of beatings in the $< 0.5, 1 >$ and $< 1, 2 >$ ranges, respectively. Later in the text, we shall refer to the normalised values of EWAIF and IWAIF as NEWAIF and NIWAIF. They are shown subsequently in Figs. 2–4 with the frequency $f_L = 480\,\text{Hz}$ and the frequency difference $\Delta f = 40\,\text{Hz}$.

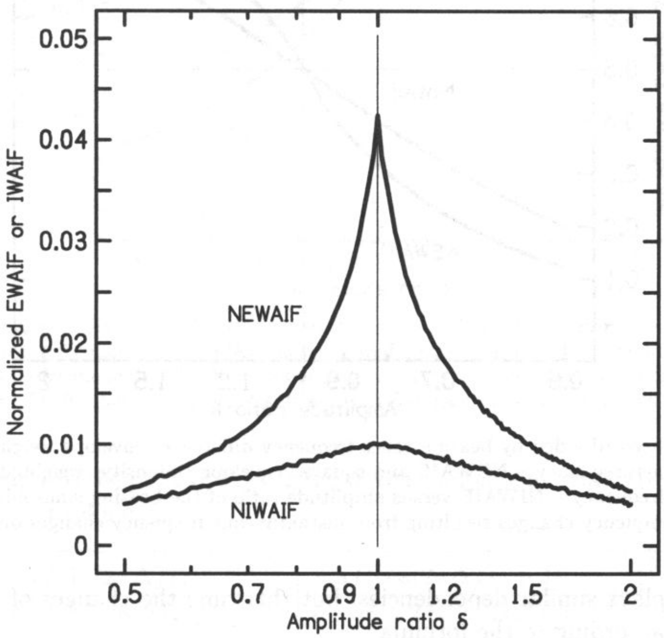


Fig. 2. Normalised (divided by beating tones frequency difference) envelope weighted average of instantaneous frequency – NEWAIF and squared envelope (intensity) weighted average of instantaneous frequency – NIWAIF versus amplitude ratio of the beating sinusoids. Numerical calculations for frequency changes resulting from the relative envelope changes only (Eq. (I.20)).

In Fig. 2 there are plots of the normalised envelope weighted average of the instantaneous frequency and of the normalised intensity weighted average of instantaneous frequency vs. δ, i.e. vs. the ratio of the two amplitudes; the frequency changes are linked to the changes of the amplitude envelope of the beatings via the formula:

$$f(t) = \sqrt{\left[\frac{\delta\Delta f(\sin(2\pi\Delta ft)}{1+\delta^2+2\delta\cos(2\pi\Delta ft)}\right]^2 + f_L^2} - f_L\,. \tag{I.20}$$

Thus, in the plots of Fig. 2 only the real part of complex instantaneous frequency (I.13) was accounted for, while the imaginary part was omitted. One may see from these plots that both relations reach maximal values at $\delta = 1$ (largest modulation depth of the beatings). It is also interesting that for the complementary pairs of the signals SL and SH these functions display the same values (parity).

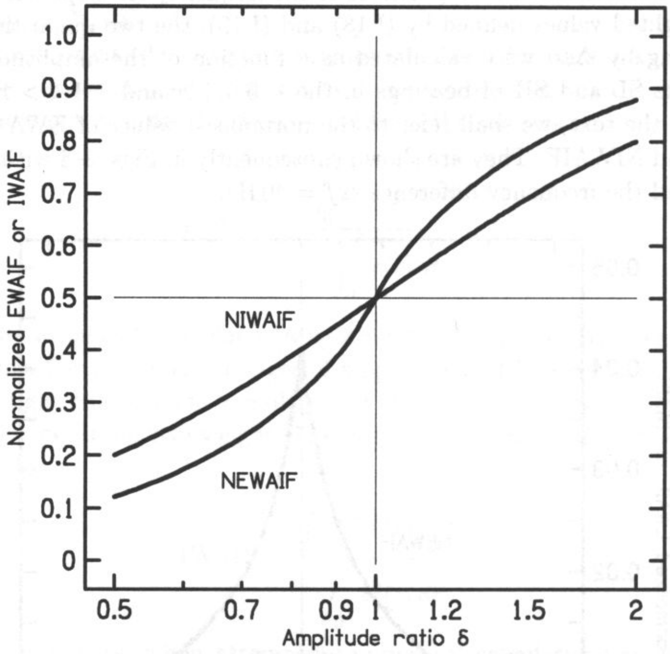


Fig. 3. Normalised (divided by beating tones frequency difference) envelope weighted average of instantaneous frequency – NEWAIF and squared envelope (intensity) weighted average of instantaneous frequency – NIWAIF versus amplitude ratio of the beating sinusoids. Numerical calculations for frequency changes resulting from instantaneous frequency changes only (Eq. (I.21)).

Figure 3 displays similar dependencies, but this time the changes of the frequency are determined according to the formula

$$f(t) = \sqrt{\left[\frac{\delta\Delta f(\cos(2\pi\Delta ft) + \delta)}{1 + \delta^2 + 2\delta\cos(2\pi\Delta ft)} + f_L\right]^2} - f_L\,. \tag{I.21}$$

In this case only the imaginary part of the frequency envelope (I.13), usually referred to as instantaneous frequency IF(t), (Eq. (I.7)), was accounted for. The coordinates of the point, with respect to which the symmetry may be discussed, can be set at $\delta = 1$ on the abscissa and at NEWAIF (NIWAIF) $= 0.5$ on the ordinate. The following relation is valid: the magnitude of NEWAIF and NIWAIF for the pair denoted as SH equals 1 minus the value of NEWAIF and NIWAIF for the pair denoted as SL. At $\delta = 1$, thus when the amplitudes of the two signals are equal (SL = SH), both functions amount to 0.5, i.e. the resultant frequency yields $f_L + 0.5\Delta f$ which corresponds to arithmetic average frequency of the two component signals of the beatings. The functions describing the graphs displayed in Fig. 3 are odd with regard to the argument $\delta = 1$.

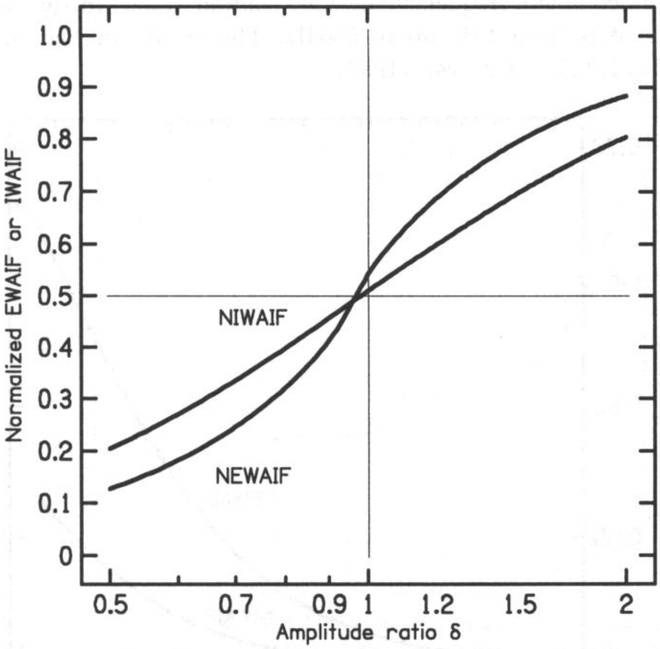


Fig. 4. Normalised (divided by beating tones frequency difference) envelope weighted average of instantaneous frequency – NEWAIF and squared envelope (intensity) weighted average of instantaneous frequency – NIWAIF versus amplitude ratio of beating sinusoids. Numerical calculations for frequency changes resulting from frequency envelope changes (Eq. (I.22)).

Next, similarly to the plots of Figs. 2 and 3, the complete formula $f(t)$ for the changes of the frequency envelope, i.e.

$$f(t) = \sqrt{\left[\frac{\delta \Delta f \sin(2\pi \Delta f t)}{1+\delta^2+2\delta\cos(2\pi\Delta f t)}\right]^2 + \left[\frac{\delta\Delta f(\cos(2\pi\Delta f t)+\delta)}{1+\delta^2+2\delta\cos(2\pi\Delta f t)} + f_L\right]^2} - f_L\,, \qquad \text{(I.22)}$$

was employed to produce the dependencies shown in Fig. 4.

The appearances of the graphs in the Figs. 3 and 4 are identical. A detailed analysis of the numerical data points out certain differences, especially in the vicinity of $\delta = 1$. The crossing point of the two curves, NEWAIF(δ) and NIWAIF(δ), is slightly shifted to the left from the value $\delta = 1$, whereas the magnitudes of the two functions are less then 0.5. With the amplitude ratio δ equal one, both the normalised functions exceed 0.5.

Hence, following the results of calculations, an asymmetry of the NEWAIF(δ) and NIWAIF(δ) curves occurs with regard to $\delta = 1$. Till now, at the special point $\delta = 1$ NEWAIF and NIWAIF were believed to be exactly equal to an average frequency of the two beating tones (normalised functions equals 0.5). The observed asymmetry results from taking into account the real part of the complex instantaneous frequency. Let us take a closer look at the dependence of NEWAIF and NIWAIF on the frequency separation Δf of the two tones. Figures 5 and 6 show the results of calculations of the two functions NEWAIF(Δf) and NIWAIF(Δf) performed for the frequency variations weighted with the complete frequency envelope contour (I.22). In the two figures $f_L = 500\,\text{Hz}$ and Δf varies from 1 Hz up to 500 Hz. The graphs in Figs. 5, 6 and 7 were obtained with $\delta = 1$, 0.5 and 2, respectively.

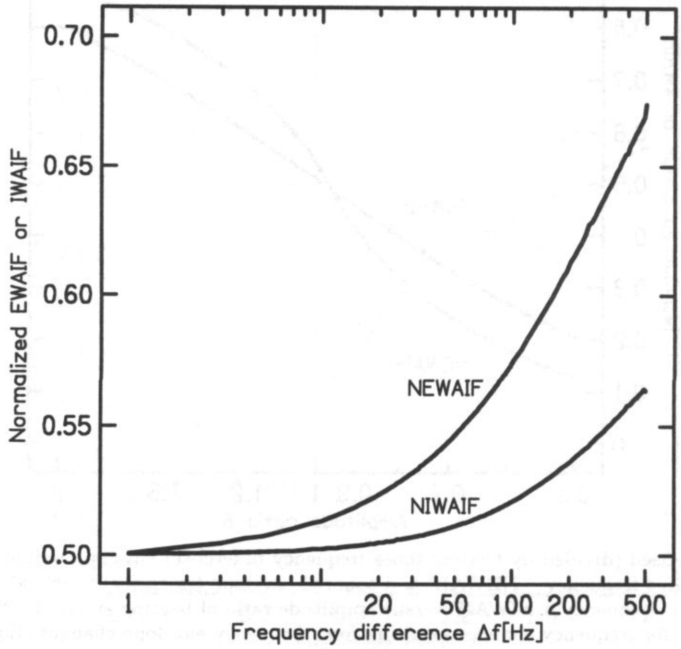


Fig. 5. Illustration of the NEWAIF and NIWAIF changes as function of the frequency difference between the beating sinusoids. The frequency changes for calculations are taken from Eq. (I.22). The lower frequency is constant and equals 500 Hz. The amplitude ratio of two components $\delta = 1$.

If there was a full symmetry for the data illustrated in Fig. 5, then the values of NEWAIF and NIWAIF should be constant and equal to 0.5, independent of the magnitude of Δf. Similarly, for the data presented in Fig. 6 and Fig. 7, the values of EWAIF and IWAIF would not be related to Δf and correspond to the appropriate ordinates read at the lowest range of frequencies (here, $\Delta f = 1\,\text{Hz}$).

The graphs in all three Figs. 5, 6 and 7, illustrate the essential problem often disregarded in the context of acoustic waveforms: the problem of a narrow band property of signals. The narrow band signal, in the case of beatings, is a signal for which $\Delta f / f_L \ll 1$, whereas the broadband criterion is the similarity of the magnitudes of the frequency difference Δf and of the lower frequency of the complex signal, i.e. $\Delta f / f_L \sim 1$. Figures 5, 6 and 7 demonstrate that a significant effect of the real part of the complex instantaneous frequency on the frequency envelope is observed when the beatings do not obey the narrow band condition (this can be extended to other signals, too). It remains an open question, if the mathematically correct calculations of the envelope and squared envelope averages of the instantaneous frequency values would, in the entire Δf range, match the subjective impression of the pitch. Eventually, these values are intended for the evaluation of the latter.

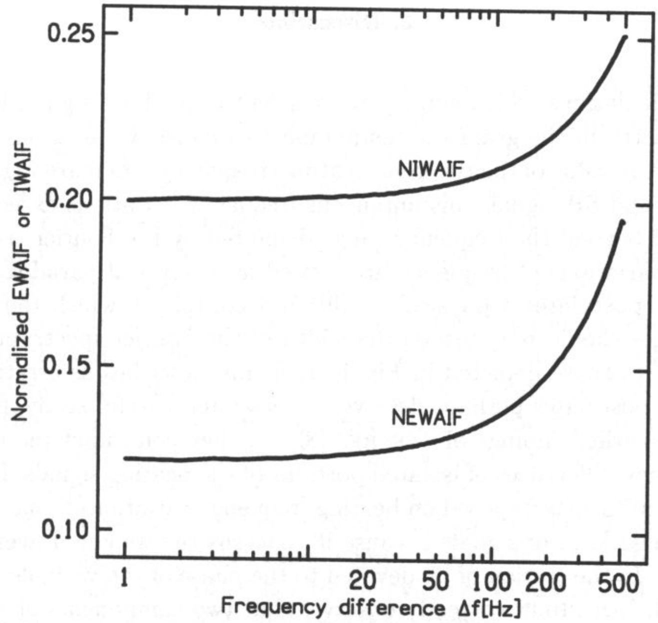


Fig. 6. Illustration of the NEWAIF and NIWAIF changes as function of the frequency difference between the beating sinusoids. The frequency changes for calculations are taken from Eq. (I.22). The lower frequency is constant and equals 500 Hz. Amplitude ratio $\delta = 0.5$.

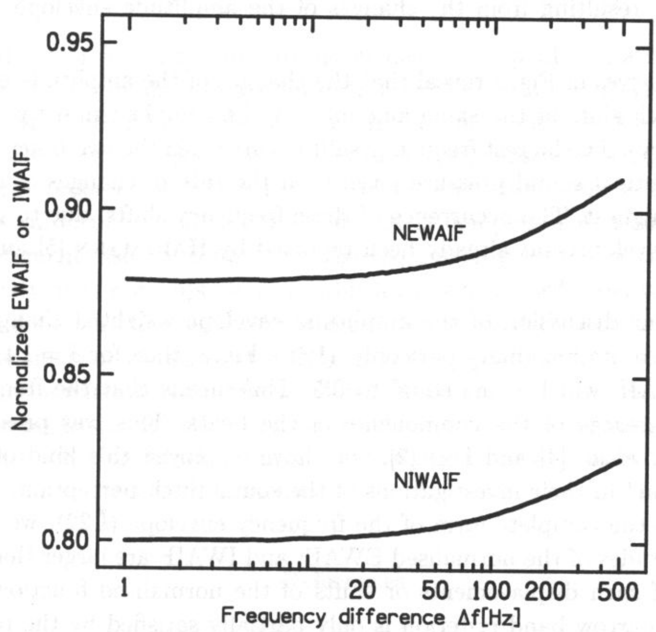


Fig. 7. Illustration of the NEWAIF and NIWAIF changes as function of the frequency difference between the beating sinusoids. The frequency changes for calculations are taken from Eq. (I.22). The lower frequency is constant and equals 500 Hz. Amplitude ratio $\delta = 2$.

5. Discussion

From the calculations which employ the concept of analytic signal, it follows, that there is a symmetry in the graphs of instantaneous frequency variations IF(t) with respect to the average value of the two-tone beatings frequency (compare Fig. 1c). However, for both the SL and SH signals, instantaneous frequency variations exceed and, in case of $\delta \approx 1$, go well beyond the frequency range delimited by the Fourier transform of the beatings. Such variations of frequency are viewed as a physical paradox by LOUGHLIN *et al.* [9] and he postulates a physical condition according to which the instantaneous frequency changes should not surpass the width of the Fourier spectrum of the signal. Changes similar to those depicted in Fig. 1c Loughlin describes as erratic. Though his suggestions and postulates perhaps deserve a special and careful study, they are not in accord with the earlier findings of JEFFRES [8], i.e. they contradict the experimentally observed frequency differences of isolated portions of the beating signals. Loughlin's idea of the physical limitations imposed on beating frequency variations remains closer to the perception of these kind of signals because it concerns the weighted average frequency of the two tones. In the section of [9] devoted to the phase of the variable amplitude and frequency signal, LOUGHLIN correctly observes that two components of the phase have to be considered, i.e. φ_F, the derivative of which yields the instantaneous frequency IF, and φ_A effected by the amplitude envelope changes (see also RUTKOWSKI [13]). Such an approach was also adopted in the present study (Eq. (I.9) and (I.10)) and the frequency patterns, resulting from the changes of the amplitude envelope, are displayed in Fig. 1b.

Normalised curves in Fig. 2 reveal that the changes of the amplitude envelope alone may cause a pitch shift of the same amounts for the complementary pairs of beating signals (symmetry). The largest frequency shift occurs when the two tones of the beating signal are of identical sound pressure level; then the rate of changes of the amplitude envelope is the highest. The occurrence of these frequency shifts, due to the changes of the amplitude envelope, has already been reported by HARTMANN [5] and ROSSING *et al.* [11].

If we limit our discussion of the amplitude envelope weighted changes of the frequency envelope to its imaginary part only, (IF) – Fig. 3, then for $\delta = 1$ the normalised EWAIF and IWAIF will become equal to 0.5. This means that the frequency will be the arithmetic average of the components of the beats. This was presumed, among others, by FETH *et al.* [4] and DAI [2], who have employed this kind of signal as an "adjustable signal" in their investigations of the sound pitch perception. However, taking into account the complete form of the frequency envelope (I.22), we found that at $\delta = 1$ the magnitudes of the normalised EWAIF and IWAIF are larger than 0.5 (Fig. 4). The existence of such displacements or shifts of the normalised functions magnitudes proves that the narrow band criterion is only partially satisfied by the beatings (compare Fig. 5). These shifts occur also with other values of the amplitude ratio (compare Figs. 6 and 7) and ought to be considered every time when the condition $f_L \gg \Delta f$ is not fulfilled.

6. Conclusions

Summing up the above analysis of physical features of beatings, it can be stated that the signal variability is mainly described by two functions: the amplitude envelope (I.5) and the frequency envelope (I.14). As the amplitude envelope, or its square, is directly related to the changes of the signal sound pressure level, the frequency envelope tells us the rate with which phase variations occur as well as the relative speed of the amplitude envelope variations. Both the rate of phase variations and the relative rate of the envelope changes exhibit the same dimension-frequency. Association of the time evolution of the amplitude envelope of beats and the variations of the frequency gives evidence (see Fig. 1) that at values of the amplitudes ratio approaching 1, the frequency variations occurring near the maximum of the amplitude envelope ought to be estimated more consequently than those in the vicinity of the envelope minimum. In the light of this finding, the concept of two sets of values, the envelope and squared envelope weighted averages of instantaneous frequency, appears to be correct and entirely justified. The performed analysis indicates, however, that it is not trivial which form of the frequency variations is used in the calculations. Using the so called complete formula for the frequency changes, i.e. the enveloped frequency (I.14), leads to the evidence of asymmetries (Fig. 4) in the course of the normalised EWAIF(δ) and IWAIF(δ) curves.

The above conclusions are quite general and may be applied to analysis of changes of the frequency envelope of arbitrary signals, featuring concurrent variations of amplitude and frequency envelopes known as MM (Mixed Modulation) or CM (Combined Modulation).

In the analysis of beatings, due attention has to be paid to narrow-band and/or wide-band aspects of signals that are important in the question of symmetry or asymmetry of the complementary pairs of two-tone complex signals. As demonstrated in section 4, when the distance between the beating components increases, the trend of the variations of the frequency envelope is to an increasing extent controlled by the rate of changes of the amplitude envelope (the real part of complex instantaneous frequency). The above statement also holds for other signals, not only for the beatings.

Unfortunately, it is not possible to establish the exact border between narrow-band and broad-band signals. Consequently, it must be accepted that each signal (this concerns especially the acoustic signals) with time-varying parameters always exhibits some departure from the narrow-band criteria. When determining the attributes of the signals variability, one ought to use the function which describes the frequency envelope of a sound. It is this function which permits to establish to what extent the narrow-band attribute determines the signal variability.

References

[1] J.N. Anantharaman, A.K. Krishnamurthy, L.L. Feth, *Intensity-weighted average of instantaneous frequency as a model for frequency discrimination*, J. Acoust. Soc. Am., **94**, 2, 723–729 (1993).

[2] H. Dai, *On the pitch of two-tone complexes*, J. Acoust. Soc. Am., **94**, 2, 730–734 (1993).

[3] L.L. Feth, *Frequency discrimination of complex periodic tones*, Perception & Psychophysics, **15**, 2, 375–378 (1974).

[4] L.L. Feth, H. O'Malley, J. Ramsey, *Pitch of unresolved, two-component complex tones*, J. Acoust. Soc. Am., **72**, 5, 1403–1412 (1982).

[5] W.M. Hartmann, *The effect of amplitude envelope on the pitch of sine wave tones*, J. Acoust. Soc. Am., **63**, 4, 1105–1113 (1978).

[6] S. Iwamiya, K. Fujiwara, *Perceived principal pitch of FM-AM tones as a function of the phase difference between frequency modulation and amplitude modulation*, J. Acoust. Soc. Jpn. (E), **6**, 3, 193–202, 1985.

[7] S. Iwamiya, S. Nishikawa, O. Kitamura, *Perceived principal pitch of FM-AM tones when the phase difference between frequency modulation and amplitude modulation is in-phase and anti-phase*, J. Acoust. Soc. Jpn. (E), **5**, 2 (1984).

[8] L.A. Jeffres, *Beating sinusoid and pitch changes*, J. Acoust. Soc. Am., **43**, 6, 1464, 1968.

[9] P.J. Loughlin, B. Tacer, *On the amplitude- and frequency-modulation decomposition of signals*, J. Acoust. Soc. Am., **100**, 3, 1594–1601 (1996).

[10] E. Ozimek, L. Jugowar, L. Rutkowski, *Problem of the instantaneous sound frequency measurement*, Archives of Acoustics, **9**, 4, 325–339 (1984).

[11] T.D.Rossing, A.J.M. Houtsma, *The effect of amplitude envelope on the pitch of short sine-wave and complex tones*, ICA 12- Toronto 1986, **B2-1**.

[12] L.Rutkowski, *A comparison of the frequency modulation transfer function with the modulation transfer function in a room*, Applied Acoustics, **49**, 4, 307–320 (1996).

[13] L. Rutkowski, *Room response for frequency change and its relation to the pitch changes*, Archives of Acoustics, **21**, 2, 201–214 (1996).

[14] L. Rutkowski, *Complex instantaneous frequency changes as a result of superposition of the 2 tones* [in Polish], OSA'94, Wrocław-Szklarska Poręba, 281–284 (1994).

[15] H.B. Voelcker, *Toward a unified theory of modulation. Part I: Phase-envelope relationships*, Proceeding of the IEEE, **54**, 3, 340–353 (1966).

[16] H.B. Voelcker, *Toward a unified theory of modulation. Part II: Zero manipulation*, Proceeding of the IEEE, **54**, 5, 735–755 (1966).

ARCHIVES OF ACOUSTICS
23, 1, 67–85 (1998)

SYMMETRY AND ASYMMETRY AS A PHYSICAL AND PERCEPTUAL FEATURE OF THE COMPLEMENTARY PAIR OF BEATING SINUSOIDS PART II. PITCH MATCHING EXPERIMENTS

L. RUTKOWSKI, A. WICHER

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

Two experiments were designed and performed in that the pitch of complementary pairs of beats signals, referred to as SL and SH, was evaluated. In experiment 1, the changeable signal consists of a pair of signals of equal sound pressure level and of the same separation of frequencies as this between SL and SH. Experiment 2 was made in the AM mode with a changeable signal whose amplitude envelope is identical with that of the SL/SH pair; the carrier frequency could be varied during the experiment. The results of the experiment 1 do not confirm the asymmetry in the evaluation of the pitch of the pairs of SL/SH signals with respect to the average frequency of the beats, a feature that has been reported earlier in the literature. On the order hand, the outcome of the experiment 2 point to certain asymmetry, moreover, the data reveal a large pitch difference of the SL/SH pairs; this difference (for which no explanation has been offered yet) is often much larger than that measured in experiment 1. The observed asymmetry, which is due to frequency variations determined with the frequency envelope (Part I) resulting from the broadband character of the beatings, does not quite justify the range of asymmetries reported in literature.

1. Introduction

From the physical properties point of view, beatings are the signal consisting of concurrent, synchronous variations of the sound pressure level (SPL) and frequency envelope [15]. The problem of perception of the beatings, regarded as an effect of superposition of a pair of tones of slightly different frequencies, mostly reduces to the definition of the resulting pitch and to the discrimination of the two tones depending on their SPL and frequency. So far, there is no doubt that a listener perceiving the pitch localizes it in the frequency range bound up by the components of the two-tone complex of beats. It has also been found that the possibility of discrimination between the two tones is ruled by their separation on the frequency scale relative to the critical bandwidth. All the other observations concerning the beats are to some degree ambiguous and the conclusions are often speculative. It seems that the problem remains in the finding of appropriate physical measures, either new ones or those being a combination of already known parameters, that should well correlate with the results of the perception of beatings. No model is available yet that could unambiguously relate a set of physical parameters and the perceived pitch of signals whose both envelope and frequency are variable in time. Some authors opt for a model of perception based on the time domain representation of a signal, others prefer a spectral mechanism of perception. Therefore, the problem arises what is more appropriate: to search the physical parameters relevant to the time domain realization of beatings, or investigating those associated with their spectral properties? IWAMIYA maintains that his experimental findings [7, 8] support the model according to which the perception of the sound envelope and frequency changes occurs independently in the process of hearing. But, he does not prompt to either the spectral or time-domain model. On the other hand, JENKINS *et al.* [11] highlight the time-variable attributes, namely the fine changes of temporal structure, which are important in pitch perception.

Considerable effort has been reported on investigations dealing with the perception of the amplitude and frequency modulated sounds. According to the arguments brought out in Part 1 [15]; the signal of beats can also be viewed as a sound featuring these two kinds of modulation. Therefore, references on the loudness and pitch perception can also be helpful in the interpretation of the phenomena observable for the beatings.

Certain difficulty in perceiving the signal frequency changes arises from the fact that the sound pressure level change itself creates an impression of frequency modification. STEVENS has demonstrated [17] that the pitch of low frequency signals decreases with the increase of SPL; within the medium range frequencies, the SPL modifies the pitch although slightly, while the pitch of the high frequency signals rises with the increase of SPL. Pitch changes described above are known as the *Stevens rule*. Similar conclusions have been reached by VERSCHUURE *et al.* [19] who analyzed the published data. These data had evidenced a large differentiation in the pitch assessment made by various listeners and the difficulty in obtaining reproducible results. It was experimentally shown that the *Stevens rule* is not a sufficiently accurate measure of the pitch changes. The absolute value of the pitch variations caused by SPL changes is small within the frequency limits of 1 to 2 kHz. This value increases with increasing as well as with decreasing frequencies. For an individual listener these changes may be positive, negative or do not exhibit monotonicity. However, only the data averaged over a large number of listeners corroborate the *Stevens rule*. Except for very high or very low frequencies, the perceived pitch changes can be insignificant in many cases. Verschuure explains the discrepancies between the results as due to the differences in the experimental procedures used by various authors.

Subsequently, IWAMIYA *et al.* [7, 9] states that the pitch perceived under simultaneous AM&FM modulation, depending on the phase shift between the two modulating functions, is determined by the amplitude weighted frequency variations. He assumes a $\{A_1(t)\}^{\alpha}$ amplitude envelope, where $A_1(t)$ describes relative changes of the amplitude envelope at the amplitude modulation index equal 1 and α is an increment which modifies physical changes of the envelope in order to obtain a loudness modulation function. The frequency shift P (measured from the carrier frequency) to the so called "principal pitch" localized by listeners is given by formula [7]

$$P = m^2 \int_0^T C(t)\{A_1(t)\}^\alpha dt \Big/ \int_0^T \{A_1(t)\}^\alpha dt. \tag{II.1}$$

The increments $\alpha = 0.3$ to 0.7 fit well to the pitch measure proposed by Iwamiya. He explains the differentiation of the values, occurring with individual listeners, by a personal ability of perception of amplitude modulation. What he also has noted is that the regression curves of the perceived pitch values vs the phase shift, mentioned above, exhibit a displacement, the latter being larger at the modulation index $m = 1$ than at $m = 0.71$ (he concludes that this is caused by envelope changes). Following Iwamiya, these displacements may be caused by timbre differences for AM+FM sounds the and sinusoidal tone. In majority, they are negative and are discriminately perceived by listeners (the pitch values are reported to be lower than those predicted by the accepted model). Iwamiya attributed the differentiation of those P shifts to a differentiated effect of the sound pressure level on the perceived pitch.

HARTMANN [6] has studied the influence of the amplitude envelope on the pitch of sinusoidal tones. He discovered that the pitch of a sinusoidal tone of an exponentially decaying amplitude envelope is higher than a tone of the same frequency but of rectangular envelope. He used signals of constant change of the envelopes rate of 1000 dB/s. The experiments have been made at selected frequencies in the frequency range from 412 to 3300 Hz. For three listeners of the group of 15 people, he has noted some differentiation of the asserted changes of the pitch. Averaging the results for these 3 listeners, he obtained the best fit value of the frequency shift equal to 16 Hz for decaying tones; this result does not depend on the tone frequency. It proves that only the envelope changes determine the pitch shift, while the frequency values of the separate tones have no effect on the observed shift. If the results are analyzed in detail for each particular listener in Hartmann's experiment, it seems that this conclusion is not entirely justified. Hartmann is attempting to find a set of physical attributes of a tonal signal of decremented amplitude envelope which properly correlate with the perceived shift of pitch. Discussing the time evolution of the spectrum, he quotes the observations of some listeners whose experience was that the pitch of the exponentially decaying tone appeared to rise in the process. Concluding, Hartmann says that the shift of the pitch does not depend on frequency. However, he underlines the speculative character of this reasoning and emphasizes the need of further investigation.

ROSSING *et al.* [14] run a similar experiment, but they used different initial reference levels, different rates of the envelope change, ranging from 500 to 8000 dB/s, with both incremented and decremented amplitude envelopes. They established unambiguously that both the envelopes, the decaying and the rising ones, produce a comparable pitch increase. This increase grows up with still higher frequencies. Such result eliminates the perception model based on the "periodicity in neural firings" predicting opposite pitch changes discriminating between up- or down-trends of the envelope amplitude. The authors come to the conclusion [14] that the experiments implicate that the envelope-dependent change of the pitch would not directly be caused by the envelope, but, preferably, may be due to a process the quantity of which, in turn, depends on the envelope; for example the averaged sound pressure level.

SCHOUTEN *et al.* [16] observed that the majority of natural sounds, like speech and sounds of musical instruments, feature a harmonic spectral structure; according to them, the joint perception of a number of spectral components, the so called *residuum*, determines the sounds pitch. There is however, as they remark, certain ambiguity in assessment of the pitch which may be due to the influence of the fine structure of the sound amplitude envelope. According to them, the pitch assessment is performed within the hearing system in the time-domain mode (*via* technique of delay lines), whilst the spectral model is less plausible.

RITSMA *et al.* [12] studied the pitch under conditions of *quasi-frequency modulation*, discovering its dependence on the modulation depth (*index of modulation*, crucial in determining the depth of the changes of the amplitude envelope, too). The measured values of the pitch were not correlated with the period of the amplitude envelope of the signals, but calculated assuming that the separation measured between two adjacent peaks of positive polarity of *temporal fine structure*, in the vicinity of the neighbouring maxima of the amplitude envelope, can be a measure of pitch.

JENKINS [11], summing up the discussion on the perception of the sound pitch, sound timbre and loudness, favours the model performing a broad-band spectral analysis inside the auditory system including a detection of envelope and detection of periodicity (based on a delay line). Such a mechanism of perception is featuring a short time response. Accepting that the spectral model is preferred in the perception processes, Jenkins argues that, on physical grounds, it is not possible to accomplish an exact frequency determination with the Fourier transform (pitch evaluation) while maintaining a fast time response.

ZWICKER [20] comments on the perception thresholds of frequency differences of tones that are essential to our analysis. Zwicker has found that the perception of frequency differences are much better for tones of constant amplitude than for a very narrow band noise, centred at the same frequency, which sounds like a tone of randomly variable amplitude. For example, the thresholds of the frequency difference perception for a narrow - band noise of carrier frequency of 1500 Hz and 10 Hz bandwidth are 6 times greater than for a tone of the same frequency. In particular, according to Zwicker, the diversions in the sensitivity of the hearing systems to perceive frequency differences are controlled by the rate and non-periodicity of the amplitude changes.

ROEDER [13] has found that the response of the hearing system to the signal of beatings follows the resultant waveform of the beatings. The ear "is not aware" of the perceived sound resulting from addition of the components. Simultaneous stimulation of the closely situated regions on the basiliar membrane is responsible for the resultant impression of a sound pitch. Such stimulation of the basiliar membrane involves overlapping of adjacent regions that effects indirectly the intermediate pitch impression. The SPL changes during beatings are perceived as a modulation of loudness.

JEFFRES [10] isolated fragments of beating sinusoids within the maximum and minimum of the amplitude envelope and could distinguish the pitch differences. For the pair of signals, in which the lower frequency tone is of smaller SPL value, Jeffres found an increase in frequency, when passing from the minimum to the maximum on the envelope, and the contrary for a pair in which the higher frequency tone is the stronger one. In the case of two signals of equal sound pressure levels, Jeffres pointed out the instantaneous phase reversal at each beats minimum. The phase reversal is of a jump wise character end involves an infinite frequency shift, but of zero duration.

The results of the referenced perception studies as well as the arguments regarding the mechanism of pitch perception only prove the complexity of the phenomenon of beats leaving many questions unresolved. Therefore, it is necessary to search for other properties of the beats signal that could contribute to a better understanding of the observed relations. The perception analysis, which will be pursuit in this part of the study, is confined to a situation in that the frequency separation of the components is so small that we can not resolve two tones of unique frequencies, but we may perceive a resultant (average) sensation of pitch or its change (beat components fit within the critical band). Attention will be focused on the responses of listeners who may testify either symmetry or asymmetry of pitch perception for the complementary pairs of beating signals. Two experiments and the results reported previously [2], provide a basis for the analysis of the problem. The interpretation of the results will also be made with reference to Part I in the scope relevant to the predefined physical parameters of beats.

2. Theorems related to the experimental procedure

DAI [2] reported the results of the perception of pitch of beatings by 3 listeners. His examination of the pitch was made in two combinations denoted SL and SH (see Fig. 1) for the following set of "fixed signals":

• two-tone complex with frequencies f_L = 480 Hz and f_H = 520 Hz (average frequency = 500 Hz),

• two-tone complex with frequencies f_L = 960 Hz and f_H = 1040 Hz (average frequency = 1000 Hz).

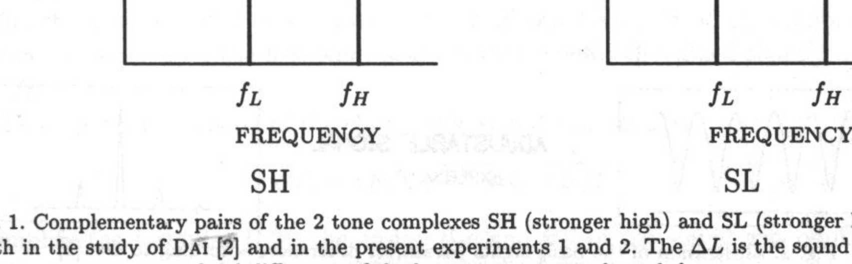


Fig. 1. Complementary pairs of the 2 tone complexes SH (stronger high) and SL (stronger low) used both in the study of DAI [2] and in the present experiments 1 and 2. The ΔL is the sound pressure level difference of the beats components f_L and f_H.

For each combination, the author checked three relative difference levels ΔL, i.e. 2.4, 3.4 and 4.4 dB.

The signal of variable parameters, the so called "adjustable signal", was a two-tone complex consisting of tones of equal sound pressure levels and a frequency difference equal to the separation of the beating frequencies. The level of the acoustic pressure was 65 dB SPL for all the signals applied in Dai's experiment, i.e. for the fixed signal and the adjustable one. The duration of the signals was 250 ms which corresponds to 10 beats at 500 Hz and 20 beats cycles at a frequency of 1000 Hz. Dai has performed his experiment by a pitch matching procedure. The listeners changed the adjustable signal frequency till perceiving the adjustable and the fixed signals as equaled in their pitches. Dai detected the occurrence of some asymmetry in the listeners valuation of the pitch for the complementary pairs of beating signals. With the aim to understand the possible causes of this asymmetry, in the present study two separate experiments were made in which the signal parameters assumed similar, if not identical, values to those reported by Dai. Experiment 1 (Sec. 3) was in principle a repetition of the experiment performed by Dai, although somewhat new methods were introduced. Experiment 2 (Sec. 4) consisted in applying as the signal of variable parameters, an amplitude modulated signal whose envelope was calculated as for the beats (the fixed signal). Due to this, the envelope changes of the two signals were identical in the course of the experiment. Basic differences regarding the signal parameters of the experiments reported in this study are shown in Fig. 2.

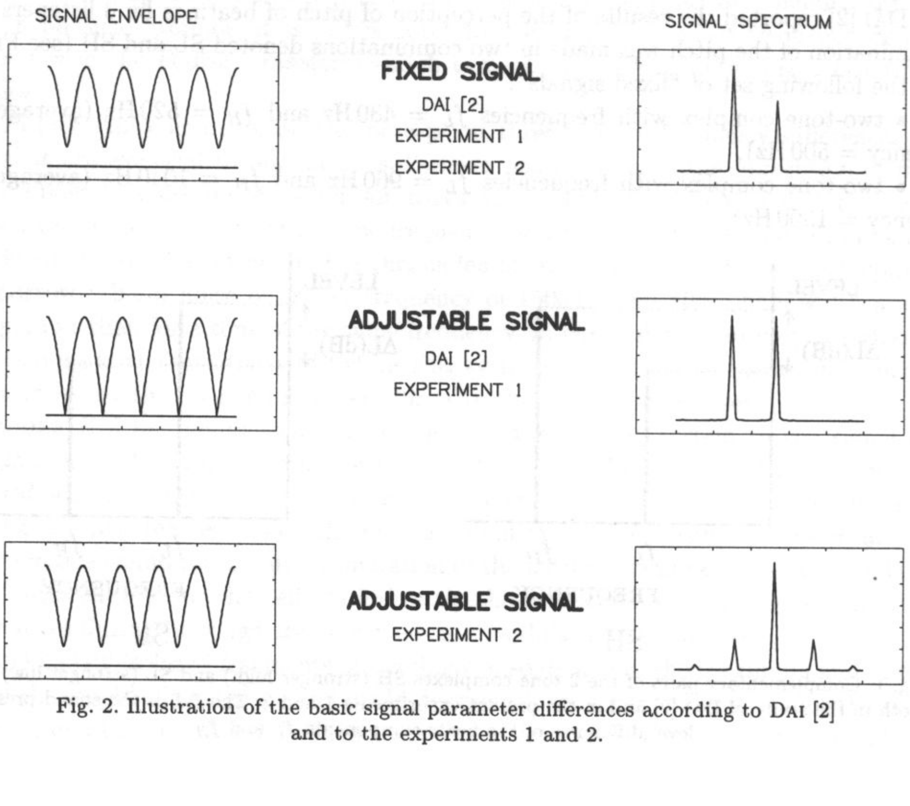


Fig. 2. Illustration of the basic signal parameter differences according to DAI [2] and to the experiments 1 and 2.

3. Experiment 1: pitch matching with equal SPL beating sinusoids

A. Procedure

The method consisted in transmitting pairs of signals as a random sequence to the listener. The parameters of one of the signals, the "fixed signal", were kept unchanged during the whole experiment. The parameters of the other one, the "changeable signal", were regulated during the experiment following the two-alternatives forced choice (2AFC) paradigm with an adaptative procedure one – down, one – up. The listener was required to decide, for each pair of the randomly launched sequence, the signal (either the first or the second one) that could be heard at a higher pitch. The message whether the listener's answer was correct or wrong was displayed each time on the monitor screen. The test series was terminated after a sequence of 40 pairs of signals. For each test at least 8 turnpoints were obtained. An intermediate value of matched frequency was calculated as an arithmetic mean of the tone's center frequency of the last 8 turnpoints. Each listener took part in 5 such sessions. The final value of frequency – f_{adj} was an average taken from those five sessions. The measurements were performed in special sound-proof rooms properly insulated from the outside noise.

B. Stimuli

The test signals were digitally generated with an instrumental setup consisting of an Array Processor DSP32C connected to a 16-bit digital-analogue converter with optical fibres (lower noise level). Next, the signals were filtered with a low-pass filter of cut-off frequency 8 kHz and a slope of trailing edge of the filter characteristics of 90 dB per octave. The described setup for the generation of the signals (Tucker Davis Technologies – USA) and the experiment sequencing was controlled with a PC computer. Each signal lasted 1000 ms including the cosine rise and decay time, both 50 ms. The time period between a successive pair of signals depended on the listeners choice.

The signals were monaurally submitted to the listeners by DT48 headphones of frequency characteristics equalised in the range of frequencies relevant to the experiments. The values of the resultant acoustic pressure of all the signals were the same and equal 65 dB SPL. Also, the remaining parameters of the signals were the same as those reported by DAI [2].

The changeable signal was a two-tone one of equal SPL ($\delta = 1$) components and the same separation on the frequency scale as that one for the "fixed signal", i.e. $\Delta f = f_H - f_L$.

The amplitude envelope of the changeable signal was equal to

$$e(t) = x_L\sqrt{2}\sqrt{1 + \cos 2\pi\Delta f t}\,. \tag{II.2}$$

To simplify the formula (II.2), the cosine rise and fall of the signal was omitted. The complex instantaneous frequency (CIF) envelope of this signal varied according to the following dependence:

$$|\mathrm{CIF}(t)| = \sqrt{\left[\frac{\Delta f \sin(2\pi\Delta f t)}{2(1+\cos(2\pi\Delta f t))}\right]^2 + \left[\frac{\Delta f}{2} + f_{L\,\mathrm{adj}}\right]^2}, \tag{II.3}$$

where $f_{L\mathrm{adj}}$ is the lower frequency value of the two-tone of equal amplitudes.

C. Subjects

Three listeners participated in the measurements, all with audiometrically normal hearing. Their age was between 21 – 24. One of the listeners was the author of this study (AW). The other one, a female (MK), had some musical education, while the third one, (PR), was a musician. Each of the listeners was instructed and practiced before the experiments; the aim was to attest their understanding of the objectives. The listeners had good practice in this kind of experiments.

D. Results

Tables 1 and 2 contain the mean values (averaged over 5 series of measurements) of the center frequency of the changeable signal of pitches matched by the listeners to the pitch of the SL and SH pairs of average frequencies of 500 Hz (Table 1) and 1000 Hz (Table 2), respectively. Standard deviations of the matching frequencies are given in brackets. δ denotes the ratio of amplitudes of the SL and SH signals for the given ΔL (level difference) values determined in decibels. Frequency values listed in tables 1 and 2 equal the arithmetic average of the frequencies of the beating tones matched with respect to the signal pitch, i.e. $(f_L + \Delta f/2)_{\mathrm{adj}}$.

Table 1. Experiment 1. The center frequencies of the equal intensity adjusted signal, (the pitch was matched to the pitch of either the SL or SH signal) and the corresponding standard error. Data for $f_{\mathrm{av}} = 500$ Hz.

	500 Hz					
ΔL	2.4 dB		3.4 dB		4.4 dB	
	SL $\delta = 0.759$	SH $\delta = 1.318$	SL $\delta = 0.676$	SH $\delta = 1.479$	SL $\delta = 0.603$	SH $\delta = 1.66$
AW	494.9 (0.3)	506.3 (0.3)	492.6 (0.4)	507.4 (0.2)	490.1 (0.2)	509.6 (0.2)
MK	492.3 (0.5)	507.7 (0.1)	490.5 (0.2)	510.3 (0.5)	488.7 (0.3)	510.2 (0.2)
PR	493.4 (0.2)	504.8 (0.2)	493.8 (0.4)	507.5 (0.3)	492.9 (0.2)	508.9 (0.3)

Table 2. Experiment 1. The center frequencies of the equal intensity adjusted signal, (the pitch was matched to the pitch of either the SL or SH signal) and the corresponding standard error. Data for $f_{av} = 1000$ Hz.

	1000 Hz					
ΔL	2.4 dB		3.4 dB		4.4 dB	
	SL $\delta = 0.759$	SH $\delta = 1.318$	SL $\delta = 0.676$	SH $\delta = 1.479$	SL $\delta = 0.603$	SH $\delta = 1.66$
AW	993.8 (0.6)	1007.9 (0.5)	990.4 (0.5)	1014.6 (0.6)	987.8 (0.4)	1019.5 (0.7)
MK	988.3 (0.6)	1010.4 (0.7)	984.5 (0.5)	1016.6 (0.6)	980.3 (0.5)	1021.6 (0.6)
PR	993.9 (0.5)	1006.8 (1.4)	992.9 (1.06)	1007.6 (0.3)	987.2 (0.8)	1010.3 (0.8)

4. Experiment 2: pitch matching with amplitude modulated adjustable signal

A. Objective of the experiment

DAI [2] and FETH *et al.* [3] have found a characteristic asymmetry in pitch matching experiments when the pitch of the complex signal having two components of unequal levels, is compared to the arithmetic average of the frequencies of the two-tone SL and SH pairs. The amplitude envelopes of the SL and SH signals are identical and, in principle, the only difference between the signals is the pattern of their instantaneous frequency variations; however the symmetry of the Fourier spectra around an average frequency is not corroborated by the symmetry of the perceived pitch values. It should be noted that in the experiments of Dai and Feth the adjustable signal for which $\Delta L = 0$ ($\delta = 1$) had an amplitude envelope different from those of the tested signals (fixed signals of the SL and SH pairs, compare Fig. 2). Therefore, in spite of the different mean frequency values of the fixed and adjustable signals (which is obvious, especially in the initial phase of the experiment), the signals were also different envelope amplitude (in each phase of the experiment). This subsidiary discrimination between the fixed and matched signal was not, most certainly, a circumstance that made the listeners task of matching the pitch of pairs of signals easier. For this reason, new modified experiments have been suggested, in that the changeable signal had exactly the same amplitude envelope as the fixed one. This idea was implemented by the technique of the amplitude modulated (AM) signal in that the envelope was calculated according to equation (I.5), hence in the same way as the envelope of the fixed signal of beatings. The carrier frequency of the AM signal was controlled by the listener during the experiment. Consequently, the signal to be matched (changeable signal) was of the form

$$x_{\mathrm{adj}}(t) = x_L\sqrt{1+\delta^2+2\delta\cos 2\pi\Delta f t}\,\cos(2\pi f_{\mathrm{adj}}t), \tag{II.4}$$

(the cosine rising and decay are omitted for simplicity) where $\Delta f = f_H - f_L$ – amplitude modulation frequency, f_{adj} – carrier frequency of the AM signal whose pitch was matched to the fixed signal. According to Eq. (I.14), the frequency envelope of the signal (II.4) was

$$|\mathrm{CIF}(t)| = \sqrt{\left[\frac{\delta\Delta f \sin(2\pi\Delta f t)}{1+\delta^2+2\delta\cos(2\pi\Delta f t)}\right]^2 + [f_{\mathrm{adj}}]^2}\,. \tag{II.5}$$

An example of the matching Fourier spectrum (II.4) is shown in Fig. 2.

B. Results

In Tables 3 and 4 average values of the carrier frequency of the AM signal (5 measurement series) are presented; the pitch was matched by the listeners (the same group which took part in the experiment 1) to the pitch of the SL and SH signals of mean frequencies 500 Hz and 1000 Hz, respectively, for Table 3 and 4.

Table 3. Experiment 2. The carrier frequencies of the amplitude modulated adjusted signal, (the pitch was matched to the pitch of either the SL or SH signal) and the corresponding standard error. Data for $f_{\mathrm{ad}} = 500$ Hz.

	500 Hz					
ΔL	2.4 dB		3.4 dB		4.4 dB	
	SL $\delta = 0.759$	SH $\delta = 1.318$	SL $\delta = 0.676$	SH $\delta = 1.479$	SL $\delta = 0.603$	SH $\delta = 1.66$
AW	492.6 (0.4)	508.1 (0.4)	491.9 (0.6)	509.8 (0.7)	490.3 (0.4)	513.4 (0.8)
MK	491.1 (0.3)	506.8 (0.2)	488.3 (0.5)	509.4 (0.6)	486.5 (0.6)	509.4 (0.3)
PR	489.9 (0.4)	501.2 (0.3)	488.7 (0.5)	504.6 (0.5)	487.3 (0.5)	507.7 (0.3)

Table 4. Experiment 2. The carrier frequencies of the amplitude modulated adjusted signal, (the pitch was matched to the pitch of either the SL or SH signal) and the corresponding standard error. Data for $f_{\mathrm{av}} = 1000$ Hz.

	1000 Hz					
ΔL	2.4 dB		3.4 dB		4.4 dB	
	SL $\delta = 0.759$	SH $\delta = 1.318$	SL $\delta = 0.676$	SH $\delta = 1.479$	SL $\delta = 0.603$	SH $\delta = 1.66$
AW	988.7 (0.9)	1019.1 (1.2)	984.5 (1.4)	1017.7 (1.1)	977.1 (0.6)	1023.9 (0.5)
MK	982.5 (0.5)	1012.7 (0.3)	981.3 (0.2)	1018.0 (0.2)	981.9 (0.3)	1022.2 (0.5)
PR	970.4 (1.1)	1008.7 (2.1)	965.9 (0.7)	1014.4 (1.05)	967.1 (0.5)	1014.5 (0.6)

5. Discussion

In Figs. 3 and 4 the results of the two experiments are displayed together with the earlier results obtained by DAI [2]. The data in Fig. 3 a–c correspond to the beats average frequency of $f_{av} = 500$ Hz and the frequency separation $\Delta f = 40$ Hz with SPL differences of the beating tones equal to 2.4 dB (Fig. 3a), 3.4 dB (Fig. 3b) and 4.4 dB (Fig. 3c). The data in Fig. 4 a–c correspond to the beats average frequency $f_{av} = 1000$ Hz and the frequency difference $\Delta f = 80$ Hz, with the SPL differences 2.4 dB (Fig. 4a), 3.4 dB (Fig. 4b) and 4.4 dB (Fig. 4c). In both figures (Figs. 3 and 4), the frequencies f_{adj} of the changeable signal matching the pitch of fixed ones are given. The values are referred to the arithmetic average of the fixed signals, so the vertical axes are in $f_{adj} - f_{av}$ [Hz] units. The positive values of such a calculated frequency correspond to the SL pairs, while the negative ones represent the SH signal pairs. Filled circles are the data points obtained by the listeners for the SL and SH pairs. The empty ones correspond to the calculated average values of the matching frequencies for the SH and SL pairs.

The displacement of these points from the dotted straight line ($f_{adj} - f_{av} = 0$) means that the responses of the listeners featured an asymmetry.

Additionally, the values of the envelope weighted average of instantaneous frequency EWAIF – dashed line, and intensity (square envelope) weighted average of instantaneous frequency IWAIF – solid line, diminished by f_{av}, are given. The two values are calculated for the frequency envelope given by Eq. (I.14) and the amplitude envelope described by Eq. (I.5) of the beatings frequency variations. These two quantities are compatible with the loss of symmetry for complementary pairs of the SL and SH signals (see Fig. 4 in Part I and Fig. 5 in Part II).

The results obtained for the beatings sinusoids of mean frequency 500 Hz in the experiment 1 point to the existence of an almost ideal symmetry of the two complementary pairs of the signals SL and SH, as reported by the listeners. The largest departures from symmetry amount to ± 0.9 Hz with an average error of ± 0.23 Hz. The mean deviation $\langle \Delta f \rangle_{\mathrm{Sub},\Delta L}$ for all the listeners (Sub) and all the values of level differences ΔL were equal to $+0.1$ Hz. DAI [2] reported a much larger deviation from symmetry for the same value of the average frequency of beatings. Namely, each of his listeners claimed frequency variations of positive sign from the symmetry related pattern. The largest reported deviation was $+4.8$ Hz at ± 0.55 Hz accuracy. The mean deviation $\langle \Delta f \rangle_{\mathrm{Sub},\Delta L}$ amounted to $+2.9$ Hz, for the Dai's listeners. The major part of the data from experiment, corresponding to 500 Hz average frequency, justifies the argument that the pitch is well correlated to the intensity weighted average of the instantaneous frequency value [1, 2], although the listener's MK results are closer to the value of the envelope weighted average of the instantaneous frequency.

The application of AM as the changeable signal (experiment 2) with the same envelope as that of the beatings just investigated resulted in a wider differentiation of the listeners responses gained at 500 Hz average beatings frequency. The listener AW localizes the pitch of the SH signal closer to envelope weighted frequency value, while for the SL pair of signals at level differences of 3.4 and 4.4 dB, his results approach the intensity weighted frequency value. Most of the data on the SL signal obtained from

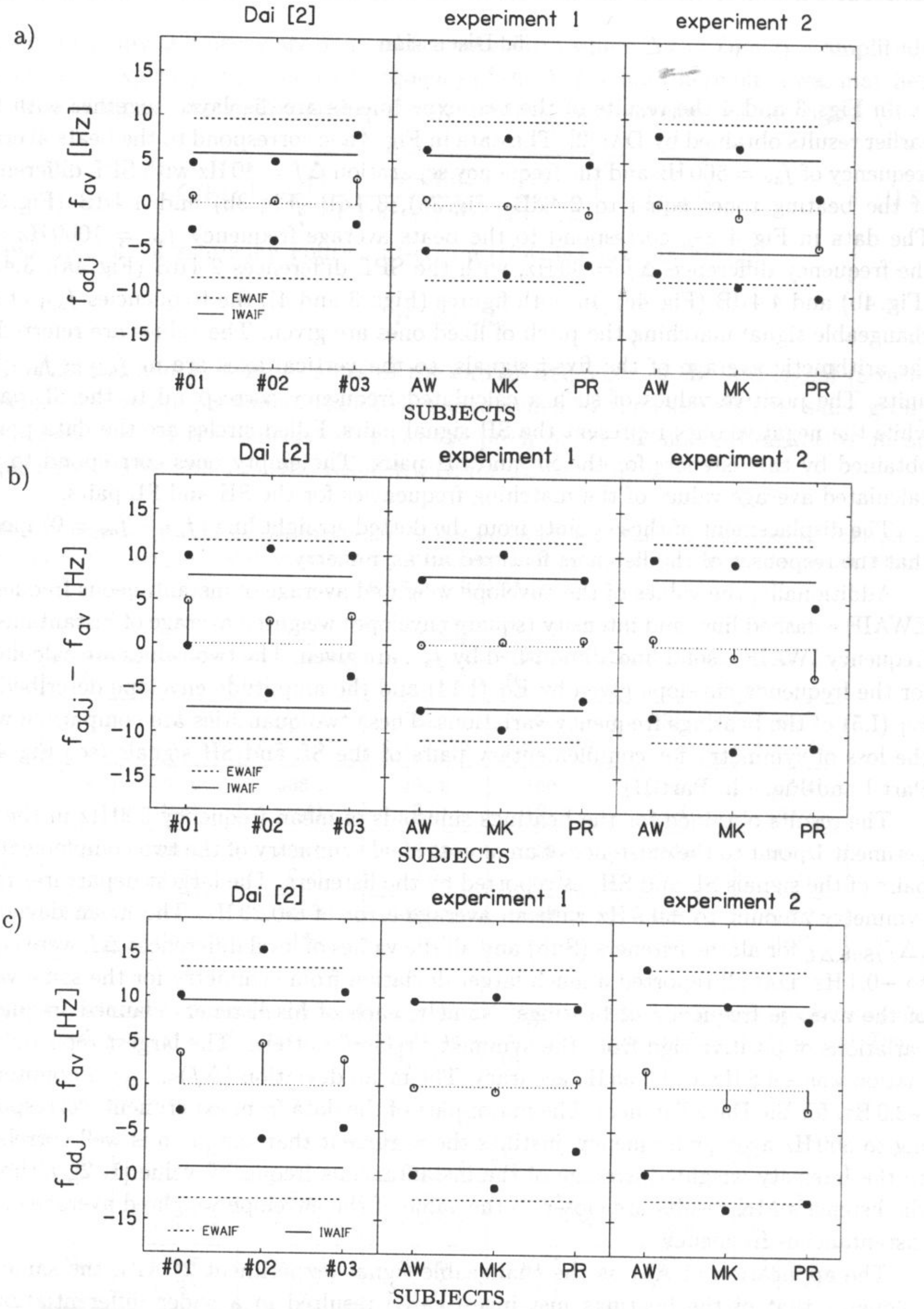


Fig. 3. Comparison of the results obtained by DAI [2] and in the experiments 1 and 2 for the average frequency of 500 Hz (closed circles) for the following SPL differences ΔL: a) 2.4 dB, b) 3.4 dB, c) 4.4 dB. Open circles – calculated average values for the complementary SL-SH pairs. EWAIF – calculated envelope weighted average of the instantaneous frequency. IWAIF – calculated intensity (squared envelope) weighted average of the instantaneous frequency.

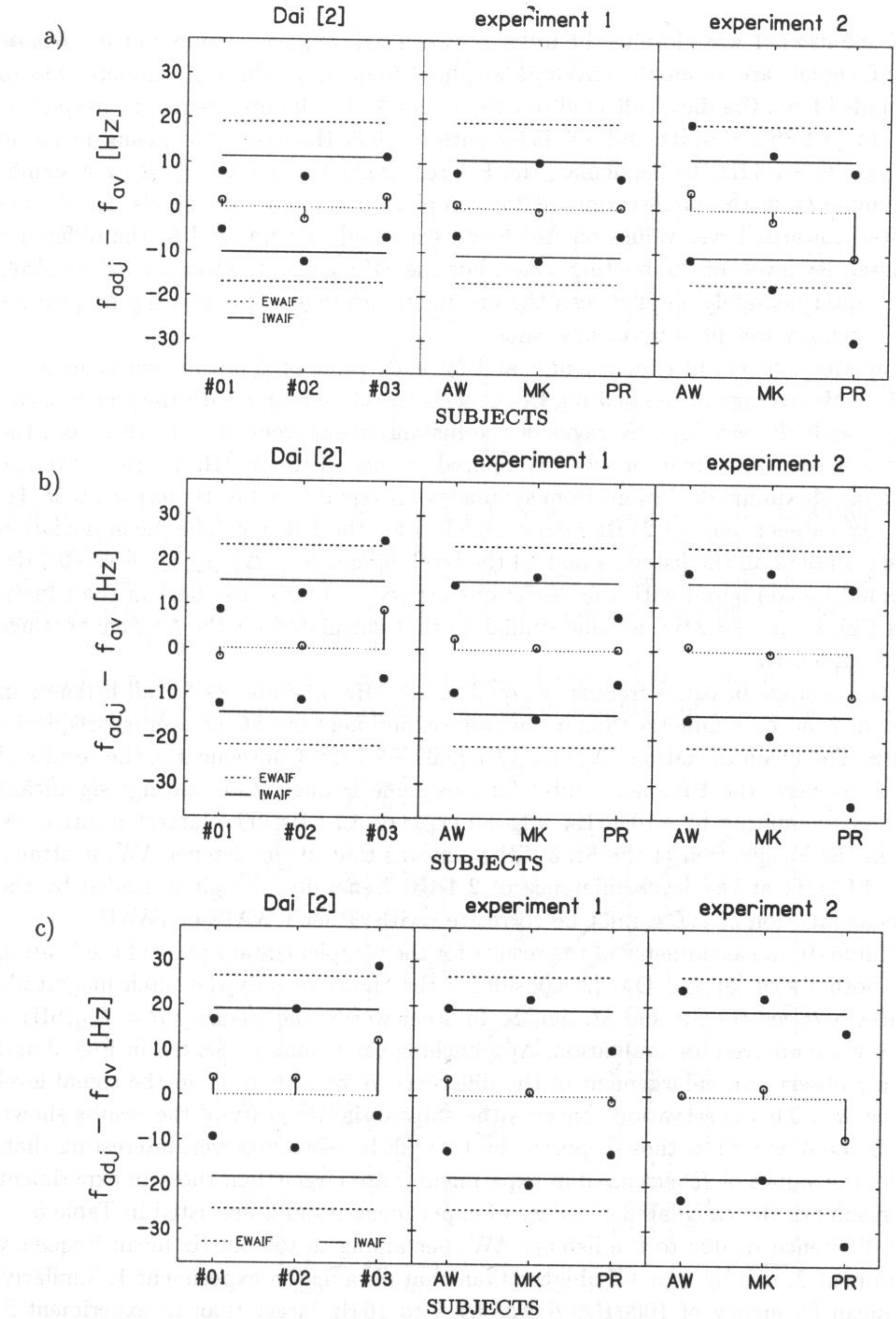


Fig. 4. Comparison of the results obtained by DAI [2] and in the experiments 1 and 2 for the average frequency of 1000 Hz (closed circles) for the following differences ΔL: a) 2.4 dB, b) 3.4 dB, c) 4.4 dB. Open circles – calculated average values for the complementary SL-SH pairs. EWAIF – calculated envelope weighted average of the instantaneous frequency. IWAIF – calculated intensity (squared envelope) weighted average of the instantaneous frequency.

the MK she-listener was close to the intensity weighted frequency value, but her results on the SL signals are about the envelope weighted frequency values. Asymmetry has to be concluded from the data collected in experiment 2. The largest frequency deviations amount to $(+1.85 \pm 0.6)$ Hz and -4.45 Hz with a ± 0.35 Hz error. The mean deviation $\langle \Delta f \rangle_{\mathrm{Sub},\Delta L}$ is -1.3 Hz. In particular, the PR responses are characterized by a significant asymmetry of the pitch estimates for complementary pairs of signals. His results are shifted towards lower values on frequency scale and unconnected to the difference in the intensity levels of the beating tones. For the SH signals, the listener PR localizes the pitch intermediately, i.e. between the arithmetic average of the beating frequencies and the intensity weighted frequency value.

For all the listeners in experiment 1 and for each value of ΔL, the results acquired at the 1000 Hz average of the beating frequencies correlate better with the square envelope than with the envelope averages of the instantaneous frequency. In this case, the symmetry is somewhat inferior when compared to that of the 500 Hz beatings average experiment. Maximum deviations from symmetry are equal to $+3.65$ Hz (error ± 0.55 Hz) for the AW listener and -1.25 Hz (error ± 0.8 Hz) for the PR one. The mean deviation over the results of all the listeners and all the level differences, $\langle \Delta f \rangle_{\mathrm{Sub},\Delta L}$ was $+0.8$ Hz. This should be compared with the deviations arrived at by the listeners in the experiment of Dai, i.e. to $+3.2$ Hz, a value similar to that calculated for the average beatings frequency of 500 Hz.

For the average beatings frequency equal to 1000 Hz, one notices for all listeners in experiment 2 more asymmetry than in the same experiment but at the average frequency of 500 Hz. The mean deviation $\langle \Delta f \rangle_{\mathrm{Sub},\Delta L}$ equals -2.7 Hz. Consequently, the results of only one listeners, the PR one, exhibit for this same frequency of beatings significant asymmetry amounting to -10.45 Hz with an error of ± 1.6 Hz. The largest positive deviation in the localization of the SL & SH pairs was that of the listener AW; it attains $(+3.9 \pm 1.05)$ Hz at the level difference of 2.4 dB. Generally, the pitch graded by the listeners as function of ΔL cannot be correlated with either EWAIF or IWAIF.

To eliminate the asymmetry of the results for the complementary pairs of the beating signals, both FETH [5] and DAI [2] considered the choice of only the pitch magnitude differences between the SH and SL signals. In other words, the quantity $R = f_{\mathrm{adj}}(\mathrm{SH}) - f_{\mathrm{adj}}(\mathrm{SL})$ was employed for evaluation. Approaching the results presented in Figs. 3 and 4, we may observe an enlargement of the difference R with increase in the signal level difference ΔL. This observation concerns the substantial majority of the results shown in Figs. 3 and 4 as well as those reported by DAI [2]. It seems however, interesting that, as a rule, the values of R measured in experiment 2 are larger than those in experiment 1. The results of the calculated R values of experiment 1 and 2 are listed in Table 5.

The difference R, due to the listener AW, pertaining to the 500 Hz mean frequency in experiment 2, was by 3 to 4 Hz higher than that obtained in experiment 1. Similarly, at the mean frequency of 1000 Hz, R was by 9 to 16 Hz larger than in experiment 2. The MK listeners results, pertaining to the mean frequency of 500 Hz, led to a difference $\Delta R = R(\mathrm{Expmt.}\,2) - R(\mathrm{Expmt.}\,1)$ equal to $0.3 - 1.4$ Hz. For the mean frequency of 1000 Hz in experiment 2, the magnitude of R was larger by 4.6 to 8 Hz, whereas, only at $\Delta L = 4.4$ dB, it was less by 1 Hz. For the average frequency of 500 Hz, the listener

Table 5. The differences between adjusted frequencies (in [Hz]) for the SH and SL complementary pairs for the experiments 1 and 2.

f_{av} [Hz]		500			1000		
ΔL [dB]		±2.4	±3.4	±4.4	±2.4	±3.4	±4.4
DAI [2]	#1	+0.8	+4.8	+3.6	+1.45	−1.8	+3.7
	#2	+0.3	+2.55	+4.65	−2.65	+0.55	+3.7
	#3	+2.85	+3.20	+2.95	+2.45	+8.95	+12.25
Exp. 1	AW	+0.6	+0.0	−0.15	+0.85	+2.5	+3.65
	MK	+0.0	+0.4	−0.55	+0.65	+0.55	+0.95
	PR	−0.9	+0.65	+0.9	+0.35	+0.25	−1.25
Exp. 2	AW	+0.35	+0.85	+1.85	+3.9	+1.1	+0.5
	MK	−1.05	−1.15	−2.05	−2.4	−0.35	+2.05
	PR	−4.45	−3.35	−2.5	−10.45	−9.85	−9.2

PR reported a difference $\Delta R = 2.2 - 4.4$ Hz; only at $\Delta L = 2.4$ dB, this difference was practically 0. Much higher differences: $\Delta R = 24 - 34$ Hz were obtained for this listener at a frequency of 1000 Hz. Relatively small values of the R – difference heard by the listener MK may result from her R(Expmt. 1) – valuations being systematically larger than those of the remaining listeners.

In Fig. 5 arranged together with the data of Fig. 4, Part I, the normalized (divided by beating tones frequency difference) envelope weighted of instantaneous frequency

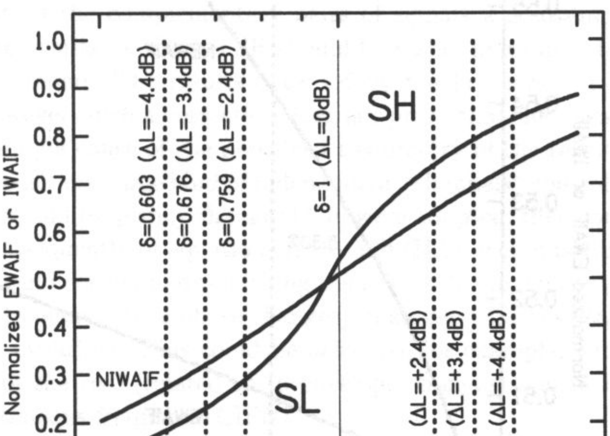


Fig. 5. Illustration of the asymmetry effect for the normalized (by dividing by beating tones frequency difference) *envelope weighted average of instantaneous frequency* (NEWAIF) and the *intensity (squared envelope) weighted average of instantaneous frequency* (NIWAIF). The signal amplitude ratios used in the experiments are denoted by vertical dotted lines. The lack of symmetry is observed for complementary pairs of the beating tones SL and SH.

NEWAIF, and squared envelope (intensity) weighted average of instantaneous frequency NIWAIF are plotted against the amplitude ratio δ; the δ – values used previously in the experiments for SL and SH pairs have been marked (dashed vertical lines).

In view of the asymmetry of the graphs shown in Fig. 5 an asymmetry for the SL/SH pairs may be anticipated. However, this asymmetry has no effect on the value of the R parameter discussed above. But we have to recall that, at $\delta = 1$ ($\Delta L = 0\,\mathrm{dB}$), i.e. for the matched signals changeable in experiment 1 and adjustable in Dai's investigations, there was some frequency shift due to the asymmetry of those graphs. Moreover, the results were referred to the arithmetic average of the two tones of the pitch matching signal. In order to estimate the frequency displacement (shift), at the two equal levels, i.e. $\Delta L = 0\,\mathrm{dB}$, the ratio $(\Delta f/f_L)\cdot 100$ [%] was introduced as an approximate measure of the bandwidth narrowness of the complex signals. For the two mean frequencies of the beats, i.e. 500 and 1000 Hz, this value was 8.33%. In Fig. 6, which is some modification of Fig. 5 (Part I), this value is marked as the vertical dashed line. The crossing points of this line with the EWAIF & IWAIF graphs provide a numerical reading of this asymmetry. The frequency shift values at $\Delta L = 0\,\mathrm{dB}$ and the mean frequency of the beats equal to 500 Hz amount to +1.65 Hz (EWAIF) and +0.38 Hz (IWAIF). At the 1000 Hz mean frequency, they are +3.27 Hz and +0.78 Hz, respectively. The correction of the pitch matching asymmetry may be accomplished by subtracting the above numbers from the listeners matched frequencies for the SL/SH pairs of signals.

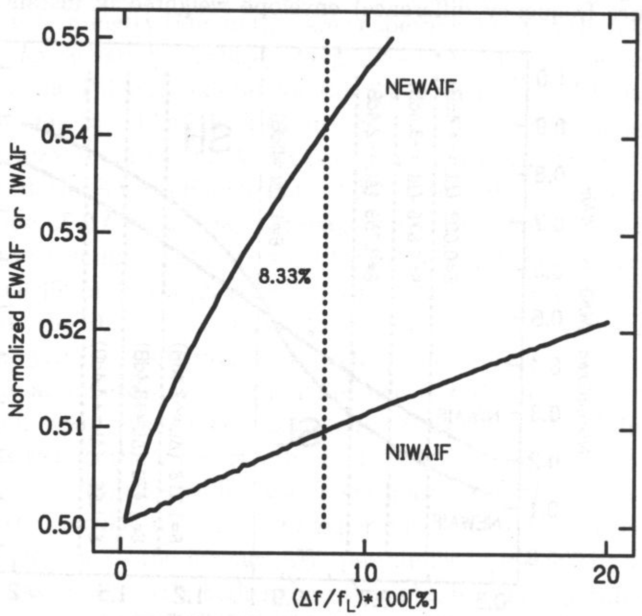


Fig. 6. Normalized (by dividing by beating tones frequency difference) *envelope weighted average of instantaneous frequency* (NEWAIF) and *intensity (squared envelope) weighted average of instantaneous frequency* (NIWAIF) shifts due to the not completely fulfiled narrow-band condition. The lower tone frequency $f_L = 480\,\mathrm{Hz}$, and SPL difference $\Delta L = 0\,\mathrm{dB}$. Δf is the frequency difference between the beating tone components. The value 8.33% (vertical dotted line) corresponds to signal parameters used in the experiments.

Using the arguments from Part I of this study, the asymmetry emerges as a consequence of the influence of the relative changes of the amplitude envelope on the frequency envelope. The greatest effect of these changes occurs at the highest rates of the amplitude envelope, i.e. for $\Delta L = 0\,\text{dB}$, which corresponds to the signals used in the experiment 1 and those of DAI [2]. In experiment 2, the envelope of the changeable signal was the same as that of the fixed one. This means that the resultant frequency envelope depended on the relative changes of the amplitude envelope, but to a lesser degree. Therefore, the asymmetry related corrections, which have to be subtracted from the matched carrier frequencies f_c of the AM signal, will be smaller.

6. Conclusions

The discussed results of the investigations carried out in experiment 1 and experiment 2 together with those of DAI [2] yield evidence of a substantial pitch differentiation by individual listeners. Contrary to the listeners, whose results have been reported by DAI [2], the listeners involved in this study have performed a pitch evaluation following consequently an individually adjusted scheme. Owing to the aforementioned differentiation, a generalization of the conclusions of these investigations is not possible. Most likely, VERSCHUURE *et al.* [19] were right when arguing that a general evaluation of the pitch should be accomplished applying a larger number of listeners.

The asymmetry reported earlier by DAI [2] was not confirmed in experiment 1, where the pitch evaluation of the complementary pairs of signals was attempted. The small differentiation of the results for SL and SH should be rather attributed to the individual preferences of the listeners. The results obtained by the listeners can be corrected as described above, however, such a correction does not prove that the physically occurring asymmetry is the unique cause of the perceived asymmetry of the results.

If, as a valid proposition, same kind of independent "channels" controlling the perception of the amplitude and frequency changes [7, 8] were adopted, then some modification of Eq. (I.14), and consequently of equations (II.3) and (II.5), would perhaps be justified. This modification would consist in diminishing the role of the imaginary part of the complex instantaneous frequency (the phase changes) in shaping the resultant values of the frequency envelope variations. This could be achieved, for example, with a summation of the magnitudes of the real (related to to envelope changes) and imaginary parts of the complex instantaneous frequency $\mathrm{CIF}(t)$.

A set of interesting data was obtained in the experiment 2. Not only an asymmetry of the SL and SH signals was discovered, but quite often, values of R (the difference of the matched frequencies) much larger than those in experiment 1 were obtained. It can be presumed that, to some extent, due to the events of both the fixed and changeable signals, the listeners could concentrate better on their experimental task in experiment 2 than in experiment 1. At the present stage, however, it is difficult to give an explicit rationale for the cause of the pitch differences perceived in these two experiments.

There is no doubt that the idea of the amplitude weighted frequency variations as a means for the evaluation of the sound pitch has been already adequately established.

The problem which remains to be solved is the choice of appropriate weighting functions, for instance those proposed by IWAMIYA [7, 9] and, may be better, correlated with the loudness. Perhaps, the suggestions by ZWICKER [20], that changes of the amplitude are responsible for the perception of the frequency separations, including their dynamical variations, ought to be reconsidered. These amplitude changes are believed [20] to obscure the perception of frequency fluctuations (smoothing effect). The frequency variations to be accounted for in "calculating" the pitch, according to the authors of the present article, should be expressed by equation (I.14) determining the frequency envelope, and modified as explained above.

Acknowledgments

The authors wish to thank Prof. E. OZIMEK for the fruitful discussions and his numerous suggestions during the experiments.

References

[1] J.N. ANANTHARAMAN, A.K. KRISHNAMURTHY, L.L. FETH, *Intensity-weighted average of instantaneous frequency as a model for frequency discrimination*, J. Acoust. Soc. Am., **94**, 2, 723–729 (1993).

[2] H. DAI, *On the pitch of two-tone complexes*, J. Acoust. Soc. Am., **94**, 2, 730–734 (1993).

[3] L.L. FETH, H. O'MALLEY, *Two-tone auditory spectral resolution*, J. Acoust. Soc. Am., **62**, 4, 940–947 (1977).

[4] L.L. FETH, *Frequency discrimination of complex periodic tones*, Perception & Psychophysics, **15**, 2, 375–378 (1974).

[5] L.L. FETH, H. O'MALLEY, J. RAMSEY, *Pitch of unresolved, two-component complex tones*, J. Acoust. Soc. Am., **72**, 5, 1403–1412 (1982).

[6] W.M. HARTMANN, *The effect of amplitude envelope on the pitch of sine wave tones*, J. Acoust. Soc. Am., **63**, 4, 1105–1113 (1978).

[7] S. IWAMIYA, K. FUJIWARA, *Perceived principal pitch of FM-AM tones as a function of the phase difference between frequency modulation and amplitude modulation*, J. Acoust. Soc. Jpn. (E), **6**, 3, 193–202, 1985.

[8] S. IWAMIYA, K. KOSUGI, O. KITAMURA, *Perceived principal pitch of vibrato tones*, J. Acoust. Soc. Jpn. (E). **4**, 2, 73–82, (1983).

[9] S. IWAMIYA, S. NISHIKAWA, O. KITAMURA, *Perceived principal pitch of FM-AM tones when the phase difference between frequency modulation and amplitude modulation is in-phase and anti-phase*, J. Acoust. Soc. Jpn. (E), **5**, 2 (1984).

[10] L.A. JEFFRES, *Beating sinusoid and pitch changes*, J. Acoust. Soc. Am., **43**, 6, 1464 (1968).

[11] R.A. JENKINS, *Perception of pitch, timbre, and loudness*, J. Acoust. Soc. Am., **33**, 11, 1550–1557.

[12] R.J. RITSMA, F.L. ENGEL, *Pitch of frequency modulated signals*, J. Acoust. Soc. Am., **36**, 2, 1637–1644 (1964).

[13] J.G. ROEDER, *The physics and psychophysics of music*, 2nd Ed., Springer, Berlin 1975.

[14] T.D. ROSSING, A.J.M. HOUTSMA, *The effect of amplitude envelope on the pitch of short sine-wave and complex tones*, ICA 12 - Toronto 1986, B2-1.

[15] L. RUTKOWSKI, *Symmetry and asymmetry as a physical and perceptual feature of the complementary pair of beating sinusoids, Part I. Amplitude and frequency envelope relations*, Arch. Acoust., **23**, 1, 51–66 (1998).

[16] J.F. Schouten, R.J. Ritsma, B.L. Cardozo, *Pitch of the residue*, J. Acoust. Soc. Am., **34**, 8, Part 2, 1418–1424 (1962).

[17] S.S. Stevens, *The relation of pitch to intensity*, J. Acoust. Soc. Am., **6**, 3, 150–154 (1935).

[18] N.J. Versfeld, A.J.M. Houtsma, *Discrimination of changes in the spectral shape of two-tone complexes*, J. Acoust. Soc. Am., **98**, 2, 807–816 (1995).

[19] J. Vershuure, A.A. van Meeteren, *The effect of intensity on pitch*, Acustica, **32**, 33–44 (1975).

[20] E. Zwicker, *Direct comparisons between the sensations produced by frequency modulation and amplitude modulation*, J. Acoust. Soc. Am., **34**, 8, Part 2, 1425–1430 (1962).

ARCHIVES OF ACOUSTICS
23, 1, 87–104 (1998)

ACOUSTIC ISSUES OF SACRAL STRUCTURES

M. NIEMAS and J. SADOWSKI

Acoustic Department of Building Research Institute
(02-656 Warszawa, 21 Ksawerów Street, Poland)

Z. ENGEL

Mechanics and Vibroacoustics Department,
Technical University of Mining and Metallurgy
(30-059 Kraków, 30 Mickiewicza Avenue, Poland)

In the paper sacral structures treated as speech-music halls were evaluated in terms of acoustic properties. The basic acoustic issues occurring at a planning stage as well as the methods of evaluation of the finished, existing and being modernized interior acoustic quality of sacral structures are discussed. Acoustic requirements refer to some type of liturgical ceremonies, and to cultural activities performed as well in those structures. The current status of church acoustics in structures built in different periods up to 1950 is presented. The methods of evaluation of the sacral interiors acoustic quality are nowadays significantly improved. They are discussed on an example of newly designed churches.

1. Status of the issue

The problem of sacral structure acoustics is a part of room acoustics. Thus it is evident that, the task of the former is twice that of the latter as illustrated after H. Kuttruff in Fig. 1 [25]. Figure 2 shows the problem of room acoustics from the viewpoint of designing and estimation of the sacral structure acoustics.

In Poland, the problem of sacral structure acoustics has been neglected contrary to other European countries where suitable attention was paid to this question.

This situation existed even after 1987 when the development of sacral buildings has been intensified. Even now, when the insufficient financial support is the only barrier for investigation, only some of the Universities and Research Institutes deal with this issue. This results in the creating of a number of churches (of different religions) with poor acoustic characteristics, because their designers are not acquainted with the results of acoustic research and pay attention only to architectural forms and functional solutions. They seem to forget that the acoustic characteristics of interiors contribute to the final assessment of the structure quality.

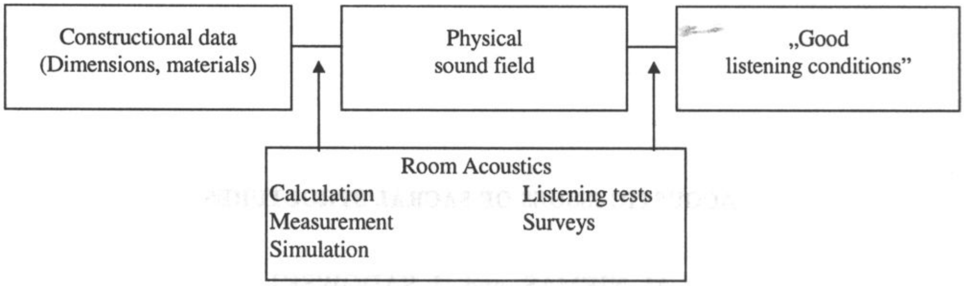


Fig. 1. Tasks and methods of room acoustics [25].

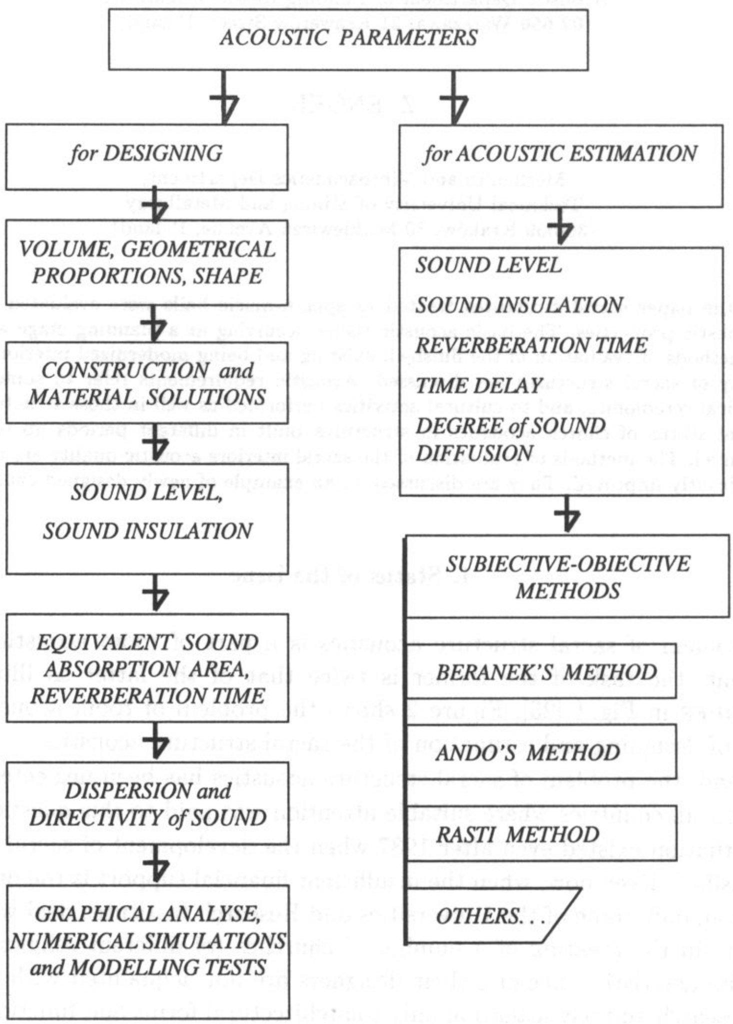


Fig. 2. Problems of room acoustics from the point of view of designing and estimation of the sacral structures acoustic.

In the year's 1975÷1985 acoustic investigations of the temples were realized at the Technical University of Gdańsk only. The first one was the room acoustical investigation of the Gdańsk-Oliva Cathedral. As the Oliwa Cathedral with its famous organ has become a world known place for church organ music, it requires a thorough acoustical investigation concerning the organ quality, the room acoustics properties and their independence. The results of the acoustical investigations were presented [9].

The next was the room acoustic investigations of the Saint Mary's church in Gdańsk built in the XIV-th century. It is the largest church in Poland. Its internal volume is about $97000\,m^3$.

Saint Mary's church in Gdańsk, due to the large volume of its interior and poor absorption, presents a serious acoustic problem. Acoustic studies on the properties of this interior and the possible acoustic corrections connected with the installation of the new great organ in the baroque style are described in [8, 39, 50, 51, 52].

Large churches differ acoustically from other multipurpose large halls by the distributed sound reinforcement systems applied for service transmission, as well as by depending on the architectural style and the highly variable sound absorption according to the large church acoustics. Some suggestions, parameters of St. Mary's Basilica in Gdańsk have been given as an example of the acoustic problem of a large church [9, 39].

Irrespective of the period of erecting the church, two groups of acoustic issues are playing an important role, namely the protection of the sacral interiors, and sometimes their vicinity, against external noise and vibrations and the acoustics of the sacral interiors.

The way of the formulation of those issues was connected over centuries with characteristic features of a given period referring to the site planning, culture of the society, building style of this period, structural and material solutions, decorative art, and availability of financial means connected mainly with the public liberality. It was also greatly influenced by historical events, invasions, wars, bondage and discrimination in the different periods of the polish history. All these factors influenced the status of sacral structures, including their acoustics (Fig. 3).

Protection of the sacral structures against noise has been formed naturally over ages and resulted from functional, structural and material solutions of these structures. The acoustics of interiors, however, resulted not only from functional solutions but also from styles governing in particular periods, the use of massive constructions, lack of sound amplification devices. The reverberation time calculated from the known Sabine's formula, and also that measured, was the only important acoustic parameter used for evaluation of the quality of the sacral structures acoustics. It was been for long years the only parameter which could be used for designing purposes as well as for the acoustic evaluation of the realized or modernized interiors. According to shape and interior – decorations, the degree of coupling of particular interiors, number of sculptures (e.g. the saints' statues), the interior form, and also the number, shape and localization of columns as well as other architectural elements a second important factor of the acoustics of interiors has been formed – the index of sound dispersion. This index, influencing the interior acoustic quality, caused that large sacral interiors with a very high reverberation time disadvantageous for its characteristics had pretty good acoustic properties.

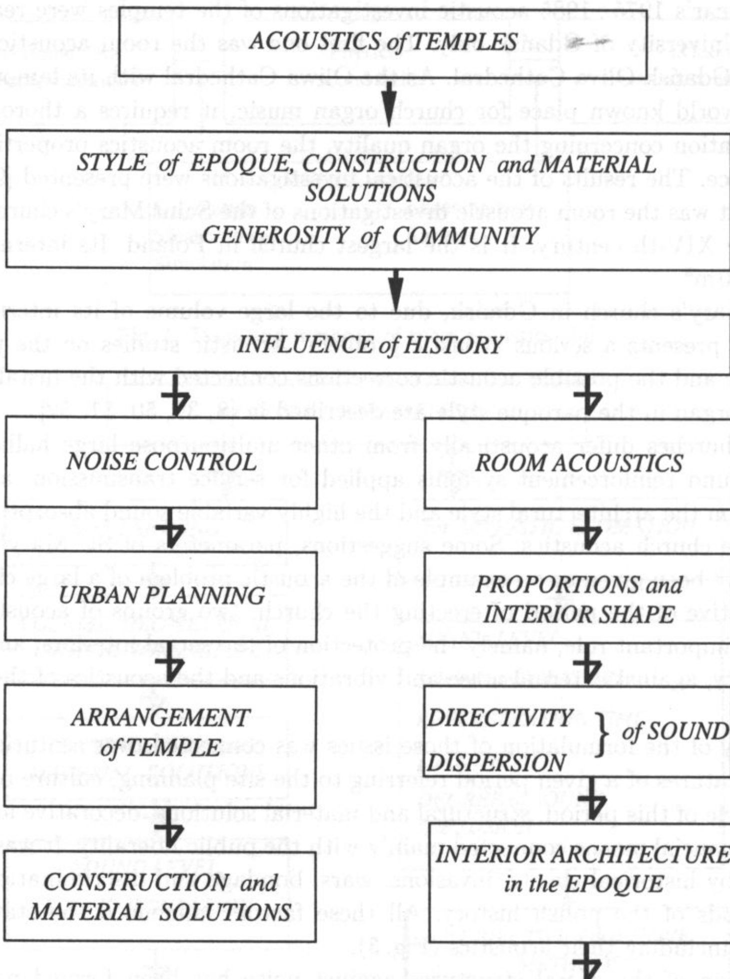


Fig. 3. Factors influencing in status of sacral structures including their acoustics.

2. Development of investigations of the room acoustics phenomenon

For a long time, the reverberation time of rooms was the most important factor of room acoustic estimation and design. Very important was the fact that it could be calculated with reasonable accuracy from room data by the simple Sabine's formula (¹).

$$T = 0.161\frac{V}{A}, \quad \mathrm{s} \tag{1}$$

where V – the room volume in m^3, A – equivalent sound absorption area, m^2.

(¹) See H. Kuttruff [25], 1991: $T = 0.163\ V/A$.

Suppose the enclosure consist of N different walls, floors, partitions with areas $S_1, S_2, ..., S_N$ etc. and with absorption coefficients $\alpha_1, \alpha_2, ..., \alpha_N$ etc. Then:

$$A = (S_1 \cdot \alpha_1 + S_2 \cdot \alpha_2 + ... + S_N \cdot \alpha_N). \tag{2}$$

In large auditoria ($V \geq 1000\,\mathrm{m}^3$) not only the wall and floor absorption but also the attenuation of sound in air should be included. This is achieved by adding a term 4 mV in the brackets of Eq. (2): m is the attenuation constant of air.

The reverberation time has been introduced into room acoustics by W.C. SABINE (1923) whose work on reverberation and other topics marked the beginning of scientific room acoustics. It can easily be measured, and nowadays it is known for many concert halls, theatres and other auditoria. Basically, reverberation has the tendency to blur the sounds produced in a hall, in particular to impair the intelligibility of speech. It might seem therefore that reverberation should be avoided as far as possible [25].

V.O. KNUDSEN (1932) after dividing the church buildings into three groups as country churches, village churches and city churches has given reverberation times optimal for the church auditorium [22] (Table 1).

Table 1. Optimal values of reverberation times for church auditoriums according to V.O. KNUDSEN [22].

Reverberation time for churches, s	Volume [m³]					
	1000	2000	4000	8000	16000	32000
Roman Catholic churches	1.5	1.6	1.7	1.8	1.9	2.1
Protestant and Jewish churches	1.3	1.4	1.5	1.6	1.7	1.8
Christian Science churches	1.1	1.2	1.3	1.4	1.5	1.6

In 1956 E. MAYER and R. THIELE [29] have carried out acoustical measurement in 31 different rooms ($V = 550$ to $22000\,\mathrm{m}^3$):

1. Measurements of the reverberation time as a function of frequency.
2. Recording of the reverberation curves with the "impulse glide method".
3. Measurements of sound pressure as a function of frequency and calculation of the frequency irregularity.
4. Recording of the directional distribution of the sound at 2 kHz and calculation "directional diffusiveness".
5. Recording of the reflections of a short pulse and measurement of the "50% energy part".

The results of these measurements show that only the methods 4 and 5, developed according to the geometrical conception of room acoustics, give more information for the judgement of the hearing conditions of a certain place in the room, than the reverberation time does. The subjective significance of the observed phenomena has still to be cleared in detail by means of hearing tests carried out by suitable electroacoustical methods. For this purpose the measured directional distributions and time sequences of the sound reflections will be very useful.

D. FITZROY (1959) [12] observed that when the average absorption was approximately the same in all the directions, the accepted formulas were quite close; thus was confirmed by his experience. But when the distribution of absorption was unequal in terms of direction and in the average absorption per square foot – as, for example, in an all – over ceiling acoustical treatment with little absorption on the side and end walls – the reverberation times, measured by Fitzroy was much longer than the Sabine and Eyring formulas predict [22, 23]. D. Fitzroy gave a reverberation times formula which in the case of nonuniform distribution of absorption seems to be more accurate than the Sabine or Eyring formulas [12, 37]:

$$T=\frac{S_x}{S}\left[\frac{0.161V}{S\ln(1-\alpha_x)}\right]+\frac{S_y}{S}\left[\frac{0.161V}{S\ln(1-\alpha_y)}\right]+\frac{S_z}{S}\left[\frac{0.161V}{S\ln(1-\alpha_z)}\right],\mathrm{s} \tag{3}$$

where V – volume of the room, m^3, $S\ln(1-\alpha)$ – equivalent absorption area of the room, m^2, S_x, S_y, S_z – areas of the walls and floors in direction x, y, z, $\alpha_x, \alpha_y, \alpha_z$ – absorption coefficients of the walls (floor) in direction x, y, z.

In Poland Andrzej RAKOWSKI (1968) and Jerzy SADOWSKI (1959, 1971) described some factors governing the acoustical quality of concert halls. The quality parameters proposed by several authors for room acoustic are discussed in the view of practical demands. The opinions of musicians on some well known concert halls and opera houses are quoted and technical parameters of halls are given [35, 37, 38]. Jerzy SADOWSKI (1971) has published a book [37] (in Polish) with the chapter 21 entitled "The Temple". It was the first book in Poland giving the guidelines of acoustical designing of temples. Some remarks on the reverberation time criterion and its connection with the acoustical properties of a room are given in the publication "Materials from the conference on measurements of reverberation phenomena in halls". The conclusion concerning the reverberation time given by Witold STRASZEWICZ (1968) was as follows: "The results of experiments lead to the conclusion that reverberation time is significant only for classification of the auditoriums as more or less reverberant. The determination of the relation between the results of the reverberation time measurements and the subjective judgement of acoustical quality prove to be difficult. The difficulty results mainly from the spatial character of the reverberation effect and from the general adoption of field diffusity as the basis of the classical measurement methods." It seems that we should dispense with the assumption and try to determine the characteristic field parameters in conditions more resembling those of the normal use of a given interior [46].

In last years, after 1990 have been some investigation results published, which are very useful for the evaluation of the room acoustic phenomenon.

The K. ŚRODECKI and A. ŚLIWIŃSKI (1991) carried out investigations in rooms of various shapes. They have demonstrated that the employment of autocorrelation and cepstrum functions enables the assessment of certain acoustical properties of rooms, i.e. of the degree of sound scattering, the possibilities of the occurence of acoustically unfavourable phenomena such as the sound colouration or echo, influencing in general a property of the room called by the authors the reverberation quality [45].

It is widely known that the measured value of reverberation time does not characterise itself the nature of the reverberation process. Thus the temporal diffusion can, together with reverberation time, be one of the basic parameters describing the acoustic properties of a room independently of many other criteria used in room acoustics.

The investigations carried out by K. ŚRODECKI (1994) [44] suggested further problems to be solved including comprehensive studies on the effect of a room on the colouration change in the signal propagated in the room by employment of the cepstrum function.

The results of L. Gerald MARSHALLS (1994) investigations show, that the importance of the early reflection portion of a sound decay process in the auditoriums is well established. Curves showing early/late sound energy rations (ELR) in the early reflection period comprise a useful way of examining energy – time data within that period. The L. Gerald MARSHALLS (1994) paper on auditorium measurements and the analysis procedure using ELR data in a period between 20 and 200 ms after the arrival of the initial signal is presented in [27]. A measurement and analysis procedure based on the early/late sound energy ratio is a potentially useful way for evaluating the acoustic response of the auditorium. The ELR procedure described in Gerald MARSHALL's paper [27] was developed for evaluating the acoustics of large spaces, such as auditoriums, concert halls, theatres, churches and so forth.

Curve C_x, shows theoretical ELR values for a pure exponential decay based on a decay time equal to that of the measured curves:

$$C_x = 10\log(e^{13.82t/RT} - 1), \tag{4}$$

where t is the ratio dividing time in seconds.

This is the theoretical equivalent of the measured without – direct – signal curve, C_{t0}, and can serve as a useful reference. For this purpose, the difference between C_{t0} and C_x is illustrated in Fig. 4.

Gilbert A. SOULODRE and John S. BRADLEY (1994) have examined the relative importance of the various subjective parameters to the overall preference of these sound fields. A series of experiments were conducted to evaluate several acoustical measures as predictors of subjective judgements. Subjects were asked to rank binaurally reproduced sound fields in terms of loudness, clarity, reverberance, bass, treble, envelopment, apparent source with, and overall preference. The binaural impulse responses were measured in several North American Concert Halls and were chosen so as to cover a range of RT, EDT, C80, G, IACC, and LF values [43] as possible as broad. The results of these tests were then correlated with the various 1/1 octave band objective measures to find the best predictor of each subjective parameter. New objective quantities correlating with loudness, clarity, bass and treble judgements were found.

The reconciliation of speech and music in sacral rooms may be realized by the distributed column sound system. This is not subject of this paper. Some information's are given by David L. KLEPPER (1993) [21].

The diffuse-field theory is used by designers to predict sound fields in rooms of every type. The fact that the theory is based on assumptions, which may limit its applicability, is often forgotten. If the theoretical assumptions do not hold in the case of a particular room for which predictions are to be done, the predictions may not be accurate.

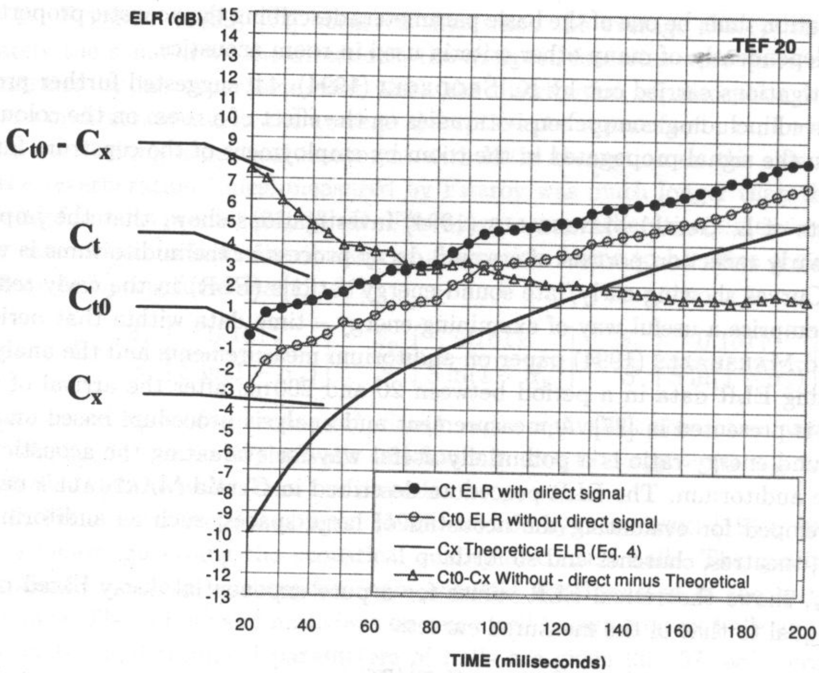


Fig. 4. ELR data curves [27].

Murray HODGSON (1996) [17] gives an objective review of what is known about the applicability of the diffuse field theory. His paper is based on the work by Kuttruff and on his own, comparing the predictions of the diffuse-field theory and the ray-tracing models. It considers two versions of the diffuse-field theory – the Eyring and Sabine versions and the prediction of both the sound decay/reverberation time and the steady state sound pressure level. It discusses applicability with respect to the following room acoustical parameters: room shape, surface absorption (spatial distribution and magnitude), surface reflection, fitting density. The author discussed how to apply the experimental results to real rooms.

Spatial information about sound fields for the room-acoustics evaluation and diagnosis were given by A. ABDON and R.W. GUY (1995) [1].

3. Acoustic requirements for sacral structures

Multiple functions of churches as well as traditionalism and ritualism and the eagerness for obtaining the architectural beauty greatly influence the construction. Almost all sacral buildings are based on the longitudinal, circular, and Greek or Roman cross maps (these are the basic forms). There are also structures, which include several forms mutually coupled. For churches with a simple construction, the acoustic requirements are easier to meet than for complicated spatial forms for which obtaining the assumed (required) acoustic properties is more difficult because these requirements should be met not only for the entire auditorium but also for each of the area individually. The detailed requirements for a sacral structure should be known to design correctly the church or evaluate acoustic terms in it.

Churches designed correctly in the acoustic terms should meet many requirements of which the most important ones are:

- conditions for undisturbed pray and meditation in the interior;
- relatively low level of sound level from external noise;
- good conditions of organ music, songs, chorus, speech listening;
- good speech intelligibility and music texts;
- proper interior volume, reverberation time (characteristics in the frequency domain and value of the reverberation time);
- uniformity of sound amplification;
- proper sound quality.

Good listening conditions in churches should not depend on the site of the sound emission beginning. Sacral interiors should be designed in a way that ensures the dissipated acoustic field. According to [37], the sound level from external noise in churches should not, exceed 25 – 35 dB to ensure undisturbed conditions for prayer and separation from the exterior world in the interior.

Table 2 presents information on the reverberation time of selected sacral interiors erected in different periods according to the various styles typical of them. High discrepancies in the reverberation time and its characteristics are seen.

Table 2. Reverberation time values in interiors of selected churches [34].

No	Name of church and localization	Volume $\cdot 10^3$ [m^3]	Reverberation time T [s] 125 [Hz]	250 [Hz]	500 [Hz]	1000 [Hz]	2000 [Hz]	4000 [Hz]
1.	Church of St. Mark in München	6	2.8	3.2	3.9	3.8	3.0	2.0
2.	Church of St. Matthew in München	9	3.7	4.0	5.9	5.8	4.9	2.0
3.	Church of St. Thomas in Lipsk	18	2.4	3.3	4.1	4.0	3.2	1.8
4.	Chapel of Blessed Kinga at Wieliczka	7.5	5.7	5.6	5.6	5.3	4.7	2.8
5.	Basilica of Our Lady in Gdańsk	100	10.7	11.1	11.4	10.5	6.9	3.7
6.	Church of Sts. Peter and Paul in Cracow	24	5.4	5.6	5.7	5.3	4.0	2.5
7.	Church of St. Charles in Tallin	19.2	4.0	4.0	5.6	6.0	5.0	3.0
8.	Church of Our Lady in Cracow	9.5	12.9	9.8	7.7	5.3	4.4	3.5

As was mentioned above, the sacral interior acoustics is greatly influenced by the architectural style of the churches, which to great degree determine the acoustic quality of the sacral structure. The church of Saint Roch in Białystok, built before World War II, and the Orthodox Church in Hajnówka (Figs. 5 and 6) built in the nineties, may be examples of churches differing from the traditional structures of the XVII and XVIII centuries.

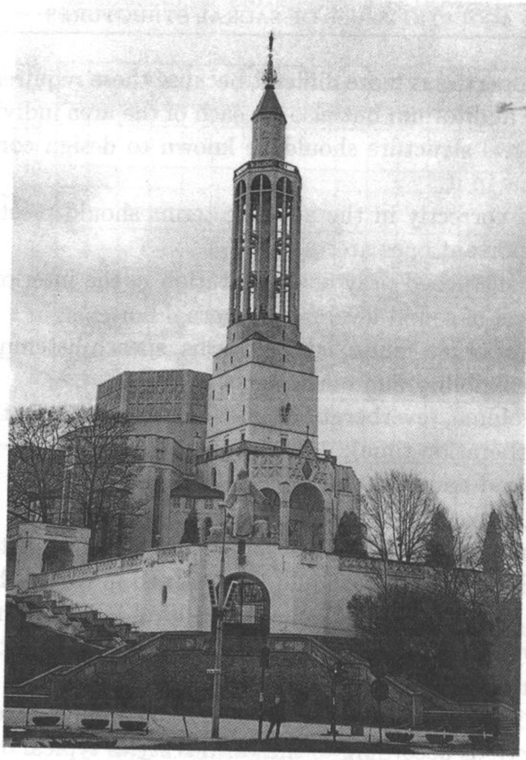


Fig. 5. Church of St. Roch in Białystok [32, 33].

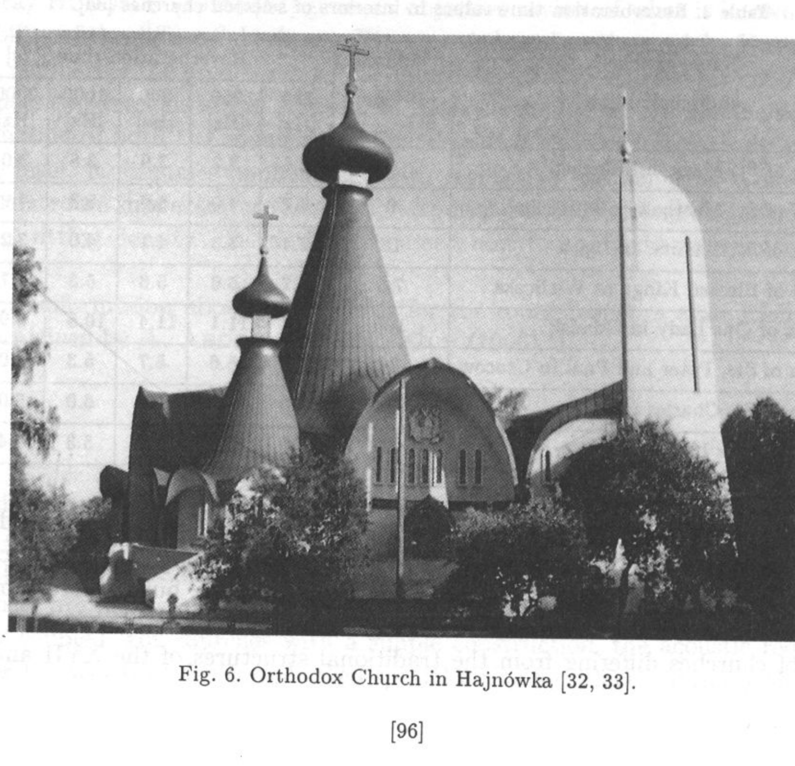


Fig. 6. Orthodox Church in Hajnówka [32, 33].

It is difficult to characterize the acoustic quality of those structures because they were not subjected to acoustic research up to now. Most certainly, the search for new architectural styles of churches observed during the last 20 years should include the acoustic requirements if the acoustics of those churches is to meet their function. There are examples of churches built in last years with of good architecture and particularly bad acoustic quality. Due to this it worthy to investigate the issue of church acoustics on examples of the sacral structures built in the last years.

4. New evaluation methods of acoustic characteristics of sacral structures

The evaluation of a sacral interior in terms of its acoustic characteristics is connected mainly with the impression made on believers (listeners) through speech, chorus or songs. This impression depends not only on the predisposition and individual characteristics of the listeners but also mainly on the acoustic characteristics of the sacral interior and the absence of external noises. It is therefore required to use evaluation methods, which combine an objective evaluation of interior acoustics (measurement or calculation), with subjective evaluation of the listeners.

Up to now no evaluation methods referring exclusively to the sacral interiors were elaborated and modified methods used of the evaluation of the auditorium acoustic properties are applied for this purpose. From amongst those methods of evaluation of the interior acoustic characteristics that have been used recently, the newest one are the methods of digital recording and processing of sound signals. They enable the recording and analysis of a response of the interior to external excitation and afterwards the estimation of the autocorrelation function, the interior impulse response, and the determination of the modulation transfer function.

Application of the autocorrelation function method and the modulation transfer function for evaluation of the quality of the sacral interior acoustics is however in a beginning phase and is limited only to several foreign research centers. Nevertheless the results of investigations are very promising. Objective measurements of sacral interior characteristics enable also the evaluation of the structure of reflections reaching the listener that depend to a great degree on the acoustic characteristic of the interior. It influences crucially the quality of the sound impression in the interior. The important parameters are: EDT – Early Decay Time, D_{50} – deutlichkeit, C_{80} – clarity, L – liveness, t_s – center time, Δ – time diffusion, RASTI – speech intelligibility.

A relationship between the above mentioned parameters of the acoustic field and the subjective attributes of the sound impression in the interior like reverberationess, reverberation quality, tone quality, intimacy, clarity and speech intelligibility was found. Thus it is possible to evaluate the acoustic properties of the investigated interior in terms of impressions made on the listener. For the evaluation of acoustic characteristics of the interiors, including sacral ones, the following methods are applied:

- **BERANEK's** (basing on analysis of the reverberation time and the first reflection);
- **ANDO's** (based on analysis of the impulse response of the interior);
- **RASTI** (based on analysis of the modulation transfer function).

The L.L. BERANEK*'s method* [6]

The first method of evaluation of auditoriums in acoustic terms was developed by the American acoustician L.L. BERANEK [6]. His evaluation scale was based on the comparison of acoustic characteristics of 54 large concert and opera halls all over the world. The method rest on points given from measured or calculated values of the reverberation time, delay time of the first reflection, the distance of the listener from the source, the kind of music performed, and the level of external noise. The halls were classified according to the on number of points obtained as follows:

A^+	– excellent	(90 – 100) points
A	– very good to excellent	(80 – 90) points
B^+	– good to very good	(70 – 80) points
B	– sufficient to good	(60 – 70) points
C	– sufficient	(50 – 60) points

Halls with a score below 50 points are not suitable for music or speech performance due to their bad acoustic quality.

The method of subjective preference acc. to ANDO [3]

For the evaluation of the acoustic field, a scale was elaborated based on the law of comparative values. Comparative tests of the subjective preference were performed according to totally independent objective parameters and in this way the optimal parameter values were obtained. Having a linear scale of preferences for each individual parameter, it is possible to apply the rule of superposition to calculate the total evaluation of the hall acoustics on the general scale of preferences.

Four parameters selected for evaluation, called the *preference factors*, are as follows:

- listening level (S_1),
- time delay (S_2),
- subsequent reverberation time (S_3),
- IAAC – maximum value of the interaural cross correlation function $f_{lr}(t)$ (S_4).

The RASTI *method*

The speech intelligibility in auditoriums, churches, etc. is of importance and serves to evaluate the acoustic quality of the interiors. Evaluation of the speech quality by the identification and evaluation of the interior effect on the sound signal received by the listeners is determined by the RASTI method (Rapid Speech Transmission Index). This method is a modification of the STI method (Speech Transmission Index) and is based on 9 measurements taken in 2 octave bands with middle frequencies of 500 and 2000 Hz. The value of the RASTI factor is calculated as follows:

$$\mathrm{RASTI} = [(S/N)_{\mathrm{app}} + 15]/30, \tag{5}$$

where $(S/N)_{\mathrm{app}}$ – sound signal to interference noise ratio.

5. Investigation of selected sacral structures

The authors of his paper performed many investigations of the acoustic properties of sacral structures. Very interesting results were obtained for the Chapel of the Blessed Kinga in the Salt Mine at Wieliczka. At the end of last century a chapel for 400 seats and 1000 standing places was established in a chamber of $7500\,m^3$.

The chapel is characterized by very good acoustic conditions; concerts of various types are often organized in this chapel. Table 3 and Fig. 7 present the reverberation time values in a frequency domain.

Table 3. Results of measurements of the reverberation time in the Chapel of the Blessed Kinga obtained by the Schroeder method [11].

Freq. [Hz]	p1-Q1 [s]	p2-Q1 [s]	p3-Q1 [s]	p4-Q1 [s]	p1-Q2 [s]	p2-Q2 [s]	p3-Q2 [s]	p4-Q2 [s]	Avg. RT [s]
125	5.0	5.6	6.1	5.9	5.9	5.5	5.5	5.8	5.66
250	5.6	5.5	5.4	5.6	5.7	5.9	5.8	5.5	5.63
500	5.6	5.4	5.5	5.6	5.8	5.9	5.7	5.7	5.65
1000	5.3	5.0	5.3	5.3	5.4	5.6	5.3	5.4	5.33
2000	4.6	4.4	4.5	4.5	4.9	5.2	4.6	4.7	4.68
4000	2.9	2.7	2.7	2.7	3.0	3.0	2.9	2.8	2.85

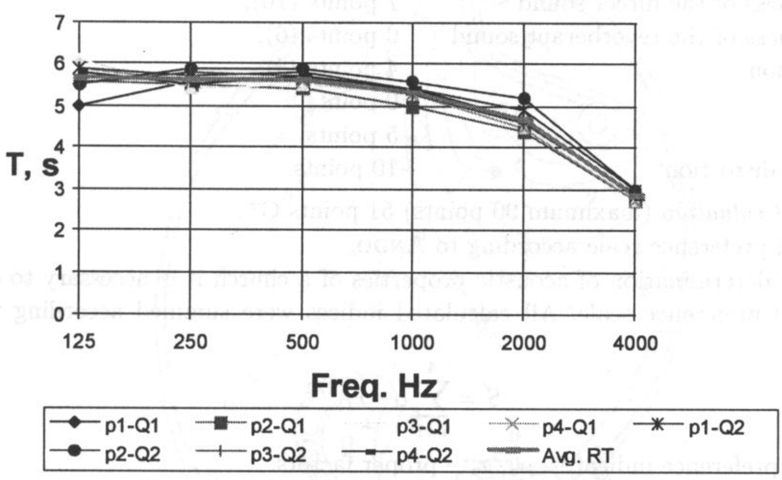


Fig. 7. Reverberation time characteristic in Chapel of St. Kinga at Wieliczka [11].

The results of the investigation obtained for the Sts. Peter and Paul Church in Cracow are presented in Table 4.

The results of investigation obtained for the modern St. John Kanty Church in Cracow are presented below. This church is the first one for which a complex evaluation was made using the methods described in point 3.

Table 4. Results of medium values of the reverberation time for the Church of Sts. Peter and Paul in Cracow [24].

Freq. [Hz]	RT [s]	σ(RT) [s]
125	5.4	0.583
250	5.6	0.473
500	5.6	0.331
1000	5.3	0.310
2000	4.0	0.363
4000	2.5	0.281

Table 5. Averaged reverberation time for the Church of St. John Kanty in Cracow [34].

Freq. [Hz]	125	250	500	1000
RT [s]	14.3	12.8	11.6	10.3
σ(RT) [s]	0.13	0.24	0.10	0.30

The general evaluation according to the BERANEK's method obtained for this church is as follow:

Intimacy	40 points (40),
Liveness	0 points (15),
Warmth	15 points (15),
Loudness of the direct sound	7 points (10),
Loudness of the reverberant sound	0 points (6),
Diffusion	4 points (4),
Echo	0 points,
Noise	−5 points,
Tonal distortion	−10 points.

General valuation (maximum 90 points) 51 points C$^+$.

General preference scale according to ANDO.

For the determination of acoustic properties of a church it is necessary to establish the general preference scale. All calculated indices were summed according to below formula:

$$S = \sum_{i=1}^{4} \alpha \cdot |x_i|^{3/2}, \tag{6}$$

where S – preference indicator, α_i, x_i – proper factors.

General values of the preference scale were calculated for two motives: organ music and speech (man voice) in 21 measurement points for the range of 125 – 1000 Hz in the octave bands and taken in isolines. The diagrams created in this way [34] illustrate the evaluation of the interior acoustic according to the ANDO's method for speech and organ music. Figure 8 presents the evaluation isolines according to ANDO for organ music for the selected frequency and Fig. 9 presents the evaluation isolines according to the RASTI method.

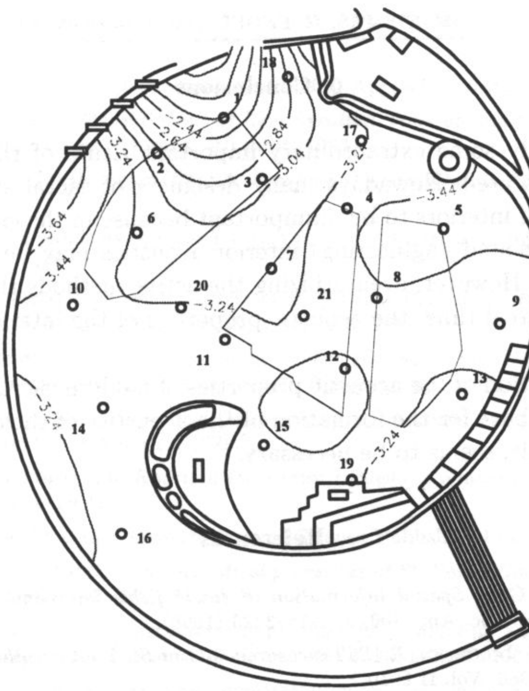


Fig. 8. Isolines of general evaluation of the interior of the St. John Kanty Church in Cracow according to the ANDO method for organ music at a frequency of 250 [Hz] [34].

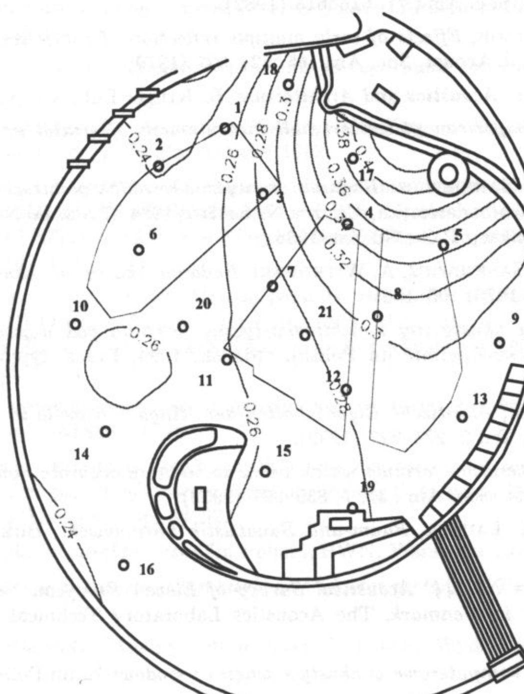


Fig. 9. Isolines of the RASTI factor for the interior of the St. John Kanty Church in Cracow [34].

6. Conclusions

In this paper only some extraordinary important issues of the acoustics of sacral structures were presented. Nowadays many designers of sacral structures believe the acoustic properties of interiors to be unimportant because in the era of the development of electroacoustics (sound engineering) interior acoustics may be improved by proper sound amplification. However, even utilizing the active methods that enable the amplification control in a real time, the acoustic properties of the interior play an important role.

Further investigation of the acoustic properties of modern sacral structures aimed at the elaboration of a base for the formation of the acoustics of those structures in terms of the designing needs, seems to be necessary.

References

[1] A. Abdon, R.W. Guy, *Spatial information of sound fields for room-acoustics evaluation and diagnosis*, J. Acoust. Soc. Am., **100**, 5, 3215–3226 (1996).

[2] J.S. Andersen, T. Jacobsen, *RASTI measurements in St. Paul's Cathedral in London*, 12 ICA Congres Toronto 1986, Vol. II E 10-2.

[3] Y. Ando, *Concert Hall Acoustics*, Springer Verlag, Berlin, Heidelberg, New York, Tokyo 1985.

[4] Y. Ando, H. Altruz, *Perception of coloration in sound fields in relation to the autocorelation function*, J. Acoust. Soc. Am., **71**, 616–618 (1982).

[5] Y. Ando, D. Gottlob, *Effects of early multiple reflection of subjective preference judgments of music sound fields*, J. Acoust. Soc. Am., **65**, 524–527 (1979).

[6] L. Beranek, *Music, Acoustics and Architecture*, R. Kruger Publ. Corp., N. York 1979.

[7] J.S. Bradley, *Comparison of concert hall measurement of spatial impression*, J. Acoust. Soc. Am., **96**, 6 (1994).

[8] G. Budzyński, M. Sankiewicz, *Badania akustyczne wielkich pomieszczeń* [in Polish], Sprawozdanie z prac naukowo-badawczych, Umowa Nr 841362/1984 Wydz. IV Nauk Technicznych PAN – Instytut Telekomunikacji P.G., Gdańsk 1984.

[9] G. Budzyński, M. Sankiewicz, A. Witkowski, *Badania Akustyczne Katedry Oliwskiej* [in Polish], XXIII OSA, Wisła 1976, 105–106.

[10] T. Dolny, *Projekt akustyczny i elektroakustyczny kościoła pod wezwaniem bł. Urszuli Ledu-chowskiej w Gdańsku-Chełmie* [in Polish], Gdańsk, 1994, Praca dyplomowa Zakład Inżynierii Dźwięku PG.

[11] Z. Engel, J. Engel, *Akustische Eigenschaften der Kinga – Kapelle in der Salzgrube Wieliczka*, DAGA'92 – Berlin, Teil 2, 225–227 (1992).

[12] D. Fitzroy, *Reverberation formula which seems to be more accurate nonuniforme distribution of absorption*, J. Acoust. Soc. Am., **31**, 7, 839–897 (1959).

[13] W. Furrer and A. Lauber, *Raum-und Bauakustik Lärmabwehr*, Birkhäser Verlag, Basel und Stuttgart 1972.

[14] A.C. Gade, *Report No. 44. Acoustical Survey of Eleven European Concert Halls a basis for discussion of halls in Denmark*, The Acoustics Laboratory Technical University of Denmark, Lyngby 1989.

[15] A. Gołaś, *Metody komputerowe w akustyce wnętrz i środowiska* [in Polish], Wyd. AGH, Kraków 1995.

[16] A. Gołaś, H. Łopacz, *Metody impulsowe w badaniach procesów wibroakustycznych* [in Polish], Wibroakustyka Maszyn i Śr., Wiedza i Życie, Warszawa 1995.

[17] M. HODGSON, *When is diffuse-field theory applicable?*, Applied Acoustics, **49**, 3, 197–207 (1996).

[18] ISO 3382-1975, *Acoustics, Measurement of Reverberation Time of Auditoria*, Revision, Measurement of the acoustic qualities of rooms, 4th draft, 1993.

[19] D.L. KLEPPER, *First Presbyterian Church Stanford, Connecticut*, J. Acoust. Soc. Am., **31**, 7, 879–881 (1959).

[20] D.L. KLEPPER, *Special problems in church design in North America*, 12 Int. Congres on Acoustics, Paper E3-2, Toronto 1986.

[21] D.L. KLEPPER, *The distributed sound system at Holy Cross Cathedral, Boston, the reconciliation of speech and music*, J. Acoust. Soc. Am., **99**, 1 417–425 (1996).

[22] V.O. KNUDSEN, *Architectural Acoustics*, John Willey and Sons, Inc. New York, Chapman and Hall, Limited London 1932, Chapter XXI. Church Buildings, pp. 462-485.

[23] V.O. KNUDSEN, C.M. HARRIS, *Acoustical Designing in Architecture*, American Institute of Physics for the Acoustical Society of America, 1978.

[24] G. KUBINYI, *Analiza akustyczna Kościoła św. Piotra i Pawła w Krakowie* [in Polish], Praca dypl. pod kier. Z. Engla, Kraków 1993.

[25] H. KUTTRUFF, *Room Acoustics*, Elsevier Applied Science Publishers Ltd., 1991.

[26] T.H. LEWERS, J.S. ANDERSON, *Some acoustical properties of St. Paul's Cathedral, London*, Journal of Sound and Vibration, **92**, 2, 285–297 (1984).

[27] L.G. MARSHALL, *An acoustics measurement program for evaluating auditoriums based on the early/late sound energy ratio*, J. Acoust. Soc. Am., **96**, 4, 2251–2261 (1994).

[28] MENG X. W.G DE BORGER and M. VAN OVERMEIRE, *Numerical analysis of the acoustical quality of a multi-purpose hall*, Building Acoustics, **2**, 3, 487–504 (1995).

[29] E. MEYER and R. THIELE, *Raum akustische Untersuchungen in zahlreichen Konzerts(len und Rundfunkstudios unter Anwendung neuerer Messvertfahren*, Akustiche Beihefte, S. Hirzel Verlag, Stuttgart 1956, pp. 425–444.

[30] E.E. MIKESKA and R.N. LANE, *Acoustical problems in two round churches*, J. Acoust. Soc. Am., **31**, 7, 857–865 (1959).

[31] M. NAWROCKI, *Charakterystyka akustyczna Kaplicy BB. Kingi w Kopalni Soli w Wieliczce* [in Polish], Praca dypl. pod kier. Z. Engla, Kraków 1992.

[32] M. NIEMAS, Z. ENGEL, J. SADOWSKI, *Zagadnienia akustyczne obiektów sakralnych* [in Polish], Konferencja Naukowo-Techniczna "Budownictwo Miast i Wsi – Budownictwo Sakralne '96, Politechnika Białostocka, ZO PZITB, ZOSARP w Białymstoku, Wydawnictwa Politechniki Białostockiej, Białystok 1996, pp. 167–174.

[33] M. NIEMAS, Z. ENGEL, J. SADOWSKI, *Metody pomiaru własności akustycznych obiektów sakralnych* [in Polish], XLIII Otwarte Seminarium z Akustyki, Materiały, vol. 2, Polskie Towarzystwo Akustyczne, Instytut Fizyki Politechniki Śląskiej, Komitet Akustyki PAN, Gliwice-Ustroń 16–20. 09. 1996, pp. 521–526.

[34] M. NIEMAS, *Zasady projektowania i oceny akustycznej pomieszczeń sakralnych* [in Polish], Praca dypl. pod kier. Z. Engla, Kraków 1994.

[35] A. RAKOWSKI, *Analiza czynników decydujących o jakości akustycznej sal koncertowych (Factors governing the acoustical quality of concert halls)* [in Polish], Państwowa Wyższa Szkoła Muzyczna w Warszawie, Zeszyty Naukowe, pp. 63–92.

[36] J. SADOWSKI, *Akustyka architektoniczna* [in Polish], PWN, Warszawa 1976.

[37] J. SADOWSKI, *Akustyka w urbanistyce, architekturze i budownictwie* [in Polish], Arkady, Warszawa 1971.

[38] J. SADOWSKI, L. WODZIŃSKI, *Akustyka pomieszczeń* [in Polish], Wyd. Komunikacyjne, Warszawa 1959.

[39] M. SANKIEWICZ, *Acoustic properties of the largest church in Poland*, Proceedings of the 2nd Meeting Architectural Acoustics and Urban Noise Protection, International Scientific Symposium Volas 1980, pp. 87–91.

[40] S. Schader, G. Schubert, *Akustiche Gesichtspunkt bei der Gestaltung von Sakralräumen*, Detail 4/81 pp. 478–482.

[41] A. Śliwiński, *On the reverberation time of rooms with small diffusity as a function of the localization of the sound source and measuring point*, Materials from the Conference on Measurements of Reverberation Phenomena in Halls, Warsaw, April, 1967, Archiwum Akustyki, **2**, 3207–221 (1968).

[42] A. Śliwiński, G.E. Budzyński, *Prospects in Modern Acoustics*, Education and Development World Scientific Publishing Co. Pte. Ltd., Teaneck NJ USA 1987.

[43] G.A. Soulodre and J.S. Bradley, *Subjective evaluation of new room acoustic measures*, J. Acoust. Soc. Am., **98**, 1 (1995).

[44] K. Środecki, *Evaluation of the Reverberation Decay Quality in Rooms Using the Autocorrelation Function and the Cepstrum Analysis*, Acustica, **80**, 216–225 (1994).

[45] K. Środecki, A. Śliwiński, *Application of correlation methods for an investigation of the acoustic field in a room*, Archives of Acoustics, **16**, 355–367 (1991).

[46] W. Straszewicz, *Some remarks on the reverberation time criterion and its connection with acoustical properties of room*, Materials from the Conference on Measurements of Reverberation Phenomena in Halls, Warsaw, April, 1967, Archiwum Akustyki, **2**, 3, 99–107 (1968).

[47] M. Szkaradowski, *Projekt centralnego nagłośnienia Bazyliki Mariackiej w Gdańsku* [in Polish], Praca dyplomowa, Zakład Inżynierii Dźwięku P.G., Gdańsk 1990.

[48] Z. Szymaniak, *Badanie akustyki Katedry Oliwskiej* [in Polish], Praca dyplomowa, Zakład Inżynierii Dźwięku P.G., Gdańsk 1973.

[49] J. Tolpilo, *Projekt akustyczny i elektroakustyczny obiektu sakralnego w Elblągu* [in Polish], Praca dyplomowa, Zakład Inżynierii Dźwięku P.G., Gdańsk 1987.

[50] A. Witkowski, *Badanie narastania dźwięku we wnętrzu Bazyliki Mariackiej w Gdańsku* [in Polish], XXXII Otwarte Seminarium z Akustyki OSA-85, Kraków 1985, pp. 31–34.

[51] A. Witkowski, *Badanie niejednorodności pola akustycznego we wnętrzu Bazyliki Mariackiej w Gdańsku* [in Polish], Mater. OSA, 1992, Kraków, pp. 167–170.

[52] A. Witkowski, *Model Bazyliki Mariackiej w Gdańsku* [in Polish], Zakład Inżynierii Dźwięku Instytutu Telekomunikacji P.G., Gdańsk 1984 [maszynopis].

[53] M. Zajączkowski, *Badanie akustyczne wielkiego wnętrza w skali 1:30* [in Polish], Praca dyplomowa, Zakład Inżynierii Dźwięku P.G., Gdańsk 1984.

[54] I. Żuchowicz, *The comparison of the reverberation time measured in halls with the use of natural and artificial sound sources*, Materials from the Conference on Measurements of Reverberation Phenomena in Halls, Warsaw, April, 1967, Archiwum Akustyki, **2**, 3, 201–207 (1968).

ARCHIVES OF ACOUSTICS
23, 1, 105–124 (1998)

PRODUCT-SOUND QUALITY: A NEW ASPECT OF MACHINERY NOISE

J. BLAUERT, U. JEKOSCH

Institut für Kommunikationsakustik
(Ruhr-Universität Bochum, D-44799 Bochum, Germany)

Quality is an issue of current attention with regard to product-sound design and assessment as it has now been widely recognised that the quality of the sound that a product makes is a significant component of the consumers' overall judgement on the product. This trivial fact has long been neglected by engineers and – much to their concern – they are now more and more pushed to take account of it in the product-development process. It is the aim of this article to provide the basis for a more differentiated view on product-sound quality than is currently common in the field. Special focus will be put on the process of perception and judgement in the context of quality assessment. It is hoped that engineers are encouraged to take on product-sound design and assessment as a generic engineering task.

1. Acoustic events and the concept of noise

"Noise" is usually defined as an unwanted acoustic event, in other words, an acoustical event that one would prefer to switch off or at least to modify – if this were possible. This simplified definition implies a basic feature of noise: To classify an acoustic event as noise, somebody must actually make the judgement "unwanted".

An acoustical event is mechanical oscillations and waves in elastic media. Consequently, it is a purely physical phenomenon. Since an acoustical event represents mechanical energy, it can apply forces and, accordingly, can have destructive effects on structures. Such effects are, for example, causing cracks in mechanical structures, degrading the function of instruments and/or causing damage to the hair cells of inner ears – such reducing the sensibility of the auditory system.

With respect to the mechanical effects of forces applied by an acoustic event, the decision of whether an acoustic event is unwanted or not, can be taken on physical or biophysical evidence. Threshold values can be found on the grounds of results of physical or biophysical measurement. Nevertheless, the judgement of experts such as mechanical engineers, acousticians, medical doctors, economists, lawyers and, ultimately, politicians is needed to establish thresholds as we find them in laws and other official regulations.

An obvious cure against mechanical, destructive effects of acoustic events is the following. Reduce the energy of the acoustical event, resp. the acoustic level. "Lower is better" – applied to the acoustic level – is indeed the most often applied recipe in noise engineering. It has the advantage of offering a straight forward, one-dimensional instruction for action, very much like the fundamental capitalistic demand to maximise profit. Yet, already in the domain of mechanical effects of acoustic events, which we are talking about at this time, one can find examples which show that the simple "linear" action of reducing the level is not always optimal. For example, when a high-Q resonance in a mechanical structure is excited by a narrow-band acoustic event, it is often much more effective to modify the tuning than to reduce the acoustic event level.

From the point of view of the effect of noise on humans it seems to be established that a permanent biophysical damage to the human ear will most likely be avoided when the level of a stationary acoustic event which the ear is exposed to for an extended period, stays below about 85 dB (A). It is, however, not the case that all acoustic problems with respect to humans are solved, once the level is sufficiently low such as to no longer be biophysically harmful to the inner ear. For this reason we shall now turn to the discussion of noise situations with levels well below 85 dB (A), in order to evaluate the question of what factors then govern the decision of whether an acoustic event is unwanted or not. We shall see that in these cases the essence of the decision is psychoacoustical and psychological rather than physical or biophysical.

2. Auditory events

Under normal circumstances a person being exposed to an audible acoustic event will hear "something". This something is, to be sure, not the (physical) acoustic event, but and outcome of auditory perception which we call auditory object or auditory event. Auditory objects, like any other objects (e.g. visual or tactile ones) have particular spatial and temporal attributes (they exist at a specific location at a specific time) as well as particular qualitative attrtibutes (in the case of auditory objects, for example, loudness, pitch and timbre). The decision of whether an acoustic event is noise is made by the listeners on their auditory perception, i.e. by judging on the characteristics of auditory objects.

This situation raises the question of whether listeners can dislike auditory events per se, i.e. by disliking specific auditory attributes or groups of attributes of the auditory events. There is no definite answer available from psychological theory, but these authors tend to say yes. The following hint may be used to support this view. Loud auditory events – like bright visual events – can be annoying in the sense of being straining to listen (look) to. Further, there seems to be a very basic kind of aesthetic quality of auditory events which makes them acoustic events pleasant to listen to, sometimes called sensory consonance.

Loudness and sensory consonance are quite basic auditory features, which require little or no abstraction. Yet, there are other, often much more important features of auditory events which prompt the listeners' judgements in terms of wanted or unwanted. As an example, consider the buzz of a mosquito circling your head when trying to find rest. How can such an auditory and, consequently, acoustic event be assessed?

3. The information aspect

Consider the physiological role of the auditory system: It is purpose of this system to gather up and process information from and about the environment, including tasks such as to identify acoustic-event sources with respect to their nature and their position and state of movement in space. Further in many species as in humans, interindividual (vocal) communication is performed through the auditory system.

With regard to auditory information gathering and processing, there are the following reasons to classify an acoustic event as noise, i.e. as being undesired: (a) An acoustic event carries information which is unwanted, undesired. (b) An acoustic event hinders the perception of desired information and is thus rated as being undesired.

Certain cars are liked by enthusiasts because of their characteristic engine acoustic events. The engine acoustic event "signals" information about the type of car. From a different point of view it is, nevertheless, noise. In more general terms, each acoustic event carries information about the acoustic-event source and its present state. The operator of a machine wants to hear the characteristics acoustic event of the machine in order to be sure about its proper functioning – and also to avoid accidents. Typical acoustic events may indicate, e.g., whether a device is switched on or off.

If an acoustic-event component has a double role of being signal and noise at the same time, depending on the point of view, a pure level reduction will have the following effect: The annoying effect of the acoustic event will be reduced and so the acoustic event will become less noisy. At the same time the function of the acoustic event as signal will be impaired and an acoustic-event situation may result which is rated to be even more undesired. It is obvious that noise problems of this kind must be very carefully evaluated. Instead of aiming at just reducing the noise level, although this is often a good thing to start with, the engineer must consider a (re)design of the complete acoustic situation – and also be aware of the particular non-acoustic environment.

When an acoustic event carries information this means that the acoustic event stands for something else, i.e. that it acts as a sign, a signal. This is, in other words, that a "meaning" is assigned to it. Actually, it is quite common that meanings are assigned to acoustic events. The most evident example is spoken language, which is actually a highly standardised code. To decipher the meaning of speech-acoustic events, the listener must know the code, i.e. the language.

In cases of acoustic events other than speech the meanings are less standardised and subject to individual interpretation. The way in which a listener interprets the meaning of response carried by an acoustic event will influence his judgement in terms of wanted – vs. unwantedness to a considerable degree. Yet, there may be inherent meaning in certain acoustic events which are understood by many people in basically the same way – obviously on a low level of abstraction. For example, particular acoustic events have been found which, when being used as warning signals, indicate such things as "Fire! Leave the house!" in contrast to "Gas alarm! Stay inside and shut the windows!"

Expectations are an important factor in this context. A coffee machine sounding like an electric razor would be rated bad although the sound might be perfectly o.k. for a razor. Obviously we expect the sound of a coffee machine to signal something very specific which is not given by the sound of a razor.

4. Cognitive factors

The interpretation of meanings assigned to auditory events requires involvement of higher processes of the central nervous system, known as cognition. A simple model of the process of judgement could be as follows. Listeners set up hypotheses with regard to the meanings of their perceived auditory events based, among other things, on their current cognitive situation. Then they test these hypotheses on the basis of all relevant information currently available to them. Besides information from the auditory modality, prior knowledge and information from other senses, e.g., vision, may be used, current actions and emotions may to be considered as well.

In the context of research on the effect of traffic noise on humans, a number of cognitive factors have been identified which play a role in the formation of hypotheses in listening and, consequently, affect both auditory perception and judgement. Some more important ones are listed in the following as examples.

Factors related to the sound sources

When judging on auditory events, listeners reflect on their specific conceptual image of the acoustic sources. For example, the auditory events stemming from acoustic sources considered dangerous tend to be rated significantly louder and noisier than those from sources considered harmless. Experimentally this has been shown by presenting pictures of different sources to subjects together with the same acoustic signal. The conceptual image of an acoustic source is of particular influence on the judgement of the auditory event attached to it, e.g., when the source behind it is a human being considered to have bad intentions, or a process considered dangerous or unhealthy. On the contrary, when acoustic sources are considered positive, such as strong, healthy, lively, beautiful, the auditory events attached to it tend to be rated lower and less disturbing.

Factors related to the situation

Depending on the situative context, auditory events are rated differently. For example, in a situation perceived as comfortable, auditory events are judged upon in a different way than in an uncomfortable situation. Aesthetics must be taken in consideration also. Of special importance is the particular state of activity (action) of the listeners, e.g., working, relaxing, sleeping. It may be required for good quality that the acoustic events change in the context of a listeners activity, e.g., to signal that an operator's commands are executed. The specific role of oral/auditory communication has to be taken into account in particular. Acoustic events which interfere here, are especially disturbing. Specific situations and activities require specific acoustical conditions.

Factor related to individuals

Personal factors, i.e. such which are related to individual listeners, are of special importance. They usually represent the dominant reason for interindividual variance and, consequently, increase the degree of subjectivity of the judgements. The understanding of personal factors is, among other things, relevant for the formation of panels of listeners which can be considered representative for the target population for certain products. Expectation is an important factor here, others are motivation, attitude, taste, economical situation, experience.

5. Auditory perception and measurement

Judgements by listeners are commonly called "subjective". Such a usage of the term "subjective" results from a lack of differentiation, however. Adequate psychometric methods may well render results which are largely independent of individual listeners and thus "objective". Hence, in psychoacoustics the term "objectivity" is preferably understood as follows. Objectivity is given when the statistical distributions of judgements are the same for each of the subjects and, insofar, independent of any individual subject. Complete objectivity is not to be expected from listening tests. In praxi, results are achieved which are more or less subjective or objective. The degree of subjectivity increases the more individual characteristics of subjects influence their individual perception and judgement.

Instrumental physics uses a very restricted definition of measurement as follows: Measurement is the quantitative assessment of the ratio of a physical quantity with respect to a reference quantity of identical dimension. A more general definition of measurement comprises any assignment of numbers to objects in such a way that relations between the numbers reflect relations between the objects. Psychoacoustic judgements can be subsumed under such a general concept of measurement, provided that the judgements map relations between attributes of percepts in a quantitative way. In experimental procedures which render this kind of results, the subjects act as a measurement device to measure their own percepts.

It is clear at this point that measurement is not restricted to instrumental physical techniques, but that it is also possible with psychological and psychophysical procedures, i.e. with the involvement of subjects. Yet, the results of tests with subjects tend to show higher variance than usually encountered with instrumental measurements.

Engineers are educated in a way which has its roots in more-than-a-hundered-year-old conceptional approaches of applied physics. Consequently, they often have an over-simplified idea of objectivity. They think that any experiment within their science is designed in such a way that the observers do not take part in the experiment in principle and that, for this reason, the process and the results of the experiment are completely independent from observers.

Engineers, in this context, are thus used to assume that only such data have any relevance in reality which have been acquired independently of any observer. A further specific feature of the conservative engineering approach is that more complex phenomena are usually dealt with in an analytic way, i.e. with the attempt of decomposing them into, preferably orthogonal, components. The idea behind this is the assumption that the original phenomena can completely be resynthesised from these components. Such a method of analysis and resynthesis has indeed been successful in many cases, see for example the Fourier analysis/synthesis in signal processing. It is, however, inadequate for problems where observers are intrinsic to the experimental procedure, as it obviously the case in psychoacoustic assessment.

There is no way out than recognising the observer as an integral part of the experiment and to accept that "absolute" objectivity cannot be achieved. Further, it has to be accounted for that observers cannot only influence the progress of the experiment, but also go through changes themselves in the course of the experiment. Observers may, for example, enhance their ability to detect certain perceptual attributes, or they may change their judgement criteria.

When dealing with problems of perception and judgement the complete system comprising the perceived object as well as the perceiving and judging observer must be dealt with an integrated way. Any concern that this would require to leave the area of rational scientific argumentation is not justified, simply because the brain as the organ of perception and judgement is a purely biological organ after all.

6. Product-sound quality: A definition

So far, in this article, we have dealt with the fact that an acoustic event can be assigned the attribute "noise", based on human judgements of negative physical effects of the acoustical events or judgements on the undesirability of the auditory events attached to the acoustic events under consideration. By calling an acoustic event a "noise", we actually put it in a certain "quality" category. In fact, the concept of "quality" has recently gained much broader relevance in acoustics. In the following, we introduce a modern definition of sound quality as appropriate in the context of industrial products, and discuss some major implications for sound-quality evaluation following from this definition.

In the following the familiar term "sound", which is often used equivocally for both the acoustic events and the auditory events attached to it, will be used for the auditory events – and only for those. In pursuit of a concept of the 2nd author product-sound quality can be defined as follows, whereby the term "product" may be associated with specific samples of a product as well as with a class of products:

"Product-sound quality is a descriptor of the adequacy of the sound attached to a product. It results from judgements upon the totality of auditory characteristics of the said sound – the judgements being performed with reference to the set of those desired features of the product which are apparent to the users in their actual cognitive, actional and emotional situation."

The prerequisite for sound quality to happen is obviously the existence of a product which emits acoustic waves, such as to initiate a percept of sound to be judged upon in terms of quality. Upon hearing something which obviously stems from the product considered, the listeners take on the task of judging upon their auditory events. They judge by using a frame of reference which, in the definition above, is called "set of desired features of the product". It goes without saying that the listeners will only judge with reference to desired features which are apparent, and that this process does evidently depend on the cognitive, actional and emotional situation given.

The involvement of cognition is easily envisaged. The listeners use to have prior knowledge of the individual product of sample or class of products under consideration. It is assumed that they judge by comparing the attributes of their auditory events with the set of desired features of the product which they have in mind. As a result of this process they finally come up with a judgement, which then, in turn, is assigned to the product as being its sound quality. Sound quality, hence, is not an inherent property of the product, but rather something which develops when listeners are auditorily exposed to the product and judge on it with respect to their desires and/or expectations in a given situational context.

Following the definition of sound quality presented above, the assignment of sound-quality to auditory events requires judgements – the judgement being performed with reference to a concept of desired features of the product. In other words, the idea is that the listeners compare their auditory events with an internal frame of reference of theirs, and thereupon issue their quality statements – be it global statements like values on a scale bad-to-good or more differentiated statement following analyses with respect to distinct quality features.

If the frame of reference would just require pleasantness (sensory consonance) of a sound as the desired quality feature, basic psychoacoustic quantities might suffice to establish a quality statement. Yet, sound-quality engineers dealing with the sound quality of industrial products will rarely get off that easily. The user and/or prospective user of a product comes up with concepts of desired features which are quite complex. Among other things, the scope of desired features is not restricted to features of the sound only, but to other features of the product as well. From an industrial product the user requires flawless and efficient functioning in the first place, but many additional items like the aesthetic appearance and the product image in a social context play an important role as well. Pleasantness of the product sound is, in fact, not very high-ranking in the hierarchy of desired features.

We have already mentioned that sounds are carriers of information, e.g., they are signals and signs of activities in the environment. In the light of this statement it becomes clear that the sound of a product has a particularly important function, that is to say, it informs the user about the presence of a product and about its state of operation in space and time. A product may, for example, indicate to the user, among other things: I am the electric razor in your hand, and I am switched on. My blades are cutting the hairs of your chin properly, but I may need a battery recharge soon.

To be able to decipher these messages of the razor, the users have to know the "language" of electric razors. They must have an idea of how a good razor sounds when functioning properly, and be able to discriminate fine shades of changes in the sound when the battery is low. They should possibly also be capable of recognising the brand of a razor by its sound. The important point here is that the customers are not only interested in the product sound per se, but that the product sound is a carrier of information for them. They certainly would prefer a pleasant sound to an unpleasant one, but, even more so, they want the "sound of quality".

It is at this point that it becomes evident why the frame of reference which customers are referring to when judging upon the quality of a sound, comprises desired auditory features as well as many kinds of desired non-auditory ones. The sound can meet the demands for auditory quality features directly. Yet, as far as non-auditory quality features of the product are concerned, the product sound can, at best, provide an indication as to which amount these may be met. To be sure, good product-sound engineers are intuitively aware of a lot of general features which high-quality product sounds should have.

7. Approaching sound-quality assessment systematically

From the point of view of product-sound engineering it is of predominant importance to analyse the process of sound-quality assignment systematically. Questions to ask are as follows: What are the constituents at each step of the development of the sound-quality assigned to a product, and how can these be evaluated and engineered in order to arrive at a predetermined product-sound quality.

Some helpful facts to be mentioned in this context are the following: There are methods available to evaluate the relationships between parameters of acoustic events and attributes of auditory events, as already mentioned above, the so-called psychometric methods. It is reasonable to make the following comment at this point: Once the relationships between acoustics parameters and auditory attributes have been established in a reliable, valid and representative way, computer algorithms can be conceived and implemented which render estimators for psychoacoustics quantities by purely instrumental methods.

Yet, it has to be kept in mind that common "peripheral" psychoacoustics, although providing basic knowledge on auditory perception, is not at all sufficient to cover the complex tasks of product-sound quality evaluation and assessment. Three important facts have to be kept in mind additionally to arrive at a more reasonable model for sound-quality evaluation: (a) The acoustic waves have a distinct source, namely, the product under consideration. (b) The process of psychoacoustic perception is augmented by a process of judgement. (c) Both perception and judgement are modified by factors originating from cognition, but also from actions, emotions, and input from non-auditory senses.

Item (a) makes clear that the origin of the acoustic event is the ultimate object of concern here, that is to say, the sound emitting product. Item (b) is introduced to denote that common psychometric measurement does not suffice to assess product-sound quality. The concept of judgement may include processes akin to psychoacoustic measurement, yet, goes far beyond it. It comprises sophisticated global as well as analytic evaluation and assessment, including complex balancing and weighing with regard to issues such as information from non-auditory senses, prior knowledge and association to actions and emotions. It is thus appropriate to regard the process of judgement as a matter of psychology rather than of just psychoacoustics. Item (c) complements the process of judgement by explicitly introducing input from non-auditory modalities, cognition, action and emotion as response-moderating factors. Note that response-moderating factors do not only influence the judgement process, but the process of auditory perception as well. Among other things, this is to make clear that already the output of auditory perception is not at all predetermined solely by the acoustic input to the auditory system, but results from a complex interaction of auditory input, non-auditory input, expectation and mood.

8. Conclusions

Product-sound engineering is certainly not an easy task, and even worse, there is no shortcut to it. Among other things it requires sufficient knowledge and experience at least in the following areas: (1) acoustics, especially related to mechanical engineering, (2) psychoacoustics, i.e. basic relationships of the acoustic input and the auditory output of perception; (3) psychologic effects and rules which govern the judgement processes through which quality statements are formed.

Further, to be able to actually engineer product sounds, product-sound engineers need potent tools, that is tools which enable them to specify design goals and, consequently, shape the product sounds in accordance to these design goals. Obviously, economical constraints have to be met during this process, as is typical for any engineering task.

Whithout doubt, the evaluation activities in the course of product-engineering start at the acoustic end of the quality cycle. It has first to be established what the contributions of the different mechanical parts of the product to its overall acoustic emission. Since the contribution add up linearly in most cases (at least with sufficient approximation), analytic methods of physical acoustics and the usual equipment for measurement of acoustic waves and vibrations can be applied to determine the individual contributions in terms of parameters of sound waves. In such a way, the acoustic input to auditory perception can be evaluated and specified.

Following purely-acoustic evaluation, the next step in the sound-quality cycle asks for an evaluation of the auditory events, i.e. of the "product sound" per se. To this end, product-sound engineers have to perform psychoacoustic experiments themselves, or to refer to the results of psychoacoustic research related to their specific engineering problems – if available. It is of paramount importance at this point to keep in mind that in most psychoacoustic research "unbiased" psychoacoustic measurement is assumed. Consequently, the results can be applied to practical problems only with great care. For some psychoacoustic quantities, for example, loudness and sharpness, modern equipment is available to mimic the process of psychoacoustic measurement electronically. As an output, these instruments render estimates of "unbiased" psychoacoustic quantities – not the psychoacoustic quantities themselves. It is important to be aware of this distinction, especially when the equipment has been standardised. Auditory perception is an individual process influenced by cognition, action and emotion, and is thus hard to be standardised in principle.

Any endeavour to evaluate product sounds by psychoacoustics only will fail in the end. To really understand and control the process of product-sound assessment specific psychologic knowledge and profound experience in the application of research methods as used in psychology are needed. It is at this point were deficiencies in the education of engineers dealing with product-sound design often become apparent. Further, quite frequently a certain shyness of engineers can be observed when it comes to psychological problems.

With regard to the evaluation of the psychology of auditory perception and judgement, the following argument is noteworthy: In the course of the perception and judgement processes a considerable reduction of information takes place. The high number of parameters necessary to represent the acoustical events at the ears of listeners is reduced to a handful of features (usually less than four) which the listeners actually refer to when coming up with a sound-quality statement. Special segmentation processes and "Gestalt" phenomena in perception as well as controllable and uncontrollable selection processes in judgement are to be evaluated and understood as well as the nature and development of the frames of reference which the listeners actually use when judging upon product sounds. As these psychological processes show profound interindividual variance, one of the important tasks of product-sound engineering is to select representative listeners for evaluation procedures – a complex task in itself. Methods for proper selection of subjects are part of the work bench of product-sound engineers. Representative listeners are not necessarily expert listeners!

Once a quality statement on a product sound is available, it will often be a challenge to improve the product sound quality – preferably but not necessarily in turn with an improvement of the product quality in toto. Improving the sound-quality means modifying those parts of the product which produce the sound. As information reduction has taken place in the processes of auditory perception and judgement, and as these processes are non-linear, time-variant and loaded with a huge amount of memory, it is not a trivial task at all to trace the quality cycle backwards. In other words, it is hard to solve this inverse problem, namely, to evaluate which attributes of the auditory events, or – going back even further along the quality cycle – which parts of the product play a role in the formation of its sound quality, and to which quality features are they finally related.

9. References

This article is based on ideas which the authors have presented at an EAA TUTORIUM which the first author had organized in Antwerp at the occasion of the FORUM-ACUSTICUM-1996 convention – and on earlier publications. The journal of the European Acoustics Association, ACUSTICA united with ACTA ACUSTICA has recently published a special issue on Product-Sound Quality (**83**, 5, September/October 1997). This issue comprises the papers presented at the EAA TUTORIUM mentioned above and contains a variety of references to the relevant literature.

ARCHIVES OF ACOUSTICS
23, 1, 115–124 (1998)

ADAPTIVE HYBRID METHODS FOR IMPROVING THE QUALITY OF ULTRASONIC IMAGES OF SKIN

M. LEWANDOWSKI, A. NOWICKI

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

This study is concerned with methods for improving the quality of micro-ultrasonographic images using digital signal processing. A method of linear adaptive filtering is presented, which makes it possible to reduce speckle in ultrasonographic images, and its non-linear modification is proposed. Quantitative results of one-dimensional simulation of these filters is proposed and the results for real micro-ultrasonographic images of skin are presented.

1. Introduction

Crawford *et al.* [1] and Bamber *et al.* [2] presented the application of the LLMMSE (Local Linear Minimum Square Error) filter for improving the quality of two-dimensional ultrasonographic images. The presented algorithm belongs to a group of adaptive methods for reducing speckle in ultrasonographic images and is an alternative to the widespread methods for averaging a line in an image or whole images. The idea of the method lies in adaptive control of the degree of smoothing, depending on the local statistical properties of the image. The need for adapting the filter to local properties of the image results directly from the assumption that the signal is nonstationary. The local statistics of the image provide information about the similarity of a given area to speckle noise and, thereby, it determines whether this area should be smoothed with a low-pass filter or left intact. The operation of the filter described depends on two parameters, characteristic for the type of tissue under study. These are the gray level variance for areas with speckle characteristics and for those that contain biological structures. These parameters are able to distinguish between (classifies) speckle areas and biological structures.

Bleck *et al.* [3] studied the variability of the parameter in question which classifies speckle areas, depending on the measuring conditions and for different pathological changes in the kidneys under examination. These studies confirmed the usefulness of the adaptive speckle reduction method, also in the case of pathological changes causing the emerging of structural inhomogeneities (quasiperiodic structures) in tissue.

The subjective physicians' perception of ultrasonographic images filtered using the adaptive LLMMSE method was also studied. CRAWFORD *et al.* [4] demonstrated that almost in every other case the structural information in the images filtered was evaluated as better than in the original images. Better legibility of images processed in this way, without at the same time losing anatomical details, was also confirmed.

2. The filtering methods applied

This section presents the filter algorithms applied for improving the quality of micro-ultrasonographic images.

The quality of proposed filter methods was determined by simulation of one-dimensional signals. These methods were also integrated within the software of the micro-ultrasonographic device, permitting their usefulness to be checked in real clinical conditions.

2.1. LLMMSE type adaptive filtering

The original scheme of the LLMMSE method applied for speckle reduction in ultrasonographic images proceeds as follows:

• The input image is scanned by a filter window. Square windows with the dimensions 3×3, 5×5 and 9×9 pixels are most often used. Figure 1 shows the scheme of scanning an image with a square window with the dimensions 3×3 pixels. It can be seen in the figure that the central point does not reach the very edge of the image, causing the output image to be smaller than the original.

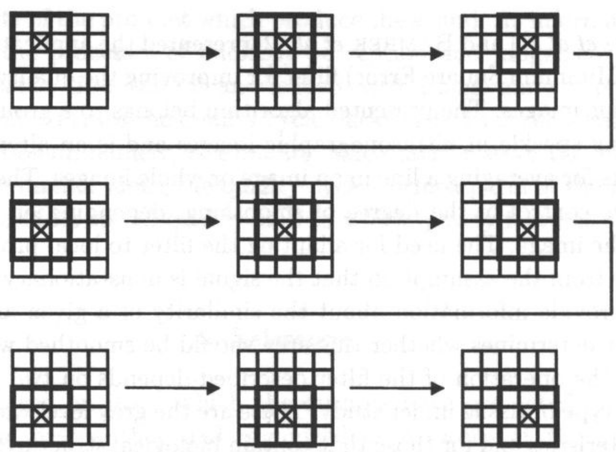


Fig. 1. An example of scanning an image with 5×5 points using a filter with the dimensions 3×3. (The central point of the window is marked with a cross).

• For every position of the window the local statistics are calculated (the grey level statistics of the pixels of the original image contained in the window): the mean $\langle x \rangle$ and the variance f. On the basis of these parameters and the parameters $f_{\text{structure}}$ and f_{speckle}

(determined arbitrarily – describing the statistical properties of the areas containing biological structures and speckle noise, respectively), the coefficient k is calculated:

$$k = \frac{(f - f_{\text{speckle}})}{(f_{\text{structure}} - f_{\text{speckle}})}, \tag{1}$$

where f is the calculated value of the variance in the window.

It is then restricted to the interval $[0, 1]$, i.e., $k = 1$ when $k > 1$ and $k = 0$ when $k < 0$.

- Finally, the output value of the filter is calculated in the following way (Fig. 2):

$$x_{\text{out}} = \langle x \rangle + k \cdot (x_{\text{in}} - \langle x \rangle), \tag{2}$$

where x_{in} is the grey level of the input image which is at the central point of the filter window, $\langle x \rangle$ is the calculated mean grey level in the window and k is the calculated weight coefficient (cf. Eq. (1)).

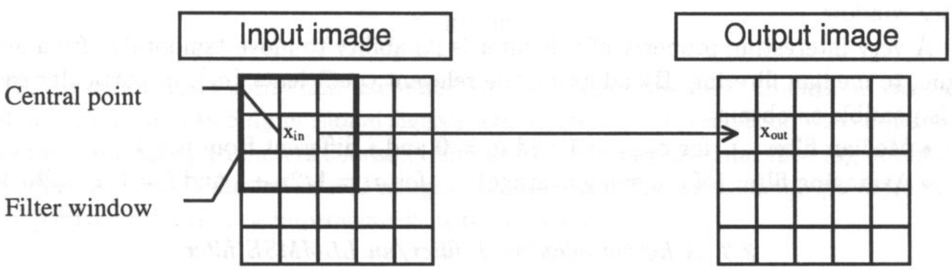


Fig. 2. The processing scheme for a single 3 × 3 window, (see Eq. (2)).

The parameters $f_{\text{structure}}$ and f_{speckle} are determined from analysis of the areas containing only biological structures and speckles, respectively, for a given type of tissue under examination. These coefficients are a criterion of similarity for the local properties of the image, and determine, through the coefficient k, the degree of smoothing of the local area BAMBER *et al.* [2].

The above algorithm makes use of the variance/mean ratio to identify the noise areas (speckle areas), and it is based on the assumption that a linear dependence is present between the variance and the mean. As was demonstrated by CRAWFORD *et al.* [1], the above assumption requires linear or logarithmic characteristics of the input circuit of the ultrasonograph. These authors examined the input characteristics of several universally used ultrasonographs and proposed methods for correcting those characteristics that did not satisfy the assumptions of LLMMSE filtering.

2.2. L-filters

L-filters can be considered a modification of a linear finite impulse response (FIR) filter. After windowing of the input data the proper ordering of samples is done and next FIR polynomial is constructed (Fig. 3).

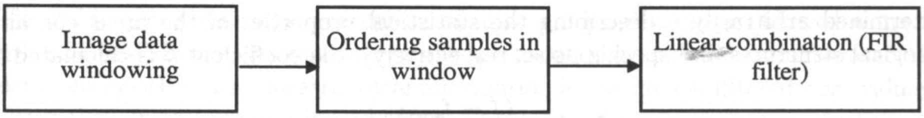


Fig. 3. A schematic representation of the working of an L-filter.

L-filters can be also considered as a generalisation of a median filter using all the signal samples instead of only the median. A basic advantage of filters of this type is their ability to cope with different types of noise.

The linear combination of FIR filtering is conducted on the ordered sequence $\{x_i\}$ of input signal samples.

$$y = \sum_{j=1}^{2n+1} a_j \cdot x_j \,, \tag{3}$$

where a_j are constant coefficients of the linear FIR filter and $2n+1$ is the length of the filter window.

A very interesting property of this filter is its ability to move "smoothly" from averaging to median filtering. By adopting the relevant coefficients $\{a_i\}$, in particular cases it is possible to obtain:

- Median filter for $a_{n+1} = 1$ and $a_i = 0$ and i different from $n+1$.
- Averaging filter (of a moving average) for $a_i = 1/2n+1$ and $i = 1, 2, ..., 2n+1$.

2.3. A hybrid adaptive L-filter/an LLMMSE filter

Since an LLMMSE type filter failed to cope with pulsed noise [5, 6], a hybrid filter was proposed, which combined the advantages of a linear and median filters. The LLMMSE algorithm was preceded by ordering signal samples (according to their values), just as was the case with L-filters. Such an approach permitted optimum filtering of nonstationary noise (a property of an LLMMSE filter) while maintaing edges within the image (a property of a median filter).

Schematically, the hybrid algorithm is as follows:

- The input image is scanned using the filter window – analogously to the case of the LLMMSE algorithm.
- In the second stage, the samples from the filter window are ordered according to the grey level of pixel. In the case of a classic median filter, a point in the output image would be the median from the sequence ordered in this way.
- For every window, the local statistics are calculated: the mean $\langle x \rangle$ and the variance f – as in the LLMMSE algorithm (sample ordering is insignificant for the parameters calculated in this step).
- Finally, the output value of the filter is calculated, just as for the LLMMSE filter, except that the mean value $\langle x \rangle$ as in formula (2) is replaced by the median x_{med} (calculated in the filter window)

$$x_{\text{out}} + k \cdot (x_{\text{in}} - x_{\text{med}}). \tag{4}$$

2.4. A modified hybrid filter

The hybrid filter presented in the previous section takes over the ability to keep details (the edge) from the median filter. It is known that most details preserved by the median filter depend on the size of the window. The process of ordering samples in the window can also be understood as a loss of information about the position of samples (in terms of time or space). In the case of ultrasonographic images in B-mode presentation, both time information (along the wave propagation axis) and space information (along the head movement direction) is lost.

The use of a weighted median is a method for preserving information about the position of samples. The weighted median filtering is obtained by a duplication of pixel values in the filter window before ordering them. As a result, it is possible to ensure a larger number of representative samples situated in the input window at a predetermined position (most frequently, the central one).

3. Results

To determine the quality and usefulness of the filters presented, simulation was carried out of one- and two-dimensional signals. One-dimensional simulation was conducted to obtain objective parameters characterizing the quality of particular filtering methods. Two-dimensional simulation of real images obtained from the ultrasonograph permitted an objective evaluation of improvement in the image quality.

3.1. One-dimensional simulation

To determine the properties of the filters presented, a simulation was performed for noise-perturbed one-dimensional signals. In the simulation, a section of a real ultrasonographic signal was used (Fig. 4). Noise with a Rayleigh distribution and pulsed noise were added to the input signal. The noise with a Rayleigh distribution in the signal of line A was a result of demodulation of a high-frequency ultrasonographic signal containing Gaussian noise. Pulsed noise was added to provide for simulation of the lack of signal or saturation occurring in the course of digital-to-analog conversion, resulting from the non-ideal nature of the clock synchronizing the processing.

The mean square signal-to-noise ratio $\mathrm{SNR_{MS}}$ was determined as a measure of signal improvement. The value of this parameter was expressed in dB, calculating

$$\mathrm{SNT_{MS}}(s,x) = 10\cdot\mathrm{LOG}\left(\frac{\sum\limits_{i=1}^{N}(x_i)^2}{\sum\limits_{i=1}^{N}(s_i-x_i)^2}\right)\ [\mathrm{dB}], \tag{5}$$

where $\{s_i\}$ is the sequence of samples of the original signal, $\{x_i\}$ is the sequence of samples of the noise-perturbed signal before (after) filtering, and N is the length of the sequence of signal samples.

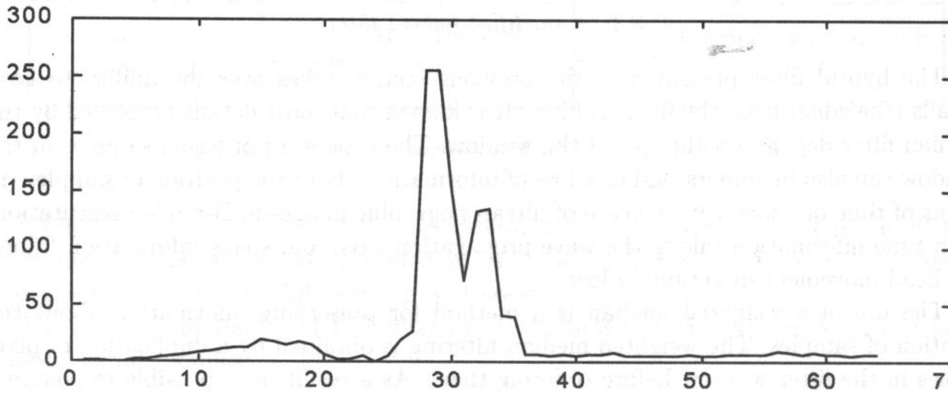


Fig. 4. A section of the signal of line A, used in one-dimensional simulation.

The parameter defined in this way was calculated for a noise-perturbed signal (i.e., before filtering) and then for the signal after filtering. An increase in the value of the ratio SNR_{MS} means signal improvement.

The results obtained for different values of the power (variance) of the added Rayleigh noise and different probabilities of pulsed noise always show the same tendency, i.e., a positive edge of the hybrid filter over the simple LLMMSE algorithm. The diagram below shows the simulation results for the variances of Rayleigh noise equal to 10, 20 and 30 and for the filter window with the length of 9. For all the results shown below, the probability of pulsed noise was 0.01.

Table 1. The results of one-dimensional simulation.

Values of the coefficient SNRMS in [dB]			
Input signal	Averaging filter	LLMMSE filter	Hybrid filter
12.86	3.15	13.84	15.54
8.03	3.41	9.63	10.99
5.77	3.43	7.27	8.17

Table 1 shows the values of the calculated coefficient SNR_{MS}, corresponding to the diagram in Fig. 5. The result of the work of the simplest averaging filter is also given as a reference. It can be seen that in this case the averaging filter very distinctly worsens the signal to noise ratio. This behaviour may be explained with two factors. Firstly, the added noise was Rayleigh rather than Gaussian noise; the averaging filter copes well with the latter. Secondly, the signal being filtered was a fast changing (high-frequency) one compared with the low cut-off frequency of the averaging filter (which was a low-pass filter) related to the length of the filter window. In the case of fast changing signals, the distortions caused by the averaging filter for an excessively long filter window may drastically worsen the signal to noise ratio.

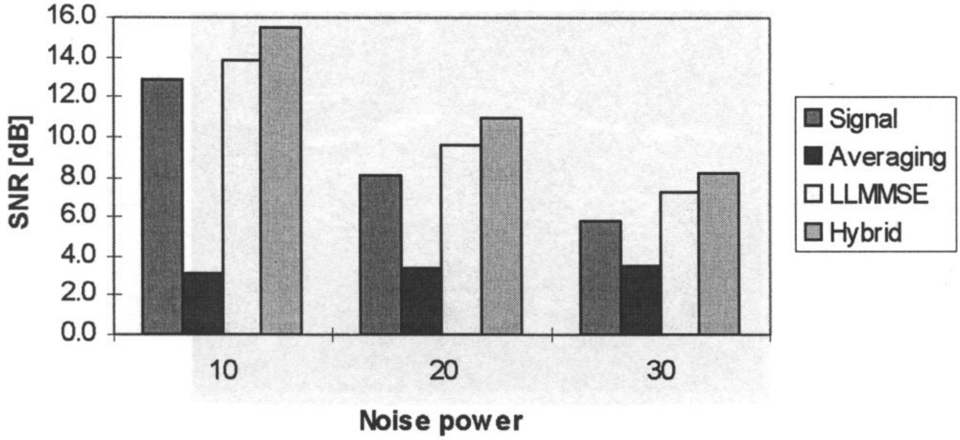


Fig. 5. The results of simulation for three different noise variances.

3.2. Two-dimensional simulation

The described filtering algorithms were applied to the series of skin images recorded from healthy volunteers and from patients with different skin melanoma e.g. superficial spreading melanoma, nodular melanoma or lentigo malignant melanoma.

The example cases are presented in the following way; the top picture is an original non-filtered image, middle one shows the same pattern after applying LLMMSE filtering and the bottom one shows the resulting image after hybrid filtering.

These photographs distinctly show the effect of the work of the filters. Both filters reduce the grain content in the image. The effect of the hybrid filter is more conspicuous, since it removes details with sizes smaller than the window length (classified as pulsed noise). The LLMMSE filtered image contains these small disturbances, therefore, it appears to be slightly sharper than the image following the hybrid filter. Both examples (Figs. 6 and 7) prove the positive impact of filters can be seen, permitting the boundaries of pathological change areas to be determined.

4. Conclusion

In micro-ultrasonographic images there are the same types of noise as in ultrasonographic ones. As an effect of this, the use of the adaptive method for speckle reduction, applied in conventional ultrasonography, brought a positive result in the case of micro-ultrasonographic images. The proposed nonlinear modification of LLMMSE filtering made it possible to improve further the signal-to-noise ratio. As the simulation on real skin and eye images indicated, both methods ensure subjective improvement in the quality of the images obtained.

a)

b)

c)

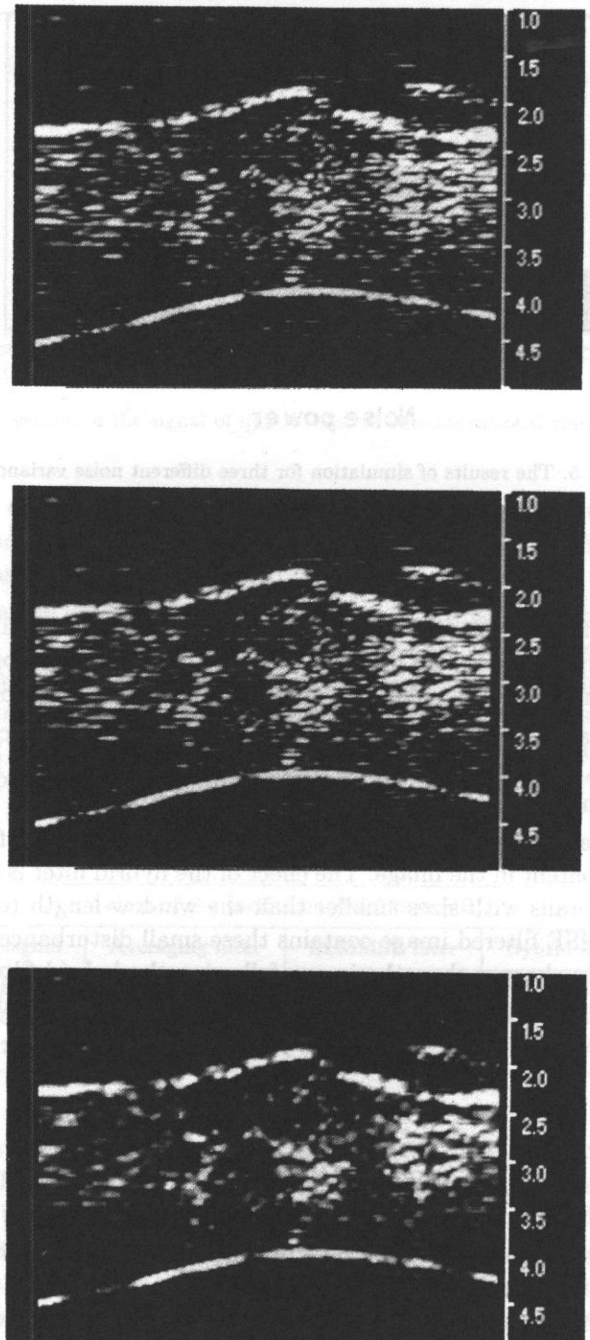


Fig. 6. Sector scan of superficial spreading melanoma, a) original image, b) after LLMMSE filtering, c) after hybrid filtering.

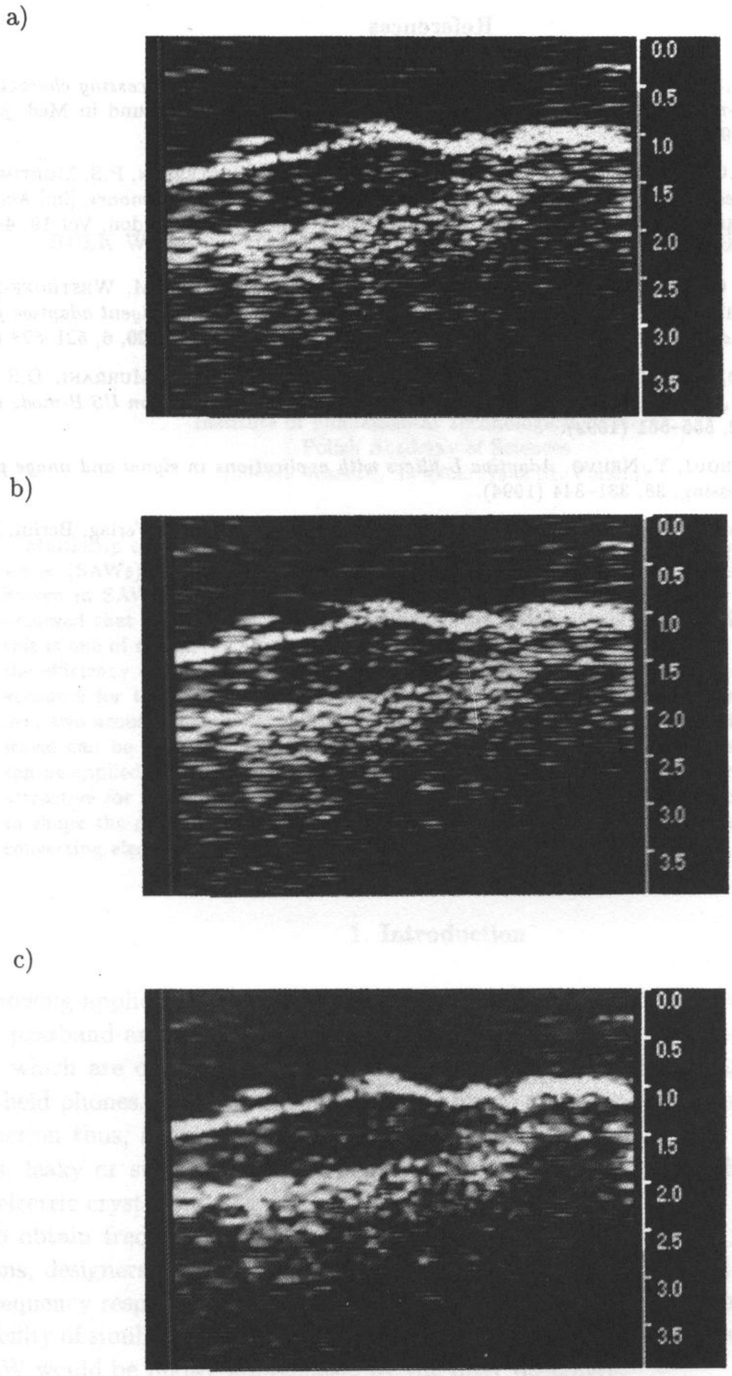


Fig. 7. Sector scan of nodular melanoma, a) original image, b) after LLMMSE filtering, c) after hybrid filtering.

References

[1] D.C. Crawford, D.S. Bell, J.C. Bamber, *Compensation for the signal processing characteristics of ultrasound B-mode Scanners in adaptive speckle reduction*, [in:] Ultrasound in Med. & Biol., Vol. 19, No. 6, 469–485, 1993.

[2] J.C. Bamber, D.C. Crawford, D.A. Bell, C.C. Harland, B.A. Gusterson, P.S. Mortimer, *Effects of speckle reduction processing on ultrasound B-mode images of skin tumours*, [in:] Acoustical Imaging, H. Ermert H., H. Marjes [Eds.], Plenum Press, New York & London, Vol. 19, 447–452, 1992.

[3] S.J. Bleck, M. Gebel, M. Hebel, S. Wagner, K. Schmidt, S. Kruip, M. Westhoff-Bleck, M. Wolf, C. Thiesemann, M. Manns, *Tissue characterization using intelligent adaptive filter in the diagnosis of diffuse and focal liver disease*, Ultrasound in Med. & Biol., **20**, 6, 521–528 (1994).

[4] D.C. Crawford, D.O. Cosgrove, E. Tohno, W.E. Svensson, B. Al-Murrani, D.S. Bell, K. Stepniewska, J.F. Bamber, *Visual impact of adaptive speckle reduction on US B-mode images*, Radiology, **183**, 2, 555–561 (1992).

[5] T. Sun, M. Gabbouj, Y. Neuvo, *Adaptive L-filters with applications in signal and image processing*, Signal Processing, 38, 331–344 (1994).

[6] T.S. Huang [Eed.], *Two-dimensional digital signal processing II*, Springer-Verlag, Berlin, Heidelberg, New York 1981.

ARCHIVES OF ACOUSTICS
23, 1, 125–137 (1998)

BULK WAVE TRANSMISSION BY A MULTISTRIP COUPLER (MSC)

E. DANICKI

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

Multistrip coupler (MSC) is a component of SAW devices that couples surface acoustic waves (SAWs) propagating in two acoustic channels on piezoelectric substrate. Little is known in SAW literature about simultaneous transmission of bulk waves by MSC; it is believed that MSC efficiently suppress bulk wave spurious signals in SAW devices, and this is one of reasons of application of MSC in SAW devices. In this paper we investigate the efficiency of bulk wave transmission by MSC. New theory of MSC is developed that accounts for the fact that MSCs comprise finite number of conducting strips spanning over two acoustic channels. Numerical analysis based on FFT algorithm shows that bulk waves can be quite efficiently transmitted by MSC. This means particularly, that MSC can be applied to couple the surface skimming bulk waves which, having large velocity, are attractive for higher frequency SAW devices. In this case, MSCs would make it possible to shape the device frequency response by two apodized interdigital transducers (IDTs) converting electric signal to ultrasonic surface waves.

1. Introduction

Growing applications of cellular telephones and mobile communication require still wider passband and higher frequency of operation of surface acoustic wave (SAW) devices, which are commonly applied for signal filtration, both in transponders, and in hand-held phones. To keep their price low, optical photolitography is prefered for their production thus, in order to get higher frequency SAW filters, one must apply faster SAWs, leaky or surface skimming bulk waves (SSBW) which are the fastest in given piezoelectric crystal substrate.

To obtain frequency characteristic of SAW devices required in modern electronic systems, designers frequently apply multistrip couplers that make it possible to shape the frequency response of the device by two apodized interdigital transducers (IDTs). A possibility of similar construction of SAW filters exploiting SSBW or leaky waves instead of SAW would be highly appreciated by the filter designers.

The theory of MSC, developed and applied in SAW literature is the coupled modes theory [1], in plane wave approximation (which is also valid in this paper because of the assumed large aperture width of acoustic channels with respect to a wavelength of bulk waves λ). Suppose the electrodes (thin metal metal strips) of MSC cover two identical acoustic channels (Fig. 1). There are two modes in such waveguide:

— symmetric mode where SAWs in both channels have equal amplitudes and phases,

— antisymmetric mode where SAWs in both channels, having equal amplitudes, have opposite phases.

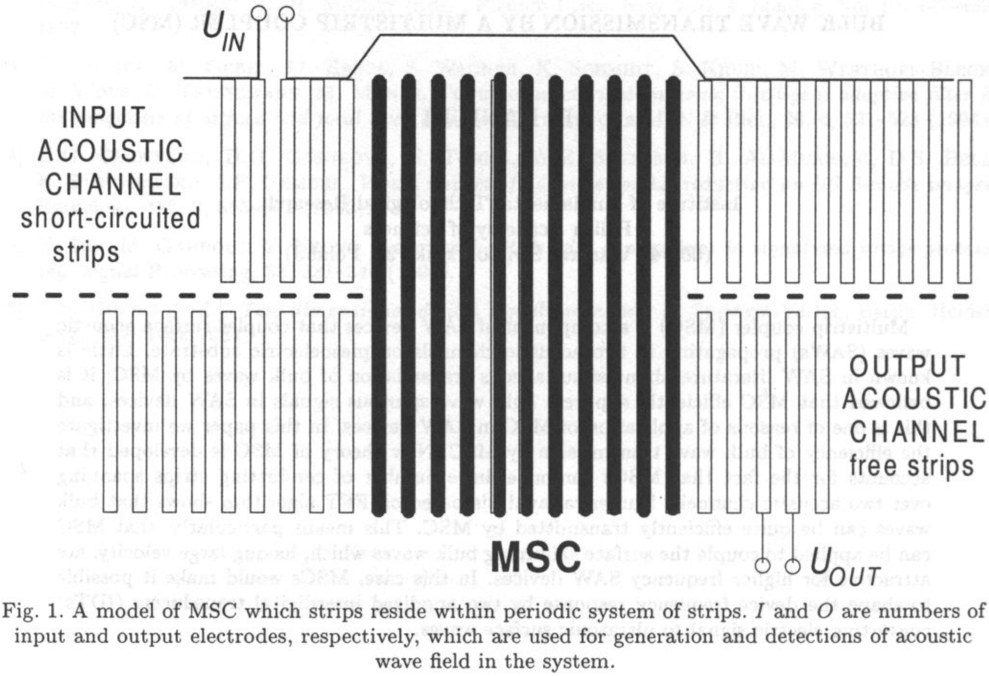


Fig. 1. A model of MSC which strips reside within periodic system of strips. I and O are numbers of input and output electrodes, respectively, which are used for generation and detections of acoustic wave field in the system.

In symmetric mode, the strip voltages accompanying the SAWs in both channels are equal, and there are no currents flowing in strips between acoustic channels. In this circumstance, the SAW in both channels propagate exactly as under free, isolated strips. And in antisymmetric mode, the strip voltages would be opposite if the corresponding strips are not connected between channels. But they are connected, thus the strip voltages must be zero and, as opposite to the previously considered case, a current flows in each strip between the channels. This means that, in both channels, the SAWs propagate in condition equivalent to short circuited, or grounded strips.

The theory of propagation of SAW under isolated, and grounded periodic strips is well developed and we can easily evaluate the SAW wave numbers in both cases which are r_v and r_o, correspondingly. These wave numbers depend on piezoelectric substrate parameters, on strip width w with respect to their period Λ, and on strip period with respect to the wavelength λ. Typically in MSCs, there are more than two strips falling in the shortest wavelength in the entire frequency band of operation of the SAW device. This is to avoid the Bragg scattering of SAWs into bulk waves. This will be assumed also in this paper, $K = 2\pi/\Lambda > 2r_o$.

The acoustic field in MSC is (Fig. 1)

$$\begin{aligned} &-\text{in upper channel} \quad A(x) = a_v e^{-jr_v x} + a_o e^{-jr_o x}, \\ &-\text{in lower channel} \quad A(x) = a_v e^{-jr_v x} - a_o e^{-jr_o x}, \end{aligned} \tag{1.1}$$

where a_v and a_o are amplitudes of SAWs having wave numbers r_v and r_o, correspondingly. Suppose that, at $x = 0$, $A(x) = 1$ in the upper channel, and $A(x) = 0$ in the lower channel, thus it must be $a_v = a_o = 1/2$ and, at $x > 0$, the SAW amplitudes are

$$\begin{aligned} |A(x)| &= \cos\frac{r_o - r_v}{2}x, \quad \text{in upper channel,} \\ |A(x)| &= \sin\frac{r_o - r_v}{2}x, \quad \text{in lower channel.} \end{aligned} \tag{1.2}$$

If we choose the length of MSC equal to $\pi/(r_o - r_v) = x$ then, at the output of MSC, the wave amplitudes will be

$$\begin{aligned} |A(x)| &= 0, \quad \text{in the upper channel,} \\ |A(x)| &= 1, \quad \text{in the lower channel,} \end{aligned}$$

what means that SAW is fully transmitted, by means of MSC, from the upper to the lower acoustic channels. This 0 dB MSC is the most frequently applied in SAW devices.

The theory of MSC working with SSBW cannot be such simple for one fundamental reason: there are not propagating modes of particular wave numbers. Instead, there are continuous spectra of bulk waves propagating in different directions inside the body, which spectra spans over the domain $(0, k_s)$, where k_s is the cut-off wave number of bulk waves (the backward bulk waves have the spectra spanning from 0 to $-k_s$). This means that there is not such simple representation of acoustic wave field under MSC as given in Eq. (1.1). Instead, we will have certain integral representation over the whole spectra of bulk waves propagating under MSC.

In this paper, we will adopt the theory of bulk wave excitation, propagation and detection in the system of periodic strips developed in [2]. The main objectives in developing theory of MSC working with SSBW are the following

1. evaluation of bulk wave field generated by certain electrode of the periodic system of strips that is supplied from external voltage source. This wave field "isonifies" the MSC placed in some distance from the exciting electrode (that is from the source of bulk waves),

2. evaluation of electric potentials and currents of strips that make MSC in the periodic system of strips and spanning over both coupled acoustic channels (Fig. 1),

3. and finally, evaluation of the voltage excited on certain strip residing in some distance of MSC in the other acoustic channel. This voltage is a measure of transmission of bulk waves from one to the other acoustic channels; without MSC, the acoustic field in the other channel would not exist.

2. Characterization of piezoelectric substrate

In this paper, we assume that strips of MSC are perfectly conducting and weightless, they influence the electric field on the substrate surface only, leaving it mechanically free, without any surface tractions. This results in purely electric boundary value problem to be considered in this paper, and only electric characterization of the body is required in the theory.

The useful and sufficient characterization is given by the so-called effective surface electric permittivity of the substrate which, originally introduced for the case of SAWs [3], was further generalized to account for bulk waves [4]. This is the characterization in spectral domain of wave number k of an electric wave field on the substrate surface

$$e^{-jkx}e^{j\omega t} \tag{2.1}$$

for given angular frequency ω, x is spatial coordinate on the substrate surface, along the wave propagation direction. It is assumed independent on the other coordinate on this plane. The wave field depends on the coordinate perpendicular to the surface (in depth of the body), and this dependence is fully accounted for in the planar characterization described below.

In piezoelectric body, the elastic waves are coupled to electric field. We will be particularly interested in electric field tangential to the substrate surface $E_{\|}$, and induction perpendicular to this surface $D_{\perp}$. For given $E_{\|}$, the corresponding induction in vacuum over the substrate can be easily evaluated, and it is useful to introduce the electric flux discontinuity at the substrate surface $\Delta D_{\perp}$. In these notations, the above mentioned generalized effective surface permittivity is the following [4]

$$\Delta D_{\perp}/E_{\|} = -j\epsilon_e \frac{k}{\sqrt{k^2}} \frac{\sqrt{k^2-k_s^2}-\beta\sqrt{k^2}}{\sqrt{k^2-k_s^2}-\alpha\sqrt{k^2}}, \tag{2.2}$$

where square roots are evaluated uniquely by applying cuts of complex k-plane as shown in Fig. 2 (in lossy media), k_s has small negative imaginary value, this makes that the corresponding branch points lay outside the real axis of k; and argumentation concerning electrostatic approximation [4] allows us to replace $\sqrt{k^2}$ with $\sqrt{k^2-o^2}$, with o small, that moves also these branch points outside the real axis of the complex k-plane.

The approximation parameters, α, β, and ϵ_e are evaluated by standard analysis of the corresponding boundary value problem for piezoelectric halfspace governed by its

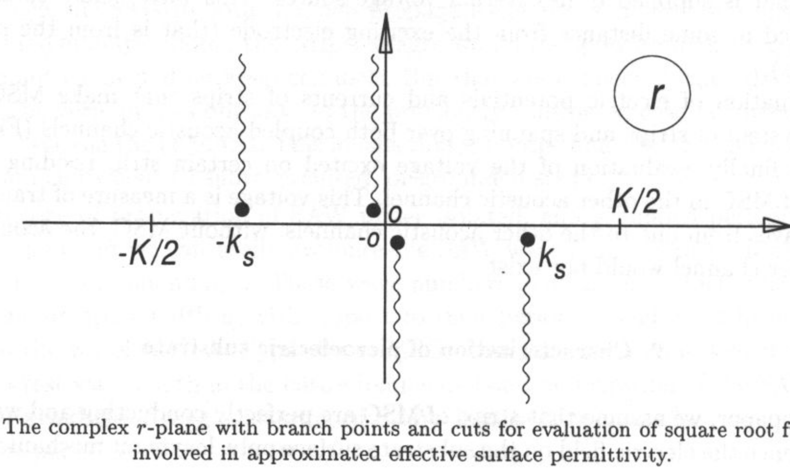


Fig. 2. The complex r-plane with branch points and cuts for evaluation of square root functions involved in approximated effective surface permittivity.

constitutive equations and the equations of motion [5]. The approximation (2.2) is the simplest one accounting for bulk waves. Generally, α and β can be complex, and $(k-k_s)^p$, $p=1/n$, $n<6$ can appear instead of $(k-k_s)^{1/2}$. In this paper we apply the approximation (2.2) because it is valid for most practical cases of piezoelectric substrates.

It results immediately from Eq. (2.2) that

$$\begin{aligned} k_v &= k_s/\sqrt{1-\beta^2}, \quad \text{provided that} \quad \beta>0,\\ k_o &= k_s/\sqrt{1-\alpha^2}, \quad \text{provided that} \quad \alpha>0, \end{aligned} \tag{2.3}$$

are the wave numbers of SAW propagating on free, and metallized substrate surface, correspondingly; the surface wave exists provided that the specified condition is satisfied.

3. Theory of periodic strips

The boundary value problem for periodic strips residing on the piezoelectric halfspace is following: find electric field in the plane of strips knowing that

$$\begin{aligned} E_{\|}(x) &= 0 \quad \text{on strips},\\ \Delta D_{\perp} &= 0 \quad \text{between strips}, \end{aligned} \tag{3.1}$$

which are the local boundary conditions on the substrate surface. The global conditions resulting from the strip interconnections, concern voltages of strips and their total currents; they will be accounted for later below.

The convenient way of solving the above problem is to expand the wave field in the Fourier series that appear natural in periodic system of strips considered here

$$\begin{aligned} E_{\|} &= \sum_{-\infty}^{\infty} E_n e^{-j(r+nK)x},\\ & \qquad\qquad\qquad\qquad r\in(-K/2,K/2),\\ \Delta D_{\perp} &= \sum_{-\infty}^{\infty} D_n e^{-j(r+nK)x}, \end{aligned} \tag{3.2}$$

where each pair (E_n, D_n) satisfies Eq. (2.2) taken for $k=r+nK$. In Eq. (3.2) we choose r in the first Brillouin zone to remove the representation redundancy.

The subsequent (BIS) expansion introduced in [6], in finite limits of m

$$\begin{aligned} D_n &= -j\epsilon_e\frac{1-\beta}{1-\alpha}\sum_m \alpha_m P_{n-m}(\cos\Delta),\\ E_n &= S_n\sum_m \alpha_m P_{n-m}(\cos\Delta), \qquad S_\nu = \begin{cases} 1, & \nu\le 0,\\ -1, & \nu<0 \end{cases} \end{aligned} \tag{3.3}$$

yields following system of equations, for $n\in(-\infty,\infty)$ and for finite number of unknowns α_m

$$\sum_m \alpha_m[1-S_{r/K+n}S_{n-m}g(r+nK)]P_{n-m}(\cos\Delta)=0, \tag{3.4}$$

where $\Delta = \pi w/\Lambda$ and P_ν is a Legendre function of rank zero, and

$$g(k) = \frac{1-\alpha}{1-\beta}\frac{\sqrt{k^2-k_s^2}-\beta\sqrt{k^2}}{\sqrt{k^2-k_s^2}-\alpha\sqrt{k^2}}, \qquad g(\pm\infty) = 1. \tag{3.5}$$

Happily, most of these equations are satisfied identically provided that we apply approximation that $g(|k| > P) = 1$ for certain large, but finite P (see discussion in [7]).

In what follows, we asume $K > 2k_s$, and applying $P \approx 1.2k_s$ that is a reasonable value for small values of α and β, and $|\alpha-\beta| \ll 1$ for all practical piezoelectric substrates, the nonzero solution to Eqs. (3.2) is following

$$\alpha_1/\alpha_0 = \frac{1-g(r)}{1+g(r)}, \qquad r \in (0, K/2) \tag{3.6}$$

with arbitrary α_0 (it will be later evaluated from the above mentioned global boundary conditions), and for $r > 0$.

It can be verified by inspection that, for $r = -\tau < 0$, the solution to α_m is α_{1-m} obtained for $\tau > 0$ from Eq. (3.4). Indeed, applying substitutions $r \to -\tau$, $n \to -n$ and $m \to 1-m$, Eq. (3.4) transforms into

$$\sum_m \alpha_{1-m}[1-(-S_{\tau/K+n})(-S_{n-m})g(-(\tau+nK))]P_{n-m} = 0$$

because $S_{-1-n} = -S_n$, $P_{-1-n} = P_n$, and $g(-k) = g(k)$.

The strip voltage which is the integral of tangential electric field, and the strip current that is proportional to the integral of electric flux discontinuity over the strip width, can be partially evaluated by summing up the field Fourier components. This yields the potential of the strip residing at $x = 0$ [8]

$$V(r) = \alpha_0\frac{\pi}{jK\sin(\pi r/K)}V_r, \qquad V_r = \sum_m \frac{\alpha_m}{\alpha_0}P_{-r/K-m}(-\cos\Delta), \tag{3.7}$$

and for next strip, at $x = \Lambda$, this potential is $V_1(r) = V(r)\exp(-jr\Lambda)$.

For further convenience, we introduce the voltage difference between two neighbouring strips $U(r) = V - V_1$. Finally, this voltage difference U, and the current flowing to the strip residing at $x = 0$ are

$$\begin{aligned} U(r) &= \alpha_0\Lambda V_r e^{-jr\Lambda/2}, \\ I(r) &= \alpha_0\Lambda\omega\epsilon I_r, \qquad I_r = \sum_m \frac{\alpha_m}{\alpha_0}P_{-r/K-m}(\cos\Delta) \end{aligned} \tag{3.8}$$

($\epsilon = \epsilon_e(1-\beta)/(1-\alpha) \approx \epsilon_e$). In the above equations, U and I depend on spectral variable $r \in (-K/2, K/2)$, and α_0 that is a function of r to be evaluated later.

The discrete set of potentials of electrodes V_m or their differences $U_m = V_m - V_{m+1}$ defined above, is the inverse Fourier transform of $U(r)$ [8]

$$U_m = \frac{1}{K}\int_{-K/2}^{K/2} U(r)e^{-jrm\Lambda}dr. \tag{3.9}$$

For known set of U_m, one can evaluate $\alpha_0(r)$. Indeed, applying

$$\alpha_0(r) = \frac{1}{\Lambda V_r} e^{jr\Lambda/2} U_m e^{jrm\Lambda} \tag{3.10}$$

in Eq. (3.8), and then applying Eq. (3.9) for k-th strip, we obtain

$$U_k = U_m \int\limits_{-K/2}^{K/2} e^{jr(k-m)\Lambda} dr/K = \delta_{km} U_m \,.$$

The same inverse Fourier transform allows us to evaluate I_n for known set of U_m, included in α_0 evaluated above

$$\begin{aligned} I_n &= y_{nm} U_m \,, \qquad y_{nm} = y_{n-m} \,, \\ y_k &= \omega\epsilon \int\limits_{-K/2}^{K/2} R(r) e^{-jr(k-1/2)\Lambda} dr/K, \\ R(r) &= \frac{I_r}{V_r} = \frac{\sum_m \alpha_m P_{-m-r/K}(\cos\Delta)}{\sum_m (-1)^m \alpha_m P_{-m-r/K}(-\cos\Delta)} \,. \end{aligned} \tag{3.11}$$

Analogously we can solve a reciprocal problem, for given set of currents flowing to strips I_n and for searched voltage differences U_m

$$\begin{aligned} U_m &= z_{mn} I_n \,, \qquad z_{mn} = z_{m-n} \,, \\ z_k &= \frac{1}{\omega\epsilon} \int\limits_{-K/2}^{K/2} \frac{1}{R(r)} e^{-jr(k+1/2)\Lambda} dr/K, \end{aligned} \tag{3.12}$$

with $R(r)$ as evaluated in the equation above; it results from the discussion below Eq. (3.6) that

$$R(-r) = -R(r). \tag{3.13}$$

4. Circuit theory of MSC

Figure 1 presents a model of MSC which electrodes are placed inside the periodic system of strips, just to make Eqs. (3.11) and (3.12) directly applicable to the model. In practical MSC, there are not strips ouside it, and electric field is slightly different at its edge strips. There are many strips in MSC however, and except the edge strips, they are correctly described by the above mentioned equations, thus small inaccuracy concerning edge strips does not change much the MSC performance.

Accordingly to Fig. 1, the current excited in this part of MSC strips which reside in the upper acoustic channel is

$$I_n = y_{nI} U_I + y_{nm} U_m \,, \qquad n, m \in \text{MSC}, \tag{4.1}$$

where U_I is the voltage difference applied to some strips outside MSC in the upper channel to excite acoustic waves that propagate towards the MSC, and U_m, I_n are voltage difference on MSC strips and currents flowing to the upper part of MSC strips. Because all strips are isolated, the corresponding current flowing to the lower part of MSC strips (the parts residing in the lower acoustic channel) is $-I_n$, and U_m is the same in both parts of strips (except, may be, the edge strips, but we neglect this). Thus the corresponding equation for the lower part of MSC strips is

$$U_k = z_{kn}(-I_n), \tag{4.2}$$

where k can take value outside MSC, for example U_O is the output voltage difference in the lower channel, or $k \in$ MSC, in this case U_k is the same as involved in Eq. (4.1). Both the above equations give the system of equations

$$[\delta_{nk} + y_{nm}z_{mk}]I_k = y_{nI}U_I\,, \tag{4.3}$$

that allows us to evaluate the MSC strip currents I_k, and subsequently to evaluate the transmitted signal to the output electrodes in the lower channel placed somewhere outside the MSC

$$U_O = z_{On}[\delta_{nk} + y_{nm}z_{mk}]^{-1}y_{kI}U_I\,, \tag{4.4}$$

δ_{mn} is the Kronecker delta, and there are sums over repeating indices, over the strips belonging to the MSC.

The output signal U_O would be zero if there was not MSC making coupling of both upper and lower acoustic channels. However, certain part of U_O is transmitted by means of strip mutual capacitances, which part of the signal is not interesting from SAW devices point of view. It is instructive however to discuss it in some details.

Purely electrostatic transmission will take place when the system is placed on dielectric substrate, which effective permittivity is that given in Eq. (2.2) with parameters $\alpha = \beta$. In this case [9]

$$R(r) = S_r P_{-r/K}(\cos\Delta)/P_{-r/K}(-\cos\Delta) \tag{4.5}$$

which is 1 for $\cos\Delta = 0$ applied here for simplicity. It is easy to evaluate the "capacitive" parts of y and z

$$\begin{aligned} I_n &= y^C_{n-m}U_m\,, \qquad y^C_k = -j\frac{\omega\epsilon}{\pi}\frac{1}{k-1/2}\,, \\ U_n &= z^C_{n-m}I_m\,, \qquad z^C_k = -j\frac{1}{\pi\omega\epsilon}\frac{1}{k-1/2}\,. \end{aligned} \tag{4.6}$$

It will be shown in next sections that the "acoustic" part of y or z are much smaller than y^C and z^C for not excessively large k, that is for strips of MSC. This means particularly, that for MSC strips $y_{mn} \approx y^C_{m-n}$ and $z_{nm} \approx z^C_{n-m}$, and subsequently that, for MSC having not too few strips (typical MSCs have 50 or more strips that is sufficient number for current considerations),

$$[y_{nm}z_{mk}] \approx [y^C_{nm}z^C_{mk}] \approx [\delta_{nk}] \tag{4.7}$$

and the inverse matrix involved in Eq. (4.4) can be put approximately equal half of the identity matrix (for typical MSC; if this is not the case, there will be certain "structural" losses in signal transmission by MSC, including losses caused by radiation of bulk waves in backward direction).

Summarizing, we see that it is easy to evaluate purely electrostatic transmission of the signal by MSC, however in practical cases the distance between input and output electrodes are chosen sufficiently large, or certain other isolation of these electrodes is applied to avoid such signal. In the rest of the paper we will neglect this signal by substracting "electrostatic" part from $R(r)$ in integrals describing y_{kI} and z_{On}.

5. Piezoelectric coupling of SAWs

This section is introduced in this paper to clarify the concept of signal transmission coefficient of MSC. The ratio U_O/U_I cannot be applied as a measure of this transmission because, due to weak piezoelectric coupling of elastic waves to electric field in piezoelectrics, even for known 0 dB coupling MSC, $U_O/U_I \ll 1$ (as shown below for the SAW case, this ratio is proportional to the piezoelectric coupling coefficient of SAW).

Applying, for simplicity reasons, that $\cos\Delta = 0$ and $K \gg r_o$, we have [8]

$$R(r) = S_r \frac{r^2 - r_v^2}{r^2 - r_o^2},$$

and without "electrostatic" contribution it is

$$R(r) - S_r R(\infty) = S_r \frac{r_o^2 - r_v^2}{r^2 - r_o^2},$$

which substituted into Eq. (3.11), allows us to evaluate the signal transmission between I-th and O-th strips in the upper acoustic channel by "acoustic" means

$$I_O = \omega\epsilon U_I \int_{-K/2}^{K/2} S_r \frac{r_o^2 - r_v^2}{r^2 - r_o^2} e^{-jr(O-I)\Lambda} dr/K.$$

For $O - I > 0$, and K sufficiently large, we can extend the integration limits to $\pm\infty$. Neglecting integration resulting from branch points at $\pm o$ ($S_r = r/\sqrt{r^2 - o^2}$) and using Jordan's lemma in the lower plane of complex r, we finally obtain

$$I_O/U_I = -j2\pi\omega\epsilon \frac{r_o^2 - r_v^2}{2r_o K} e^{-jr_o(O-I-1/2)\Lambda}. \tag{5.1}$$

Similar considerations allow us to evaluate the signal transmission between I-th and O-th strips in the lower system of isolated strips

$$U_O/I_I = -j\frac{2\pi}{\omega\epsilon} \frac{r_v^2 - r_o^2}{2r_v K} e^{-jr_v(O-I-1/2)\Lambda}. \tag{5.2}$$

Finally we arrive at the following averaged nondimensional parameter that characterizes the electric signal transmission, by "acoustic" means of elastic waves in piezoelectrics, between equally distant strips in upper, and lower channels (in short-circuited, and free systems of strips, correspondingly)

$$\kappa = \left\| \left(\frac{I_O}{U_I}\right)_{\text{short}} \left(\frac{U_O}{I_I}\right)_{\text{free}} \right\|^{1/2} = (r_o - r_v)\Lambda \tag{5.3}$$

where indices "short" and "free" mark the cases of evaluation of corresponding terms. The term $(r_o - r_v)/r_o$ is known in SAW literature as the piezoelectric coupling of SAWs in periodic system of strips; it is usually small quantity, of an order of 1% or less. The above signal transmission coefficient is proportional to this coupling coefficient.

This result shows that, when evaluating strip voltages and currents of MSC strips which are not very distant, we can neglect contributions of elastic waves, replacing y, z by y^C, z^C.

6. MSC for SAWs

Equations (5.1), (5.2), taken for n, m instead of O, I, and applied in Eq. (4.4) with the inverse matrix approximated by the identity matrix divided by 2, allow us to obtain the following relation for signal transmission between upper and lower acoustic channels by means of elastic waves and the MSC

$$\begin{aligned} U_O/U_I &= \frac{1}{2} z_{On} y_{nI}, \\ |U_O/U_I| &= \frac{1}{2}\left[(r_o - r_v)\Lambda\right]^2 \left| \sum_{\text{MSC}} e^{-j(r_o - r_v)n\Lambda} \right| \approx (r_o - r_v)\Lambda \left| \sin\frac{r_o - r_v}{2} N\Lambda \right|, \end{aligned}$$

where N is the number of strips in MSC.

In relation to the average transmission coefficient κ, the SAW transmission by MSC is

$$T = \frac{1}{\kappa}|U_O/U_I| = \left|\sin\frac{r_o - r_v}{2} N\Lambda\right|, \tag{6.1}$$

which T characterizes only the MSC, and it is independent of parameters of piezoelectric substrate. The length of MSC is $x = N\Lambda$, and if x is chosen such that $\sin x(r_o - r_v)/2 = 1$ than $T = 1$ and we have 0 dB MSC, in agreement with coupled modes theory presented in Introduction.

7. Transmission of bulk waves

7.1. The wave field of bulk waves

As opposite to the SAW case, in this section we consider somewhat artificial piezoelectric substrate that does not support SAWs at all; only bulk waves can propagate in the system. In such cases $\alpha < 0$ and $\beta < 0$ in Eq. (2.2) and in consequence, $R(r)$ has no a pole neither a zero in the integration domain in Eq. (3.11) and (3.12). This allows us to evaluate these integrals numerically, by applying the algorithm of fast Fourier transform, for example

$$\int_{-K/2}^{K/2} R(r)e^{-jr(k-1/2)\Lambda}dr/K$$

$$\approx \frac{1}{M}\sum_{l=0}^{M-1}\left\{\begin{array}{ll} R(Kl/M)e^{j\pi l/M}, & l<M/2 \\ R(K(M-l)/M)e^{-j\pi l/M}, & l\geq M/2 \end{array}\right\} e^{-j2\pi kl/M}. \tag{7.1}$$

The transmission of bulk wave signal from I-th strip to O-th strip in the same channel can be evaluated using Eqs. (3.11) and (3.12), with $R(r)$ appropriately modified to elliminate the signal resulting from purely electrostatic interactions between strips. This can be done by substracting term evaluated after substitution $\alpha = \beta = 0$ in Eq. (2.2).

Figure 3 shows the evaluated $|I_O/U_I|$ in the upper channel (with short circuited stips, curve a), and $|U_O/I_I|$ in lower channel of isolated strips (curve b), for different values of $O - I > 0$ and for $K = 2.5k_s$. These bulk wave signals decay with the distance between strips ($|O-I|\Lambda$). This is due to bulk wave diffraction making the acoustic field weaker at the substrate surface with growing distance from the source, which surface field interact with the output strip.

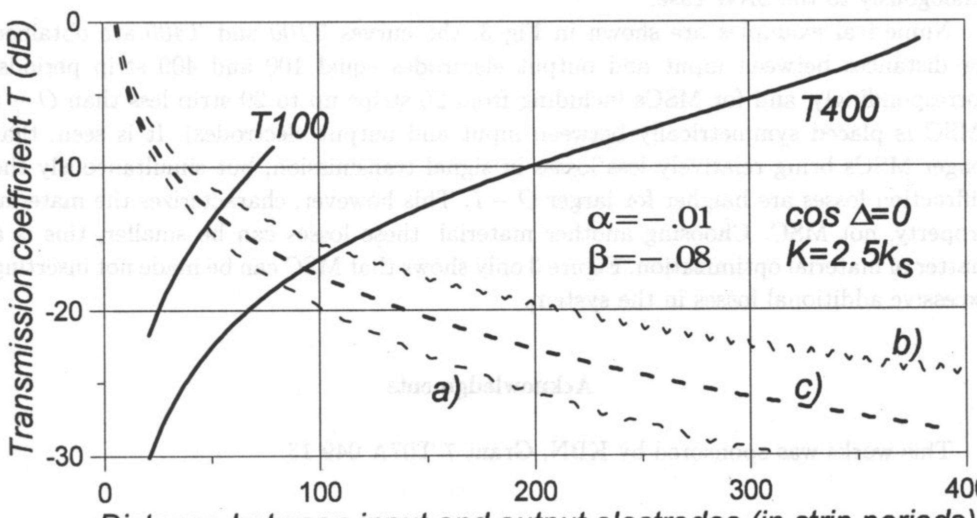


Fig. 3. Bulk wave transmission in acoustic channels of a) short-circuited strips and b) isolated strips, and c) their average κ. Curves present relative decay of the detected signal dependent on the distance (in strip periods) between input and output electrodes. *T100* and *T400* are the transmission coefficient of bulk waves by MSC, as dependent on the number of strips it includes, for $O - I = 100$ and $O - I = 400$ correspondingly; MSC includes from 20 to $O - I - 20$ strips placed symmetrically between I-th and O-th strips.

Analogously to the SAW case, we define the coupling coefficient of bulk waves by Eq. (5.3), with I_O/U_I and U_O/I_I evaluated numerically as discussed above (curve c).

This "coupling coefficient" κ includes the effect of SAW diffraction between input and output strips thus, when applied in the relation for bulk wave transmission by MSC, both these effects will be removed from the transmission coefficient characterizing only MSC (and not piezoelectric coupling coefficient or bulk wave diffraction), analogously to the SAW case considered previously.

7.2. Relative bulk wave transmission by MSC

We again assume weak piezoelectric substrate and sufficiently long MSC to make the inverse matrix in Eq. (4.4) equal to $[\delta_{nk}]/2$. Subsequently neglecting purely electrostatic interaction between input and output electrodes and MSC strips, we obtain

$$U_O/U_I = \sum_{\text{MSC strips}} z_{O-m} y_{m-I} \tag{7.2}$$

where y and z are evaluated numerically using FFT.

Finally, the transmission coefficient of bulk waves by MSC is defined by

$$T = \frac{1}{\kappa}|U_O/U_I|, \qquad \kappa = \left\| \left(\frac{I_O}{U_I}\right)_{\text{short}} \left(\frac{U_O}{I_I}\right)_{\text{free}} \right\|^{1/2}, \tag{7.3}$$

analogously to the SAW case.

Numerical examples are shown in Fig. 3: the curves *T100* and *T400* are obtained for distances between input and output electrodes equal 100 and 400 strip periods, correspondingly, and for MSCs including from 20 strips up to 20 strip less than $O - I$ (MSC is placed symmetrically between input and output electrodes). It is seen, that longer MSCs bring relatively less losses in signal transmission, but simultaneously the diffraction losses are haigher for larger $O - I$. This however, characterizes the material property, not MSC. Choosing another material, these losses can be smaller, this is a matter of material optimization. Figure 3 only shows that MSC can be made not inserting excessive additional losses in the system.

Acknowledgments

This works was sponsored by KBN, Grant 7 T07A 049 13.

References

[1] D.P. MORGAN, *Surface-wave devices for signal processing*, Chapt. 5, Elsevier, Amsterdam 1991.

[2] E. DANICKI, *Influence of bulk-wave generation on SAW filter performence*, J. Techn. Phys., **21**, 405–420 (1980).

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

[4] E. DANICKI, *New theory of SSBW*, 1980 IEEE Ultras. Symp. Proc., 235–239 (1980).

[5] E. DANICKI, W. LAPRUS, *Singularities of the planar Green function in the spectral domain*, 1996 IEEE Ultras. Symp. Proc., 225–228 (1996).

[6] K. Bløtekjær, K.A. Ingebrigtsen, H. Skeie, *A method for analyzing waves in structures consisting of metal strips on dispersive media*, IEEE Trans., **ED-20**, 1135–1138 (1973).

[7] E. Danicki, *Bending of a layered isotropic plate with periodic disbondings*, Trans. ASME, **61**, 612–617 (1994).

[8] E. Danicki, *Generation and Bragg reflection of surface acoustic waves in nearly periodic system of elastic metal strips on piezoelectric half-space*, J. Acoust. Soc. Am., **93**, 116–131 (1993).

[9] E. Danicki, A. Kawalec, *Metallization-ratio dependent higher overtone operation of IDT*, 1996 IEEE Ultras. Symp. Proc., 309–312 (1996).

ARCHIVES OF ACOUSTICS
23, 1, 139–148 (1998)

ULTRASONIC CAMERA FOR FINGER RIDGE PATTERN IMAGING (*)

W. BICZ (1), D. BANASIAK (2), P. BRUCIAK, Z. GUMIENNY (3),
S. GUMULIŃSKI, D. KOSZ, A. KRYSIAK, W. KUCZYŃSKI,
M. PLUTA (3) and G. RABIEJ

Przedsiębiorstwo Badawczo-Produkcyjne
OPTEL Sp. z o.o.
(50-212 Wrocław, ul. Otwarta 10A)

This paper describes the design of an ultrasonic camera with a resolution of 0.1 mm. The camera makes it possible to observe the near surface structures of solid objects and is suitable for finger ridge pattern imaging (i.e. a pattern reflected in the fingerprint). The device can be used for biometric identification of individuals (for access verification). It can also be employed for all other sorts of structures with ultrasonically detectable changes in the near surface structure, both natural and artificial ones (e.g. created for information recording). The paper describes the current version of the camera and the physical basis of its operation. Perspectives of further development of the device are also presented.

1. Introduction

During the last few years a new area of engineering science has been established whose products are likely to create a large market in the near future. It has been called "biometrics". The pioneers of this new domain intend to construct devices which would allow the identification of a person on the basis of his/her "biological" characteristics: voice, dynamics of movements, features of the face and other parts of the body, retina or iris pattern. However, most promising seems to be the possibility of fingertip structure recognition (this structure is reflected in the fingerprint pattern). It is well known that the finger ridge pattern is different for each individual and does not change over the life time. Touching of a sensor surface is a simple act. Many inventors of biometric devices hope to develop a button which would "know" by whom it has been pressed and which finger has been used. The button used for unlocking a door would obviously let in only

(*) The work described herein is the effect of many years of research and development at Optel.

(1) Author of the paper and coordinator of the project.

(2) Also at the Technical University of Wrocław, Institute of Engineering Cybernetics, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław.

(3) Also at the Technical University of Wrocław, Institute of Physics, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław.

authorized people. Since such a device is likely to find numerous applications, one can envisage a rapid development of the market for biometric devices [1, 2, 3].

Systems for the ridge pattern imaging with optical acquisition of the data have been investigated for a number of years. They show "live" fingerprint images directly from a finger without the need for ink and paper which have been traditionally used by policemen since Galton's times [3, 4]. Systems with optical data acquisition, however, have a number of drawbacks: the direct image of the fingertip is of very low contrast and it is easier to see the dirt on the finger than its ridge pattern. Moreover, methods employing the reflection from a surface are very sensitive to grease, dirt, and water. A three dimensional image is difficult to create and does not provide satisfactory results for damaged fingers [2]. Furthermore, no method allows to decide easily whether the object under observation is a real finger, an imitation, or perhaps a greasy residue of a finger on the sensor. The description of a typical optical fingerprint imaging system is given in [6].

Hence, it should not be surprising that there has been interest in alternative methods of ridge pattern imaging. For instance, Constantine Tsikosa proposed a capacitive method [7], recently developed further by SGS-Thomson [8] and Siemens [3, 9]. So far only prototypes of such devices have been presented and little is known about their practical usefulness.

2. Perspectives of the ultrasonic devices development

In 1986, the author of this paper proposed a method based on the ultrasonic data acquisition [14]. This approach allows to distinguish between real fingers and any imitations. Furthermore, it is not sensitive to any dirt, grease etc. There is also a completely new perspective unthinkable in the case of other methods. It is possible to create a device with a surface reacting to a finger touch (or a number of fingers) which would be able to decide where the finger has been placed, identify it and register its movements. Such a device would not have any moving parts and could replace today's keyboards, mice, graphic pads, and fingerprint identification systems, though these are not all of its potential applications. To complete the picture, it is worthwhile knowing that it is feasible to create a device that is small, inexpensive (a kind of chip), and could really fit in a button. Such a device would have another interesting property. It would enable us to devise a system for remote people identification (through a network) which cannot be cheated, even if a person sitting at a remote terminal has all the technical means to carry out a fraud.

A number of papers have been published describing our method [10–13], a few patents have been granted and a few other patents are pending [14–16] (the owner of the patents and commercial rights to the device is Sonident, Vaduz). This work is aimed at a brief presentation of the key aspects of the method employed by us which have not been described in detail in the previous papers. The paper is also intended to present the subject to the readership of "Archives of Acoustics".

3. Principle of the ultrasonic camera operation

The operation of our devices is based on a phenomenon which apparently has not been employed so far by anyone and, perhaps, not even noticed (to the best of our knowledge). It can be summarized in the form of the following rule:

Consider a surface of a solid object on which another object has been placed, so that the contact between the two objects is not ideal, i.e. there are some inhomogeneities. The sound wave which reaches such a place does not only pass from one environment to the other one, get reflected and diffracted in the contact area as described by the classical theory but it also is subject to some additional scattering and transformation to a different kind of waves. This phenomenon is the effect of a disturbance of the sound propagation conditions in the contact area between the two objects, hence it will be referred to as the contact scattering. It is sure that this kind of scattering results not only from the contact area of the two objects but also from the area near the objects' surfaces (henceforth it will be referred to as the near surface structure). It is likely that for this reason the contact scattering is strongly dependent on the substance of the objects.

Experiments show that the transition of the wave from one environment to the other one may practically not occur at all and only the contact scattering and generation of other types of waves (that is particularly conspicuous for transversal waves) are observed. It is likely that the disturbances of the wave occurring in the contact areas are mainly in phase (the phase front is spatially distorted) and are responsible for the observed contact scattering. At the moment, research aimed at a theory adequately describing this phenomenon is being carried out. We shall devote further publications to this subject.

4. Design of the ultrasonic camera

Employing the phenomenon described in the previous section, we have designed a device for measuring and analysing signals resulting from the contact scattering of objects placed on a plastic window. The device is designed mainly for the near surface observation of the finger ridge patterns. A detailed description of our device has been presented in the aforementioned papers. For all those readers who are not familiar with the subject, we offer a brief description:

An acoustic wave is sent in the direction of the surface on which an object has been placed (see Fig. 6). Signals scattered by the object are received by the transducer (T) that is moving along a circular trajectory whose axis is perpendicular to the contact surface ($x - y$). The same element can be used both as emitter and receiver. Alternatively, it is possible to employ a number of fixed transducers, instead of a single moving transducer.

For the analysis of the object with a resolution of about 0.1 mm, it is necessary to collect the scattered signal data from about 256 different angles. At the moment, our device sends a short pulse in each of the 256 directions and receives the impulse response (in the case of a finger, the signal spectrum is in the range from 4 to 16 MHz and depends on the device design). Figure 1 shows the set of impulse responses for a small ball, whereas

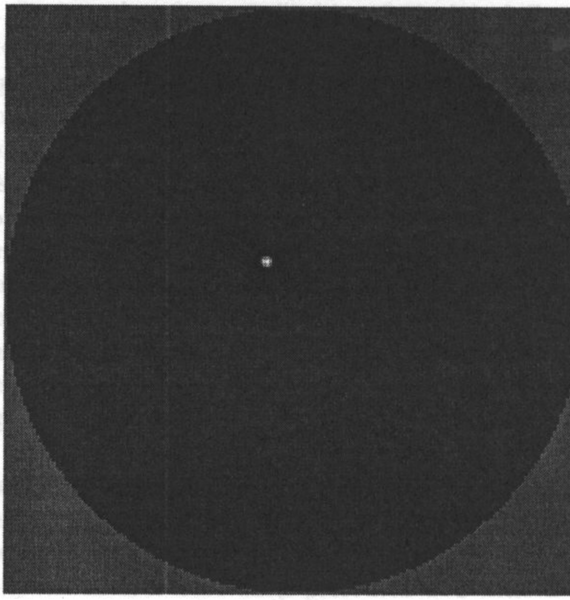


Fig. 1. Reconstruction of a ball.

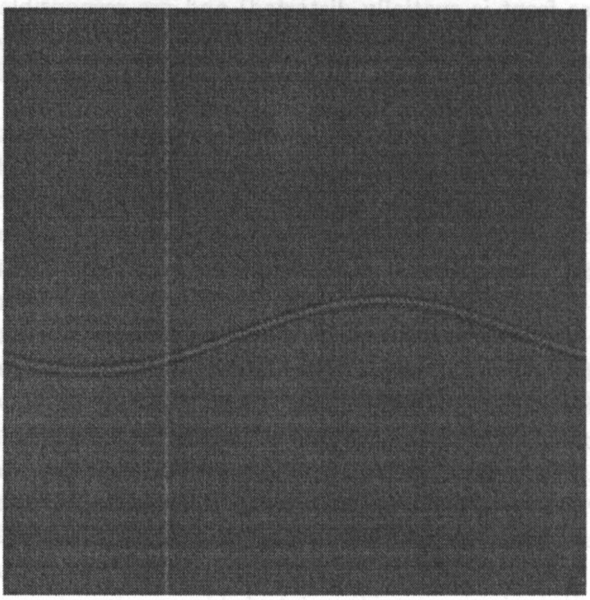


Fig. 2. Impulse response of a ball.

Fig. 3 shows those for a finger (the vertical axis corresponds to time, the horizontal on to the angle, the value of the signal is represented by the grey level). In order to obtain the observed structure from the collected data, a reconstruction procedure is used which

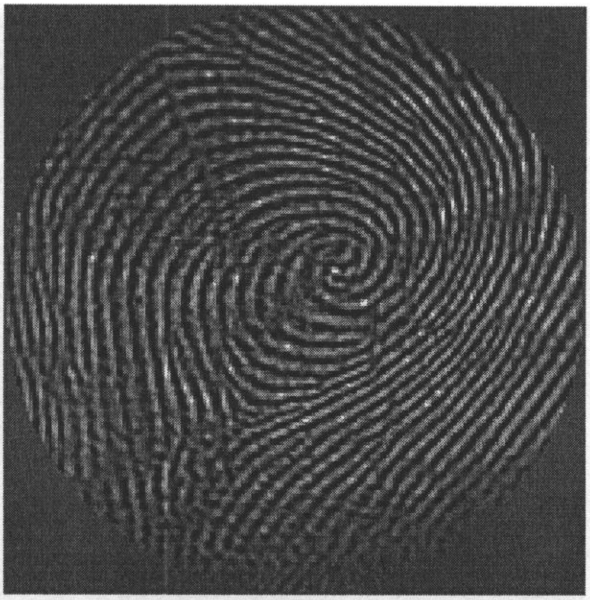


Fig. 3. Reconstruction of a finger.

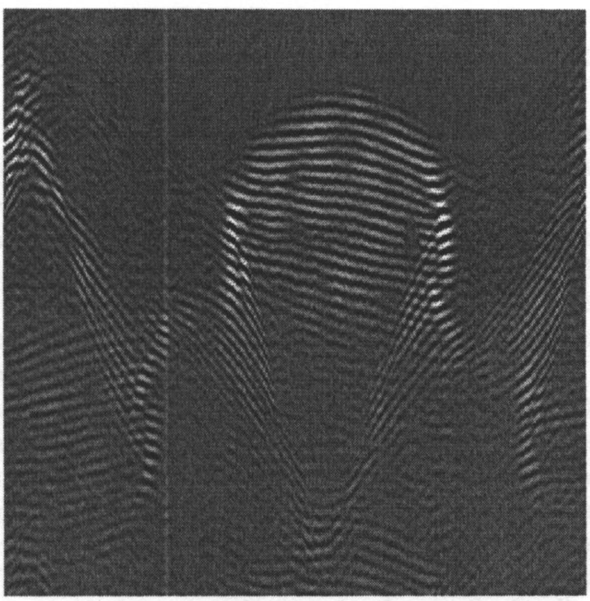


Fig. 4. Impulse response of a finger.

is similar to the methods used in ultrasonic reflection tomography. A set of programs aimed at achieving high quality and high speed reconstruction have been written. The algorithms developed at Optel enable image reconstruction based on a set of 256 impulse

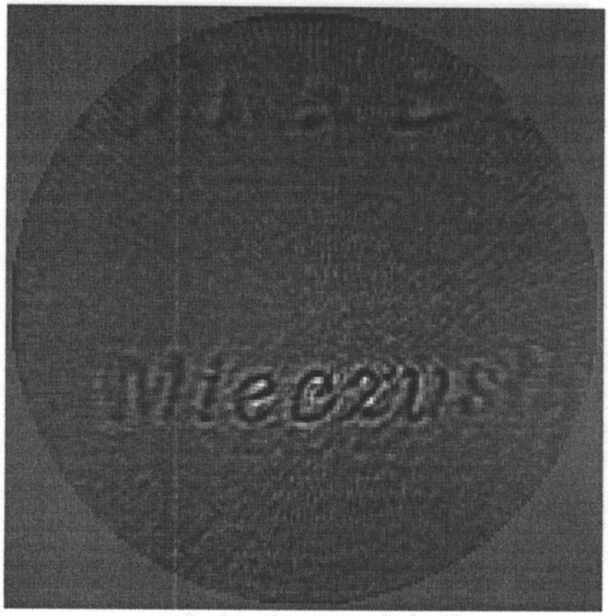


Fig. 5. Reconstruction of a stamp.

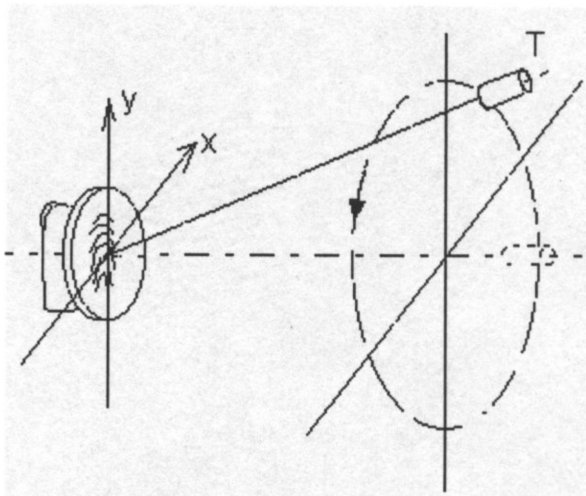


Fig. 6. Schematic diagram of the device.

responses each containing 181 samples in about 20 ms (using a standard PC based on the Cyrix 6x86 P200+ processor). The reconstructions for the impulse response given in Figs. 1 and 3 are presented in Figs. 2 and 4, respectively. Figure 5 shows the image of a stamp placed on the sensitive surface of the device. A photograph of the current version of the device is presented in Fig. 7.

Fig. 7. Appearance of the camera.

5. Technical solutions employed in the camera's design

The use of the contact scattering phenomenon discovered by us and the computer tomography methods were not sufficient to construct an ultrasonic camera. We had to solve a few other problems:

In order to obtain the required resolution, it was necessary to develop a device which, having a relatively small diameter, would emit a gaussian ultrasonic beam of high amplitude and would have a high sensitivity as receiver. A device which meets these requirements has been developed and patented [16]; we intend to present its construction in a separate paper.

It was also necessary to develop a transducer which would be able to emit a short pulse and would have the required bandwidth (4 – 16 MHz) as receiver. Moreover, its phase function was required to have the smallest possible variance. It was also important that such a transducer would be cheap and would have repeatable parameters (the final effect of our research is expected to be a device suitable for mass production at a reasonable price). The researchers at Optel managed to develop a transducer which has a completely new design (a patent application has been submitted). It is able to emit very short pulses (in the range of 20 ns – see Fig. 8) and has a very wide bandwidth as receiver (ca 4 – 25 MHz). The amplitude of the signal emitted by the new transducers is about two times higher than those of classical pulse transducers. Their sensitivity as receivers is however slightly lower giving a comparable result in the measurement cycle. Nevertheless, the idea of those new transducers opens up new possibilities for

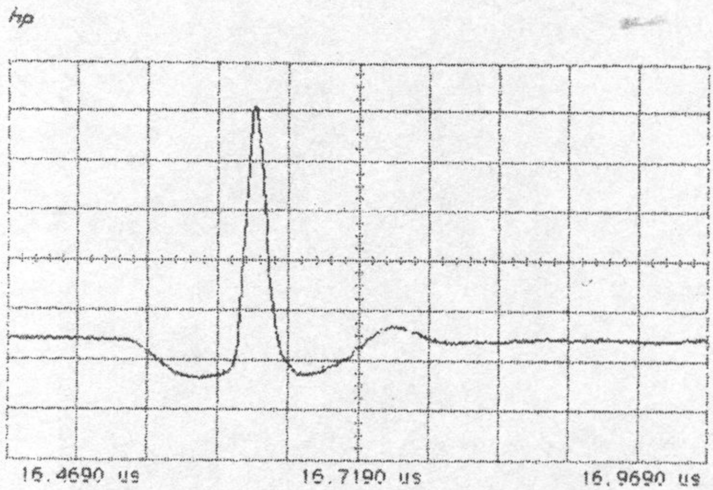


Fig. 8. The shape of a pulse generated by our transducer.

designing ultrasonic transducers and a significant improvement of their parameters can be expected; we wish to devote a separate paper to this subject.

The design of our ultrasonic camera would not have been possible, unless we had not developed our own electronic circuitry which includes the transceiver circuit and an oscilloscope card. These elements are also based on our own original ideas: the pulse generator is capable of generating pulses as short as 20 ns with an amplitude of ca 600 V; the receiver has a sensitivity of 5 μV for frequencies within the range of 4 – 16 MHz, and a dynamic range of 60 dB. The oscilloscope card makes it possible to sample at up to 200 MS/s and is specifically dedicated for processing sets of ultrasonic signals (it satisfies some strict timing parameters).

It should also be noted that our ultrasonic camera would not be of much use if there were no methods for the finger ridge pattern analysis. Also in this area we have some original solutions, though they are perhaps of less interest to the readers of this journal. It is however worthy of notice that the algorithms which have been developed allow not only fingerprint recognition but also a significant compression of the fingerprint data. For example, a finger ridge pattern can be synthesized from the information contained in as few as 100 bytes.

6. Observations with the use of the camera

• Objects of similar structure but made of different substances give significantly different signals (both in amplitude and character). The structure of the objects is nevertheless visible. Hence, it is possible to distinguish between "real" and "artificial" fingers.

• Spreading gel on the surface of an object, soaking it in water or covering with dirt does not cause significant changes of the signal.

• A fingerprint is hardly noticeable because its signal level is at least by 30 dB lower than that given by a real finger (in contrast to this, for optical devices this level does not change significantly). The above observation is also true when soot or metal powder is used in order to enhance the fingerprint.

• A fingerprint left on a thick (ca 0.5 mm) layer of gel or grease is noticeable but it is very different when observed directly.

• Fingers with a damaged surface still give a relatively clear image. Their internal structure seems to be visible, since the phenomenon on which our observations are based applies to the near surface layer.

7. Future work

In the near future, we plan to develop a new model of the camera that will be based on fixed transducers and will be capable of showing "live" pictures of objects at 25 frames per second. It will be a kind of a "real-time" ultrasonic camera which can see the near surface structures of objects placed on its sensitive surface. The camera will contain its own electronic circuit for reconstruction and will have an output for a standard monitor. The camera used at present is based on a moving transducer and can produce a few frames per second. It also needs a computer which performs the signal processing and displays the image. In 1998, we plan to develop an integrated version of the device. Eventually, we hope to implement it in a kind of chip.

References

[1] E. Newham, *The Biometrics Report*, SJB Services.

[2] D. Mohrmann, *Biometrie als Quantensprung?*, W&S, 7, p. 28 (1997).

[3] H. Mueller, *Der Koerper als Passwort*, PC Magazin, p. 256, Januar 1998.

[4] F. Galton, *Finger prints*, published in 1892 by Macmillan and Co., London and New York, reprint: Da Capo Press, New York 1965.

[5] H.C. Lee, R.E. Gaensslen, *Advances in Fingerprint Technology*, Elsevier, 1991.

[6] L.H. Andersen, Peter Juergensen, *Fingerprint verification – for use in identity verification systems*, Master Degree Work, Aalborg University 1993.

[7] Siemens AG, EP 0 041 693 A1, 1981.

[8] *Fingerprint-Chips*, Markt&Technik, 50, p. 46 (1997).

[9] *Der Fingertip-Sensor*, Blick durch die Wirtschaft, (18/12/97, p. 6), FAZ Verlag, Frankfurt am Main.

[10] Z. Gumienny, M. Pluta, W. Bicz, D. Kosz, *Rekonstrukcja struktur dwuwymiarowych z zastosowaniem algorytmu FFT*, Akustyka Molekularna i Kwantowa, **15** (1994).

[11] Z. Gumienny, M. Pluta, W. Bicz, D. Kosz, *Synthetic aperture acoustic microscope for evaluation of fingertip peripheral skin structure*, SPIE, 2390.

[12] W. Bicz, Z. Gumienny, M. Pluta, *Ultrasonic sensor for fingerprint recognition*, SPIE, 2634.

[13] Z. Gumienny, M. Pluta, W. Bicz, D. Kosz, *Ultrasonic setup for fingerprint patterns detection and evaluation*, Acoustical Imaging, **22** (1996).

[14] Sonident, World Patent WO-A-8 705 790.

[15] Sonident, US Patent Nr. 5258922.

[16] Sonident, US Patent Nr. 5515298.

[17] B. FAY, *Strukturuntersuchungen mit Hilfe der Ultraschallrueckstreuung*, in K. Brendel (Ed.) Stand und Entwicklungstendenzen der Ultraschallmesstechnik (Vortraege des 64. PTB-Seminars am 23./24.04.1986).

[18] G.T. HERMAN, A.K. LOUIS, F. NATERRER [Eds.], *Mathematical methods in tomography*, Proceedings, Springer Verlag, Oberwolfach 1990.

[19] A.C. KAK, M. SLANEY, *Principles of computerized tomographic imaging*, IEEE Press, 1988.

ARCHIVES OF ACOUSTICS
23, 1, 149–150 (1998)

Przedsiębiorstwo Badawczo-Produkcyjne
OPTEL Sp. z o.o. ul. Otwarta 10a
PL-50-212 Wrocław
tel. +48 (071)329 68 53
email: optel@optel.com.pl.

www.optel.com.pl.

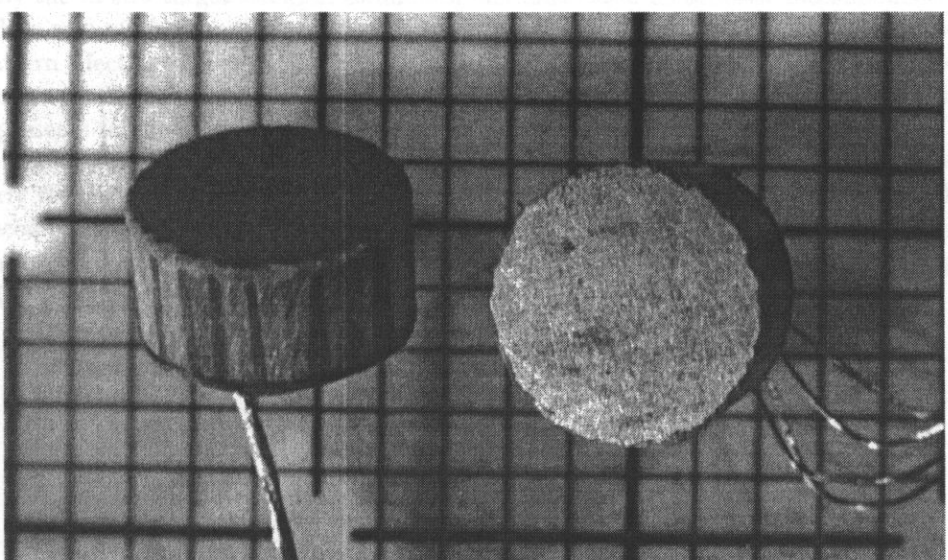
We offer a new type of

Ultrasonic Transducers with a very short impulse response

for nondestructive testing and all other purposes, where broadband transfer function, high amplitude of the pulse, high sensitivity and low cost are required

The idea of this type of transducer (a patent application has been submitted) was born as a spin-off from our biometric project (see our paper in this issue). The new transducers offer:

— **broadband transfer function** (4 – 25 MHz);
— **very short pulse response** (up to 20 ns);
— **high signal amplitude and high sensitivity**;
— **thickness** of not more than 2 mm;
— operation at **high temperatures**;
— **very low cost**.



Apart from our patented transducers, we are able to produce piezo composite transducers up to 3 MHz main frequency.

We also offer all electronic devices that are necessary to produce, receive and analyse ultrasonic signals:

— **pulse generator** with high amplitude (600 V), short pulse (20 ns) and low internal impedance (~ 1 Ohm) with an integrated send-receive switch;

— **many types of amplifiers** (2 – 60 MHz) with low noise (5 μV);

— **oscilloscope card** (ISA) with 80 MSPS and many features that are important for ultrasonic measurements (large post-trigger time, sophisticated synchronisation possibilities for automatic measurement).

We are able to offer very competitive prices. Our team of researchers has many years of **experience in developing acoustic devices, electronics and software**. We can provide the **leading edge ultrasonic technology and "know-how"**. If you are looking for someone who can help you find a solution – do not hesitate – contact us:

Optel sp. z o.o. R&D company
ul. Otwarta 10a PL 50-212 Wrocław
Tel.: 48 (071) 3296853 Fax.: 48 (071) 3296852
email: optel@optel.com.pl www.optel.com.pl

ARCHIVES OF ACOUSTICS
23, 1, 151–153 (1998)

C H R O N I C L E

DISSERTATIONS

Inverse problems in fish target strength estimation [in Polish]
by MAREK MOSZYŃSKI,
Department of Acoustics, Technical University of Gdańsk, 18 November 1997
Politechnika Gdańska, Biblioteka Główna,
ul. Narutowicza 11/12, 80-952 Gdańsk (Poland)

Fish target strength estimation constitutes one of the main problem in the hydroacoustic methods of fish populations assessment and monitoring. Knowledge of target strength and its equivalent value in absolute units – backscattering cross-section – is needed for proper scaling of measurements data from echo integration surveys. In the thesis the hydroacoustic methods of fish abundance estimation are analyzed. Particularly the *in situ* target strength estimation methods based on acoustic measurements of fish in natural environment are described in the context of statistical removing of beam pattern effect in a indirect way by processing of the collection of fish echoes. These indirect methods solves the inverse problem under question in which the unknown function represents target strength probability distribution function (PDF).

Newly introduced by the author fish target strength estimation methods, particularly the method based on Discrete Mellin Transformation (DMT) with singular value decomposition (SVD) and regularization techniques of Fredholm integral equation, reduces shortages if methods in use and guarantees obtaining more reliable target strength PDF estimates less susceptible to artifacts and other deformations.

The acoustic system for measurement of the nonlinearity parameter of the water medium using parabolic model [in Polish]

by KRZYSZTOF ZACHARIASZ,
Department of Acoustics, Technical University of Gdańsk, 16 December 1997
Politechnika Gdańska, Biblioteka Główna,
ul. G. Narutowicza 11/12, 80-952 Gdańsk (Poland)
Supervisor A. Stepnowski

The thesis deals with the finite amplitude method for the determination of the acoustic nonlinearity parameter for liquids, especially for water medium. Parameter is determined by fitting the theoretical model of the acoustic system to the results of measurements of the nonlinear distortions which are carried out in the nearfield of a circular ultrasonic transducer with sinusoidal excitation. These distortions are studied as a function of the distance from the source using a planar receiving transducer with a diameter equal to that of the source. The existing acoustic methods employ measurements of the fundamental and second harmonic components of the acoustic pressure and the quasilinear theoretical models. The low accuracy of these models and necessity of the assumption that the transducers are pistonlike lead to significant errors in the estimates. The theoretical model developed for the proposed method involves parabolic approximation of nonlinear wave equation. Direct numerical solutions of the KZK equation are used to determine the averaged acoustic pressure on the receiving transducer (obtained by the integration over the receiver aperture). The parabolic model enables the higher excitation levels, thus allowing to consider the additional harmonic components and to increase the dynamic range of the measurements. The developed method allows also to take into account the actual source characteristics for radiating transducer those are reconstructed by the backward projection of the measured acoustic pressure field (for the fundamental harmonic component).

The ability of the introduced method to precisely determine of the parameter B/A is demonstrated *via* comparison with experiments conducted in distilled water and ethylene glycol. Some results of performed calculations and measurements for the first four harmonic components of the averaged pressure are given to illustrate the capability of the method and to estimate its accuracy. The results obtained confirm the better accuracy of the method that uses the parabolic model.

Composite ultrasonic transducers for medical application [in Polish]
by KRZYSZTOF KYCIA,
Ultrasonic Department, Institute of Fundamental Technological Research,
Polish Academy of Sciences, 22 January 1998
Zakład Ultradźwięków Instytutu Podstawowych Problemów Techniki
Polskiej Akademii Nauk,
ul. Świętokrzyska 21, 00-049 Warszawa (Poland)
Supervisor Andrzej Nowicki

The purpose and manner of conducting the process of asymptotic homogenization of a type 1–3 type composite structure are presented. The formulation of the homogenization process was reduced to numerical static analysis of an elementary symmetry cell of the composite with generalised forces applied at the boundaries of material phases. It was demonstrated that the effective values of the material tensors of the composite depended not only on the tensors of the component materials, but also on variability course of the aforementioned tensors over the volume of the solid of an elementary symmetry unit of the composite. The acoustic parameters of composite transducers were calculated using finite element method including the homogenisation process. The results of experimental measurement were close to hose obtained theoretically. These transducers utilising the complementary properties of a piezoelectric ceramic and piezopolymer. Coupling constant can be 20% larger than those of the piezoceramic, while the acoustic impedance almost reaching the range of the piezopolymers. Acoustic impedance of composite is close to tissue so it recommends them for using in medical ultrasonic imaging. Composites are also over 50% much broadband than conventional piezoceramic. It is very important in many ultrasonic system when high-frequency pulses of short duration are frequently used. The manufacturing process of 2–5 MHz composite transducers were presented.