On the Influence of the Number and Phase of Harmonics on the Perceptibility of the Pitch of Complex Signals

by F. A. Bilsen *

Applied Physics Department, Delft University of Technology, Delft, The Netherlands
and Research Laboratory of Electronics, Massachusetts Institute of Technology,
Cambridge, Mass.

Dedicated to Prof. Dr. C. W. Kosten on his 60th birthday

Summary

Using white noise to mask a 6%-change in pitch, information was obtained on the interaction between harmonics of a complex signal with respect to the perceptibility of pitch (residue pitch, periodicity pitch or musical pitch). Signals composed of all harmonics above the second with equal amplitude in the audio range show a pitch perceptibility that is independent of the phases of the harmonics and is about equal to the perceptibility of pitch for only two successive harmonics, if the harmonic number of the two components is around 4. For two-component signals, pitch and pitch perceptibility are essentially equal for dichotic presentation (one component to one ear and the other component to the other ear) and monotic presentation (both components to the same ear). Pitch perceptibility depends on phase for signals composed of three or more harmonics, if the harmonic number is, roughly, larger than 8. These results explain some disagreement in literature on the influence of phase with respect to residue pitch. Further, they shed some light on the phenomenon of spectral dominance with respect to pitch.

Über den Einfluß der Anzahl und der Phase von Harmonischen auf die Wahrnehmbarkeit der Tonhöhe komplexer Signale

Zusammenfassung

Weißes Rauschen wurde zur Maskierung einer 6%igen Tonhöhenänderung benutzt. Hieraus erhielt man Informationen über die Wechselwirkung zwischen den Harmonischen eines komplexen Signals in bezug auf die Tonhöhenwahrnehmbarkeit (Residuentonhöhe, musikalische oder periodische Tonhöhe). Signale, die oberhalb der zweiten Harmonischen alle anderen Harmonischen im Hörbereich enthalten, zeigen eine Tonhöhenwahrnehmbarkeit, die unabhängig von den Phasenlagen der Harmonischen ist und die etwa der Tonhöhenwahrnehmbarkeit für nur zwei aufeinanderfolgende Harmonische gleicht, wenn die Ordnungszahl der Harmonischen bei 4 liegt. Für Signale, die aus zwei Komponenten bestehen, ist die Tonhöhe und die Tonhöhenwahrnehmbarkeit bei binauraler Darbietung (je eine Komponente auf ein Ohr) und monauraler Darbietung (beide Komponenten auf dasselbe Ohr) im wesentlichen gleich. Die Tonhöhenwahrnehmbarkeit ist phasenabhängig für Signale, die aus drei oder mehr Komponenten bestehen, wenn die Ordnungszahl der Harmonischen etwa größer als 8 ist. Diese Ergebnisse erklären einige Diskrepanzen in der Literatur über den Einfluß der Phase auf die Residuentonhöhe. Weiter erhellen sie etwas das Phänomen der spektralen Dominanz bezüglich der Tonhöhe.

L'influence du nombre et de la phase des harmoniques sur la perceptibilité de la hauteur de signaux complexes

Sommaire

Utilisant un bruit pur pour masquer une variation de 6% de la hauteur, on a obtenu une information sur l'interaction entre les harmoniques d'un signal complexe au point de vue de la perceptibilité de la hauteur (hauteur du résidu, hauteur de la périodicité ou hauteur musicale). Des signaux composés de tous les harmoniques au dessus du second, avec une amplitude égale dans la gamme audible, montrent une perceptibilité de la hauteur qui est indépendante des phases des harmoniques et est à peu près égale à la perceptibilité de la hauteur pour seulement deux harmoniques successifs, si le nombre d'harmoniques des deux composantes est voisin de 4. Pour des signaux de deux composantes, la hauteur et la perceptibilité de la hauteur sont essentiellement égales pour une présentation dichotique (une composante à une oreille et l'autre composante à l'autre oreille) et une présentation monotique (les deux composantes à la même oreille). La perceptibilité de la hauteur dépend de la phase pour des signaux composés de trois harmoniques ou plus, si le nombre des harmoniques est, en gros, plus grand que huit. Ces résultats expliquent quelques divergences dans les publications sur l'influence de la phase à l'égard de la hauteur du résidu. De plus, ils répandent quelque lumière sur le phénomène de l'importance du spectre en ce qui concerne la hauteur.

* Present address: Applied Physics Department, Delft University of Technology, Delft, The Netherlands.

1. Introduction

Periodic signals, like most sounds produced by musical instruments, are able to bring about a sensation of pitch. Sometimes, the pitch equals a pure tone pitch evoked by the fundamental (FOURIER) component of a sound signal, that is physically present as a separate (resolved) entity in the vibration pattern of the inner ear (basilar membrane).

But in most cases the overall pitch sensation, that is the pitch that one perceives if one does not listen analytically (i. e. does not try to distinguish partial tones), is not associated with the sensation of a pure tone; its value, however, roughly corresponds to the pitch of a pure tone having a frequency value equal to the fundamental frequency. In these cases (that correspond to normal musical behaviour) no energy at the fundamental frequency needs to be present in the signal. (For example, music or speech reproduced by small portable-radios lacking low frequencies is not essentially deteriorated.)

This overall pitch, that has been subject of study for more than a hundred years, currently is called residue pitch (SCHOUTEN [1]), periodicity pitch (LICKLIDER [2]), or musical pitch (HOUTSMA and GOLDSTEIN [3]). Unless otherwise stated, in the present study, we will simply use the word "pitch" to indicate this particular mode of pitch perception. For rather complete historical reviews we refer to PLOMB [4], RITSMA [5] or SMALL [6] [1].

In the past, several theories have been developed to explain the existence of this pitch. A possible explanation based on the place theory of pitch perception and the generation of a difference tone by non-linear distortion in the hearing organ (FLETCHER [7], BÉKÉSY [8]) has been ruled out by experiments on the first effect of pitch shift (SCHOUTEN [9], DE BOER [10], SCHOUTEN et al. [11]) and masking experiments (LICKLIDER [2], THURLOW and SMALL [12], SMALL and CAMPBELL [13], PATTERSON [14]).

Most aspects of the pitch behaviour, in particular the first effect of pitch shift, could be successfully accounted for by various models of mechanical and neural processing in the peripheral auditory pathway on, basically, the temporal fine structure or spectral fine structure belonging to the so-called dominant spectral region of a sound signal (DE BOER [10], SCHOUTEN et al. [11], RITSMA [15], BILSEN [16], WALLISER [17], V. D. BRINK [18], SMOORENBURG [19]).

Recently, important experiments by HOUTSMA and GOLDSTEIN [3] revealed that pitch cannot possibly be the result of a proces of autocorrelation at a single place in the peripheral auditory pathway. Because of the fact that dichotic presentation of a two-component stimulus (i. e. one frequency component to one ear and the other – successive harmonic – component to the other ear) delivers essentially the same pitch behaviour as monotic presentation, it is concluded that pitch is extracted from more than a single place at higher (central) centres in the nervous system.

In the present study, we tried to deal with some features of residue pitch, particularly in connection with the following questions:

(1) How many frequency components, being successive harmonics of a missing fundamental, are needed to evoke an optimal pitch sensation? On the one side, it is known that multi-component signals are able to evoke a strong effect. On the other side, only two components seem to provide sufficient residue pitch information (SMOORENBURG [19], HOUTSMA and GOLDSTEIN [3]). For our understanding of the pitch extracting features of the auditory system it is important to know the minimum number of components that provide optimal pitch.

(2) Is the perception of pitch dependent on the phases of the frequency components? On the one side, pitch has been reported to be strongly dependent on the phases of harmonics (MATHES and MILLER [20], LICKLIDER [21], RITSMA and ENGEL [22]). On the other side, experiments by PATTERSON [23] and the results with dichotic two-tone complexes (HOUTSMA and GOLDSTEIN [3]) suggest phase independency. Is there an explanation for this divergence in opinions?

(3) Is there any indication to find why harmonics around, roughly spoken, the fourth harmonic are dominant in the perception of pitch (RITSMA [15])?

To answer these questions in a quantitative way we performed experiments that provide data on the "strength", "clarity" or "perceptibility" of the pitch. Methods like those used before in experiments on the perceptibility of repetition pitch (BILSEN and RITSMA [24]) will be described in the next section.

2. Experimental methods

In order to obtain a proper quantitative measure for the perceptibility of pitch, we assumed that pitch

[1] After submission of the manuscript my attention was drawn to two articles on pitch by TERHARDT (Acustica **26** [1972], 173), in which he introduces the name "virtual pitch".

perceptibility can be expressed in terms of presentage correct recognition of a melody or a simple two-note interval produced by a periodic signal in the presence of background noise. As a first attempt to quantify this idea we resorted to a simple masking experiment in which the level of white masking noise needed to mask a small change in pitch was measured. Thus, it was assumed that higher tolerance to masking noise implies greater pitch perceptibility (see also BILSEN and RITSMA [24]).

The experimental paradigm and definitions are represented in Fig. 1 for the particular case of a two-component complex signal. In two-alternative forced-choice experiments, subjects were presented with pairs of two-, three-, or multi-component

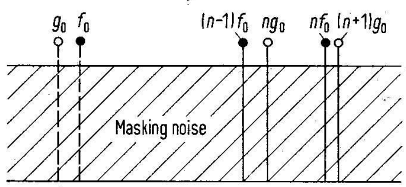

Fig. 1. The experimental paradigm, for two-component signals.

stimuli of 400 ms in duration and with fundamental frequencies f_0 and g_0 6% apart in random order. For each stimulus pair, a subject had to decide whether he heard the pitch going up or down. In the special case of two-component stimuli (sketched in Fig. 1), stimuli could be composed in such a way that the first stimulus of a pair consisted of, for example, the $(n-1)$th and n-th harmonic of f_0, and the second stimulus of the n-th and $(n+1)$th harmonic of g_0 $(=f_0-0.06\,f_0)$, or vice versa (similar configurations were used in the experiments by PATTERSON [14] and SMOORENBURG [19]). Thus, we could be certain the subjects really were listening to periodicity pitch and not to separate components, when they, for example, heard the pitch going down while the components went up in frequency. In one of the experiments, we used the method and setup described by HOUTSMA and GOLDSTEIN [3], in which the harmonics were randomized by means of a PDP 4-computer.

The masking noise level for which the subjects got 75% correct responses was defined as the masked threshold of pitch. In the graphs this threshold is plotted relative to the corresponding threshold for multi-component signals with equal-intensity components and same fundamental frequency. In other words, the perceptibility of pitch for two- or three-component signals relative to multi-component signals is expressed by the level difference ΔL between the masked threshold of pitch for the two- or three-component signal and the masked threshold of pitch for the corresponding multi-component signal.

The signals used were derived from periodic noise and a periodic pulse. The periodic noise consists of a segment of pseudo-gaussian digital noise, a so-called maximum length sequence, that is generated periodically by a digital shift register with modulo-two added feedback. Two oscilloscope pictures of such a signal are reproduced in Fig. 2. The

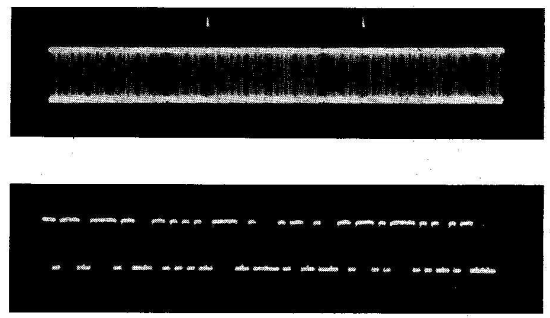

Fig. 2. Oscilloscope pictures of periodic noise; upper one: three periods marked by the trigger pulses; lower one: an enlarged section of one period.

upper trace gives three complete periodes (three sequences) marked by trigger pulses at the beginning of each period. The lower trace shows an enlarged part of one sequence.

Both the periodic pulse and periodic noise have a spectrum consisting of even and uneven harmonics with equal intensity in the audio range. The phase angles of the components of the periodic pulse are all zero; on the contrary, these of the periodic noise are pseudo-randomly distributed. For a detailed description of periodic noise we refer, for example, to [25], [26].

Three-component signals are obtained from periodic noise and the periodic pulse by appropriate bandpass filtering (three Allison 2 BR filters in series).

Two-component signals having fundamental frequencies 6% apart were obtained from four free running sine oscillators.

During the experiments the subjects were seated in a sound-proofed booth. The signals were presented by headphone at an average sensation level of about 40 dB.

3. Experimental results

3.1. Multi-component signals without fundamental

For the range of fundamental frequencies investigated, viz. 50 till 500 Hz, none of the three subjects could measure a significant difference in the masked threshold of pitch for the periodic noise and periodic pulse composed of all harmonics above the second. From this result we conclude that, for multi-component signals, the phases of the components have no influence on the perceptibility of pitch.

The timbre, however, of such sounds definitely is dependent on the phases of the components. Up to fundamental frequencies of about 100 Hz, periodic noise has a noisy character and the periodic pulse sounds as a rattle. From 100 Hz till about 700 Hz the signals are still distinguishable, the periodic pulse having a piercing sound character, the periodic noise producing a more hissing sound. Above about 700 Hz the two signals sound alike (Bilsen [25]). Compare also Licklider [21], Schroeder [27], and Plomp [28].

Also the capability of the hearing organ to distinguish separate components of these kinds of signals highly depends on the phases of the harmonics (Duifhuis [29]).

3.2. *Three-component signals*

The difference ΔL between the levels of white masking noise needed to mask the pitch of a three-component signal and a corresponding multi-component signal (the corresponding frequency components having equal intensity) has been plotted in Fig. 3. The horizontal axis represents the centre frequency of the passband, equal to the frequency value of the n-th harmonic.

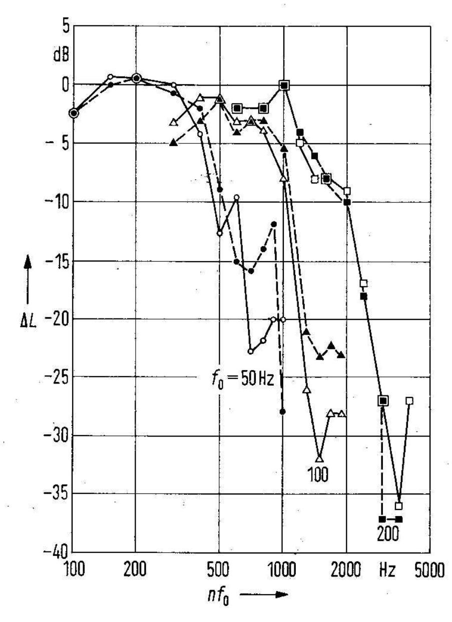

Fig. 3. The masked threshold level difference ΔL as a function of $n\,f_0$, for a three-component signal. Dashed lines: periodic noise; solid lines: periodic pulse.

Three sets of curves have been obtained, for three different values of fundamental frequency, viz. 50, 100, and 200 Hz. The solid line represents ΔL for the periodic pulse, the dashed line belongs to the periodic noise. Two experimental facts are noteworthy.

First, ΔL reaches its maximal value, namely about 0 dB, for about the 3rd, 4th, and 5th harmonic (for example, note that the curves $f_0 = 50$ Hz have their maximum at 200 Hz; this means that the largest amount of masking noise is tolerable for the 3rd, 4th, and 5th harmonic. For $f_0 = 100$ Hz, the curve has its maximum at $n\,f_0 \approx 500$ Hz, i. e. the 5th harmonic). Remember that, according to our definition, $\Delta L = 0$ means that the same amount of masking noise is needed to mask the pitch of the corresponding (unfiltered) multi-component signal.

So this fact suggests that the existence of the so-called dominant frequency region for pitch perception, according to Ritsma [15] positioned at the 3rd, 4th, and 5th harmonic, may be accounted for by optimal pitch perceptibility in that region.

Secondly. The phases of the components influence pitch perception only in the harmonic range above, roughly, $n \approx 8$. Possibly this is due to the influence of combination tones, the intensity of which depends on the phases of the primary, generating, frequencies (Goldstein [30]). These combination tones are in a lower, more dominant, frequency range than the actual stimulus and thus may predominantly determine the perceptibility of pitch.

To investigate this further, experiments are in progress with three-component signals produced by three phase-locked oscillators. The intensity of combination tones of the type $f_1 - k\,(f_2 - f_1)$ is measured in relation to the perceptibility of pitch as a function of the phase of one of the three stimulus components. It turns out that perceptibility of pitch is optimal or minimal when the amplitudes of lower combination tones are optimal or minimal respectively.

3.3. *Two-component signals*

The experiments with two, not phase-locked, harmonics delivered the results presented in Fig. 4. Qualitatively, these results are not much different from the three-component results. Even only two components appear to evoke an optimal pitch perceptibility in the region of the 3rd, 4th and 5th harmonic.

Recently, Houtsma and Goldstein [3] showed that essentially the same pitch behaviour is found

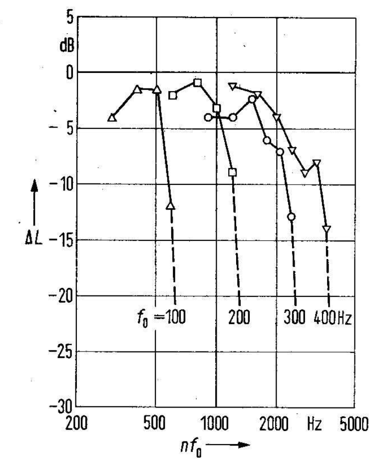

Fig. 4. The masked threshold level difference ΔL as a function of $n f_0$ for a two-component signal.

when, instead of presenting both harmonics to the same ear (monotic presentation), one presents one component to one ear and the other component to the other ear (dichotic presentation). This new finding is confirmed by the pitch perceptibility measurements, shown in Fig. 5. The curves repre-

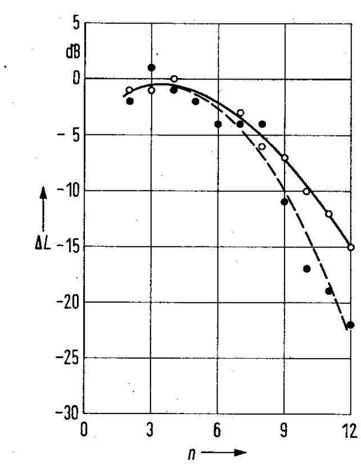

Fig. 5. The masked threshold level difference ΔL as a function of n for a two-component signal.
●●● Monotic signal,
○○○ dichotic signal.

sent ΔL as a function of the harmonic number n, for a fundamental frequency $f_0 = 200$ Hz. Closed circles indicate monotic presentation; open circles dichotic presentation. For low n-values the curves coincide; for higher n-values, pitch perception seems to be higher for dichotic than for monotic presentation.

4. Final conclusions and discussion

From the foregoing results the following final conclusions can be drawn.

1) The perceptibility of pitch (residue pitch, periodicity pitch or musical pitch) of a multi-component signal with harmonics two and higher, is not influenced by the phases of the harmonics.
2) Signals consisting of two (or three) lower successive harmonics, around about the 4th harmonic, provide optimal perceptibility of pitch. This result, to a certain degree, explains why the pitch of (an)harmonic signals is predominantly determined by the spectral region around, roughly spoken, the 4th harmonic (spectral dominance).
3) In the dominant region, pitch perceptibility of two-component signals is equal for monotic and dichotic presentation. There is phase independency.
4) Signals consisting of three harmonics above about the eighth show a perceptibility of pitch that is considerably dependent on the phases of these harmonics. Other experiments reveal this fact to be strongly related to the phase-dependent existence region of combination tones.

The first conclusion is explainable by the conclusions 2) and 3) together. Apparently the pitch (perceptibility) of a multi-component signal is completely determined by pairs of two lower (dominant) harmonics. Because two harmonics do not show phase sensitivity, multi-component signals with energy in the dominant region are not phase sensitive either. The disagreement in literature ([20], [21] and [23]) on phase influences can be understood in the light of conclusion 1), 2) and 3). Whether phase dependence is observed, is dependent on which and how many components are used to generate the residue pitch.

Acknowledgements

The main part of the experiments reported is based on a master's thesis by J. Blank, Delft University of Technology. The author is indebted to the staff of the Communication Biophysics Laboratory of M.I.T. for providing the computer facilities for one of the experiments. Financial support has been obtained from the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

(Received May 3rd, 1972.)

References

[1] Schouten, J. F., The residue and the mechanism of hearing. Proc. Kon. Ned. Akad. v. Wetensch. **43** [1940], 991.

[2] LICKLIDER, J. C. R., Periodicity pitch and place pitch. J. Acoust. Soc. Amer. **26** [1954], 945.
[3] HOUTSMA, A. J. M. and GOLDSTEIN, J. L., The central origin of the pitch of complex tones: Evidence from musical interval recognition. J. Acoust. Soc. Amer. **51** [1972], 520.
[4] PLOMP, R., Experiments on tone perception. Doctoral dissertation, University of Utrecht 1966, p. 88.
[5] RITSMA, R. J., Periodicity detection, in "Frequency analysis and periodicity detection in hearing". Eds.: R. PLOMP and G. F. SMOORENBURG. Sythoff, Leiden 1970, p. 250.
[6] SMALL JR., A. M., Periodicity pitch, in "Foundations of modern auditory theory". Ed.: J. V. TOBIAS. Academic Press, New York 1970.
[7] FLETCHER, H., The physical criterion for determining the pitch of a musical tone. Phys. Rev. **23** [1924], 427.
[8] VON BÉKÉSY, G., The missing fundamental and periodicity detection in hearing. J. Acoust. Soc. Amer. **51** [1972], 631.
[9] SCHOUTEN, J. F., The perception of pitch. Philips Tech. Rev. **5** [1940], 286.
[10] DE BOER, E., On the residue in hearing. Doctoral dissertation, University of Amsterdam 1956.
[11] SCHOUTEN, J. F., RITSMA, R. J. and CARDOZO, B. L., Pitch of the residue. J. Acoust. Soc. Amer. **34** [1962], 1418.
[12] THURLOW, W. R. and SMALL, A. M., Pitch perception for certain periodic auditory stimuli. J. Acoust. Soc. Amer. **27** [1955], 132.
[13] SMALL, A. M. and CAMPBELL, R. A., Masking of pulsed tones by bands of noise. J. Acoust. Soc. Amer. **33** [1961], 1570.
[14] PATTERSON, R. D., Noise masking of a change in residue pitch. J. Acoust. Soc. Amer. **45** [1969], 1520.
[15] RITSMA, R. J., Frequencies dominant in the perception of the pitch of complex sounds. J. Acoust. Soc. Amer. **42** [1967], 191.
[16] BILSEN, F. A., Repetition pitch; its implication for hearing theory and room acoustics, in "Frequency analysis and periodicity detection in hearing". Eds.: R. PLOMP and G. F. SMOORENBURG. Sythoff, Leiden 1970, p. 291.
[17] WALLISER, K., Zusammenhänge zwischen dem Schallreiz und der Periodentonhöhe. Acustica **21** [1969], 319.
[18] V. D. BRINK, G., Two experiments on pitch perception: Diplacusis of harmonic AM signals and pitch of inharmonic AM signals. J. Acoust. Soc. Amer. **48** [1970], 1355.
[19] SMOORENBURG, G. F., Pitch perception of two-frequency stimuli. J. Acoust. Soc. Amer. **48** [1970], 924.
[20] MATHES, R. C. and MILLER, R. L., Phase effects in monaural perception. J. Acoust. Soc. Amer. **19** [1947], 780.
[21] LICKLIDER, J. C. R., Effects of changes in the phase pattern upon the sound of a 16-harmonic tone. J. Acoust. Soc. Amer. **29** [1957], 780.
[22] RITSMA, R. J. and ENGEL, F. L., Pitch of frequency modulated signals. J. Acoust. Soc. Amer. **36** [1964], 1637.
[23] PATTERSON, R. D., Residue pitch as a function of the number and relative phase of the component sinusoids. J. Acoust. Soc. Amer. **48** [1970], 76.
[24] BILSEN, F. A. and RITSMA, R. J., Some parameters influencing the perceptibility of pitch. J. Acoust. Soc. Amer. **47** [1970], 469.
[25] BILSEN, F. A., On the interaction of a sound with its repetitions. Doctoral dissertation, Delft University 1968, p. 45.
[26] GOLOMB, S. W., Shift register sequences. Holden-Day Inc., San Francisco 1967.
[27] SCHROEDER, M. R., New results concerning monaural phase sensitivity. J. Acoust. Soc. Amer. **31** [1959], 1579.
[28] PLOMP, R., Timbre as a multidimensional attribute of complex tones, in "Frequency analysis and periodicity detection in hearing". Eds.: R. PLOMP and G. F. SMOORENBURG. Sythoff, Leiden 1970, p. 397.
[29] DUIFHUIS, H., Perceptual analysis of sound. Doctoral dissertation, University of Eindhoven 1972.
[30] GOLDSTEIN, J. L., Aural combination tones, in "Frequency analysis and periodicity detection in hearing". Eds.: R. PLOMP and G. F. SMOORENBURG. Sythoff, Leiden 1970, p. 230.