EINDHOVEN UNIVERSITY OF TECHNOLOGY

Edge pitch of complex sounds

Citation for published version (APA):
Kef, P. A., & Kloos, S. J. (1992). *Edge pitch of complex sounds: influences of loudness-level, fundamental frequency and edge frequency on the perception of the upper-edge frequency of complex sounds*. (IPO-Rapport; Vol. 857). Instituut voor Perceptie Onderzoek (IPO).

Document status and date:
Published: 24/06/1992

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Download date: 17. Nov. 2023

Institute for Perception Research
PO Box 513, 5600 MB Eindhoven

24.06.1992

Rapport no. 857

Edge pitch of complex sounds

P.A. Kef
S.J. Kloos

Edge pitch of complex sounds

Influences of loudness-level, fundamental frequency and edge frequency on the perception of the upper-edge frequency of complex sounds.

Peter Kef
Sebastiaan Kloos
Hogeschool Eindhoven
Applied Physics
Eindhoven, June 1992

Edge pitch of complex sounds

Influences of loudness-level, fundamental frequency and edge frequency on the perception of the upper-edge frequency of complex sounds.

Authors : Peter Kef
Sebastiaan Kloos

Mentors : Prof. Dr. A.J.M. Houtsma (IPO)
Dr. A. Kohlrausch (IPO)
Ir. C. Schumm (Hogeschool Eindhoven)

Autograph :

Company : Philips
Institute for Perception Research (IPO)

Institute : Hogeschool Eindhoven
Applied Physics

Eindhoven, June 1992

Preface

As a finishing of our study at the 'Hogeschool Eindhoven' we performed an experimental study on the perception of the upper-edge pitch of complex tones. This research is a continuation of earlier work from Kohlrausch and Houtsma (1991), and it can be split into two different parts. One part investigated the dependence of the sound pressure level of the complex on the upper-edge pitch. The other part investigated the subject's ability to perceive upper-edge pitch as a function of the frequency of the upper edge.
Peter supervised project 1, and worked out the results of this part, chapter 5.1 and the first part of the summary and discussion (chapter 6). Sebastiaan supervised the second project, and worked out the results of this part, chapter 5.2 and the second part of the summary and discussion (chapter 6). Furthermore Peter has written down chapter 1, 3, and 4, while Sebastiaan has written down chapter 2, and processed all the data.
Finally, we would like to thank the company, Philips and the Institute for Perception Research (IPO, Prof. Dr. H. Bouma), for giving us the opportunity to finish our study at IPO. Moreover we like to thank Marcel van der Heijden and Niek Versfeld for being subjects in our experiments. Especially we like to thank Adrian Houtsma and Armin Kohlrausch for being our coaches, and also for being subjects in our experiments.

Summary

Complex tones with a low-pass spectrum can produce a pitch sensation related to the upper-edge frequency. This upper-edge tone has been studied previously as a function of the order of the highest harmonic, the fundamental frequency of the complex and the phase values of the harmonics (Kohlrausch and Houtsma, 1991). In the part of the present study, the influence of the intensity of a complex signal on the pitch has been investigated in the first part of the experiments (project 1). In the second part (project 2) the influence of the frequency of the upper-edge tone and the frequency of the fundamental on the perceptivity of the edge tone has been investigated.
In the first project we measured the accuracy of the matches to complex tones and pure tones. The accuracy for pure tones (averaged over six subjects and four levels) is about 0.3 % (percentage of the nominal frequency), and increases slightly with decreasing sound level. For the complex signal the accuracy is about a factor 3 worse. The increase of the accuracy for decreasing levels is stronger than for the pure tones.
In this project we also measured the shift of the pitch as a function of the level. For a pure tone there was a small positive shift with increasing level, in the range of previously published results (Terhardt, 1979). For a complex signal there was a strong increase of the shift with the level. The main reason for this slope can be found in partial masking.
In the second project we found that the edge pitch for complex signals with low fundamentals (< 100 Hz) is perceived similar to, though somewhat less accurate, than the shift of pure tones. For complexes with a fundamental frequency of 200 Hz, three subjects could hear the edge pitch up to a frequency of about 3.5 kHz. Two other subjects were able to hear the pitch up to about 7.2 kHz. These results for very high frequencies show that there is a combined effect of upper-edge frequency and fundamental frequency on the subject ability to perceive the edge pitch at frequencies about 3 kHz.

Contents

1 Introduction

We finished our study at Philips Research Laboratory, Eindhoven, at the Institute for Perception Research (IPO), in the group hearing and speech. In this group one not only deals with the pure physical aspects of sound, but mainly with the human perception of sound. We are entering the world of psychoacoustics. This interdisciplinary field involves the academic disciplines of physics, biology, psychology, music, audiology, and engineering, and utilizes principles from each of them. This combined effect has greatly helped our understanding of sound perception which has increased substantially in recent years.

Loudness, pitch, timbre, and duration are four perceptual attributes used to describe sound, especially musical sound. These attributes depend in a rather complex way on measurable quantities such as sound pressure, frequency, spectrum of partials, duration, and envelope.

The relationship between the subjective attributes of sound and physical quantities is the central problem of psychoacoustics and it has received a great deal of attention in recent years. As a continuation of earlier work (Kohlrausch and Houtsma, 1991), we performed a study of how people perceive the upper-edge tone of a complex signal. In this report there has been made a split into two projects. The complex signal used in the first project, consisted of the first twenty harmonics of a low fundamental $f_0 \approx$ 50 Hz, and therefore had an upper-edge frequency around 1 kHz. Human listeners can hear this upper-edge frequency in the form of a clear tone-like pitch. We were especially interested in the dependence of this pitch on the sound pressure level and the spectral composition of the complex signal. The perceive pitch of a sinusoidal (pure) tone depends on the sound pressure level (Terhardt, 1979). For low frequencies ($\approx<1000$ Hz) there is a negative pitch shift with increasing sound pressure level. For high frequencies ($\approx>2000$ Hz) there is a positive pitch shift with increasing sound pressure level. These shifts are averaged shifts over several subjects. Shifts are subject dependent, and large intersubject variations are no exception. We measured how the level of the complex influenced the perception of the edge tone of complex signals.

The same kind of harmonic complexes were used for the second project, in which we tried to locate the highest frequency at which the pitch was still perceivable. For a normal pure tone this will be in the region between twelve and twenty kHz, depending on the age of the subject.

We used six subjects, including the two authors, to perform several tests. The results

of the tests will be used to test some existing hypothesis. If necessary these hypothesis will be adjusted to give more clearness in our understanding of the way our ear works.
In this report the reader will find some background anatomy about and functioning of the ear in chapter two, and in chapter three there will be a more specific description about pitch perception. Chapter four describes the way we did our experiments, after which the results follow in chapter five. Finally chapter six contains the discussion.

2 Hearing

The human auditory system is complex in structure and remarkable in function. Not only does it respond to a wide range of stimuli, but it can precisely identify the pitch and timbre (quality) of a sound and even the direction of the source. Much of the hearing function is performed by the organ we call the ear, but recent research has emphasized how much hearing depends on the data processing that occurs in the central nervous system.

2.1 Structure of the ear

To describe the ear we first divide it into three sections: the outer ear, the middle ear, and the inner ear (see Fig. 2.1)

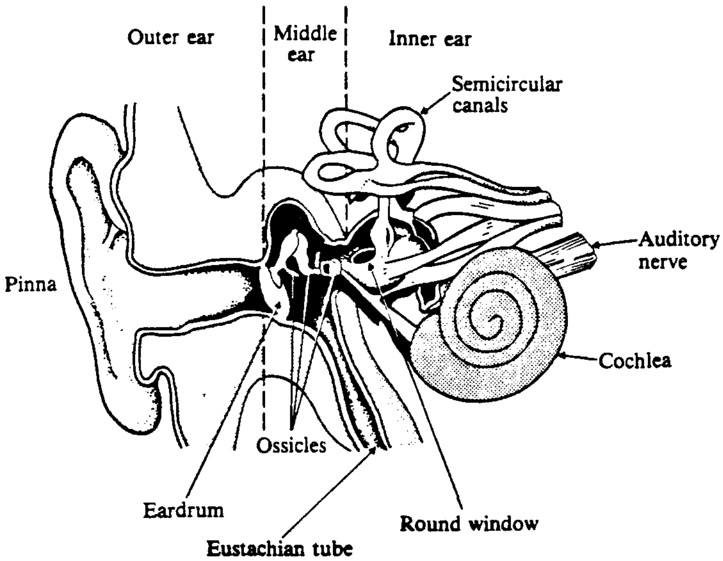


Fig. 2.1: *A schematic diagram of the ear, showing outer, middle, and inner regions. This drawing is not to scale; for purposes of illustration, the middle ear and inner ear have been enlarged (Rossing, 1990).*

The *outer ear* consists of the external *pinna* and the *ear canal* (meatus), which is terminated by the *eardrum* (tympanum).

The *middle ear* begins with the eardrum, to which three small bones called *ossicles* are attached. The eardrum changes the pressure variations of incoming sound waves into mechanical vibrations to be transmitted via the ossicles to the inner ear.

The ossicles perform a very important function in the hearing process. Together they act as a lever, which changes the pressure exerted by a sound wave on the eardrum into a much greater pressure on the oval window of the inner ear. This pressure gain is further increased by the effective area reduction when going from the eardrum to the oval window of the cochlea, so that together the pressure can increase by a factor 30. Another function of the small bones and their attached ligaments and muscles is to protect the inner ear from loud noises and sudden pressure chances.
Since the eardrum makes an airtight seal between the middle ear and the outer parts of the ear, it is necessary to provide pressure equalization. This is established by the *Eustachian tube*, which connects the middle ear to the oral cavity.
The complex *inner ear* contains the *semicircular canals* and the *cochlea.* The semicircular canals, although anatomically part of the ear, contribute little or nothing to hearing. They are the body's horizontal-vertical and rotation detectors necessary for balance. The spiral cochlea contains the mechanisms for transforming pressure variations into properly coded neural impulses. The cross section of the cochlea is divided into 3 distinct chambers that run the entire length (see Fig. 2.2): the *scala vestibuli*, the *scala tympani*, and the *cochlear duct*. The cochlea is filled with two different liquids, called perilymph (in the scala vestibuli and the scala tympani) and the endolymph (in the cochlear duct), and is surrounded by rigid bony walls. The two liquids are kept separate by two membranes, the *Reissner's membrane* and the *basilar membrane*.

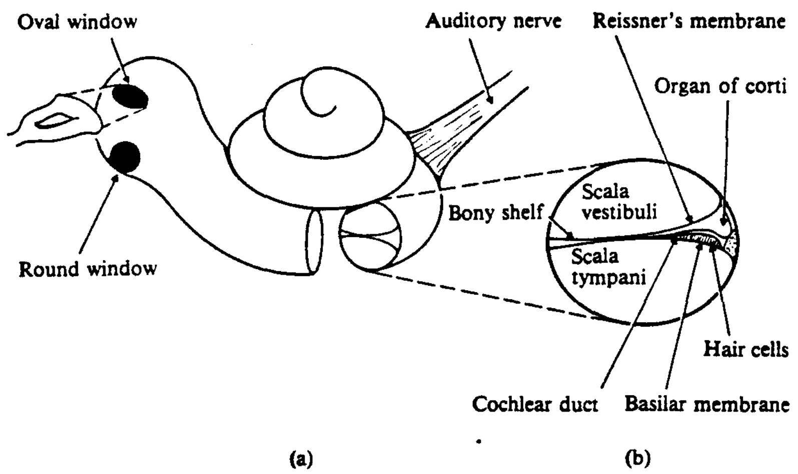


Fig. 2.2: *A schematic diagram of (a) the cochlea; (b) a section cut from the cochlea (Rossing, 1990)*

Resting on the basilar membrane is the delicate and complex *organ of Corti*. This "seat of hearing" contains several rows of tiny *hair cells* to which are attached nerve fibers. A single row of inner hair cells contains about 7000 cells, whereas about 24,000 outer hair cells occur in several rows. Each hair cell has about 40 to 1000 small hairs at its apex. The inner cells are the actual sound or vibrator transducers. The outer hair cells appear to play a role in the efferent system and are probably responsible for the active filtering processes observed in the cochlea.

2.2 Cochlear mechanisms

We have already discussed the manner in which the conductive mechanism influences the signal and transmit it to the inner ear. In this paragraph we will concentrate upon the sensory mechanism. The cochlea may be conceived of as a transducer that converts the vibratory stimulus into a form usable by the nervous system. About the way this happens are two main theories, the *place theory* and the *time* or *periodicity* theory. In the next two subparagraphs we will discuss these two theories.

2.2.1 Place theory

The place or resonance theory is an old theory based upon Ohm's auditory "law" (Ohm, 1843) and Müller's doctrine of specific nerve energies. Ohm's auditory law states that the ear performs a Fourier analysis upon complex periodic sound. It breaks the complex wave down into its components regardless of their phase relationships. The modern version of this resonance theory began with *Helmholtz resonance theory* (Helmholtz, 1863). It assumes that the basilar membrane is composed of a series of segments, each of which resonates in response to a particular frequency. Thus, an incoming stimulus results in the vibration of those part of the basilar membrane whose natural frequencies correspond to the components of the stimulus. Since these resonators are arranged by place along the cochlear partition, the precise place of the vibrating segment would signal the existence of a component at the natural frequency of that location.

Such a strict theory described above is faced with several serious problems. The first problem is the sharp frequency tuning in the inner ear. The theory demands that the basilar membrane contain segments which are under differing amounts of tension.

However, Békésy (1960) demonstrated that the basilar membrane is under no tension at all. A second problem is the perception of the missing fundamental, the phenomenon in which the presence of higher harmonics of a tone (e.g. 800, 1000, and 1200 Hz) results in the perception of the fundamental frequency (200 Hz) even though the latter is not physically present.

2.2.2 Periodicity or time theory

The periodicity or time theory states that the cochlea is not frequency-sensitive along its length, but rather that all parts respond to all frequencies. The job of the hair cells is simply to transmit all parameters of the stimulus waveform to the auditory nerve.
Since a neuron can only respond in a digital manner (all or none), the only way in which it can transmit frequency information is to discharge the same number of times per second as the frequency of the stimulus. Time theory thus presumes that the auditory nerve fibers can fire fast enough to represent this information. There is no problem at low frequencies; however, an upper limit on the number of discharges per second is imposed by the absolute refractory time of the neuron. The absolute refractory time is the time required after discharging for the cell to reestablish the polarization it needs to fire again; it lasts about 1 msec. The 1 msec absolute refractory time corresponds to a maximum frequency of 1000 Hz. Thus, simple time theory is hard pressed to explain how sounds higher in frequency than about 1000 Hz can be transmitted by the auditory nerve and perceived by the listener.
A second problem of this theory is that the damage to the apical parts of the cochlea results in high frequency hearing loss. This is contradictory to the time theory.
In order to agree with these problem the frequency theory was modified with the volley principle advanced by Wever. Instead of suggesting that each neuron must carry all the information burden, the volley principle states that groups of fibers cooperate to represent the stimulus frequency in the auditory nerve. This is shown in figure 2.3.

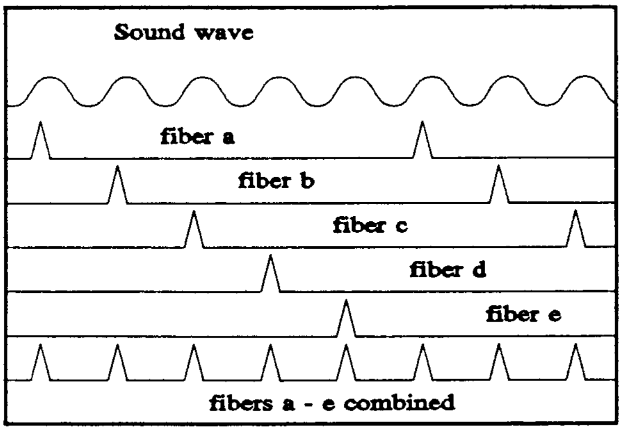


Fig. 2.3: *Diagrammatic representation of the volley principle*

The sinusoid (sound wave) at the top of the figure has a frequency too high to be represented by a series of spike discharges from a single auditory nerve fiber. Instead fibers work together in groups so that the total response of the group (at the bottom of the figure) corresponds to the temporal waveform of the sound signal.

2.3 Sound pressure and loudness

In this paragraph we will discuss the subjective quality of loudness and the physical parameters that determine it. Related to the sound pressure are the sound power emitted by the source and the sound intensity (the rate of energy flow per unit area in a sound wave). The sound pressure can be measured directly, however, and our ears respond to sound pressure.

2.3.1 Sound pressure level

In a sound wave there are extremely small periodic variations in atmospheric pressure to which our ears respond in a rather complex manner. The minimum pressure fluctuation to which the ear can respond is less then a billionth (10^{-9}) of the atmospheric pressure. This threshold of audibility, which varies from person to person, corresponds to an rms sound pressure value of about 2×10^{-5} N/m^2 at a frequency of 1000 Hz. The threshold of pain corresponds to a pressure amplitude approximately one million times greater, but still 1/1000 of the atmospheric pressure. Because of the wide range of pressure stimuli, it is convenient to measure sound pressure on a logarithmic scale, called the *decibel* (dB) *scale*. The decibel scale of sound is defined by comparing sounds to a reference sound with a pressure amplitude $p_0 = 2 \times 10^{-5}$ N/m^2, to which a sound pressure level of 0 dB is assigned (see formula 2.1).

$$L_p = 20 \log p/p_0 \qquad \text{......... (2.1)}$$

Sound pressure levels of a number of sounds are given in Table 2.1

Table 2.1: *Typical sound levels*

Jet takeoff (60 m)	120 dB	
Construction site	110 dB	*Intolerable*
Shout (1.5 m)	100 dB	
Heavy truck (15 m)	90 dB	*Very noisy*
Urban street	80 dB	
Automobile interior	70 dB	*Noisy*
Normal conversation (1 m)	60 dB	
Office, classroom	50 dB	*Moderate*
Living room	40 dB	
Bedroom at night	30 dB	*Quiet*
Broadcast studio	20 dB	
Rustling leaves	10 dB	*Barely audible*
	0 dB	

2.3.2 Loudness level

Although a sound with a greater L_p usually sounds louder, this is not always the case. The sensitivity of the ear varies with the frequency and the quality of the sound. Figure 2.4 shows curves of equal *loudness level* for pure tones. These curves demonstrate the relative insensitivity of the ear to sounds of low frequency and moderately low intensity levels.

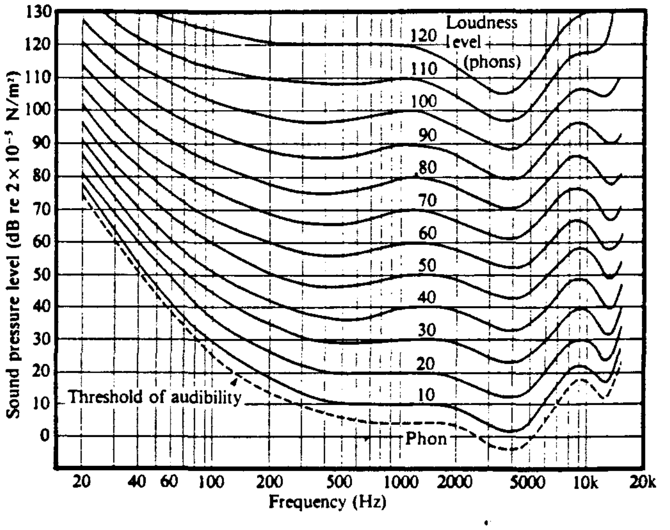


Fig. 2.4: *Equal-loudness curves for pure tones (frontal incidence). The loudness levels are expressed in phons (Fletcher and Munson, 1933).*

Hearing sensitivity reaches a maximum between 3500 and 4000 Hz, which is near the first resonance frequency of the outer ear canal, and a second relative maximum around 13 kHz, the frequency of the second resonance.

2.3.3 Loudness of pure tones

In an effort to obtain a quantity proportional to the loudness sensation, a loudness scale was developed in which the unit loudness is called the *sone*. The sone is defined as the loudness of a 1000 Hz tone at a sound level of 40 dB (a loudness level of 40 phons).

For loudness levels of 40 phons or greater, the relationship between loudness S in sones and loudness level L_s in phons recommended by the International Standards Organization (ISO) is:

$$S = 2^{(L_s-40)/10} \qquad \text{..........(2.2)}$$

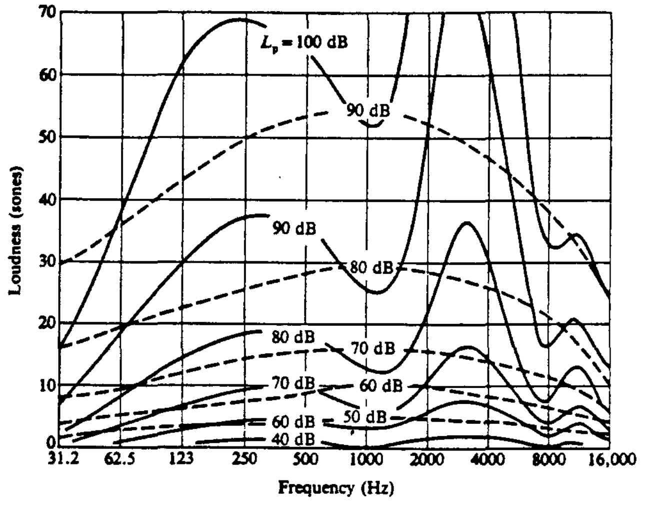


Fig. 2.5: *Subjective loudness of pure tones (solid curves) and "musical" tones with five harmonics (dashed curves) as a function of frequency and sound level (L_p)(Rossing, 1990).*

One way to represent the loudness graphically is to combine the above expression for loudness with the curves of equal loudness given in figure 2.4. The solid curves in figure 2.5 make use of eq. 2.2 to give the loudness of pure tones at different frequencies.

2.3.4 Loudness of complex tones: critical bands

As we pointed out in table 2.1 and figure 2.4, loudness depends mainly on the sound pressure but it also varies with frequency, spectrum, and duration. Broadband sounds seem louder for instance, than pure tones or narrowband noise having the same sound pressure level. Figure 2.6 illustrates the dependence of loudness of broadband noise on bandwidth with fixed sound pressure level and center frequency. Note that the loudness is not affected until the bandwidth exceeds a *critical bandwidth*, 160 Hz for this frequency.

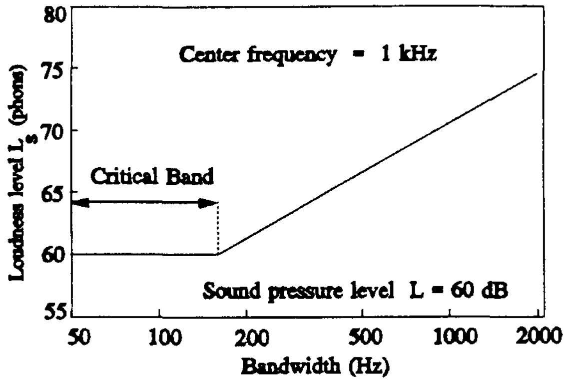


Fig. 2.6: *The effect of bandwidth on loudness*

The concept of critical bands is of great importance in our understanding of the mechanism of hearing. They have considerable significance in experiments on pitch discrimination, loudness, musical consonance, etc. Each critical band may be regarded as a data unit on the basilar membrane. About 24 critical bands span the audible frequency range. The critical bandwidth varies with center frequency, as shown in figure 2.7

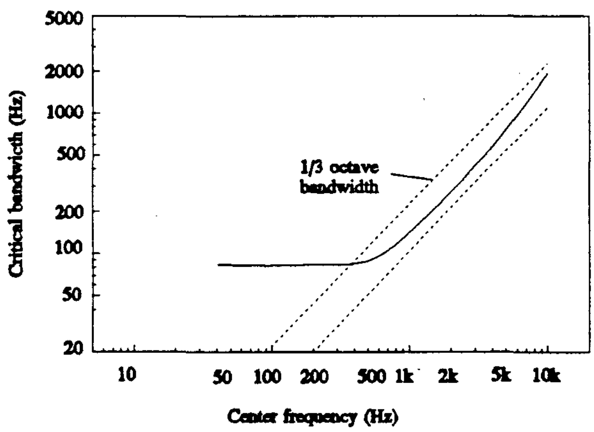


Fig. 2.7: *Critical bandwidth as a function of the critical band center frequency.*

A complete curve of the difference limen obtained by frequency modulation would show that the ear can perceive frequency changes of about 1/30 of the critical bandwidth. One suggestion, supported by experimental evidence, is that the ear may recover frequency from the long-term combined rate at which the nerves fire. Even though each nerve usually fires with a repetition rate far below the detected frequency, if there are many nerves firing at low rates, the collection of all the nerve firings will yield "volleys" generated with a repetition rate equal to the frequency.

2.4 Masking

When the ear is exposed to two or more different tones, it is a common experience that one may mask the other. Masking is probably best explained as an upward shift in the hearing threshold of the weaker tone by the louder and depends on the frequencies of the two tones. Pure tones, complex sounds, narrow and broad bands of noise all show differences in there ability to mask other sounds. Masking of one sound can even be caused by another sound that occurs a split second before or after the masked sound. From the many masking experiments that have been performed some interesting conclusions can be drawn:

1 Pure tones close together in frequency mask each other more than tones widely separated in frequency.

2 A pure tone masks tones of higher frequency more effectively than tones of lower frequency.

3 The greater the intensity of the masking tone, the broader the range of frequencies it can mask.

4 Masking by a narrow band of noise shows many of the features as masking by a pure tone.

5 Masking of tones by broadband ("white") noise shows an approximately linear relationship between masking and noise level. Broadband noise masks tones of all frequencies.

Some of the conclusions about masking just stated can be understood by considering the way in which pure tones excite the basilar membrane. High-frequency tones excite the basilar membrane near the oval window, whereas low-frequency tones create their greatest amplitude at the far end near the apex. The excitation due to a pure tone is asymmetrical having a tail that extends toward the high frequency end as shown in figure 2.8. Thus it is easier to mask a tone of higher frequency than one of lower frequency. As the intensity of the masking tone increases, a greater part of its tail has an amplitude sufficient to mask tones of higher frequency.

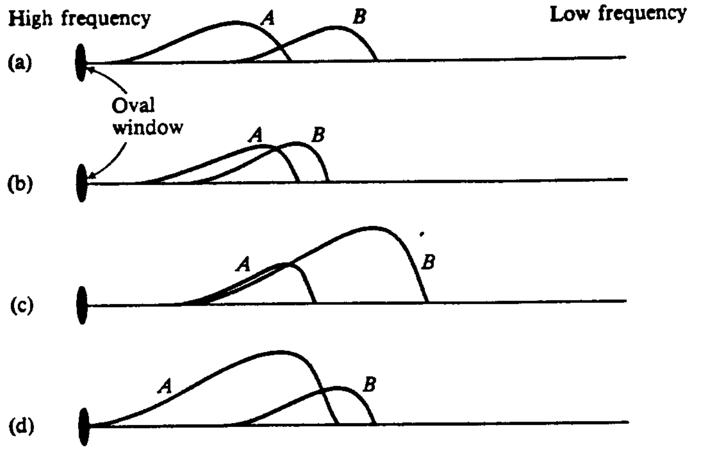


Fig. 2.8: *Simplified response of the basilar membrane for two pure tones A and B (a) The excitations barely overlap; little masking occurs. (b) There is an appreciable overlap; tone B masks tone A and somewhat more then the reverse. (c) The more intense tone B almost completely masks the higher-frequency tone A. (d) The more intense tone A does not completely mask the lower-frequency tone B (Rossing, 1990).*

3 Pitch

Pitch has been defined as that characteristic of a sound that makes it sound high or low or that determines its position on the musical scale. For a pure tone, the pitch is determined mainly by the frequency, although the pitch of a pure tone may also change with sound level. The pitch of complex sounds also depends on the spectrum (timbre) of the sound and its duration. In the following paragraphs, we will go deeper into some of these subjects. In fact, the pitch of complex sounds has been one of the most interesting objects of study in psychoacoustics for quite a few years.

3.1 Pitch discrimination

The ability to distinguish between two nearly equal stimuli is often characterized, in psychophysical studies, by a *difference limen* or *just noticeable difference (jnd)*. Two stimuli will be judged "the same" if they differ by less than a jnd. This judgement depends somewhat on the musical training of the listener and to some extent on the method of measurements. The jnd for pitch has been found to depend on the frequency, the sound level, the duration of the tone, and the suddenness of the frequency change (Zwicker, Flottorp, and Stevens, 1957). Figure 3.1 shows the average jnd (of four subjects) for pure tones at a sound level of 80 dB.

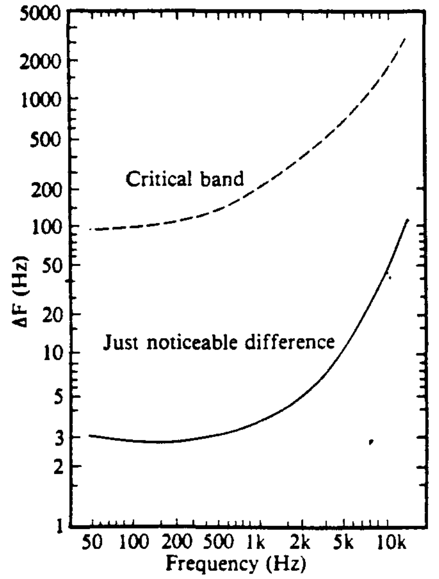


Fig. 3.1: *Just noticeable difference (jnd) in frequency determined by modulating the frequency of a tone at 4 Hz. Note that the jnd at each frequency is nearly a constant percentage of the critical bandwidth (Zwicker, Flottorp, and Stevens, 1957).*

By comparing the upper and lower curves in figure 3.1 we can see that critical bandwidth is roughly equal to 30 jnd's at all frequencies. This remarkable result *suggests* that the same mechanism in the ear is responsible for critical bands and for pitch discrimination. It is quite likely related to regions of excitation along the basilar membrane.

3.2 pitch and intensity

Early experiments (Stevens, 1935) on pitch versus sound level reported substantially larger pitch dependence on sound level then more recent studies do (Ward, 1970). Tones of low frequency were found to fall in pitch with increasing intensity, tones of high frequency rise in pitch with increasing intensity. Tones of middle frequency (≈1-2 kHz) show little change (Terhardt, 1979). New experiments showed that the average effect is small, even for pure tones, and varies from observer to observer. Figure 3.2 shows the pitch shifts of pure tones with frequencies from 200 Hz to 6000 Hz averaged over 15 subjects.

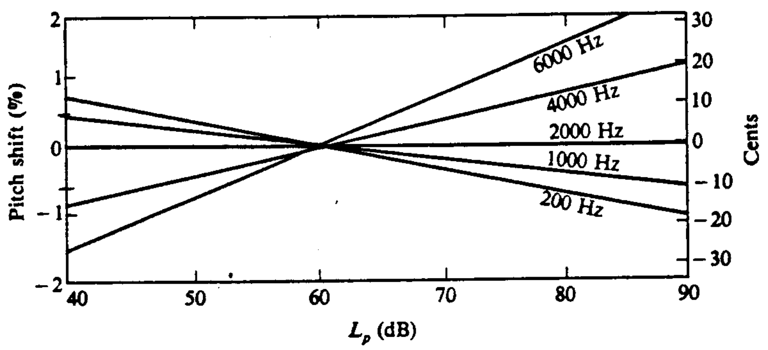


Fig. 3.2: *Pitch shift of pure tones as a function of sound pressure level. Shifts are shown both as percentage of the reference frequency and in the musical measure cents (100 cent = 1 semitone). The curves are based on data from 15 subjects (Terhardt, 1979).*

The small pitch changes shown in figure 3.2 as well as the larger changes described by early investigators, are valid for pure tones. Less is known about the effect for complex tones. Studies with musical instruments have generally shown very small pitch changes with intensity (around 1% maximum for an increase from 65 to 95 dB) (Rossing, 1990).

It is fortunate for performing musicians and listeners alike that the change in pitch with sound level is much less than was reported from early experiments with pure tones. Musical performance would be very difficult if substantial changes of pitch occurred during changes in dynamic level.

3.3 Maximum audible frequency

For young persons with unimpaired hearing the upper limit of hearing lies in the region of ≈24000 Hz. For older people this boundary lies in the region of ≈12000 Hz. Unlike the lower limit, which is in the region of 15 Hz, the upper limit is a terminal point of sensitivity; beyond it nothing at all is heard.

About the upper limit of the edge pitch of a complex signal not much is known. There will probably be some masking effect, so that the limit may be lower.

3.4 Edge frequency of complex tones

There are many ways to generate a signal that has an upper-edge frequency. In the early days noise signals were used which were cut off sharply at a certain frequency. Later on experiments have been done with signals containing several sinusoidal components. In our experiments we have used signals which contained all harmonics of the fundamental frequency f_0 up to a certain maximum. No harmonics were missing. These harmonics all had the same amplitude and "zero" starting phase, and can mathematically be represented by the expression;

$$S(t) = \sum_{k=1}^{n} \sin(k\omega t) \quad , \qquad \text{..........(3.1)}$$

when n is the total number of harmonics, k is the harmonic number, and $\omega = 2\pi f_0$.

This expression can be written as:

$$S(t) = Im\left(\sum_{k=1}^{n} e^{jk\omega t}\right) \quad , \qquad \text{..........(3.2)}$$

this can be written as:

$$S(t) = Im\left(\frac{1 - e^{j\bar{n}\omega t}}{1 - e^{j\omega t}}\right) \quad , \qquad \text{..........(3.3)}$$

with: $\bar{n} = n + 1$.

Which is equal to:

$$S(t) = Im\left(\frac{e^{j\frac{\pi}{2}\omega t}\left(e^{-j\frac{\pi}{2}\omega t} - e^{j\frac{\pi}{2}\omega t}\right)}{e^{j\frac{\omega}{2}t}\left(e^{-j\frac{\omega}{2}t} - e^{j\frac{\omega}{2}t}\right)} \right) \quad , \qquad \text{..........(3.4)}$$

rewriting gives:

$$S(t) = Im\left(e^{j\frac{\pi}{2}\omega t}\right) \frac{\sin\left(\frac{n+1}{2}\omega t\right)}{\sin\left(\frac{\omega}{2}t\right)} \quad , \qquad \text{..........(3.5)}$$

which is the same as:

$$S(t) = \frac{\sin\left(\frac{n}{2}\omega t\right) \cdot \sin\left(\frac{n+1}{2}\omega t\right)}{\sin\left(\frac{\omega}{2}t\right)} \quad , \qquad \text{..........(3.6)}$$

so eventually we get

$$S(t) = \frac{\cos\left(\frac{\omega}{2}t\right) - \cos\left(\left(n+\frac{1}{2}\right)\omega t\right)}{2\sin\left(\frac{\omega}{2}t\right)} \quad , \qquad \text{..........(3.7)}$$

In this last formula we see three components of which two have a small frequency ($f_0/2$) and one has a large frequency (f_0*(n+0.5)), this large frequency represents the frequency of the upper edge! In figure 3.3 the three components are separately shown as function of the time (f(t)), in 3.4 a plot of the "total complex signal" (S(t)) is shown.

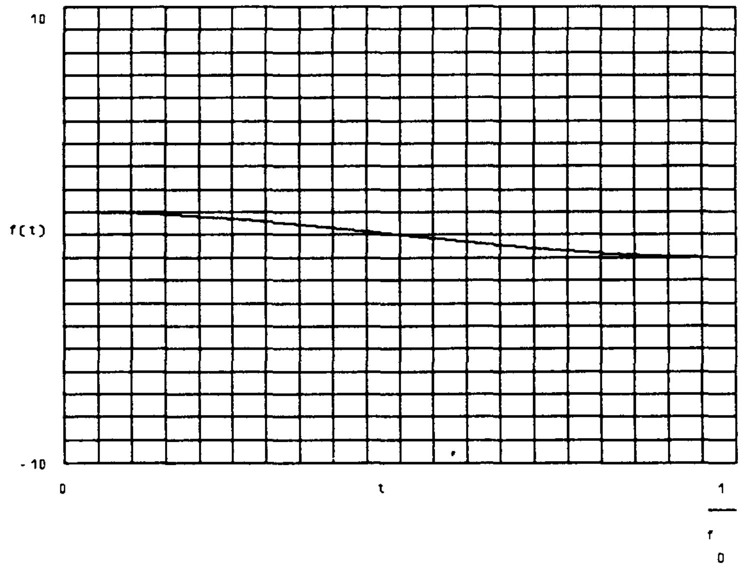


Fig. 3.3a: *First component of (3.7)*

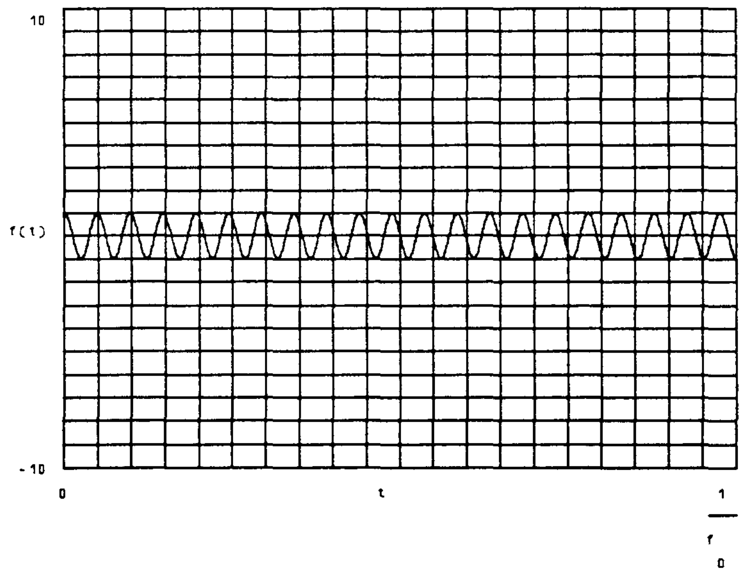


Fig. 3.3b: *Second component of (3.7)*

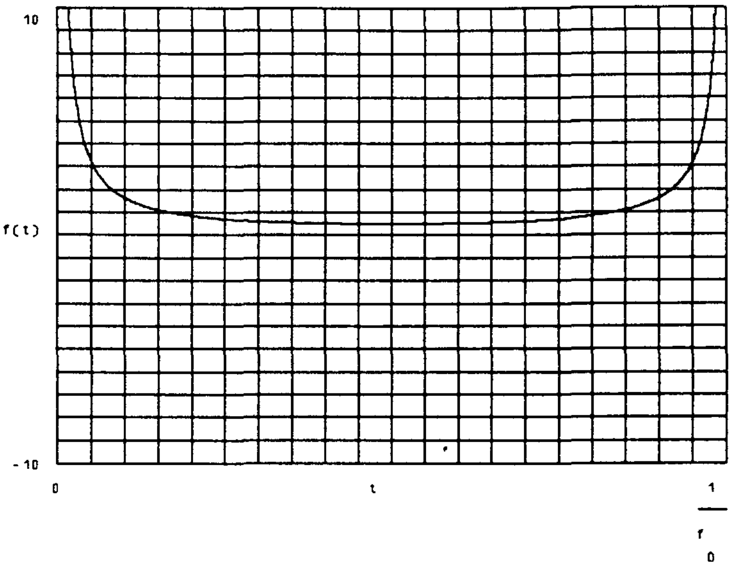


Fig. 3.3c: *Third component of (3.7)*

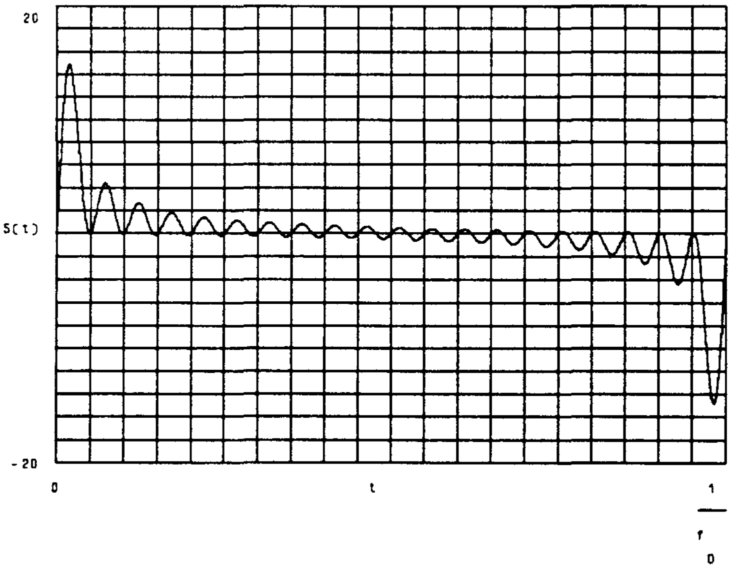


Fig. 3.4: *Total complex signal S(t) for n = 20, (3.7)*

The main difference between these stimuli and band-limited noise is found in the temporal waveform, the pure tone parts of the waveform alternate with peaks, and their relative duration within each period depends on the harmonic number of the edge component, as shown in figure 3.3 and 3.4. In figure 3.5 this relation is illustrated again, yet by two signals. The left signal has a fundamental frequency of 100 Hz and the highest harmonic (n) has order 20, while for the other signal, with a fundamental of 250 Hz, the order is only 8. The upper edge is thus 2050 Hz, respectively 2125 Hz.

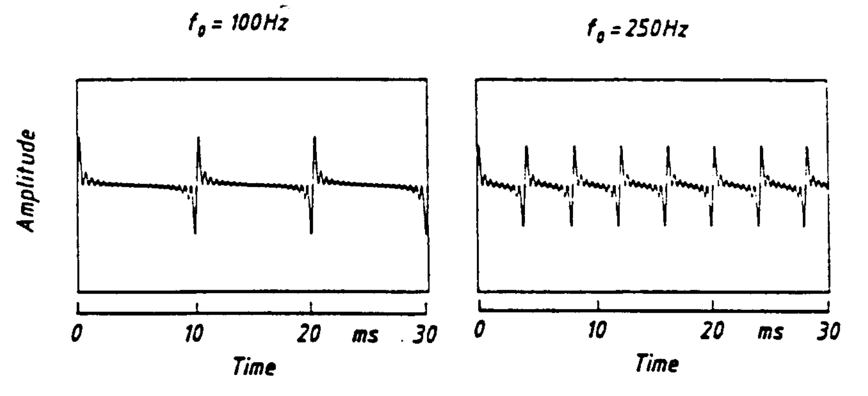


Fig. 3.5: *Time functions of two harmonic complexes with an upper edge frequency of about 2 kHz. The left-hand panel shows a complex with 100 Hz fundamental frequency, consisting of 20 harmonics, the right-hand panel shows a complex with 250 Hz fundamental and 8 harmonics (Kohlrausch and Houtsma, 1991).*

An alternative way to emphasize the spectro-temporal properties of complex tones is a short-time Fourier analysis. The two panels in figure 3.6 show such an analysis for the two sounds from figure 3.5.

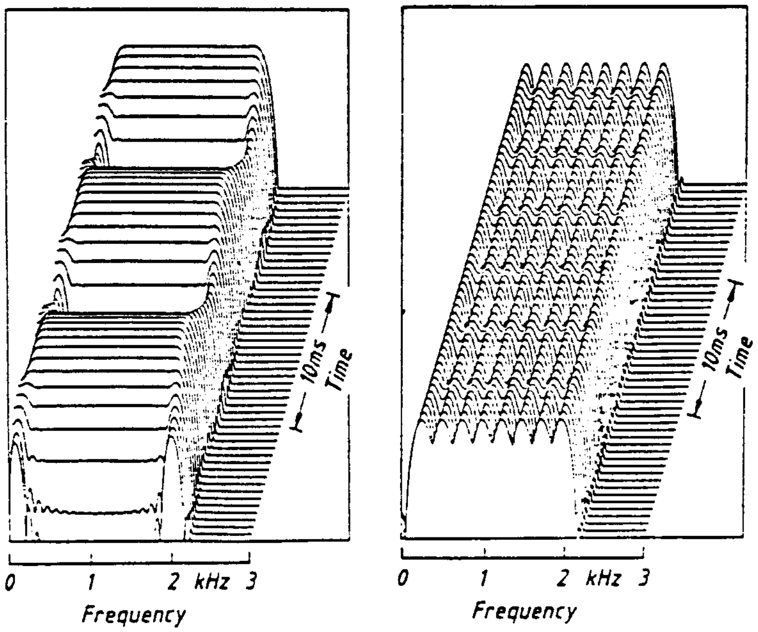


Fig. 3.6: *Short-time spectra of the two sounds from fig 3.5. The spectra are calculated using a Hanning window of 10 ms duration (Kohlrausch and Houtsma, 1991).*

The short-time spectra in both panels are calculated using a Hanning window of 10 ms duration. A Hanning window means a multiplication with a raised cosine, so that the borders at the window have a low weightfactor. This means that if the top of the window is placed between the peaks of the signal, the Fourier analysis shows mainly the fundamental and edge frequency. The time axis, running from front to back, covers approximately 25 ms. The temporal shift between adjacent spectra is 0.5 ms. Calculating one spectrum, however, takes at least 10 ms (f_0 = 100 Hz), so that the Fourier analysis is not real time. The left-hand.panel for the complex with 100 Hz fundamental reveals two different spectral patterns alternating in time. One is a flat spectrum showing the spectral extent of the complex from 100 to 2000 Hz. This pattern results from those time points where the window is centered on a pulse of the temporal waveform. The other pattern (visible e.g. in the first spectrum) shows

distinct maxima at the edge frequencies, slightly amplitude modulated in time. This pattern results from window positions in between the peaks.
Such an alternating pattern is only seen if the window covers no more than one period of the sounds. In the left-hand panel, the total window length is exactly equal to one period. The effective window duration, however, is shorter due to the raised-cosine ramps of the Hanning window. In the right-hand panel of figure 3.6, the same calculation is performed for the complex with 250 Hz fundamental. In this case, the window length is 2.5 times the period of the time signal and therefore, the short-time spectra emphasize the individual harmonics of the complex. In other words, the creation of a spectral peak at the edge frequency as shown in the left-hand panel of figure 3.6 will depend on the effective time constance of the analyzing system, or at the duration of the Hanning window (Kohlrausch and Houtsma, 1991).
One of our projects will try to give more insight of how the ear uses a similar mechanism as described in this paragraph. This as a continuation of the work done by Kohlrausch and Houtsma (1991).

4 Method

4.1 Acoustic stimuli

The whole project can be split into two separated projects. In the first project the influence of the sound pressure level on the perceived edge pitch was investigated. The test sounds consisted of equal-amplitude sinusoids with a common fundamental frequency f_0. All 20 components below the upper-edge frequency were present in the complex, and had zero starting phase. To avoid memory effects during the experiment, complexes of five different fundamental frequencies were used: 46, 48, 50, 52 and 54 Hz. The stimuli had a duration of 500 ms and were shaped with 20 ms raised-cosine ramps. They were presented diotically to the subjects via headphones (TDH 49) at different sound levels (10-70 dB SL).

Sinusoidal comparison tones had the same duration (500 ms) and alternated with the test sounds with a 500-ms silent interval in between. The frequency of the comparison tone was adjusted by the subjects with an unmarked ten-turn potentiometer, which controlled the frequency of an oscillator (Philips PM5190). The potentiometer received a random offset at the beginning of each new measurement in order to avoid systematic errors. There was no temporal limit for the matching procedure, but the time for a typical match was about 30 s for a trained person and up to 120 s for an untrained person. The finally adjusted frequency was taken as the data point for further evaluation.

Six subjects, among whom were the two authors, participated in this first project. At each complex-sound level, they performed six matches for each of the five complex signals (upper edge frequencies of 943, 984, 1025, 1066 and 1107 Hz). Each subject had to perform pitch matches for at least four different sound levels of the complex tone, at two different sound levels of the comparison tone, 55 dB SPL and 41 dB SPL. For one subject, AH, some data from a previous experiment have been used. The conditions of this previous experiment where the same, except for the level of the comparison tone, which was at 50 dB SPL.

In an additional reference experiment, the influence of the sound pressure level on the perceived pitch for a sinusoidal test signal was investigated. Tones were generated at five frequencies, 920, 960, 1000, 1040, and 1080 Hz. This reference experiment (tone-on-tone match) was performed for at least four different sound levels of the test sinusoid, and one fixed level of the comparison tone, 55 dB SPL.

From the six matches per complex-tone level and edge frequency, the mean, the shift to the nominal edge frequency, and the unbiased standard deviation were calculated.
The second project used the same kind of complexes, but the parameter now was the upper-edge frequency in the range from 2 kHz to 8 kHz. Complexes with these upper-edge frequencies were calculated for fundamental frequencies of 50, 100, and 200 Hz. The total number of harmonics was thus dependent on the upper-edge and the fundamental frequency. The stimuli had a duration of 500 ms and were presented diotically to the subjects via insert earphones (Etymōtic), which had a flat transfer function up to the highest frequencies used. The average sound pressure level of the complexes was 60 dB.
The sinusoidal comparison tones had the same duration (500 ms) and alternated with the test sounds with a 500-ms silent interval in between. The frequency of the comparison tone was adjusted by the subjects the same way as in the first project. The sound pressure level of the comparison tone was approximately 55 dB, so that its perceived loudness was similar to that of the edge tone.
In this second project again a reference experiment was included in which a pure tone had to be matched to the sinusoidal comparison tone. The frequency of this tone was also in the range 2 kHz to 8 kHz.
Five subjects participated in the experiments of this second project and performed 10 matches at each frequency for all four stimuli (pure-tone, f_0= 50, 100 and 200 Hz). For three subjects data from earlier experimental runs could be used. From the 10 matches the unbiased standard deviation was calculated. Evident mismatches (e.g. octave confusions) were omitted.

4.2 Used set-up

The complex stimuli were calculated by a computer and converted by a 16-bit D/A converter at a sampling rate of 20 kHz for the first project and 25 kHz for the second project. Next the complex signal went through three amplifiers and a lowpass filter with a cutoff frequency of 9 kHz. After the filter it was added to the comparison tone and split into a stereo signal. The comparison tone was generated by an analog sine generator, the frequency of which was controlled by the computer.

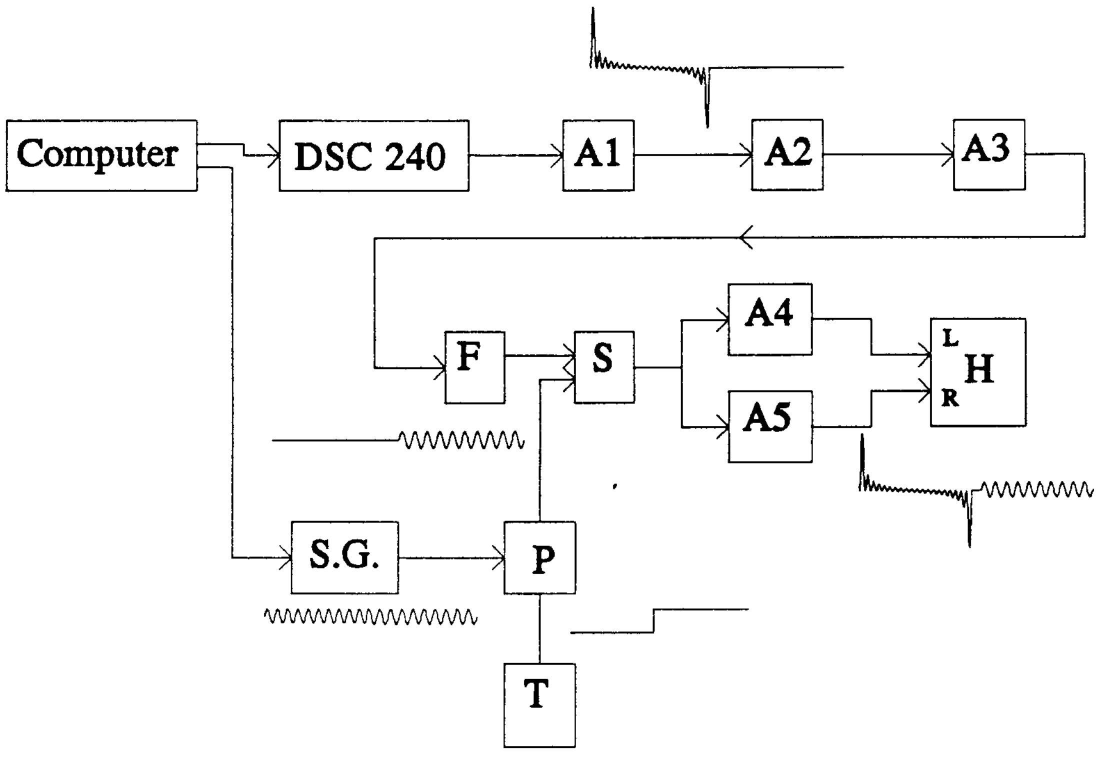


Fig 4.1 *Used equipment and set-up, with:* *DSC 240 = Digital Sound corporation D/A converter*
A1 = Analog amplifier PM 5170
A2 = DSC 240 analog amplifier
A3-5 = DSC 240 digital amplifier
F = Kemo vbf8 filter
S.G. = Sine Generator
P = Varidac port
T= Parallelle timer
S= Summator
H= Headphones
C= Computer

5 Results

We can describe the results separately for the two projects. The first project investigates the perception of the edge-pitch as a function of the sound pressure level, and the second project deals with the perception of the edge pitch as a function of the upper-edge frequency.

5.1 Project 1

In this first project we used the stimuli with the upper edge frequency around 1 kHz. The experimental data contain information about the *accuracy* of the matches and about the *absolute shift*, i.e. the difference between the nominal edge frequency of the complex and the averaged matched frequency. We investigated if there was a certain (low) level of the complex signal, below which the upper-edge frequency couldn't be perceived anymore. We also investigated the perception of pitch as a function of the sound pressure level of the complex signal. In the following sub-paragraphs we will discuss the results with respect to both aspects. Therefore we did two experiments with the complex signal, and one reference experiment with a pure tone signal.

5.1.1 Accuracy as a function of the sound pressure level

For the first experiment we used complex signals with f_0= 46, 48, 50, 52 and 54 Hz, containing the first 20 harmonics, which all had a zero starting phase. This gives five different upper-edge frequencies: 943, 984, 1025, 1066, and 1107 Hz. The various complex-tone levels are always adjusted relative to the subject's absolute threshold. Each subject first adjusted the absolute threshold for the complex, which is referred to as 0 dB SL (sensation level). The comparison tone had a fixed absolute sound pressure level of 55 dB SPL.
In the first two figures, 5.1 a and b, we see the standard deviation as a function of the level. The standard deviation is calculated in the following way: for all five upper-edge frequencies we calculated the standard deviation over the six data points as a percentage of the mean. These five standard deviations are averaged so that one average standard deviation is the result. Each panel shows results for three listeners.

It is clear that for some subjects (AK, PK, SK) there is only little change in the standard deviation with decreasing level. The standard deviation for these three subjects varies between 0.25 and 0.38 % (AK), 1.10 and 1.75 % (PK) and between 0.70 and 1.45 % (SK). For the other subjects (AH, MH, NV) there is a *stronger* level dependence; if the level of the complex sound decreases, the standard deviation increases by a factor 3 (MH) to 10 (NV) compared with the smallest standard deviation at higher levels. For some subjects the standard deviation also increases at the higher levels, although the effect is smaller than at low levels.

We see that for the complex-tone match two subjects (AH and AK) are able to adjust the comparison tone down to levels of 10 dB SL of the complex tone, three subjects (MH, NV and PK) down to 20 dB SL and one subject (SK) down to 30 dB SL. These differences could be due to different amounts of musical training, and to different experience with this kind of complex tones and other pitch experiments in general.

To see whether the effect of increasing standard deviation at decreasing level is somewhat dependent on the level of the comparison tone, we decreased the sound pressure level of the comparison tone. The decrease was 14 dB, all other parameters remained the same as in the first experiment. The results of this experiment are shown in figure 5.1 c and d. We see that the standard deviation at the lower levels of the complex is smaller than it was in the first experiment, a decrease from on average 1.2% to 0.7% (at the lowest level measured). We also see that the increase at low levels has become smaller.

One of the motivations for this experiment was to investigate how the perception of the upper edge changes, and probably disappears at low levels. But before we can say anything about the influence of the complex signal on the increase of the standard deviation at lower levels, we should first perform the same experiment for a pure tone, the reference experiment. In this experiment, the test sound is a sinusoid, presented at various levels, while the comparison tone is a sinusoid with a constant level of 55 dB SPL. These results are shown in figure 5.1 e and f, and again we see an increasing standard deviation with a decreasing sound pressure level. Notice that for this pure tone the standard deviation is about a factor 2 lower than for the complex tone.

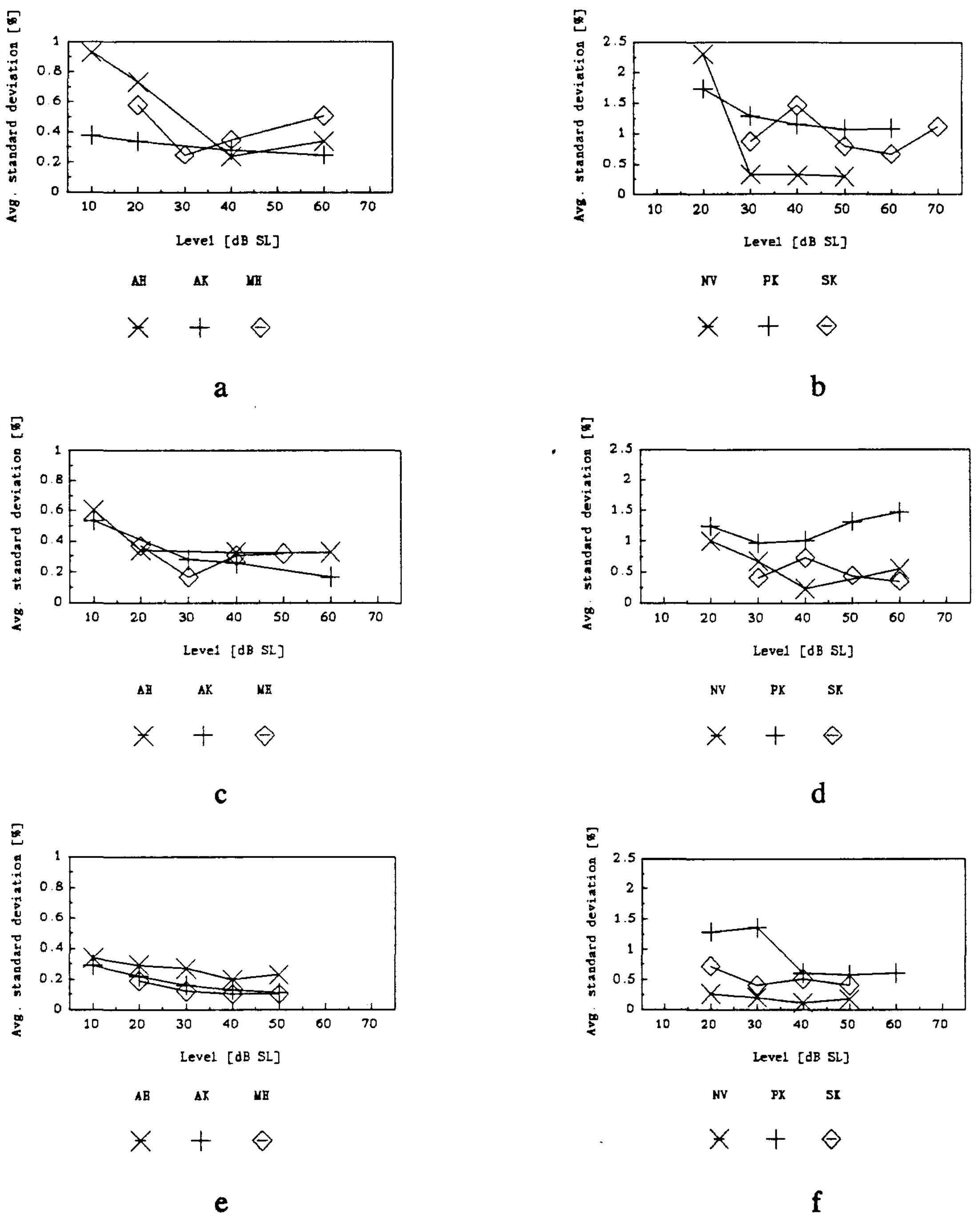
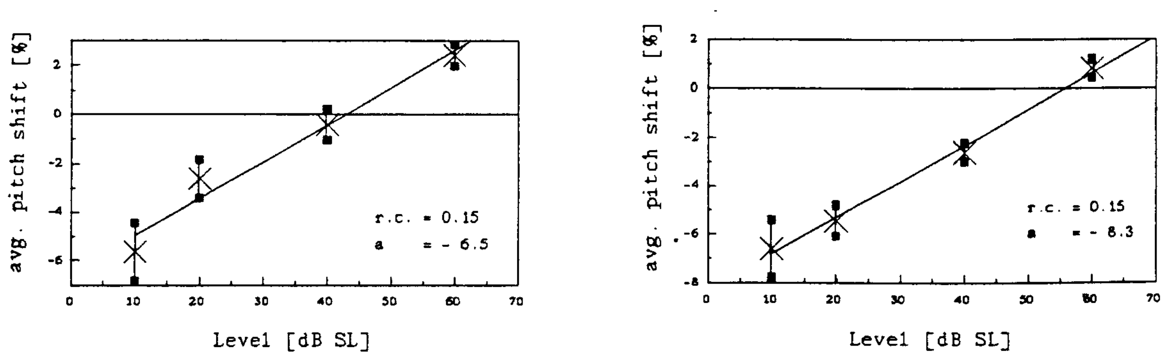


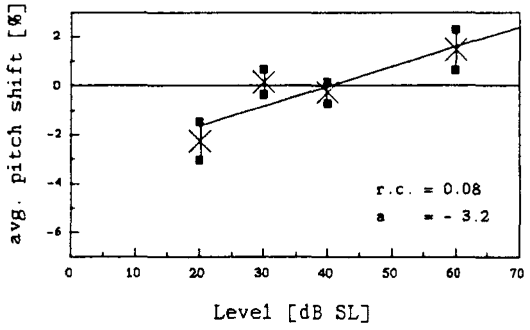
Fig. 5.1: *A and b, average standard deviation of adjustments of a pure tone's frequency to a complex signal with an upper-edge around 1 kHz, as a function of the level of the complex signal. The level of the comparison tone is 55 dB; c and d, the same, but now with the level of the comparison tone at 41 dB; e and f, the same, but now for a tone-on-tone match. Level of the comparison tone at 55 dB.*

5.1.2 Influence of the sound pressure level on the averaged matched frequency

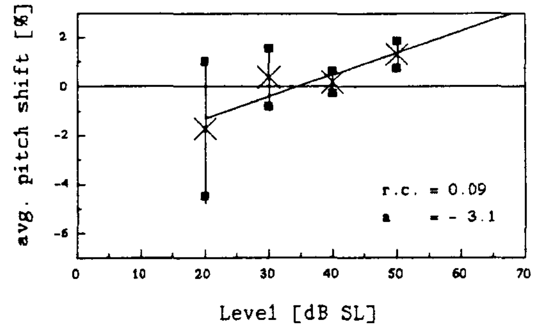
In the next part, in which we use the same data as in the first part, the relative shift between the five nominal edge frequencies and the five means will be presented as a function of the level. Because we use five different fundamental frequencies, we also have five different upper edge frequencies; 943, 984, 1025, 1066, and 1107 Hz. The relative shift is now defined as follows. Each data point, six for each of the five frequencies, is expressed as the normalized difference times 100 %. The average over all 30 data points is the relative shift. The standard deviation of all 30 data points is plotted as an error bar in figure 5.2 - 5.4. If we compare these error bars with the standard deviations of the first part (figure 5.1 à - f), we see that some subjects show rather large differences. This difference is due to the fact that, for some subjects, the perception of the upper-edge frequency in this range depends on the frequency. Therefore each upper-edge frequency can be adjusted quite well (part 1), but with different relative shifts for the five frequencies. As a consequence, our way of averaging results leads to large standard deviations (part 2). Still we chose for this conservative way of presenting our data, because it is not exactly clear in what way each subject is frequency dependent.
We plot in figure 5.2 a - f the relative shifts from the first experiment as a function of the sound level. We see in each panel a positive slope, i.e. the shift increases with the level. The slope is between 0.15 and 0.08 % per dB, with an average of 0.12% per dB. This *is* essential different from figure 3.2, in which we see the dependence of a pure tone's pitch on the sound pressure level. For a pure tone with a frequency of 1 kHz the slope is slightly negative.



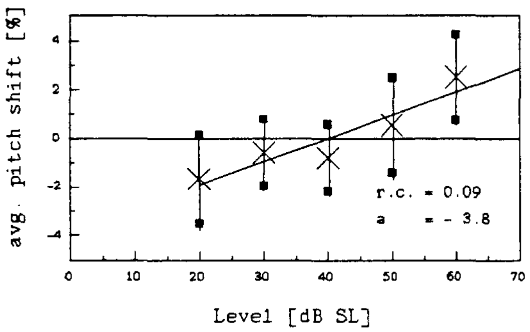
a b



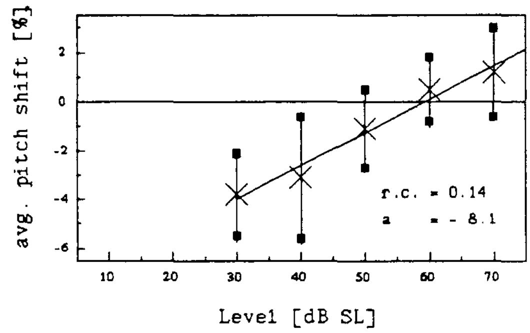
c



d



e



f

Fig. 5.2 a - f: *Average, relative pitch shift between the nominal edge frequency of the complex signal and the adjusted frequency of the comparison tone (55 dB). Panel a represents results of subject AH, b of subject AK, c of subject MH, d of subject NV, e of subject PK, and panel f of subject SK. With r.c. = slope, and a = intersection with the y-axis.*

In a similar way the results of the reference experiment can be plotted. In this experiment the test sound was a sinusoid at different sound pressure levels. These results are shown in figure 5.3 a - f. The first thing we notice is the small standard deviation, similar to that in figure 5.1 e - f. This means that for this kind of stimuli, there exists no frequency dependence of the shift in this small spectral range. We see a linear slope in the range between 0.08 and -0.02 % per dB, with an average of 0.03 % per dB. This *is not* essentially different from the results of Terhardt (1979), but it *is* essentially different from the results of figure 5.2 a - f. This means that, with respect to the level dependence of the pitch, the upper-edge pitch perception is not equal to pitch perception of a pure tone.

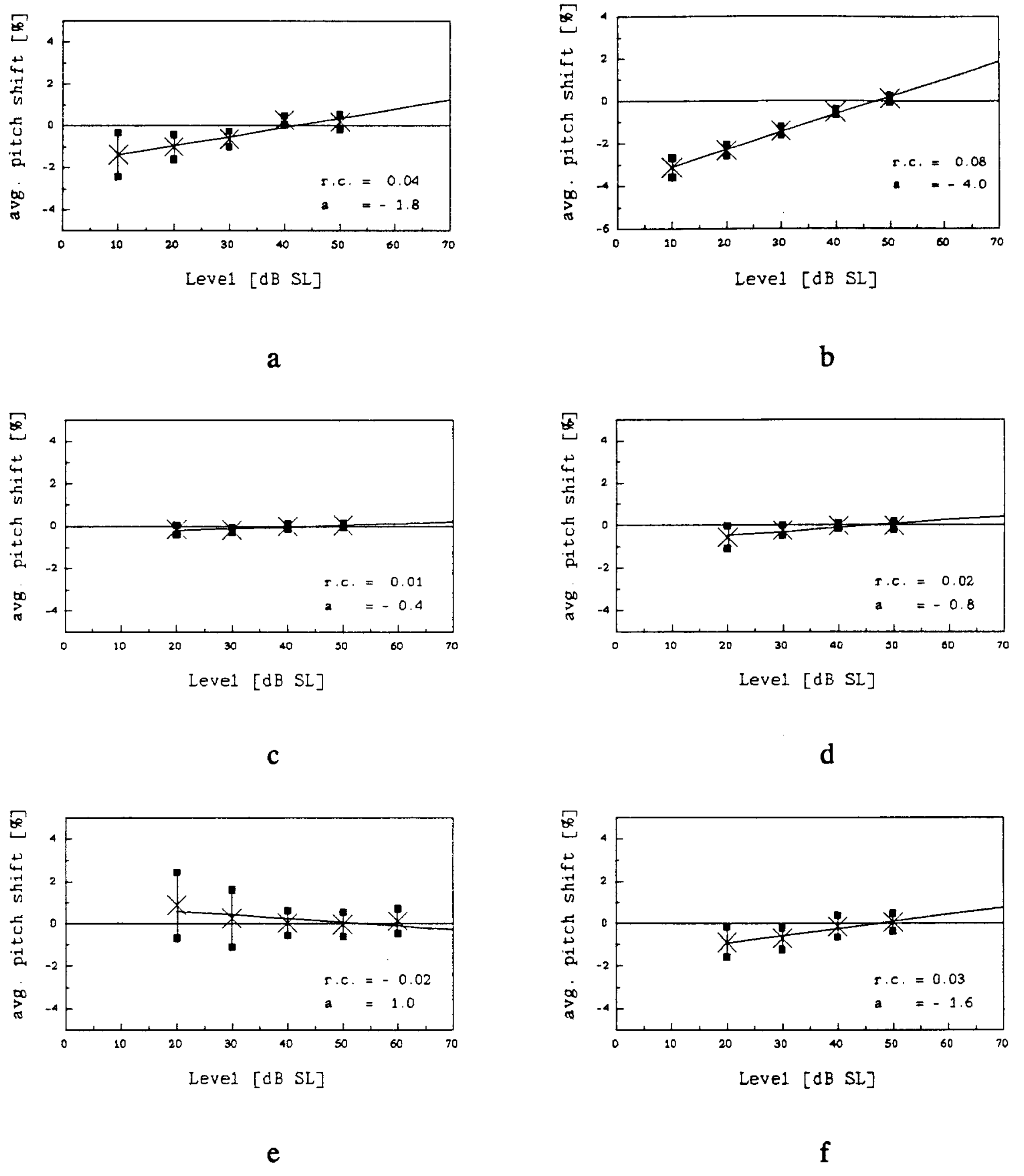


Fig. 5.3 a - f: *Average, relative pitch shift between the nominal frequency of a pure tone and the adjusted frequency of the comparison tone (55 dB). Panel a represents results of subject AH, b of subject AK, c of subject MH, d of subject NV, e of subject PK, and panel f of subject SK. With r.c. = slope, and a = intersection with the y-axis.*

Finally, in this project, we repeated the complex-tone match, but with the sinusoidal comparison tone at a lower level. A lower level of the comparison tone means that the perception of the pitch of the comparison tone changes. The way it changes is subject dependent and can be read from figure 5.3. What we thus expect is a horizontal shift of the graphs of figure 5.2, the amount of which should be related to the shifts of the pure tones. The results are plotted in figure 5.4 a - f.

As in the first experiment we see again large positive slopes, although they are not exactly the same as in fig 5.2. The slopes go from 0.19 to 0.09 % per dB, the average is 0.13 % per dB. This average value is almost the same as 0.12 % per dB from the first experiment.

As an example, we will relate the results of the two experiments with each other for the subject with the strongest level dependence (AK). The slope from fig. 5.3 b is 0.08 % per dB. The difference between the two levels of the comparison tones for the complex-signal experiments was 14 dB. 14 * 0.08 = 1.12 % (vertical shift). If we now compare the two graphs of the complex signals (fig. 5.2 b and 5.4 b), at for instance 40 dB, we find a value of 1.8 % per dB, which is obviously more. The same is found for the other subject's, although the difference for some subject's is larger. There is however one clear exception. The negative shift, for the pure tone, for subject PK should also lead to a negative shift between the two complex-signal experiments. This is not the case, the shift is 0.5 % per dB. Notice, that this is, however, the smallest shift. One of the main problems at comparing both experiments is the mutual difference in the slopes of one subject. This leads to choosing "a" point. Because of the different slopes, it is not possible to compare both experiment with the complex in a quantitative way.

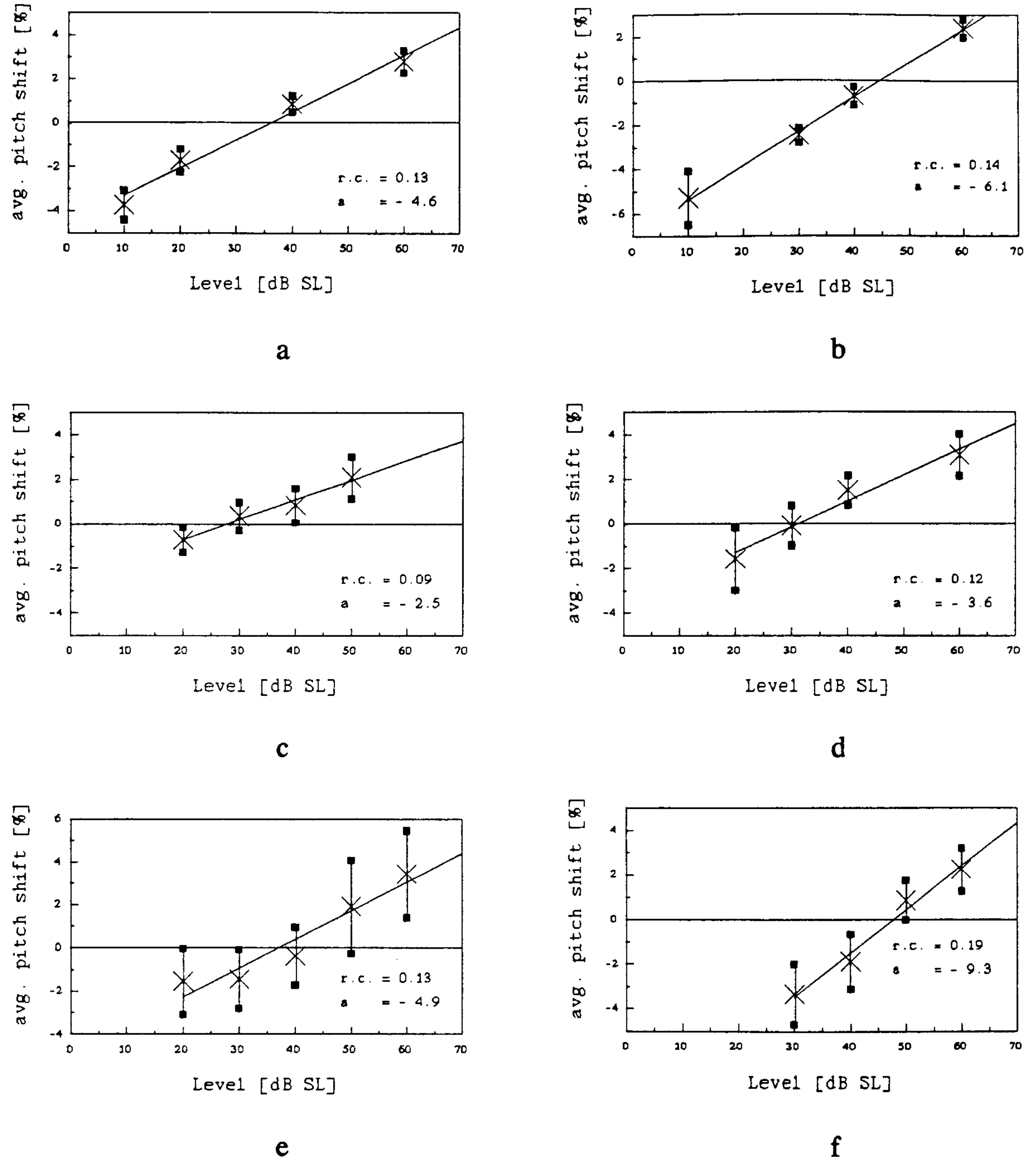


Fig. 5.4 a - f: *Average, relative pitch shift between the nominal edge frequency of the complex signal and the adjusted frequency of the comparison tone (41 dB). Panel a represents the results of subject AH, b of subject AK, c of subject MH, d of subject NV, e of subject PK, and panel f of subject SK. With r.c. = slope, and a = intersection with the y-axis.*

5.2 Project 2

In the second project we examined how clearly the upper-edge frequency of a high-frequency complex signal could be perceived by the subjects. As indication for the accuracy of the percept we used the standard deviation of ten repeated matches to the same complex.

5.2.1 Upper-edge frequency

In a pilot experiment we investigated how clearly the pitch of a sinusoidal tone could be perceived by the subjects. The stimuli which were used had frequencies between 2 - 8 kHz with steps of 400 Hz. In figure 5.5, the accuracy of the matches to this sinusoidal tone is shown for the five subjects. The standard deviations are shown in percentage of the matched frequency.

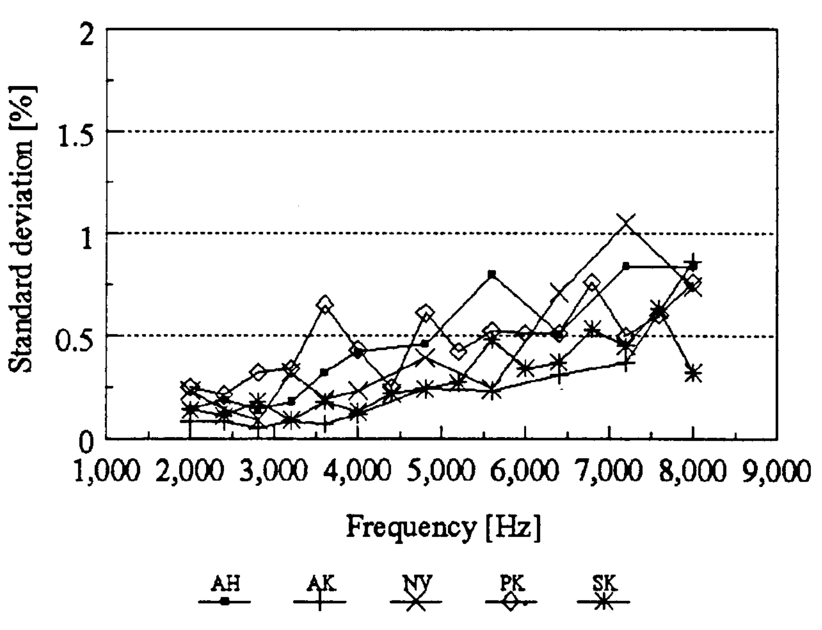


Fig. 5.5: *The accuracy of pitch matches to sinusoidal tones with frequencies between 2 and 8 kHz.*

All subjects were able to adjust the frequency over the entire range with a high accuracy. The standard deviation increased from about 0.2% of the matched frequency near 2 kHz to about 0.8% near 8 kHz.
In the second experiment a complex signal with a fundamental frequency of 50 Hz was used. The highest component of the complex had a frequency in the range 2 to 8 kHz. Figure 5.6 shows the results for the five subjects. Compared with the previous figure (tone-on-tone matches) there is a great variation between the subjects.

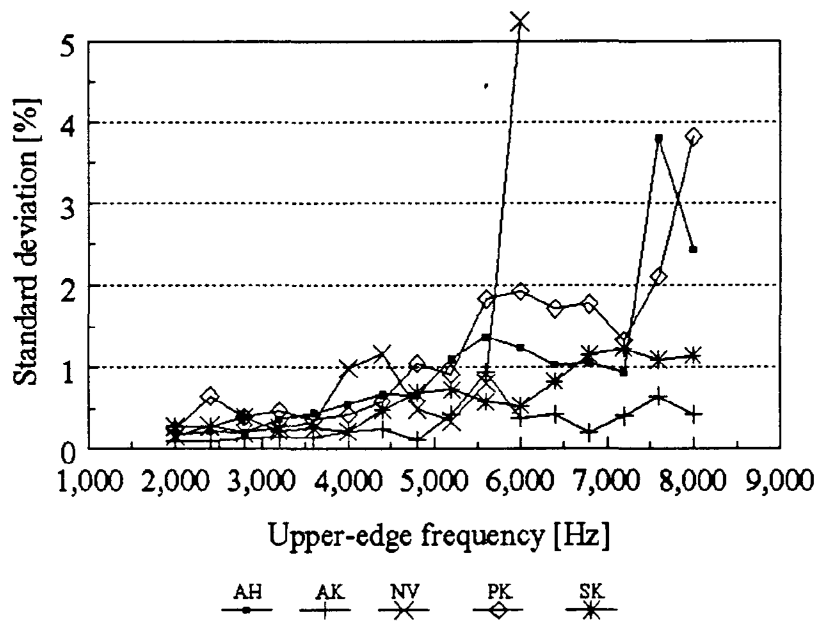


Fig 5.6: *The accuracy of matches to a complex signal with a fundamental frequency of 50 Hz as a function of the upper-edge frequency.*

In this experiment two subjects (SK and AK) were able to match the upper-edge tone up to a frequency of 8 kHz with an accuracy of better then 1%. Subjects AH and PK could do this similarly well, but the standard deviation rises sharply above about 7200 Hz for both subjects. The last subject (NV) could not match the upper-edge above a frequency of about 5600 Hz.
In the third experiment the fundamental frequency of the complex sound was increased with one octave from 50 Hz to 100 Hz. Therefore the number of components decreased with a factor 2 and the duration of a period decreased from 20

ms to 10 ms. Figure 5.7 shows the results of the 5 subjects.

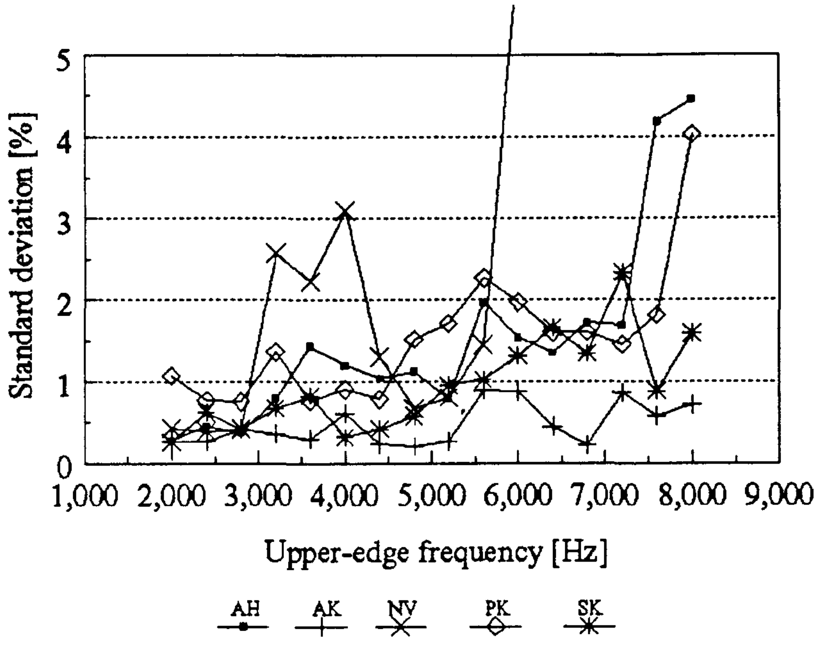


Fig 5.7: *The accuracy of matches to a complex signal with a fundamental frequency of 100 Hz as a function of the upper-edge frequency. The data point at 5600 Hz falls outside the range of the plot (6.1 %).*

The results are similar to the results with the fundamental frequency of 50 Hz but in general less accurate. Again subjects SK and AK could hear the upper-edge frequency over the entire range with an accuracy of less then 2 % for subject SK and still less then 1 % for subject AK. The results of the other three subject are very similar to the results for a fundamental frequency of 50 Hz. Subject AH and PK again were able to match up to about 7200 Hz while from about 7200 Hz the standard deviation increases rapidly. Subject NV again couldn't match the upper-edge above about 5600 Hz. Remarkable is that he could hardly match the upper-edge frequency in the range of 3 - 4.5 kHz while he could do it well in the range of 4.5 - 5.6 kHz. This effect is also seen with the fundamental frequency of 50 Hz though in a less clear way.
In the fourth and last experiment the fundamental frequency was increased again one octave from 100 to 200 Hz. Therefore the number of components decreases again with a factor 2 and the duration of a period decreased from 10 ms to 5 ms. Figure 5.8 shows these results.

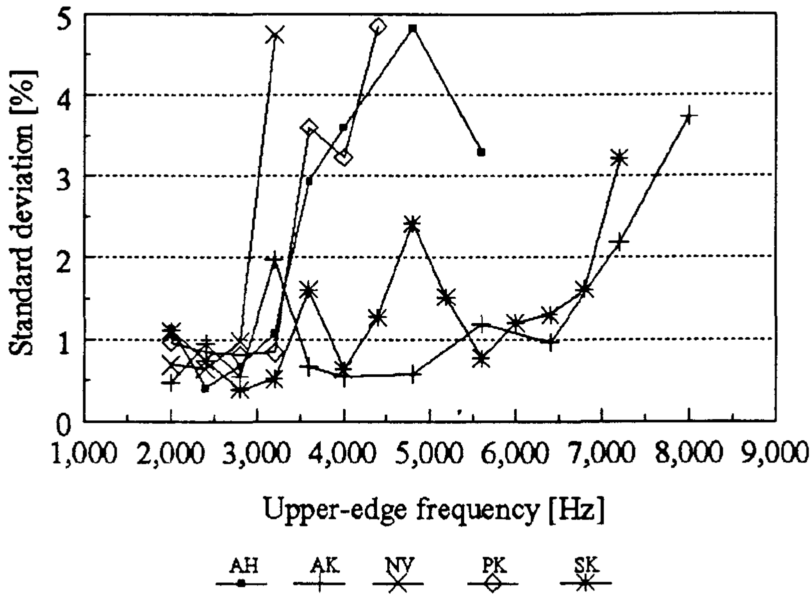


Fig 5.8: *The accuracy of matches to a complex signal with a fundamental frequency of 200 Hz as a function of the upper-edge frequency.*

The results for 200 Hz are quite different to those for 50 Hz and 100 Hz fundamental. None of the subjects was able to match the upper-edge frequency over the entire range below an accuracy of 3 - 4 % though subject AK could almost do it. Subject SK could match the upper-edge frequency still up to about 7 kHz. Remarkable for this subject is that some frequencies, around 4 and 6 kHz, could be matched with a relative high accuracy while at others, the accuracy was much less (around 3.5 and 4.8 kHz). Subjects NV, AH, and PK could not match the edge tone in this experiment above about 3 kHz to about 3.5 kHz.

5.2.2 Quality of results

To see how accurate the results are, a confidence interval of the standard deviation can be calculated. In this case the parameters that are used are ten matches to calculate the standard deviation (s) and a confidence interval of 50 %, corresponding to the interval between the interquartiles. In the table below it is shown how this confidence interval can be calculated.

Table 5.1 *Determination of a confidence interval for the standard deviation*

Standard deviation = s
Confidence interval γ = 0.5 (50 %)
Number of samples n = 10

Two constants, c_1 and c_2, should be choosen from the table of chi-square distribution with n - 1 degrees of freedom. Below a summary of this table is shown for n = 10.

Summary of chi-square distribution

γ	=	0.95	0.90	0.80	0.50
c_1	=	2.70	3.33	4.17	5.90
c_2	=	19.0	16.9	14.7	11.4

Below a summary of the calculation of the confidence interval is shown.

Compute $(n - 1)s^2$ where s^2 is the variance of the sample.

Compute $k_1 = \sqrt{\{(n - 1)s^2/c_1\}}$ and $k_2 = \sqrt{\{(n - 1)s^2/c_2\}}$.

The confidence interval is now: CONF $\{k_2 \leq \sigma \leq k_1\}$

This confidence interval is proportional to the standard deviation and the interdependence is in this experiment as follows: $k_1 = 1.24$ s and $k_2 = 0.89$ s

As example, the confidence interval for an upper-edge frequency of 4.8 kHz in figure 5.8, for subject SK would be:

$k_1 = 1.24 \times s = 1.24 \times 2.5 = 3.1$

$k_2 = 0.89 \times 2.5 = 2.2$

$2.2 \leq \sigma \leq 3.1$

From the same figure, this would be for subject AK at 4.8 kHz:

$k_1 = 1.24 \times 0.6 = 0.7$

$k_2 = 0.89 \times 0.6 = 0.5$

$0.5 \leq \sigma \leq 0.7$

6 Discussion

This chapter gives the conclusions which can be drawn from the results presented in chapter 5. In paragraph 6.1 we give some conclusions of the first project (chapter 5.1). In paragraph 6.2 we give some conclusions of project 2 (chapter 5.2), and together with hypotheses published by Kohlrausch and Houtsma (1991).

6.1 Discussion of project 1

In chapter 5.1 we started to show some accuracy measurements, where the accuracy of the pitch match to a signal (complex or pure tone) was measured as a function of the signal level. For a pure tone the standard deviations are between 0.1 and 0.3 % (percentage of the nominal frequency), dependent on the subject. The highest standard deviation at low signal levels is typically a factor 2 larger than the best values at high levels. The edge pitch of complex signals, however, can be matched with an accuracy between 0.2 and 2.5 %, if the comparison tone is presented at a level of 55 dB SPL. It is also clear that there is a decreasing accuracy at a decreasing level, i.e. the standard deviation increases by a factor up to 10. If we decrease the level of the comparison tone down to 41 dB, the standard deviation decreases to values between 0.2 and 1.5 %. The effect of decreasing accuracy with decreasing level also becomes smaller, and the standard deviation increases maximally by a factor 4. The fact that the complex signal has an accuracy worse than for a pure tone is expected, because the former not only has an edge tone but also a fundamental pitch. In summary these results show that the edge tone of complex sounds can be matched also at low levels (10 to 30 dB above the absolute threshold). But as for a pure tone, the standard deviation increases (accuracy decreases) with decreasing level. This increase is in general stronger for the edge pitch then for a pure tone.
In chapter 5.2 the shift as a function of the level has been investigated. The results we found for a pure tone agree with the existing literature (Terhardt, 1979), although the average of our slopes for an 1-kHz pure tone is slightly positive instead of slightly negative. For the complex signals, on the other hand, we found for all conditions and subjects a strongly positive slope. Similar relations between the pitch of pure tones of level were observed by Terhardt and Fastl (1979) and Houtsma (1981).

Terhardt and Fastl measured the pitch of a sinusoid with a fixed level in the presence of a low-pass noise with a cutoff frequency below the sinusoidal frequency. When the level of the noise was increased, the pitch of the pure tone increased. For a frequency of 3.8 kHz the average slope of the pitch shift was about 0.1 % per dB noise level.
A similar slope was reported by Houtsma (1981) for sinusoidal frequencies 600, 1000, and 1400 Hz. In this study, a strong inter-subject difference with respect to the slope was measured.
An explanation is partial masking. In figure 6.1, a schematic plot indicates how through partial masking the pitch of the pure tone could be shifted to high frequencies. In panel a, the excitation created by a sinusoid on the basilar membrane, is plotted (see also figure 2.8). We assume that the perceived pitch is derived from the weighted mean. In b the excitation for a pure tone and low pass noise is sketched. Now, a part of the pure-tone excitation is masked by the noise (partial masking). The remaining weighted mean of the pure tone is now shifted toward higher frequencies and thus the perceived pitch is higher.

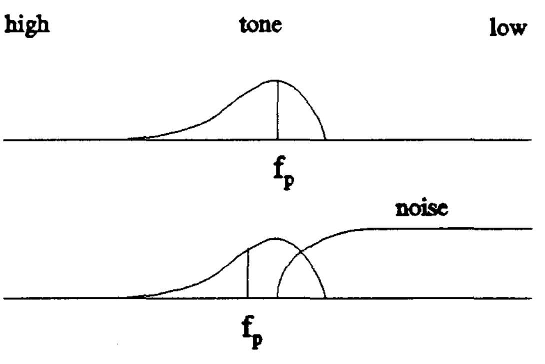


Fig. 6.1: *The upper plot shows the perception of a pure tone on the basilar membrane (see also fig. 2.8).*
The lower plot shows the same pure tone, except for the right side (lower frequency) which is now masked.

A same kind of explanation can also be made for our complex signals. The upper-edge tone is always partially masked by the components of lower frequency. If we increase the level, both the lower (masking) components and the edge tone become louder. Since the masking effect becomes stronger at higher levels ("upward spread of masking"), a pitch shift towards higher frequencies is expected.

6.2 Discussion of project 2

The results of chapter 5.2 showed that, in general, it became harder to match the upper-edge tone, when the fundamental frequency was increased, or when the frequency of the upper-edge was increased. For a complex of 50 Hz, the frequency dependence of the accuracy was similar to that of a pure tone. For both, the standard deviation increases slightly between 2 and 4 kHz and strong up to 8 kHz. For a pure tone this increase of relative accuracy is usually explained by a decrease of phase locking In the auditory nerve, such a decrease begins around 1 kHz (Rose, Brugge, Anderson & Hind, 1967).

At low frequencies, pitch information for a sinusoid can be derived from the place of excitation (cf. figure 6.1), as well as from timing (phase-locked) information in the auditory nerve. At higher frequencies only place information is available. The parallel increase of the accuracy for the 50 Hz complex (at least up 6 kHz) and the sinusoid suggests that in both cases similar restrictions for an accurate pitch match apply at high frequencies.

For complexes with a higher fundamental, we see two effects. In general the matches become less accurate and especially at 200 Hz, matches can be performed only within a limited spectral range. If we increase the fundamental frequency, the duration of the period decreases. Thus the pure-tone parts in the waveform become shorter and, for the central nervous system, it is more difficult to extract the pitch information from these patterns. This effect and the loss of phase locking, in combination, may be the reason why, for f_0 = 200 Hz, most of the subjects are unable to perceive the edge pitch at high frequencies.

References

Békésy, G. von (1960), *Experiments in Hearing*, New York: McGraw-Hill book company.

Fletcher, H., and Munson, W. A. (1933), *Loudness, Definition, Measurement and Calculation*, J. Acoust. Soc. Am. 5:82-108.

Houtsma, A. J. M. (1981), *Noise-induced shifts in the pitch of pure and complex tones*, J. Acoust. Soc. Am. 70:1661-1668.

Helmholtz, H. L. F. von (1863), *Die Lehre von den Tonemfindungen als Physiologische Grundlage für die Theorie der Musik* (F. Vieweg & Sohn, Braunschweig), 1 st ed.

International Standard Organization, Standards r131, 26, R357 and R454 (International Organization for Standardization, 1 Rue de Narembé, Geneva, Switzerland).

Kohlrausch, A., and Houtsma, A. J. M. (1991), *Edge pitch of harmonic complex tones*, IPO Annual Progress Report 26:39-49.

Ohm, G. S. (1843), *Ueber die Definition des Tones, nebst daran geknüpfter Theorie der Sirene und ähnlicher tonbildender Vorrichtungen*, Ann. d. Phys., 59: 497-565.

Rose, J.E., Brugge, J. F., Anderson, D. J., and Hind, J. E. (1967), *Phase-locked responce to low-frequency tones insingle auditory nerve fibers on the squirrel monkey*, Journal of Neurophysiology, 30:769-793.

Rossing, T. D. (1990), *The science of sound*, Addison-Wesley Publishing Company, 2nd ed.

Stevens, S. S. (1935), *The Relation of Pitch to Intensity*, J. Acoust. Soc. Am. 6:150.

Terhardt, E. (1979), *Calculating Virtual Pitch*, Hearing Research 1:155-182.

Terhardt, E. and Fastl, H. (1971), *Zum Einfluß von Störtönen und Störgeräuschen auf die Tonhöhe von Sinustönen*, Acustica 25:53-61.

Ward, W. D. (1970), *Musical Perception*, in Foundations of Modern Auditory Theorie. Ed., J. V. Tobias. New York: Academic.

Zwicker, E., G. Flottorp, and S. S. Stevens (1957), *Critical Bandwith in Loudness Summation*, J. Acoust. Soc. Am. 29:548.