EXPERIMENTS ON TONE PERCEPTION

R. PLOMP

1966
INSTITUTE FOR PERCEPTION RVO-TNO
NATIONAL DEFENSE RESEARCH ORGANIZATION TNO
SOESTERBERG - THE NETHERLANDS

PREFACE

Several times in the history of science there has been a competition between two seemingly mutually exclusive theories. Such has been the case in the study of the ear's capacity to analyze sounds: on the one hand the perception of tones is considered to be governed by frequency analysis, on the other hand to be governed by periodicity analysis. In this study, Plomp has made an attempt to present a synthesis of these points of view in which the theories are treated as different aspects of one mechanism.

The experiments performed to give a foundation for this point of view are all of a psychophysical nature. This sets a limit to a description of the actual mechanism of the transport of information from the cochlea to the higher centres. The hearing organ is considered as a black box. A number of important properties of this black box could be investigated by comparison of the physical input and the response of the subject. Although in general psychophysical experiments in a laboratory are rather remote from practical situations, the attempt has been made in the experiments reported here to keep the stimuli as close as possible to as what happens in practice, without sacrificing the analytical power of the method.

The use of electronic equipment in experimentation has played an important rôle in the development of theories. Considering the limited variety of sounds which can be made by the mechanical means available in the foregoing century, it is understandable that no conclusive answers could be given to certain questions. It was very fortunate that the development of highly specialised electronic equipment inside the Institute was possible. Nevertheless, looking back at this time from our current familiarity with very specialised equipment, one must admire these observers who by careful listening derived many fundamental properties of the hearing organ.

One other side also needs mentioning. The psychological nature of consonance and dissonance, in the musical sense of the words, of complex tones could be related to the functional properties of the hearing organ. In analyzing this, much help was obtained from the psychologists in the Institute. The cooperation between the disciplines of physics and psychology has made a valuable contribution here to the understanding of hearing. The fundamental structure of the Institute in which several disciplines are brought together under one roof, as introduced by its founder Prof. Dr. M.A.Bouman, has shown its usefulness here in a very clear way.

It is to be hoped that Plomp, who obtained with this study his Ph.D. at the University of Utrecht, can follow his way of thinking in further analyses of the hearing of the sounds of speech and music.

PIETER L. WALRAVEN

ACKNOWLEDGEMENTS

This study was prepared under the guidance of Prof. Dr. M.A.Bouman, former director of the Institute. I wish to express my sincere thanks for the many stimulating discussions of the experiments and their interpretation.

The investigations described in Chapter 4 were performed in cooperation with Dr.W.J.M.Levelt. A great deal of the special equipment was built by A.M.Mimpen, who also carried out most of the experiments. J.M.Schipper and E.Agterhuis designed the apparatus used for the statistical analysis of musical chords. A.J.M. van de Beld, F.R. Bogtstra, T.Houtgast, L.C.W.Pols, and H.J.M.Steeneken participated in the measurements. The manuscript was read by Dr.J.L. Goldstein from U.S.A., presently on leave at the Institute for Perception Research in Eindhoven. The French and German translations of the summary and the text of the grammophone record were made by Mrs.E.A. van Setten-de Bruin and Mr. and Mrs. Huigen-Brudereck, respectively. A.Huigen also took care of the illustrations. The demonstrations on the grammophone record were made by H.J.M.Steeneken. I am very grateful to all of them for their contributions to this study.

The investigations described in Chapter 4 were sponsored financially by the Netherlands Organization for the Advancement of Pure Research (Z.W.O.).

Most of the contents of Chapters 2, 3, and 4 have been published in the following papers:

R.Plomp: The Ear as a Frequency Analyzer, *J. Acoust. Soc. Am. 36,* 1628-1636 (1964).

R.Plomp: Detectability Threshold for Combination Tones, *J. Acoust. Soc. Am. 37,* 1110-1123 (1965).

R.Plomp and W.J.M.Levelt: Tonal Consonance and Critical Bandwidth, *J. Acoust. Soc. Am. 38,* 548-560 (1965).

CONTENTS

"Probleme der Technik und Probleme der Physiologie sind oft sehr verwandt. Die Technik stellt uns die Aufgabe, gewisse Zwecke zu erreichen, und lässt innerhalb bestimmter Grenzen die Wahl der Mittel frei. In der Physiologie hingegen finden wir gewisse Zwecke erreicht und haben nach den Mitteln zu forschen, welche wirklich zur Anwendung gekommen sind."

Ernst Mach, 1863

1

INTRODUCTION

The statement that wondering is the origin of science is particularly verified when we recognize the remarkable achievement of the human ear in perceiving and analyzing periodic sound waves. The most impressive demonstration of this power is the way in which the ear discriminates simultaneous tones produced by, for example, the musical instruments of an orchestra. Apparently, the addition of the various periodic sound waves produced by the violins, flutes, oboes, etc. is no obstacle to hearing the tones of the instruments individually. With little effort, we are able to recognize the different tones by their characteristic timbre and pitch.

On first sight, it might seem rather simple to present an adequate model of the way in which these superimposed sound waves are perceived and analyzed by the hearing organ. In the same way as we use frequency analyzers to find out the sinusoidal components of a complex sound, we could imagine that the ear is provided with a large number of band-pass filters tuned to different frequencies. Assuming a correspondence between filter frequency and perceived pitch, this analyzing mechanism would explain how simultaneous tones are discriminated by the ear. We are not surprised to learn that this elegant model has played a very important rôle in the history of hearing theory. Its greatest promotor was the eminent German scientist H.L.F. von Helmholtz. He published this theory in his book "Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik". Soon after its publication in 1863, this work was recognized as a "classic" and its influence on the development of hearing theory during the last century has been great. Helmholtz's comparison of the ear's frequency-analyzing mechanism with the strings of a piano strongly appealed to the imagination. Although the physiological basis of this parallelism did not stand the test, in fact, most recent hearing theories may be considered as adaptations of Helmholtz's conception.

There are, however, serious objections against this model. Physics has proved that, usually, periodic sound waves are composed of a large number of sinusoidal vibrations with frequency ratios 1:2:3:4:5..., called fundamental and harmonics. If it is correct to consider the ear as a frequency analyzer with numerous sharp filters, this implies that any periodic sound wave must be heard as a sum of sinusoidal components with their corresponding pitches. Although it is possible indeed, under favourable conditions, to distinguish some lower harmonics, we never hear them when listening to a concert. In that case, the sound of each instrument is perceived as a single tone with one definite pitch equal to the pitch of the fundamental, and a timbre related to the relative amplitudes of the harmonics. Moreover, experiments have demonstrated that, even when the fundamental is very faint or even absent, still the same pitch is heard.

These phenomena contradict the conception of the ear as a frequency analyzer. They suggest that the hearing organ uses periodicity rather than frequency as a basis of pitch, since a periodic sound wave is characterized by one unambiguous period equal to the period of the (present or absent) fundamental. This opinion, too, has had fervent supporters among scientists involved in auditory research, although their number and influence have been much smaller than of the adherents of the frequency-analyzer theory.

So there are two quite different view-points on the problem of how periodic sound waves are perceived. Both views are very successful indeed in describing one aspect of tone perception but encounter difficulties in explaining the whole problem. Frequency analysis can explain the fact that we are able to discriminate different tones sounding simultaneously but does not answer the question why the sounds produced by musical instruments manifest themselves as a unity with a single pitch. Periodicity analysis, on the contrary, accounts for the latter phenomenon very neatly but does not indicate how simultaneous tones can be heard separately. In both views, many questions arise in relating timbre to either the relative amplitudes of the harmonics or the waveform of the periodic sound. Therefore, we may conclude that neither frequency analysis nor periodicity analysis give a satisfactory explanation of all basic aspects of tone perception.

In view of these difficulties, the question can be raised whether we are forced indeed to choose between the two types of analysis. The opinion has grown, in particular during the last years, that the dilemma between frequency and periodicity analysis as an exclusive description of the way in which simultaneous tones are perceived is not correct. Perhaps we have to give up our efforts to mold all experimental data into the framework of one of these theories but rather should look for a synthesis of both. On first sight, it might seem that by this approach more questions are raised than answered. There is no model available to show how the contradictory claims of discriminating frequencies and periodicities can be integrated. This is no reason, however, not to make a serious attempt to find a new description of the ear's analyzing mechanism in which all data show to better advantage.

In the writer's opinion, the experimental evidence justifies the hypothesis that tone perception indeed involves both frequency analysis and periodicity analysis. However, we need much more data than are available at the moment to find out the contribution of each of these processes and the way in which they are related. It is the purpose of this study to present the results of some recent experiments which may be of value to answer these questions.

OUTLINE OF THIS STUDY

The experiments reported in the following chapters are all of a psychophysical nature. This means that they deal with the relation between the physical parameters of tones and the psychological parameters of their corresponding sensations. In this way, the over-all characteristics of the hearing organ, considered as a "black box", can be investigated. This knowledge can be used to find out a consistent model of the hearing process. Since this model has to meet also the anatomical and physiological data on the organ of hearing, its development will be successful only if all data available have been taken into account.

The experiments are all closely related to the main topics of Helmholtz's book mentioned above. His attention was especially drawn to the frequency-analyzing power of the ear and the origin of timbre, combination tones, beats, and musical consonance. Although nowadays these subjects are no longer in the centre of interest, nevertheless they have to be considered as basic problems in tone perception. More knowledge about each of these subjects may help us in obtaining a better view on the whole problem.

The merits of the frequency-analyzer theory can only be evaluated adequately if we know to what extent the hearing organ is able to discriminate simultaneous tones. Therefore, this question is treated first (Chapter 2).

When the ear is presented with two *loud* tones, one or more secondary tones may be audible. The appearance of these tones, known as combination tones, cannot be explained satisfactorily on the basis of both frequency analysis and periodicity analysis. For that reason, Helmholtz introduced an additional mechanism, namely nonlinear distortion, to account for combination tones. Experiments on the audibility of these tones give us information about the conditions under which they appear (Chapter 3).

For small frequency differences between two simultaneous simple tones, slow beats are heard, changing into a roughness sensation for increasing frequency difference. Helmholtz related this roughness to dissonance and, by taking into account also the beats between harmonics, he was able to indicate why consonant tone intervals are marked by simple frequency ratios as 1:2, 2:3, 3:4, etc. Chapter 4 is devoted to this subject. Not only some experiments on the evaluation of tone intervals are treated, but also some statistical analyses of the tone structure of chords of two musical compositions are given.

The beats show that interference occurs for small frequency differences. This phenomenon is not compatible with a sharp frequency-analyzing mechanism but supports the assumption that simultaneous tones are only perceived separately when their frequency distance exceeds a critical value. However, beats also occur for tone intervals slightly different from consonant tone intervals with frequency ratios 1:2, 2:3, 3:4, etc., respectively. Helmholtz interpreted these beats as products of nonlinear distortion of the ear and this view is still the most current one. The question whether this opinion is right or not led to the experiments described in Chapter 5.

Helmholtz gave little attention to the fact that complex tones have only one definite pitch, equal to the pitch of the fundamental. This problem occupies us in the next chapter (Chapter 6).

Finally, we ask how the experimental results, which are formulated in "black box" terms, can be used in obtaining a better insight in the way in which tone perception is actually achieved by the hearing organ. This discussion (Chapter 8) is preceded by an exposition of the present status of our physiological knowledge of the ear, as far as tone perception is concerned (Chapter 7). In Chapter 8, also the origin of timbre is treated. This chapter concludes with the description of how both frequency analysis and periodicity analysis may be considered as essential aspects of the hearing process.

STIMULI AND TERMINOLOGY USED

It might be useful to include in this introduction some general remarks on the stimuli applied in the experiments and the terms used for describing them.

Nearly all stimuli are *periodic sound waves*, which means that the sound pressure p at any point varies as a function of time t in such a way that

$$p=p(t)=p(t+nT),$$

with $n=1,2,3,\ldots$. T is called the *period* of the sound wave. After Fourier's famous theorem, such a periodic function can be expanded into a series of sinusoids:

$$p(t) = a_0 + \sum_{n=1}^{\infty} a_n \sin 2\pi nft + \sum_{n=1}^{\infty} b_n \cos 2\pi nft,$$

with

$$a_0 = \frac{1}{2\pi}\int_0^T p(t)dt \quad (=0 \text{ for sound waves}),$$

$$a_n = \frac{1}{\pi}\int_0^T p(t) \sin 2\pi nft.dt,$$

and

$$b_n = \frac{1}{\pi}\int_0^T p(t) \cos 2\pi nft.dt.$$

This means that the function $p(t)$ can be described mathematically as the sum of an infinite number of sinusoids with periods of T, $\frac{1}{2}T$, $\frac{1}{3}T$, etc., respectively; the amplitudes of these sinusoids are determined by the *waveform* of the function over the time interval T.

The number of periods per second of a sinusoidal sound wave is called the *frequency* in *cps* (cycles per second); numerically, the frequency is equal to the reciprocal of the period. We shall use the term frequency only for describing sinusoids and not as a name for the reciprocal of the period of a nonsinusoidal sound wave in order to avoid misunderstandings about the meaning of the term *frequency analyzer*. Therefore, the number of periods per second of a periodic nonsinusoidal sound wave shall not be denoted in cps but in *pps*.

The definitions mentioned thus far are given in physical terms and do not say anything about the way in which periodic sound waves are perceived. We are accustomed, however, to describe these waves also in terms related to our auditory sensations. The result of this is that terms as *tone, fundamental*, and *harmonic* are used in two different meanings. With respect to the physical properties of the sound, the term *tone* can be considered as a synonym for *periodic sound wave*. We shall apply many times the expressions *simple tone* and *complex tone* as denotations for a sinusoidal and a nonsinusoidal periodic sound wave, respectively. The complex tone can be regarded as the sum of simple tones, called *harmonics*, with frequencies f, $2f$, $3f$, etc. (see above); the n-th harmonic has a frequency nf. The first harmonic is often called the *fundamental*. The expression: *a complex tone of f cps* shall be used, for the sake of brevity, instead of: a complex tone with a fundamental of frequency f cps. Also the term *partial* is used, both as a synonym for harmonic and as a name for the sinusoidal components of an inharmonic complex of tones.

The second meaning of *tone* refers exclusively to our sensation of a sound, irrespective of its physical properties. In this case, the only criterion for applying the term *tone* for a sound sensation is that it is qualified by a *pitch*. By *pitch* is meant that attribute of auditory sensation in terms of which sounds may be ordered on a *musical scale*, or, otherwise stated, that attribute that constitutes *melody*; wider definitions of pitch, as can be found in literature, give rise to confusion and have to be abandoned. The writer hopes that it will be always clear from the context whether the term *tone* is used in its first or second meaning. To avoid misunderstandings, the expressions *simple tone, complex tone, harmonic*, and *fundamental* shall not be used as descriptions of our sensation but exclusively of the stimulus.

The ambiguity of the term *tone* does not exist for its attributes *pitch, loudness*, and *timbre*. Although there is, for simple tones, a clear relationship between *pitch* and *frequency*, the words should not be exchanged: frequency refers to the sound stimulus and pitch to the auditory sensation. The same holds for *sound-pressure level* and *loudness level*. *Sound-pressure level*, in which the strength of the sound is expressed in this study, is defined as

$$L = 20 \log_{10} \frac{p}{p_o}$$

with p=effective sound pressure, p_o=effective reference sound pressure of 2.10^{-4} dyn/cm^2, and L=sound-pressure level in dB (decibel). Often, the sound-pressure level of a stimulus is given relative to the threshold of hearing for that stimulus; in that case, the expression *sensation level* is used. The *loudness level* of a sound, expressed in *phons*, is

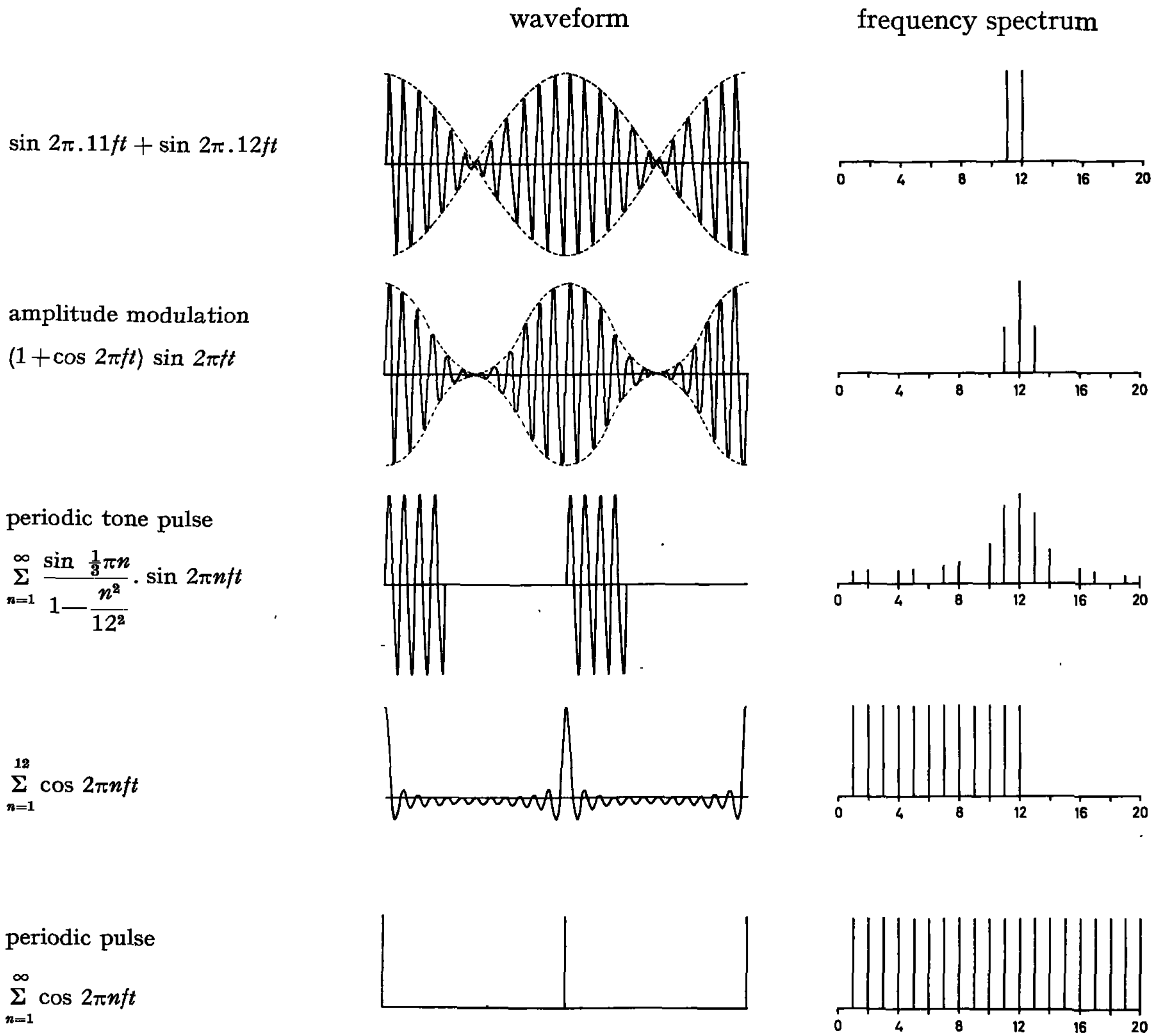


FIGURE 1. Waveforms and frequency spectra of some periodic sound waves.

numerically equal to the sound-pressure level (in dB) of a simple tone of 1000 cps of the same loudness as that sound. *Timbre* refers to the difference in auditory sensation between tones with the same pitch and loudness; it depends primarily upon the *frequency spectrum* of the stimulus.

Finally, in Figure 1 the waveforms and frequency spectra of some periodic sound waves discussed in this study are represented.

2

LIMIT OF THE EAR'S FREQUENCY-ANALYZING POWER

The conception of the ear as a frequency analyzer implies that it must be possible to hear the harmonics of a periodic sound wave individually. As was mentioned already in Chapter 1, this is the case indeed, although it appears to be rather difficult. This phenomenon, known as *Ohm's acoustical law*, may be considered as a strong support of the frequency-analyzer theory.

It is remarkable that the limit of the ear's analyzing power was never investigated systematically. We may expect that more knowledge of this limit will help us in evaluating the rôle of the underlying mechanism. For this reason, some experiments about the conditions for discriminating the partials of a complex of tones were carried out. In addition to investigating periodic sound waves, we also considered stimuli consisting of a series of inharmonic partials and consisting of only two simple tones, respectively. Before these experiments are described, a short historical review is given.

HISTORICAL REVIEW

Although some statements of Aristotle might be interpreted as an indication that he already was acquainted with the fact that more than one tone can be distinguished in the sound of a vibrating string, study of this phenomenon first began in the seventeenth century. We owe the first clear description of it to the French scientist Mersenne. In the "Traité des instrumens", a supplement to his "Harmonie universelle" (1636), Mersenne asserted that "the string struck and sounded freely makes at least five sounds at the same time, the first of which is the natural sound of the string and serves as the foundation for the rest..." All these sounds "follow the ratio of the numbers 1, 2, 3, 4, 5, for one hears four sounds other than the natural one, the first of which is the octave above, the second is the twelfth, the third is the fifteenth, and the fourth is the major seventeenth." Then there is "a fifth one higher yet, that I can hear particularly toward the end of the natural sound, and at other times a little after the beginning; it makes the major twentieth with the natural sound." This means that Mersenne not only distinguished the first five harmonics but, under favourable conditions, also the seventh one.

Mersenne also pointed out that the third and fifth harmonics are often more easily heard than the second and fourth ones. Similar remarks were made by Sorge (1745), Müller (1840), and others. Helmholtz (1863, Chapter 4) generalized this opinion by stating that as a rule the odd harmonics are better distinguished than the even ones. This view was shared by Appunn (1868) and Stumpf (1890, pp. 231-243). The terms *fundamental* and *harmonics* were introduced by Sauveur (1704).

In an attempt to explain some experiments carried out by Seebeck with a siren, Ohm formulated his famous definition of tone, stating that a tone with frequency f is only heard when the complex sound contains $\sin(2\pi ft+\varphi)$ as a component (Ohm, 1843). As a means to decide whether a periodic sound wave contains this component or not, Ohm introduced Fourier's theorem in acoustics.

Helmholtz accepted Ohm's definition and used it as a basis of his theory on the perception of tones. His formulation of it may be considered at the same time as an extension of this definition, and reads: "Every motion of the air, which corresponds to a composite mass of musical tones, is, according to Ohm's law, capable of being analysed into a sum of simple pendular vibrations, and to each such single simple vibration corresponds a simple tone, sensible to the ear, and having a pitch determined by the periodic time of the corresponding motion of the air." (Helmholtz, 1863).

In his book, Helmholtz dwells at length on various ways to facilitate the hearing of harmonics. He recommends to listen first to a simple tone with the same pitch in order to direct the attention to a particular partial. About the results, he says: "The second, fourth, and eighth partials are higher Octaves of the prime, the sixth partial an Octave above the third partial, that is, the Twelfth of the prime; and some practice is required for distinguishing these. Among the uneven partials, which are more easily distinguished, the first place must be assigned, from its usual loudness, to the third partial, the Twelfth of the prime, or the Fifth of its first higher Octave. Then follows the fifth partial as the major Third of the prime, and generally very faint, the seventh partial as the minor Seventh of the second higher Octave of the prime.... If the same experiments are tried with an harmonium in one of its louder stops, the seventh partial will generally be well heard, and sometimes even the ninth... Using thin strings, which have loud upper partials, I have thus been able to recognise the partials separately, up to the sixteenth. Those which lie still higher are too near to each other in pitch for the ear to separate them readily."

Some further information on the audibility of harmonics was given by Brandt (1861) and Stumpf (1890, pp. 231-243). Using gut-strings, Brandt could hear the first seven or eight partials separately, but with brass strings he distinguished them up to about the thirteenth. In his opinion, the latter limit, contrary to the first one, is due to the small distances between the higher partials rather than to their intensities. Stumpf stated that, as a consequence of fusion, the lower harmonics as the second and third ones are often more difficult to hear than the harmonics of higher order, particularly the seventh and the ninth ones. In some cases, he heard partials of much higher order, one time even the twenty-seventh partial.

As far as the writer knows, these are the only more or less quantitative investigations on the audibility of harmonics. They all date from the time that it was still impossible to measure sound-pressure levels, so we do not know how far the observations were influenced by the relative loudness of the various partials. Moreover, the investigations do not give information about any possible dependence of the distinguishability of harmonics upon frequency.

Besides, there are a few studies in which the question of the analyzing power of the ear was approached differently. Baley (1915) did experiments on the localization of each of a series of simple tones presented simultaneously, part of them to the left and part of them to the right ear. It appeared that three observers, including Stumpf, were able to identify, with only a few errors, which of ten tones with frequency separation of about a musical fourth were presented to each ear. In similar experiments with six tones, separated by whole tones, no localization of the partials was possible. As was indicated by Baley, distinct perception of a tone appeared to be a condition for its localization. So the smallest relative frequency separation for which tones can be distinguished in the way described must lie between about 0.3 and 0.1.

Recently, Thurlow and Rawlings (1959) investigated the accuracy with which the number of simultaneously sounding simple tones can be determined. They presented various one-, two-, and three-tone stimuli to subjects who had to judge how many tones were present. The authors comment that, although Ohm's acoustical law would suggest high accuracy, the discrimination appeared to be rather poor, even when the tones were widely spaced in frequency.

Finally, some investigations of the minimum frequency separation in a two-tone sound for which two pitches could be distinguished are relevant to this problem. These investigations include the fragmentary observations by Bosanquet (1881), Stumpf (1890, pp. 319-324, 480-484), Krueger (1900), and Hauge (1931), and the studies of Schaefer and Guttmann (1903) and of Thurlow and Bernstein (1957). Apparently due to the vagueness of the criteria of judgement used, there are large differences in the experimental results. The shaded area of Figure 2 indicates the limits between which all available data points spread.

EXPERIMENTS

Analysis of a complex tone

At first, experiments were performed in which a complex tone was used as a stimulus. In order to give all harmonics "equal chances" to be distinguished, the signal contained the first n harmonics all at the same loudness level.

Method. As was mentioned above, the discrimination of partials is a difficult task,

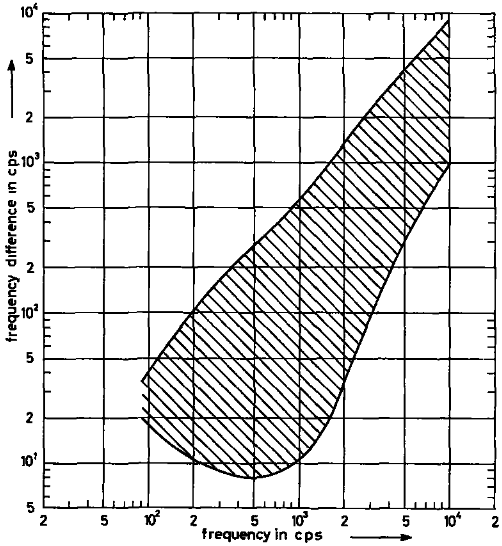


FIGURE 2. The shaded area indicates the spread of all available data for the minimum frequency separation in a two-tone stimulus for which two pitches could be distinguished.

so it is necessary to facilitate it as much as possible by drawing the subject's attention to the pitch of the partial to be heard. Moreover, we need a reliable criterion of judgement. After some trials, a method was chosen that satisfied both requirements.

This method is illustrated in Figure 3. The subject has at his disposal a switch with three positions, indicated schematically at the left. In the middle position, the complex tone is heard. The partials in this position are represented by equidistant lines. In the other two positions of the switch, simple tones are presented, one at the same frequency as one of the partials, and the other one at a frequency midway between this partial and the adjacent higher or lower one. In the example of Figure 3, the upper position corresponds with a tone coinciding with the fourth partial and the lower position with a tone between the fourth and the fifth. The subject is allowed to switch freely from one position to another, so he may listen successively to the three stimuli in any order. He has to judge in a forced-choice procedure which of the two auxiliary tones coincides with one of the partials of the complex tone, in this case the fourth one. He knows that there is always an *a priori* chance of 0.5 to give a correct answer. The *a posteriori* chance on a correct response varies between 1 and 0.5, corresponding with 100% audibility of the partial and 0% (response completely guessed), respectively.

Apparatus and procedure. Figure 4 represents a block diagram of the apparatus used. The complex tone was produced by a function generator consisting of a series of 60 shift-register elements (Philips circuit blocks B8 920 01) connected in a ring and each

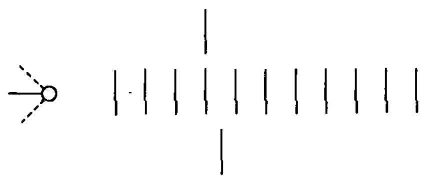


FIGURE 3. Illustration of the method used to determine the number of distinguishable harmonics of a complex tone.

loaded by a potentiometer. Only one of these potentiometers is drawing current at a time. By applying a shift pulse to all elements, the next potentiometer takes over drawing current, and so on. All the sliders of the potentiometers are interconnected through diodes. The potential at this junction is determined over each time interval between two shift pulses by the position of the slider of that potentiometer that draws current at the moment. When periodic shift pulses are applied, a periodic function is generated consisting of 60 successive and adjustable potential levels. This function generator works properly up to a pulse repetition rate of 120 kpps, so the highest repetition frequency of the function obtainable is 2000 cps. The shift pulses are initiated by oscillator 1 (Hewlett-Packard 200 CD).

The periodic signal used (see Figure 1) was

$$f(t) = \sum_{n=1}^{12} \cos 2\pi n f t.$$

After computing the values of $f(t)$ for $t = \frac{1}{60} \cdot \frac{1}{f}, \frac{2}{60} \cdot \frac{1}{f}, \frac{3}{60} \cdot \frac{1}{f}, \ldots, \frac{60}{60} \cdot \frac{1}{f}$, respectively,

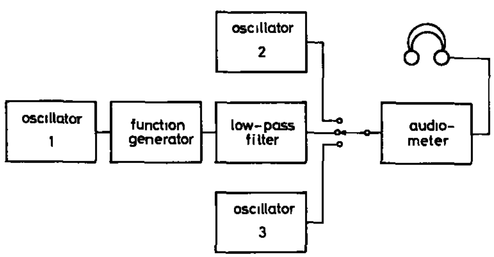


FIGURE 4. Block diagram of the apparatus.

the 60 sliders were adjusted, with the aid of a cathode-ray oscilloscope, to produce potentials proportional to these values. It was checked with a frequency analyzer that the electric output signal of the function generator contained the first 12 harmonics of f with amplitude differences not exceeding 0.5 dB. The signal passed through a variable low-pass filter in order to eliminate all frequency components above the twelfth partial.

The amplitude of the output signals of oscillators 2 and 3 (Hewlett-Packard 200 AB) was always equal to the amplitude of the partials of the complex signal. The audiometer (Peekel D4) was slightly modified in order to give it, in combination with the telephone (Standard Telephones and Cables 4026A), a frequency-response curve corresponding to the equal-loudness contour of about 60 phons. The complex tones were presented monaurally to the test subjects at this loudness level.

The following procedure was used to determine how many partials of a complex tone of frequency f could be distinguished. The experimenter adjusted oscillator 2 at $\frac{1}{2}mf$ and oscillator 3 at $\frac{1}{2}(m+1)f$, m being an integral number, and the subject was asked to judge, by operating the three-position switch, which of the two simple tones coincided with a component of the complex tone. In all cases, the subject was forced to make a choice. He was told after each decision whether his response had been correct or not and then the oscillators 2 and 3 were readjusted, using another value of m. Successively, all relevant values of m were tested in random order.

Since it appeared that for most values of f the responses for $m>16$ were given by guess, generally no higher m-values were involved. This means that usually only the first 8 harmonics were examined. The frequency of oscillator 2 coincided for even values of m with one of the partials and the frequency of oscillator 3 for odd values, so in each trial there was an *a priori* chance of 0.5 that the upper position of the switch corresponded with a correct response.

In a typical test session, the following values of f were presented in random order: 44, 64, 88, 125, 175, 250, 350, 500, 700, 1000, 1400, and 2000 cps. There was a pause after each response during which the experimenter re-adjusted oscillators 2 and 3, so the experiments did not fatigue the subjects. The frequencies were adjusted very accurately by means of an electronic counter. The measurements were repeated five times with different orders of f to avoid any influence of order on the results. During a test session, two decisions were made concerning the audibility of each of the partials considered (for instance, the fourth partial of 250 cps was once tested with oscillator 2 adjusted at 875 cps and oscillator 3 at 1000 cps, and once with the oscillators adjusted at 1000 and 1125 cps, respectively). So in 5 test sessions, 10 responses concerning each of the partials were obtained.

The stimuli were presented monaurally at a loudness level of 60 phons. Two subjects participated in the experiments, both well-trained (AJMvdB and RP). As a comparison of the total number of correct responses for each of the five test sessions showed, the influence of learning during the experiments was negligible.

Both the experimenter and the observer were seated in a soundproof room. The observer could not see the apparatus, having only the three-position switch at his disposal. The subject was allowed to operate the switch as long as he wanted for making his decision.

Results. At first, the percentage of correct responses for each of the first 8 harmonics was computed. This percentage, averaged over all test frequencies from 64 up to 2000 cps, is reproduced in Figure 5 as a function of partial number. (Since for $f=44$ cps only the fundamental was distinguished properly, these data were excluded.) The graph shows that the data points for both test subjects can be fitted readily by a single monotonically decreasing curve. For each of the test

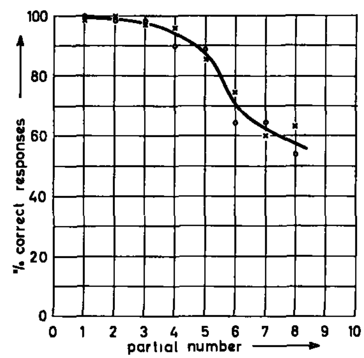


FIGURE 5. Percentage of correct responses as a function of partial number, averaged over all frequencies f, with exception of 44 cps. (× subject AJMvdB, o subject RP.)

frequencies separately, the data agreed with similar curves, although the spread of the points was larger. This indicates that the audibility of partials can be expressed by one number, namely the highest partial which can just be distinguished. We define the value of the abscissa corresponding with 75% correct responses (=5.7 in Figure 5) as a measure of this critical number.

In the same way, the number of partials that can be discriminated was determined graphically for all test frequencies individually. The result is plotted in Figure 6. Since the experimental results of the two subjects were very similar, their responses were added. The upper and lower ends of the vertical dashes indicate the partial numbers corresponding with 58 and 92% correct responses, respectively. If the transition range between

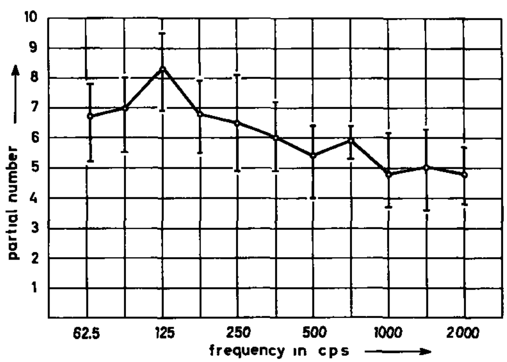


FIGURE 6. Number of distinguishable harmonics of a complex tone as a function of the frequency of the fundamental, averaged over two subjects. The points correspond with 75%, the upper and lower ends of the vertical dashes with 58 and 92% correct responses, respectively.

100 and 50% correct responses may be considered as an integrated Gaussian distribution, the dashes represent its standard deviation.

The experimental data can also be plotted in another way, more directly related to the ear's frequency-analyzing power. This may be illustrated by an example. As Figure 6 shows, for $f=250$ cps the number of partials which could be heard separately (75% correct responses) was 6.5. This means that beyond $6.5 \times 250 = 1625$ cps partials must be separated by more than 250 cps in order to be distinguished. This critical frequency difference can be considered as a measure of the ear's analyzing power at 1625 cps. On the basis of the data points in Figure 6, the corresponding values were computed for all test frequencies. The result is plotted in Figure 7. The vertical scale represents the fundamental frequency of the complex tones and the horizontal scale the frequency beyond which the harmonics could not be distinguished. Interpreted in the other way, the graph represents, as a function of frequency, the minimum frequency difference between a harmonic and the adjacent ones required to hear that harmonic separately. The horizontal dashes correspond with the vertical ones in Figure 6. Since for $f=44$ cps only the fundamental was distinguished, the corresponding data point is given with some restriction. It is not improbable that this tone was heard correctly only because it is the first partial, coinciding in pitch with the complex tone as a whole. So the just-perceptible frequency difference may be

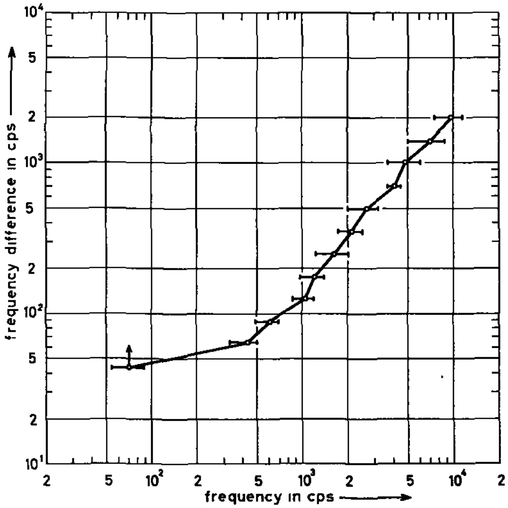


FIGURE 7. Frequency difference between a harmonic of a complex tone and the adjacent ones required to hear that harmonic separately, averaged over two subjects. The points correspond with 75%, the right and left ends of the horizontal dashes with 58 and 92% correct responses, respectively.

larger than is indicated by the data point, which is symbolized by the vertical arrow.

Discussion. As Figure 5 shows, the experimental results do not support the view, mentioned in the historical review, that the odd harmonics of a complex tone are more audible than the even ones. This may be a consequence of the fact that the present results were obtained by a much more critical procedure than was used in previous observations. The second and fourth harmonics, respectively, are the octave and double octave of the fundamental, so they are more readily overlooked than the other harmonics. This is especially true when complex tones are used for comparison, as usually has been done.

The same graph indicates that, on the average, only the first 6 harmonics were distinguished individually. This number depends slightly on frequency, varying between a maximum of 8 for $f=125$ cps and 5 for $f=1000$ cps and higher (Figure 6). This result agrees rather well with the statements of some investigators cited in the review. Since Helmholtz, however, also larger numbers of perceptible partials have been mentioned. To understand this discrepancy, we must realize that in all those cases the relative intensities of the partials were not known. The audibility of particular higher harmonics may have been due to their having higher intensity than adjacent partials.

The curve of Figure 7 represents, as a function of frequency, the minimum frequency separation between a harmonic of a complex tone and the adjacent partials required to discriminate that harmonic. So it is attractive to interpret the curve as the representation of the frequency-dependent bandwidth of the analyzing mechanism of the ear.

We have to be somewhat cautious with this conclusion because the experiments were carried out with complex tones. This implies that the frequencies of the partials were all multiples of frequency f and that fixed phase relations between these components existed. So the question can be raised how far the experimental results depend upon these factors. It is not impossible that the correct responses of the test subjects were (partly) due to their ability to recognize frequency ratios. According to this idea, the subjects could have compared the pitch of each of the tones produced by oscillators 2 and 3 with the pitch of the complex tone and then decided which of the simple tones had a frequency equal to a multiple of f. Although the subjects denied making such comparisons, we may not conclude that this was not done unconsciously. Moreover, the experiments do not allow us to conclude that the frequency analysis of inharmonic complexes of tones is governed by the same principles as of complex tones. For these reasons, a second series of experiments was carried out using an inharmonic complex of tones as the stimulus.

Analysis of an inharmonic complex of tones

The stimulus in the preceding experiments consisted of 12 simple tones with frequency ratios 1:2:3:....:12. In the present case 12 simple tones were used also, but now with frequency ratios 1.00:1.95:2.85:3.70:4.50:5.25:5.95:6.60:7.20:7.75:8.25:8.75. Different stimuli with lowest partials of 62.5, 88, 125, 175, 250, 350, 500, 700, 1000, 1400, and 2000 cps, respectively, were examined.

The tones were produced by separate sine-wave oscillators. In order to avoid the re-adjustment of 12 oscillators over and over during each test session, the complex signals were recorded on magnetic tape. Having a recorder with two identical tracks to our disposal (Revox Stereo E 36), one track was used for the "even" partials and the other one for the "odd" partials; the outputs were interconnected. The substitution of oscillator 1 and function generator (see Figure 4) by the tape recorder was the only difference between the apparatus used in this experiment and in the previous one. The stimuli were presented again at a loudness level of about 60 phons. The experimental procedure was the same as in the former case. Both observers took part in the new experiment.

In Figure 8, the experimental results are reproduced in a way similar to Figure 7. The data points are obtained by plotting for

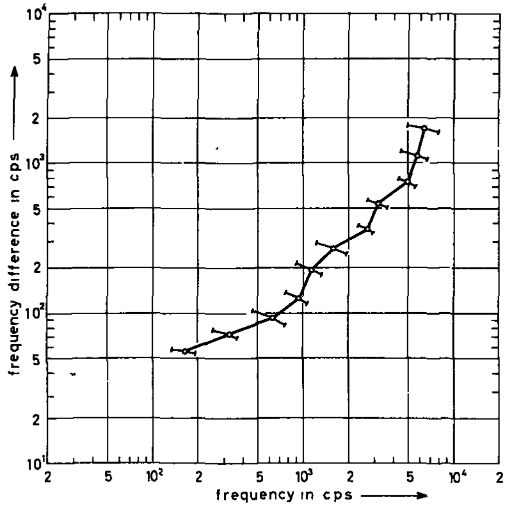


FIGURE 8. Frequency difference between a partial of an inharmonic complex of tones and the adjacent ones required to hear that partial separately, averaged over two subjects. The points correspond with 75%, the right and left ends of the dashes with 58 and 92% correct responses, respectively.

each stimulus the percentage of correct responses as a function of the frequency of the partials, after which was determined (1) the frequency corresponding with 75% correct responses and (2) the frequency difference, by interpolation, between the partials at that frequency. The ends of the dashes correspond again with 58 and 92% correct responses. Since the frequency differences between adjacent partials were not equal in this experiment, these dashes are, contrary to Figure 7, not horizontal.

The data points obtained with harmonic and inharmonic complexes of tones are represented together in Figure 9. (The crosses are taken from a third experiment, described below.) We see that the critical frequency differences determined for complex tones agree very well with the corresponding data points for inharmonic complexes of tones. Both data can be fitted by one smooth curve (dashes). This strongly supports the view that in both cases the same frequency-analyzing mechanism is involved and answers the question put forward at the end of the preceding section.

During recent years, the concept of critical band has proved to be of basic significance in hearing. Many investigators have shown that the ear's behaviour in response to acoustic stimuli with a frequency spectrum exceeding the critical band is different from its behaviour when stimuli not exceeding this band are used (*e.g.* Feldtkeller and Zwicker, 1956; Zwicker *et al.*, 1957; Scharf, 1961). This applies to the absolute and masked hearing thresholds, the loudness level, and the ear's

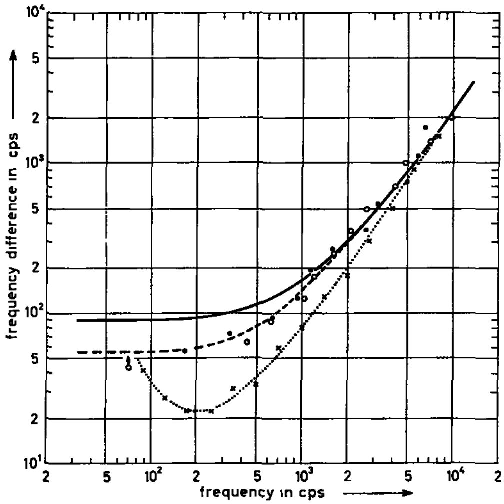


FIGURE 9. Frequency difference between the partials of tone complexes required to hear them separately. The open points represent the values for complex tones (Figure 7), the solid points for inharmonic complexes of tones (Figure 8), and the crosses for two-tone stimuli (Figure 11). The solid curve represents the critical bandwidth as a function of frequency, after Zwicker *et al.* (1957). The dashed curve and the dotted curve fit the data points concerning multitone and two-tone stimuli, respectively.

sensitivity to phase.

In view of this fact, it is of interest to ascertain whether the results of the present experiment on the frequency-analyzing power of the ear are also in harmony with the width of the critical band. For this reason, the solid curve of Figure 9 was drawn, representing the critical bandwidth as a function of frequency after Zwicker *et al.* (1957). It is evident that above 1000 cps the curve is in excellent agreement with the data points derived from the multitone experiments. Below 1000 cps, these points would predict somewhat smaller bandwidths. It is questionable, however, whether we have to attach any value to this discrepancy. Below 500 cps, not many data are available in the critical-bandwidth experiments nor in our experiments either. Moreover, there are some other investigations resulting in smaller values of the critical band in the low-frequency range than represented by the solid curve (Zwicker, 1952; Plomp and Bouman, 1959). Also Greenwood (1961b) has made assumptions in the same direction. It seems for these reasons to be justified to conclude that the ear can discriminate the partials of a multitone stimulus only when their frequency separation exceeds critical bandwidth.

A similar conclusion was drawn by Versteegh (1954) in a study on frequency modulation of simple tones. He determined the minimum modulation rate for which the stimulus sounds like a complex of separate tones. This rate is numerically equal to the frequency distance between the partials in the stimulus. The minimum rate was roughly one half of the critical-bandwidth curve reproduced in Figure 9. The interpretation of Versteegh's results, however, seems to be rather complicated because the amplitudes of the partials were not equal and because the criterion used does not imply that the partials could indeed be heard separately.

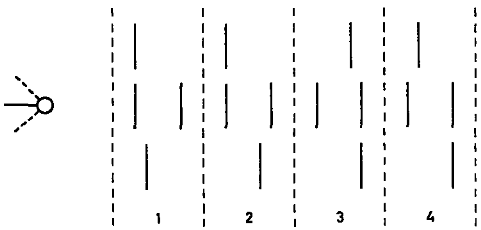


FIGURE 10. Illustration of the method used to determine the frequency difference between two simultaneous tones required to distinguish their pitches.

Pitch discrimination of two simultaneous tones

After considering the case of many tones, it is also of interest to study the extreme condition of only two simple tones. As was mentioned in the historical review, there is a large spread in the available data, which may be due to the vagueness of the criteria of judgement used (Figure 2). In view of this fact, various criteria were tested for stability of results. A method very similar to the one used in the preceding experiments was chosen.

As is illustrated in Figure 10, the test subject again has a switch with three positions at his disposal. The middle position presents a stimulus consisting of two tones with a frequency difference of Δf. In order to investigate whether or not the pitches of these tones can be heard separately, the upper and lower switch positions are used to present four different combinations of comparison tones. For each combination, one of these tones coincides in frequency with one of the stimulus tones and the other comparison tone differs $\frac{1}{4}\Delta f$ from the lower (conditions 1 and 3) or the higher (conditions 2 and

4) stimulus tone. The subject has to judge in a forced-choice procedure which of the combination tones coincides in frequency with a component of the complex stimulus.

The frequencies of the comparison tones were chosen for two reasons: (1) tones with frequencies below or above the two-tone stimulus must be avoided because they are easily identified as such, and (2) a tone midway between the stimulus tones must be avoided too because the stimulus itself has for small values of Δf an apparent pitch corresponding to this frequency.

The apparatus used in the preceding experiments (Figure 4) was used also in the present one, except for a substitution of oscillator 1+ function generator+low-pass filter by two sine-wave oscillators (Hewlett-Packard 200 AB) in parallel. All tones were presented monaurally at a loudness level of about 50 phons to the same subjects as in the other experiments.

Tests were carried out using two-tone stimuli with centre frequencies of 88, 125, 175, ..., 4000, 5600, and 8000 cps. Preliminary trials were made in order to determine the range of relevant Δf-values for each test frequency. In the final experiments, always five different frequency separations were involved.

Only one test frequency was examined in one session by presenting, in random order, stimuli with different values of Δf and different combinations of comparison tones. Since we used five Δf-values and four conditions (Figure 10), each session consisted of 20 trials. An electronic counter was used to adjust the oscillators accurately.

The results of these experiments are represented graphically in Figure 11 as a function of centre frequency, averaged over both subjects. The points correspond with the frequency differences for which 75% correct responses was obtained, the ends of the vertical dashes with 58 and 92% correct responses.

The results are also reproduced in Figure 9, showing that up to about 4000 cps the individual pitches of two-tone stimuli were distinguished for smaller frequency differen-

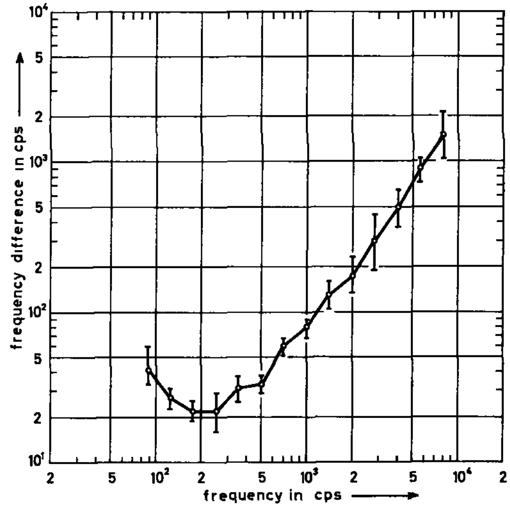


FIGURE 11. Frequency difference between two simultaneous tones required to distinguish two pitches, plotted as a function of centre frequency, and averaged over two subjects. The points correspond with 75%, the lower and upper ends of the vertical dashes with 58 and 92% correct responses, respectively.

ces than were found for multitone stimuli. This discrepancy seems to be incompatible with the conception that the same frequency-analyzing mechanism is involved in both cases. An explanation of this discrepancy will be proposed in Chapter 8.

Masking pattern of a complex tone

The experiments communicated in the preceding sections dealt with the question of how large the frequency difference between simple tones must be in order to hear them separately. As was mentioned in the historical review, this way of investigating the ear's analyzing power is rather unexplored. It is much more common, however, to study the selectivity of the auditory system by masking experiments. Therefore, it seemed to be of interest to determine also the masking pattern of a complex tone and to look for any relation with the results of the previous experiments. Except for the extreme condition of two tones, examined by Greenwood (1961a, 1961b), no relevant data were found in literature.

The masking pattern of a complex tone of 500 cps was investigated by measuring, as a function of frequency, the threshold shift of a short simple-tone pulse presented immediately after the complex-tone pulse. This method was chosen in order to avoid interference of the two signals (beats) when presented simultaneously. The threshold-shift values were determined using a two-alternative forced-choice procedure.

A block diagram of the apparatus is reproduced in Figure 12. The function generator produced a signal consisting of the first 12 harmonics of 500 cps, all with the same amplitude. Oscillator 2 (Hewlett-Packard 200 CD) produced the test tone. Both the stimulus tone and the test tone passed through a two-channel electronic gate, adjusted to produce stimulus pulses of 200 msec duration, in pairs, with a pause of 1 sec; at the moment that one of these two stimulus pulses stopped, a short test pulse of 20 msec started. There was always an *a priori* chance of 0.5 that the first or the second stimulus pulse was followed by the test pulse. The two types of pulses were amplified separately before they were presented to the subject via the same audiometer and headphone as used in the preceding experiments. The subject had at his disposal two pushbuttons to be pressed alternately according to his decision whether the first or the second stimulus pulse was followed by the test pulse. Two electromechanical counters recorded automatically the total number of responses and the number

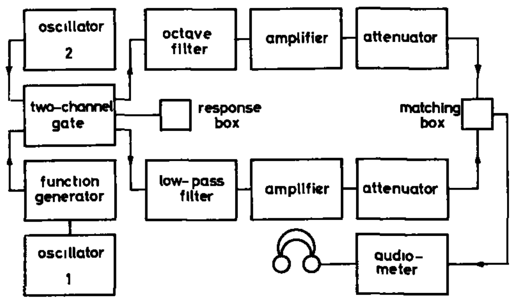


FIGURE 12. Block diagram of the apparatus used for the masking experiments.

of correct responses, respectively. The counters were housed in the same box ("response box") as the pushbuttons, so the subject could notice immediately after his response whether it had been correct or not. The pairs of stimulus pulses were presented with pauses of 5 sec.

The threshold shift was measured with this apparatus for all multiples of 50 cps between 300 and 4000 cps. Since the masking effects decrease rapidly (see Plomp, 1964), it was desirable to use the shortest test pulse possible. This limit is determined by the spread of energy for short durations of a tone pulse. As masking was measured at frequencies 50 cps apart, it was required to use test pulses with a bandwidth not exceeding this value, so a duration of 20 msec was chosen. The test-pulse channel was provided with an octave band-pass filter (Rohde & Schwarz PBO BN 4920) in order to eliminate clicks at the beginning and the end of the pulses. The low-pass filter in the channel of the stimulus tone was adjusted at a cut-off frequency of 7000 cps.

The experiments were carried out monaurally. Firstly, the absolute hearing threshold of the test pulse, defined as the attenuator value for 75% correct responses, was determined for a particular frequency. Next, the stimulus pulses were presented at a loudness level of about 60 phons and again the threshold of the test pulse was measured. The difference in attenuator values represents the threshold shift due to the stimulus pulse. Each threshold-shift determination lasted 6 to 10 minutes, so the measurement over the whole range of 300 to 4000 cps took about 10 hours. The frequency of the test tone was adjusted accurately by means of an electronic counter. In order to minimize the effects of individual irregularities in hearing threshold, the experiments were carried out for four subjects, the subjects of the previous experiments included.

Figure 13 represents the results, averaged over the subjects. We see that only the first 5 harmonics are manifested separately in the masking pattern. This limit agrees with the number of distinguishable partials for $f=$ 500 cps (Figure 6). It demonstrates that there is an excellent agreement between these quite different approaches of the limit of the ear's frequency-analyzing power and supports the view that this limit is determined by the

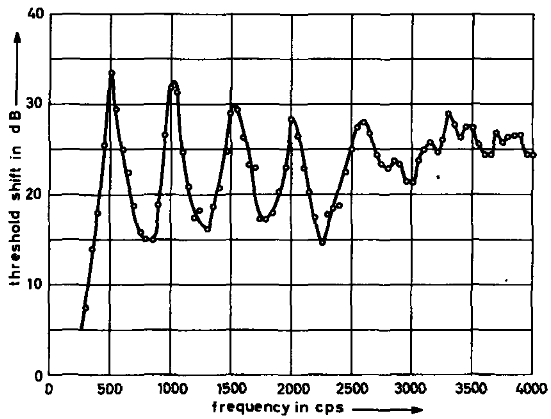


FIGURE 13. Masking of a tone pulse of 20 msec presented immediately after a 200-msec pulse of a complex tone consisting of the first twelve harmonics of 500 cps. The points give the values averaged for four subjects.

critical bandwidth. A possible explanation of the fact that the peaks are smaller than critical bandwidth is given in Chapter 8.

CONCLUSIONS

1. Even under most favourable conditions, not more than the first 5 to 8 harmonics of a complex tone were perceived individually.

2. This limit agrees with critical bandwidth, that is to say harmonics are distinguished only if their frequency separation exceeds this bandwidth.

3. This limit was also found for the discrimination of the partials of inharmonic complexes of simple tones.

4. The masking pattern of a complex tone affirms this rôle of critical bandwidth as the limit of the ear's frequency-analyzing power.

5. The individual pitches in a two-tone sound can be distinguished for smaller frequency differences than a critical bandwidth, particularly at low frequencies.

REFERENCES

Appunn, G.: Ueber die Helmholtz'sche Lehre von den Tonempfindungen als Grundlage für die Theorie der Musik, nebst Beschreibung einiger, zum Theil ganz neuer Apparate, welche zur Erläuterung und zum Beweis dieser Theorie geeignet sind, *Bericht der Wetterauischen Gesellschaft für die gesammte Naturkunde zu Hanau über den Zeitabschnitt vom 14. October 1863 bis 31. December 1867,* 73-90 (1868).

Baley, S.: Versuche über die Lokalisation beim dichotischen Hören (Mit einem Anhang von C. Stumpf), *Z. Psychol. 70,* 347-372 (1915). [= *Beitr. Akust. Musikwiss. 8,* 83-108 (1915).]

Bosanquet, R. H. M.: On the Beats of Consonanoes of the Form *h*:1, *Phil. Mag. 5th Ser. 11,* 420-436 (1881).

Brandt, S.: Ueber Verschiedenheit des Klanges (Klangfarbe), *Ann. Phys. Chem. 112,* 324-336 (1861).

Feldtkeller, R., and E. Zwicker: *Das Ohr als Nachrichtenempfänger,* S. Hirzel Verlag, Stuttgart (1956).

Greenwood, D. D.: Auditory Masking and the Critical Band, *J. Acoust. Soc. Am. 33,* 484-502 (1961a).

Greenwood, D. D.: Critical Bandwidth and the Frequency Coordinates of the Basilar Membrane, *J. Acoust. Soc. Am. 33,* 1344-1356 (1961b).

Hauge, I. B.: An Investigation of the Phenomena Connected with the Beating Complex, *Psychol. Monogr. 41, No. 4,* 32-73 (1931).

Helmholtz, H. L. F. von: *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik,* F. Vieweg & Sohn, Braunschweig, 1st Ed. (1863). The quotations are from the English translation by A. J. Ellis: *On the Sensations of Tone as a Physiological Basis for the Theory of Music,* 2nd Ed. (1885), reprinted by Dover Publications, Inc., New York, pp. 33, 49-51 (1954).

Krueger, F.: Beobachtungen an Zweiklängen, *Phil. Studien 16,* 307-379 (1900).

Mersenne, M.: *Traité des instrumens,* Sebastian Cramoisy, Paris, Book IV, Prop. 11, pp. 208, 209 (1636). English translation of the quotations taken from C. Truesdell: *The Rational Mechanics of Flexible or Elastic Bodies, 1638-1788,* Leonhardi Euleri Opera Omnia Ser. II, Vol. 11, Part 2, Orell Füssli, Zürich, p. 32 (1960). Other passages on hearing of harmonics in the works of Mersenne are given by H. Ludwig: *Marin Mersenne und seine Musiklehre,* Buchhandlung des Waisenhauses G.m.b. H., Halle/Saale and Berlin, p. 41 (1935).

Müller, J.: *Handbuch der Physiologie des Menschen,* Vol. 2, Verlag J. Hölscher, Coblenz, p. 136 (1840).

Ohm, G. S.: Ueber die Definition des Tones, nebst

daran geknüpfter Theorie der Sirene und ähnlicher tonbildender Vorrichtungen, *Ann. Phys. Chem. 59*, 513-565 (1843).
Plomp, R.: Rate of Decay of Auditory Sensation, *J. Acoust. Soc. Am. 36*, 277-282 (1964).
Plomp, R., and M.A. Bouman: Relation between Hearing Threshold and Duration for Tone Pulses, *J. Acoust. Soc. Am. 31*, 749-758 (1959).
Sauveur, J.: Systême general des intervalles des sons, & son application à tous les systêmes & à tous les instrumens de musique, *Mém. Acad. Sci. Paris 1701*, 297-364 (1704).
Schaefer, K.L., and A. Guttmann: Ueber die Unterschiedsempfindlichkeit für gleichzeitige Töne, *Z. Psychol. Physiol. Sinnesorgane 32*, 87-97 (1903). [= *Beitr. Akust. Musikwiss. 4*, 51-61 (1909).]
Scharf, B.: Complex Sounds and Critical Bands, *Psychol. Bull. 58*, 205-217 (1961).
Sorge, G.A.: *Vorgemach der musicalischen Composition*, Verlag des Autoris, Lobenstein, p. 14 (1745).
Stumpf, C.: *Tonpsychologie*, Vol. 2, Verlag S. Hirzel, Leipzig (1890).
Thurlow, W.R., and S. Bernstein: Simultaneous Two-Tone Pitch Discrimination, *J. Acoust. Soc. Am. 29*, 515-519 (1957).
Thurlow, W.R., and I.L. Rawlings: Discrimination of Number of Simultaneously Sounding Tones, *J. Acoust. Soc. Am. 31*, 1332-1336 (1959).
Versteegh, R.M.: *Frequentie-modulatie en het menselijke gehoororgaan*, Doctoral Dissertation, University of Utrecht (1954).
Zwicker, E.: Die Grenzen der Hörbarkeit der Amplitudenmodulation und der Frequenzmodulation eines Tones, *Acustica 2, Akust. Beihefte*, 125-133 (1952).
Zwicker, E., G. Flottorp, and S.S. Stevens: Critical Band Width in Loudness Summation, *J. Acoust. Soc. Am. 29*, 548-557 (1957).

3

DETECTABILITY THRESHOLD FOR COMBINATION TONES

We saw in the preceding chapter that the ear is able to distinguish, to a certain extent, the partials of a complex of tones. It is known, however, that sometimes tones can be heard which are not a part of the sound. For example, stimulating the ear with two loud simple tones of 1000 and 1200 cps, respectively, gives rise to a clear secondary tone of 200 cps.

This phenomenon presents a serious problem for both the frequency and periodicity principle in tone perception. In particular during the nineteenth century, the enigmatic character of these *combination tones*, as they were called by Vieth in 1805, challenged many scientists to more or less extensive studies about their occurrence and origin. Nowadays the opinion proposed by Helmholtz (1856) that they are caused by non-linear processes in the hearing organ is universally accepted.

During the last decades little research about the conditions under which combination tones are audible has been done. This means that our knowledge concerning the significance of these tones in hearing largely depends upon investigations dating from the time that one was not able to measure sound-pressure levels, so this knowledge is much more qualitative than quantitative. As a consequence of this fact, we have no clear notion of the relative importance of the various combination tones predicted on theoretical grounds. This lack of data is especially unfavourable in view of the fact that the existence of particular combination tones is frequently used to explain other hearing phenomena, such as beats of mistuned consonances and the pitch of complex tones with "missing fundamental".

In this chapter, the results of some experiments on the audibility of combination tones as a function of the frequencies of the primary tones are treated. They mainly concern the minimum sound-pressure level required to perceive particular combination tones. Since these tones were studied extensively in the nineteenth century, the description of our experiments is preceded by a survey of the most important results of former investigations.

HISTORICAL REVIEW

The phenomenon

The first communications about a third tone, audible during simultaneous sounding of two tones, are made by the German organist Sorge (1744), his French colleague Romieu (1751), and the Italian violinist and composer Tartini (1754). As Jones (1935) has shown, it is likely that, with respect to the *date* of the discovery, the order of these names must be reversed, so that not unjustly the third tone is sometimes named after Tartini.

It is not surprising that the discoverers, being musicians, detected the new tones for consonant intervals, given by simple frequency ratios as 2:3, 3:4, 4:5, and 5:6. On the relative frequency of the third tone, their statements are not unanimous. Most of the pitches communicated by Tartini correspond with a frequency one octave above the frequency difference of the primary tones. Romieu mentioned the greatest common divisor of their frequencies, whereas Sorge even stated that for three primary tones with frequency ratio 3:4:5 the relative frequency 1 as well as 2 can be observed. The possibility of two simultaneous secondary tones was confirmed by Young (1800), who ascertained that for the frequency ratio 4:5 a tone with the relative frequency 3 is heard as distinctly as a tone with frequency 1.

Weber communicated in 1829 the results of some experiments on the audibility of combination tones, made by the French baron Blein two years before. Blein examined, using strings, a number of intervals within the octave; he found for most of the intervals one combination tone, but in a number of cases also a second one could be detected. The frequencies of these tones, originally given in two tables, are reproduced graphically in Figure 14 as a function of h/l, with h=frequency of the higher primary tone and l=frequency of the lower tone. As we see, the frequencies of all these combination tones agree with $h-l$ and $2l-h$.

In connection with Weber's paper, Hällström (1832) published shortly after it the results of a comparable investigation, already carried out in 1819 with a violin. Figure 15 is plotted from his tables. Hällström distinguished on the basis of his findings combination tones of the first, second, third, fourth order, etc., defined by $D_1=h-l$, $D_2=l-D_1$ $(=2l-h)$, $D_3=h-D_2$ $(=2h-2l)$, $D_4=D_2-D_1$ $(=3l-2h)$, respectively. This derivation demonstrates Hällström's supposition that first-order combination tones may give rise to the production of second-order

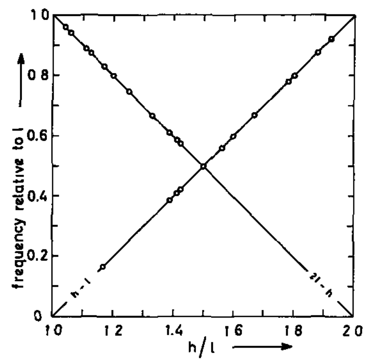


FIGURE 14. Relative frequencies of combination tones heard by Blein (Weber, 1829) for l=256 cps as a function of h/l (h=frequency of the higher primary tone and l=frequency of the lower tone). The open points in Figures 14, 15, 17, and 18 represent the tones heard, and the straight lines the loci of possible values of the relevant combination tones.

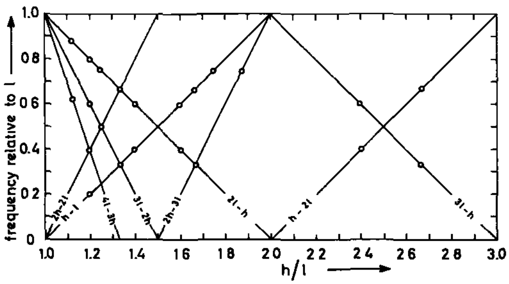


FIGURE 15. Relative frequencies of combination tones heard by Hällström (1832) for l between 512 and 1440 cps.

combination tones, these tones to third-order combination tones, etc.

Some years later, Ohm (1839) criticized this derivation of higher-order combination tones and suggested that they also can be considered as first-order combination tones of particular harmonics of the primary tones, as indicated above by the terms between brackets. Helmholtz (1856) submitted for this reason the phenomenon to a new examination, taking much care to use sinusoidal primary tones. He heard in this case, listening to a large number of different intervals, only the combination tone $D_1=h-l$; for complex tones, however, also Hällström's D_2 could be observed. Helmholtz considered this result as an affirmation of Ohm's supposition, although, in view of the appearance of beats for mistuned consonances like 2:3 and 3:4 (see Chapter 5), he admitted that for simple primary tones also higher-order combination tones are produced. In his opinion, however, these tones cannot be heard individually; they manifest themselves only by beats of mistuned consonances. Helmholtz observed that the loudness of the combination tone increases more rapidly than the loudness of the primary tones. Moreover, he discovered, led by his theory on the origin of combination tones, the existence of the summation tone $h+l$, although being much weaker than the difference tone $h-l$.

König published in 1876 a detailed study on the beats and combination tones produced by two simple primary tones of considerable loudness (tuning forks on resonance boxes). We can extract from his tables the frequency ranges over which he could detect particular combination tones. The result is reproduced in Figure 16. The frequencies of these combination tones agree with $h-l$ and $2l-h$ for $h/l<2$, $h-2l$ and $3l-h$ for $2<h/l<3$, $h-3l$ and $4l-h$ for $3<h/l<4$. König did not deny the existence of other combination tones, but in his opinion they are "extraordinarily weaker" than the tones mentioned.

At the end of the nineteenth century, the phenomenon still was in the centre of interest. Meyer (1896) examined a number of tone intervals with simple frequency ratios, most

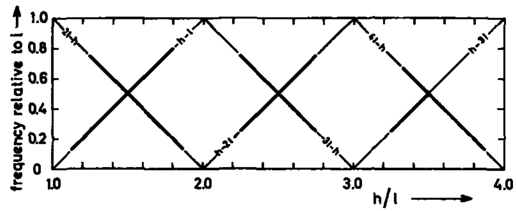


FIGURE 16. Frequency ranges over which combination tones were heard by König (1876) for $l=256$, 512, and 1024 cps.

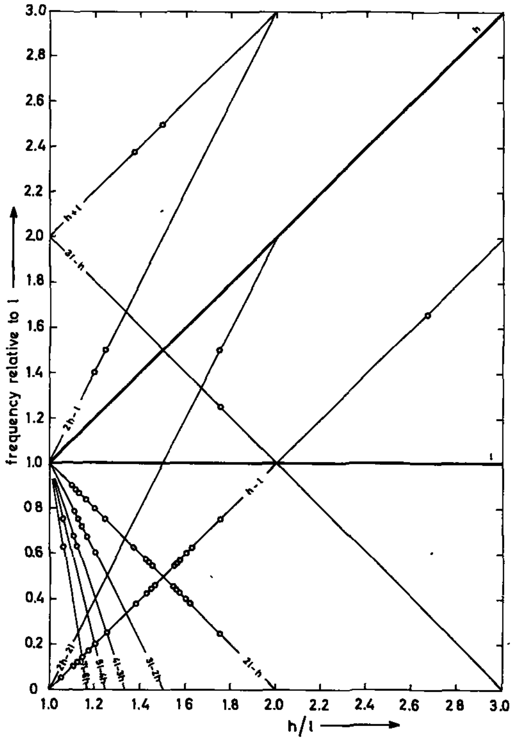


FIGURE 17. Relative frequencies of combination tones determined by Meyer (1896) for l between 400 and 1600 cps.

of them within the octave, using tuning forks on resonance boxes as König did. The combination tones that he and some other observers could detect are given in Figure 17. Meyer noticed that in many cases the loudness of $2l-h$ exceeds the loudness of $h-l$. Combination tones with frequencies between the frequencies of the primary tones, if audible, were always much weaker than combination tones below l.

A much more extensive study was published by Krueger (1900) a few years later. He presented numerous tone intervals with lower frequencies of 256, 512, and 1024 cps, respectively, to 9 subjects in total, also using tuning forks.

The results of these observations are very complex, due to the fact that many secondary tones were heard. The pitch of a large part of these tones could hardly be distinguished, resulting in vague statements concerning their frequencies. Krueger communicated the individual observations for $l=512$ cps and $h/l<2$ and from this long table the points of Figure 18 are taken. He concluded on the basis of these data that four to five orders of

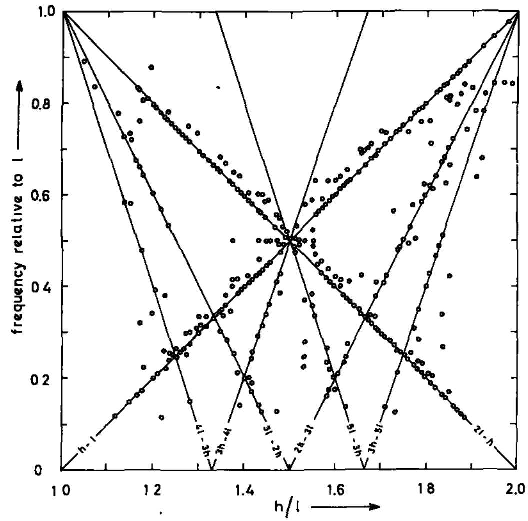


FIGURE 18. Relative frequencies of combination tones determined by Krueger (1900) for $l=512$ cps.

combination tones must be distinguished, the frequencies of which can be determined by subtracting the frequencies of the primary tones and the combination tones in the following way (for $h/l<2$): $D_1=h-l$, $D_2=l-D_1= =2l-h$, $D_3=|D_2-D_1|=|3l-2h|$, $D_4= =|D_3-D_1|=|4l-3h|$ for $h<\frac{3}{2}l$ and $D_4= =|D_2-D_3|=|5l-3h|$ for $h>\frac{3}{2}l$. As Figure 18 shows, not all noticed combination tones are included in this rule; Krueger considered the remaining tones as so-called *intertones*, owing to interference of adjacent combination tones satisfying the rule. Concerning the relative intensities, he concluded that (1) combination tones are fainter for larger intervals; (2) for $h/l<2$ the loudness of D_1 and D_2 is larger than of D_3 and D_4; (3) combination tones with frequencies between the primary tones (so D_1 for $h/l>2$ and D_2 for $h/l>3$) are faint; (4) intervals with coinciding combination tones (*e.g.* 2:3 with $D_1=D_2=D_4$) give the loudest combination tones; (5) the combination tone $h+l$ is not always heard.

Krueger used these results as a basis of a new theory on musical consonance. Stumpf (1910), rejecting this theory, repeated Krueger's experiments for this reason. He, for the first time, presented the results graphically; they are reproduced in Figure 19. We see that he was not able to detect most of Krueger's higher-order combination tones. After Stumpf, only $h-l$ and $2l-h$ are clearly audible over a wide frequency range. Some other combination tones could also be observed for small differences between the frequencies of the primary tones, but they were very faint. In accordance with Krueger, he noticed that very high loudness levels, as used by König, are not favourable to distinguish faint combination tones.

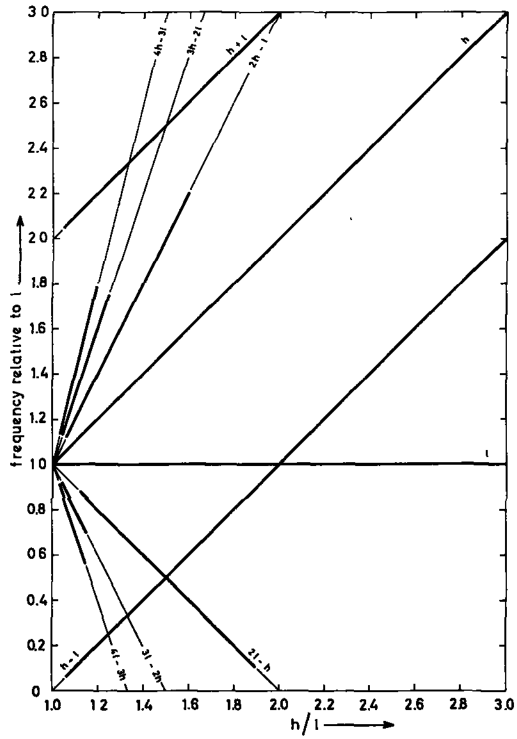


FIGURE 19. Frequency ranges over which combination tones were heard by Stumpf (1910) for l between 150 and 2000 cps.

The studies treated may be considered as the most important publications on the detectability of combination tones up to 1920. There are many other papers giving additional information, some of which are mentioned in the next section. Those papers also confirm our conclusion that all investigators accepted the production of both $h-l$ and $2l-h$, with the exception of Helmholtz, but they disagreed more or less profoundly on the existence of other combination tones.

This disagreement may be related to the fact that they had no equipment to control the sound-pressure level of the primary tones. This situation was altered after about 1920 by the development of electronic amplifiers and high-quality electro-acoustic transducers. No repetition of the former experiments, however, has been carried out up to now. The interest of workers in this field was primarily directed to the percentage of amplitude distortion of music and isolated musical chords that can be tolerated by the ear without being noticed (Janovsky, 1929; von Braunmühl and Weber, 1937; Weitbrecht, 1950; Haar, 1952; Feldtkeller, 1952, 1954; Zwicker and Spindler, 1953), the dependence of the loudness of $h-l$ and $2l-h$ upon the sound-pressure level of the primary tones (von Békésy, 1934; Kuhl, 1939; Zwicker, 1955), and the waveform of the cochlear microphonics of cats and guinea pigs at high sound-pressure levels (Newman *et al.*, 1937; Wever *et al.*, 1940a, 1940b).

Explanations proposed

Although it is not the purpose of this chapter to go into the subject how combination tones are created in the ear (see Chapters 7 and 8), it makes sense to discuss the question how the explanations proposed refer to the frequency and periodicity principle in tone perception. In fact, most of these explanations were in terms of one or the other of these models of the ear's analyzing mechanism.

Originally, combination tones were related to the beats audible for small frequency differences between the tones. The adherents of this view assumed that, for increasing frequency difference, the beats cannot be heard individually but are perceived as a new tone. This so-called *beat-tone theory*, already suggested by Romieu (1751), was accepted universally during more than a century, and adopted by investigators as Lagrange (1759), Young (1800), Chladni (1802), and Weber (1829). Although their statements are not always clear on this point, the underlying proposition was that the beat tone is given by the greatest common divisor of the frequencies of the primary tones (so tone intervals with relative frequencies of 2:3, 3:4, 4:5, 5:6, etc. give rise to a combination tone with relative frequency 1). This means that the periodicity of the superimposed sinusoidal waves of the two tones was considered as resulting in a new tone with a corresponding pitch. Therefore, this theory was in accordance with the conception of the ear as a periodicity detector.

Hällström (1832) demonstrated, however, that this opinion could not be correct: generally, the beat tone is not given by the greatest common divisor but by the frequency difference $h-l$. This led to a modification of the rôle of periodicity: the combination tone corresponds to the number of maxima of the envelope of the superimposed sinusoids. Hällström explained the existence of more than one combination tone at the same time by accepting that the interference of lower-order beat tones gives rise to new ones, defined by his D_1, D_2, D_3, and D_4, respectively, as we saw above.

It is remarkable that, in his paper on the definition of tone (1843), Ohm did not confront his definition with the beat-tone explanation of combination tones that was accepted by him four years earlier (1839). Apparently, he did not realize the contradiction between the condition that a sinusoidal sound wave must be present for hearing a corresponding tone and the beat-tone theory.

Helmholtz, however, clearly saw the incompatibility of Ohm's acoustical law and the beat-tone theory and therefore abandoned the last (Helmholtz, 1856). He presented an entirely new and revolutionary explanation of combination tones, based on the assumption that only for infinitesimally small amplitudes a linear relation between the pressure and the displacement exists for an elastic body. He showed that a quadratic term gives rise to the combination tones $h-l$ and $h+l$, and the harmonics $2h$ and $2l$. It is surprising that Helmholtz explained Hällström's higher-order combination tones, as far as they also appear for simple primary tones, mainly as due to quadratic distortion of lower-order combination tones and not as direct products of higher-order terms in the equation of motion.

It may be convenient to show how a nonlinear relation between pressure and displacement gives rise to combination tones. Let this relation be represented by the equation

$$d=a_1p+a_2p^2+a_3p^3+\dots,$$

with p=pressure and d=displacement. Suppose that

$$p=\hat{p}\cos 2\pi ht+\hat{p}\cos 2\pi lt,$$

with h=frequency of the higher tone and l=frequency of the lower tone, both of the same amplitude $\hat{p}$. Then

$$d=a_1\hat{p}(\cos 2\pi ht+\cos 2\pi lt)+a_2\hat{p}^2(\cos 2\pi ht+\cos 2\pi lt)^2 \\ +a_3\hat{p}^3(\cos 2\pi ht+\cos 2\pi lt)^3+\dots$$

Applying simple trigonometric rules, this equation can be written as

$$\begin{aligned} d=\;&a_1\hat{p}(\cos 2\pi ht+\cos 2\pi lt)+a_2\hat{p}^2\{1+\tfrac{1}{2}\cos 4\pi ht+\\ &+\tfrac{1}{2}\cos 4\pi lt+\cos 2\pi(h+l)t+\cos 2\pi(h-l)t\}+\\ &+a_3\hat{p}^3\{\tfrac{9}{4}\cos 2\pi ht+\tfrac{9}{4}\cos 2\pi lt+\tfrac{1}{4}\cos 6\pi ht+\\ &+\tfrac{1}{4}\cos 6\pi lt+\tfrac{3}{4}\cos 2\pi(2h+l)t+\tfrac{3}{4}\cos 2\pi(h+2l)t+\\ &+\tfrac{3}{4}\cos 2\pi(2h-l)t+\tfrac{3}{4}\cos 2\pi(2l-h)t\}+\dots \end{aligned}$$

As we see, the quadratic term a_2p^2 introduces the combination tones $2h$, $2l$, $h+l$, and $h-l$; the cubic term a_3p^3 the tones $3h$, $3l$, $2h+l$, $h+2l$, $2h-l$, and $2l-h$. Generally, we have that the term a_kp^k introduces new combination tones with frequencies given by $|mh\pm nl|$ with $m+n=k$ (m and n both positive integral numbers). We will refer to these tones as k-th order combination tones.

It is of interest to point to the fact that the amplitude of combination tones, for the relation between d and p considered, does not depend upon the frequencies of the primary tones.

Helmholtz distinguished between objective and subjective combination tones. The first type exists objectively, independent of the hearing organ, and can be reinforced by correctly tuned resonators. Subjective combination tones, however, are not detectable with physical means, but have their origin in the human ear. The fact that he could establish the production of objective combination tones by his siren was used by him as an argument for his explanation of subjective ones.

This attempt of Helmholtz to reconcile the existence of combination tones with his concept of the ear as a frequency analyzer by introducing distortion was criticized by several others. So König (1876), especially in view of the beats for slightly mistuned consonances (see Chapter 5), preferred the beat-tone theory to explain the combination tones represented in Figure 16. In his opinion, Helmholtz's theory may only hold for other very weak combination tones.

This preference for the beat-tone theory above the distortion theory was shared more or less by many other workers, as Preyer (1879), Wundt (1880), Dennert (1887), Voigt (1890), and Hermann (1891). Their most important arguments were: (1) the phenomenon of the so-called *interruption tone* corresponding in pitch with the rate in which a tone is interrupted; if the ear is able to perceive these periodic amplitude variations as a tone, why not also the beats between two simple tones? (2) many investigators could only hear summation tones for complex primary tones, in which case they could be interpreted as difference tones of harmonics; (3) no high sound-pressure levels were required to hear combination tones.

In the long run, however, the distortion theory of Helmholtz was accepted universally. This was partly caused by more evidence on objective combination tones (Lummer, 1887; Rücker and Edser, 1895; Schaefer, 1905; Waetzmann, 1906), but the nearly general acceptance of the frequency principle in tone perception, so strongly propagated by Helmholtz, may be considered as the most important reason.

As we saw above, the only combination tones that Helmholtz could distinguish clearly for simple primary tones were $h-l$ and $h+l$. His theory was especially developed to explain these tones. So the question remained how other combination tones, whose existence was shown irrefutable by other workers, had to be explained in the framework of his theory. Helmholtz himself suggested that they had to be considered as new combination tones of the primary tones and first-order combination tones (Helmholtz, 1863). Bosanquet (1881) criticized this approach, calling attention to the fact that higher-order combination tones can be explained directly by the primary tones when the relation between pressure and displacement consists of not only a linear and a quadratic term but also of higher-order terms. This assumption was supported by experiments of Waetzmann (1910) and Peterson (1915), who both found that, for example, $2l-h$ cannot be considered as produced *via* the distortion tones $2l$ or $h-l$, but must have its own origin.

By the introduction of electronics in the first decades of this century, the phenomenon of nonlinear distortion became so familiar that the existence of combination tones lost much of its enigmatic character.

EXPERIMENTS

The detectability threshold for combination tones can be investigated as a function of the frequencies of l and h, both variable over a wide range. In order to restrict the number of measurements, it is necessary to examine a selection of l- and h-values. The attempt was made to base the following experiments on a representative set of these values.

Detectability threshold for $h-l$

The low difference tone produced by two primary tones of much higher frequency is the combination tone most frequently heard. For this reason, the first series of experiments dealt with the threshold of detectability for $h-l$ with $h-l \ll \frac{1}{2}(h+l)$.

The experimental set-up was simple. Two separated channels were used for the production of the primary tones, each consisting of a sine-wave oscillator (Hewlett-Packard 200 AB), power amplifier, attenuator with steps of 1 and 10 dB, and loudspeaker. In this way, the creation of combination tones by the equipment itself was avoided. Both loudspeakers were situated at a distance of about 35 cm from the subject's ear.

The subject was seated in a soundproof room that was equipped with soundabsorbing walls. The experimenter, who was seated in another room, connected the loudspeakers to their signal sources by means of a pushbutton. The test subject noted detection by operating a signal light in the experimenter's room.

The following experimental procedure was used. At first, the hearing threshold for each of the primary tones was determined. The experimenter presented tone pulses of about 0.5 sec by operating the pushbutton and successively crossed the hearing threshold for decreasing and increasing sound-pressure level (steps of 2 dB). Then, the two tones were given the same sensation level and sound pulses consisting of both tones simultaneously were presented. The detectability threshold of $h-l$, defined as the sensation level of the primary tones for which that combination tone can just be detected, was first determined roughly by increasing the sensation level of the primary tones in steps of 5 dB, and after that more accurately by using steps of 2 dB.

In each test session, only $\frac{1}{2}(h+l)$ was varied, maintaining $h-l$ constant. The detectability threshold was determined for the different values of $\frac{1}{2}(h+l)$ in random order. Four observers with normal hearing thresholds were used, all having much experience in psychoacoustic experiments. As the experimental set-up indicates, the investigation was carried out monaurally.

The detectability threshold was measured for $h-l=100$, 200, and 400 cps, respectively, with $\frac{1}{2}(h+l)$-values between $3.5(h-l)$ and 8000 cps. All tests were carried out twice, on different days, in order to determine the reliability of the data points and the influence of learning effects.

The experimental results for each of the three frequencies of the combination tone are plotted in Figure 20. The points represent

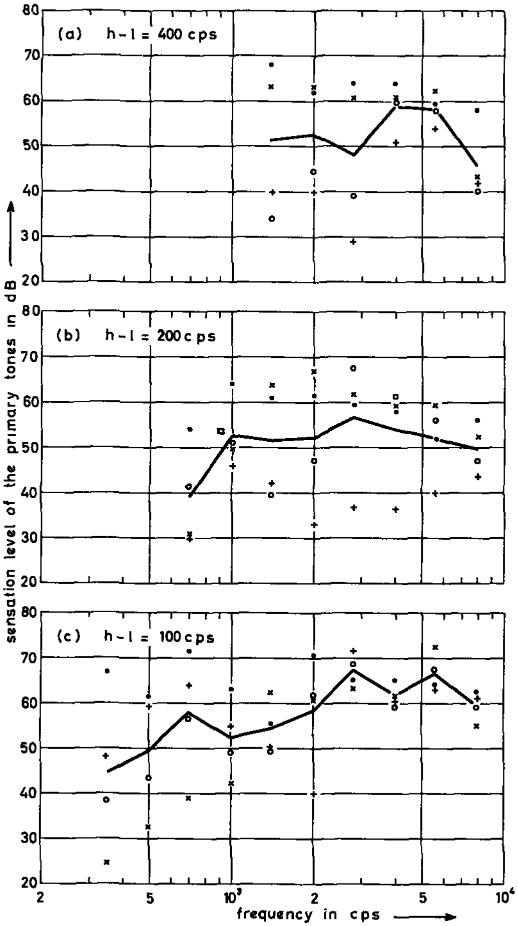


FIGURE 20. Detectability threshold for combination tone $h-l$ as a function of $\frac{1}{2}(h+l)$. The different symbols correspond with four subjects. The solid lines represent the mean value.

the mean values of the two determinations for each of the subjects.

An analysis of the original data showed a small difference of 1.1 dB between the detectability threshold averaged over all first and second determinations, respectively, which may be neglected. The standard deviation of all individual differences is 4.5 dB, which presents a measure of the reliability of the data points.

As Figure 20 shows, there is a large spread in the data points of the four subjects. Therefore, the results do not permit us to say whether the detectability threshold depends upon frequency or not. Averaged over all values of $\frac{1}{2}(h+l)$, the detectability threshold is 57.4, 51.0 and 52.6 dB for $h-l=100$, 200, and 400 cps, respectively. The corresponding values of the standard deviation are 11.1, 11.0, and 11.5 dB.

Summarizing these results, we may conclude that, for the subjects and test frequencies involved, the combination tone $h-l$ becomes audible for sensation levels of the primary tones of about 51-57 dB, with a standard deviation of about 11 dB.

Detectability threshold for combination tones of 800+1000 and 800+1400 cps

The experiments of the preceding section indicated a large spread of the data points. As only four subjects were involved, it appeared to be desirable to investigate this spread more carefully, using a much larger group of subjects and also including higher orders of combination tones.

Two pairs of primary tones were examined in these experiments, namely 800+1000 and 800+1400 cps. For both intervals, the detectability threshold for the combination tones of 200, 400, and 600 cps was determined. The intervals chosen have the attractive property that for 800+1000 cps $h-l=$ $=200$ cps and $2l-h=600$ cps, while 800+1400 cps gives the reverse values $h-l=600$ cps and $2l-h=200$ cps. Furthermore, in the first case the 400-cps combination tone corresponds to $2h-2l$ and $3l-2h$, and in the second case to $2h-3l$ and $4l-2h$.

The apparatus and experimental procedure used were the same as described in the preceding section, with one exception. To facilitate that the subject indeed listened successively to each of the combination tones to be detected, he had a switch with three positions at his disposal. In the middle position, the primary tones were presented; in the other positions, a tone of the same frequency as the relevant combination tone and a tone about 10% higher in frequency were presented, respectively, both of low sensation level. So, for each level of the primary tones the subject had to ascertain, by operating the switch, whether or not an audible combination tone coincided in frequency with one of the auxiliary tones. In this way, reliable responses were obtained.

A total number of 18 subjects, all with normal hearing thresholds, was tested monaurally. The subjects were trained in distinguishing the different combination tones in the complex sound. As there were 6 detectability thresholds to be determined, the total number of subjects was divided in 3 groups of 6 subjects; a different order of threshold determinations, corresponding with a Latin square, was used for the subjects in each group. The experiments were carried out twice for all subjects during the same test session.

Table 1 represents the individual results of the subjects, averaged over the two determinations (subjects 1 to 4 are the same as in the previous experiments). An analysis of variance of the original data showed that it is allowed to use these mean values as a basis for further calculations. This may be demonstrated by the fact that the difference between the detectability thresholds, averaged over all first and second determinations, respec-

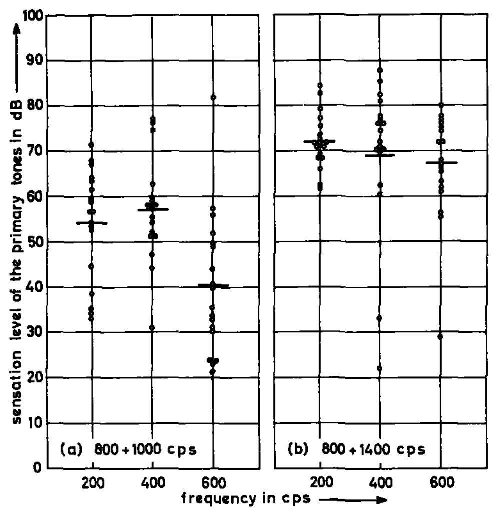


FIGURE 21. Detectability threshold for combination tones of 200, 400, and 600 cps, produced by the tone intervals 800+1000 and 800+1400 cps, respectively. The horizontal dashes represent the mean value of the eighteen subjects.

subject	detectability threshold in dB					
	800+1000 cps			800+1400 cps		
	200 cps	400 cps	600 cps	200 cps	400 cps	600 cps
1	33	47	24	71	85.5	76
2	54	58	30	72.5	77.5	55.5
3	61.5	76.5	35.5	79	88	66.5
4	35	51	32.5	75.5	74.5	56.5
5	64	77	33.5	84.5	81	72
6	56.5	57.5	44	62.5	62.5	72
7	56.5	55.5	52	77	77	63.5
8	44.5	59.5	49	68.5	76	77
9	38.5	44	21.5	66	22	29
10	34.5	31	23	62	60.5	62
11	52.5	51	31	69.5	33	67
12	67	59	49.5	82.5	82.5	77.5
13	59.5	58.5	40	71.5	76.5	80
14	59	54	41	71	71.5	75.5
15	63	58.5	56	72	70.5	65.5
16	67.5	62.5	57.5	68.5	72.5	68
17	53.5	52	24	72	70.5	61
18	71.5	74.5	82	73	70	75
mean value	54.0	57.1	40.3	72.1	68.9	67.2
standard deviation	12.1	11.3	15.5	6.0	16.4	11.6

TABLE 1. Detectability threshold for combination tones of 200, 400, and 600 cps, produced by the intervals 800+1000 cps and 800+1400 cps.

tively, is only 0.5 dB. The standard deviation of all individual differences is 4.1 dB, comparable with the corresponding value in the preceding section (4.5 dB).

The data of Table 1 are reproduced graphically in Figure 21. The horizontal dashes correspond with the mean values of the detectability thresholds. The threshold value for 200 cps with $l+h=800+1000$ cps is the only one that can be compared directly with the results of the preceding section. In Figure 20(b), this threshold value is plotted (small square). As we see, this value agrees rather well with the other data points.

Table 2 represents the results of an analysis of variance for the data in Table 1. From these tables, the following conclusions can be drawn:

source	sum of squares	degrees of freedom	mean squares	ν_1/ν_2	F	p
observers O	8644	17	508.5	17/17	2.94	<.05
intervals I	9728	1	9728	1/2	22.4	<.05
combination tones C	2254	2	1127	2/2	2.59	n.s.
O×I	2938	17	172.8	17/34	3.09	<.01
O×C	3172	34	93.3	34/34	1.67	n.s.
I×C	870	2	435	2/34	7.79	<.01
O×I×C	1899	34	55.8			

TABLE 2. Results of an analysis of variance for the data in Table 1.

1. *Concerning the spread around the mean values in Table 1.* (a) The standard deviation of the detectability threshold varies for the 6 combination tones between 6.0 and 16.4 dB, with an average value of 12.2 dB, comparable with the corresponding values in the preceding section. (b) The corresponding variance can be explained for 61% by differences in the over-all susceptibility to distortion between the subjects. (c) Also the difference in susceptibility for the two intervals is not the same for all subjects; this explains 21% of the variance around the mean values. (d) Finally, 11% of the variance may be caused by the fact that the difference in susceptibility for the 3 combination tones of 200, 400, and 600 cps, respectively, varies from subject to subject. Of the sources of variance, mentioned in (b), (c), and (d), only the first two are significant.

2. *Concerning the differences between the mean values in Table 1.* (a) There is a significant interval effect; that is to say, the detectability thresholds for combination tones of 800+1000 cps are distinctly lower than of 800+1400 cps. (b) There is no over-all significant difference between the detectability thresholds of the 3 combination tones 200, 400, and 600 cps. An analysis of variance using the order ($h-l$, $2l-h$, and higher orders together, respectively) instead of the frequency of the combination tones as a source of variance also indicated that this order cannot be considered as a significant source. (c) There exists a significant interaction of intervals and combination tones; this interaction is due to the fact that only for the interval 800+1000 cps the detectability threshold of the combination tone of 600 cps is significantly lower than of 200 and 400 cps. The Scheffé-test indicated that a difference beyond 10 dB between mean values in Table 1 is significant at 1%-level.

Audibility of combination tones as a function of sound-pressure level and h/l

The preceding sections only give information on the detectability threshold of combination tones for some restricted relations of the primary tones. It was of interest to investigate in connection with these experiments the audibility of combination tones as a function of the ratio between h and l. This will permit a comparison with the results of earlier investigations as discussed in the historical review and represented graphically in Figures 14-19.

For this reason, a new series of experiments was carried out with l=1000 cps and h continuously variable between 1000 and 3000 cps. In this case, not the detectability threshold was determined, but the frequency range over which particular combination tones were audible for different sound-pressure levels of the primary tones.

The same two separated tone channels as in the preceding experiments were used to produce the primary tones. The frequencies of these tones were adjusted by the experimenter. The sound-pressure level at the entrance of the external auditory canal was controlled continuously with a condensor microphone and sound-level meter. The subject had a switch with two positions at his disposal. In one position, the primary tones were presented and in the other position a third tone was introduced, with a frequency adjustable by the subject. Both the experimenter and the observer were seated in the same soundproof room with sound-absorbing walls.

The experimental procedure was as follows. The experimenter had a graph as reproduced in Figure 22 before him. In this graph, the frequencies of all combination tones $|nh \pm ml|$ with $n+m$=2, 3, 4,..., 8 are plotted as a function of h/l. At first, both l and h were tuned to 1000 cps, and the sound-pressure levels were adjusted separately to, for

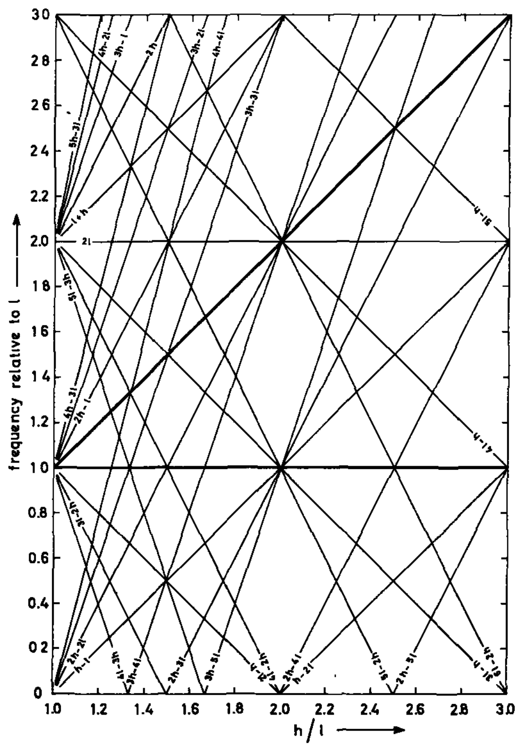


FIGURE 22. This graph represents the relative frequencies of all combination tones $|nh \pm ml|$, with $n+m$=2, 3, 4,..., 8, as a function of h/l.

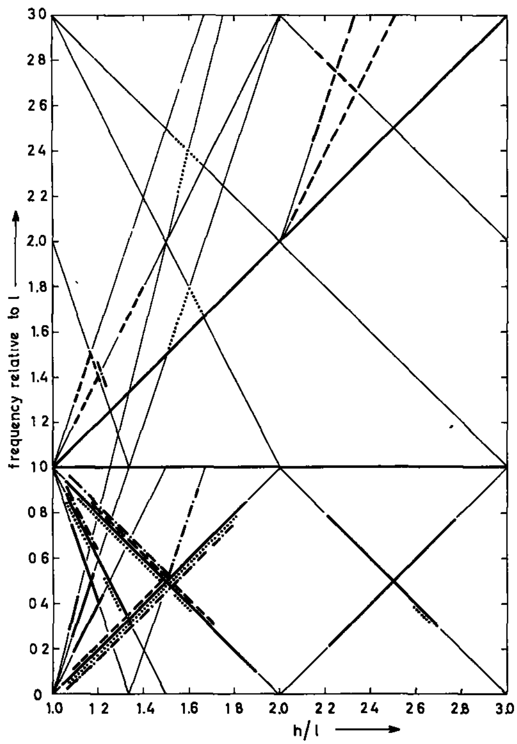


FIGURE 23. Frequency ranges over which combination tones were heard by four subjects for l=1000 cps and 80 dB sound-pressure level of the primary tones. In Figures 23-27, the same type of line is used for each subject. When a combination tone was heard by more than one subject, the corresponding lines are shifted in parallel.

example, 60 dB. Then the experimenter shifted h to a higher frequency, unknown to the subject, and asked him to tune, by listening successively to the primary tones and the auxiliary tone, this third tone to the frequency of a combination tone that he

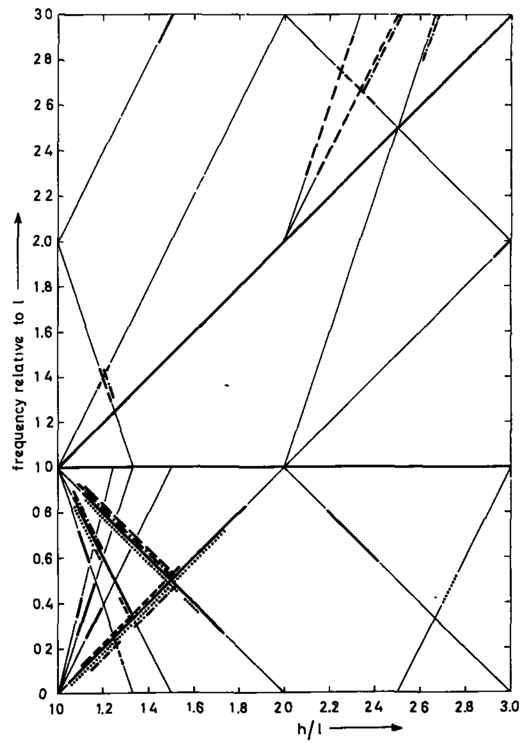


FIGURE 24. Frequency ranges over which combination tones were heard by four subjects for l=1000 cps and 70 dB sound-pressure level of the primary tones.

could distinguish. In order to restrict the total time of searching, the subject's attention was directed successively to the frequency ranges in which a combination tone might be expected after Figure 22. This procedure was repeated for a large number of frequencies of h up to 3000 cps, including all h-values for which two or more combination tones co-

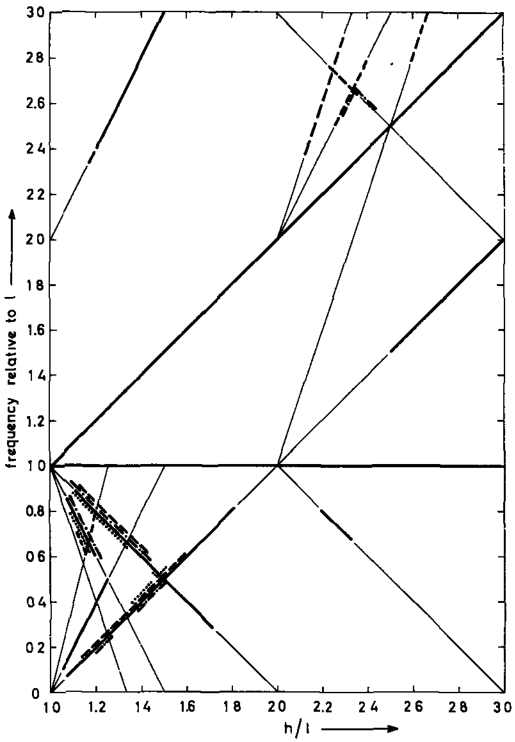


FIGURE 25. Frequency ranges over which combination tones were heard by four subjects for $l=1000$ cps and 60 dB sound-pressure level of the primary tones.

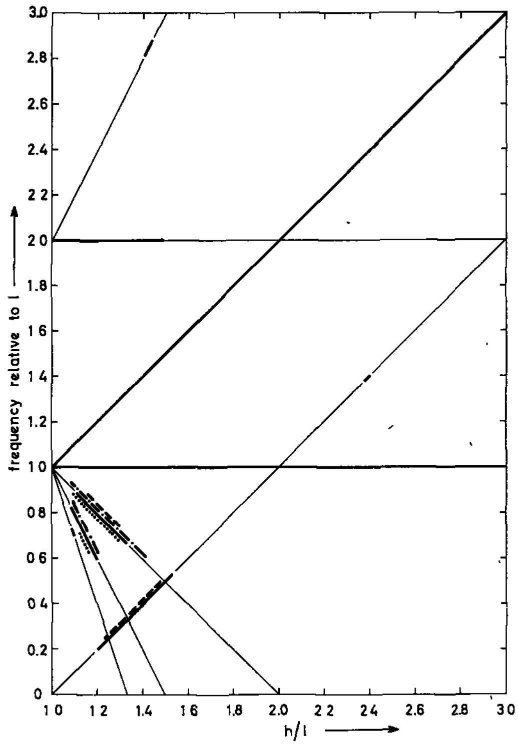


FIGURE 26. Frequency ranges over which combination tones were heard by four subjects for $l=1000$ cps and 50 dB sound-pressure level of the primary tones.

incide. For each of the audible combination tones, the range of h-values was determined over which this tone could be heard. During the experiment, care was taken for a constant sound-pressure level of h.

The experiments were carried out monaurally for four subjects, the same as employed before. Successively, in different test sessions, the sound-pressure level was 80, 70, 60, 50, and 40 dB.

The results are plotted in Figures 23 to 27. The thick lines, partly shifted in parallel when a combination tone was heard by more than one subject, correspond with the frequency ranges over which particular combination tones were audible. The same

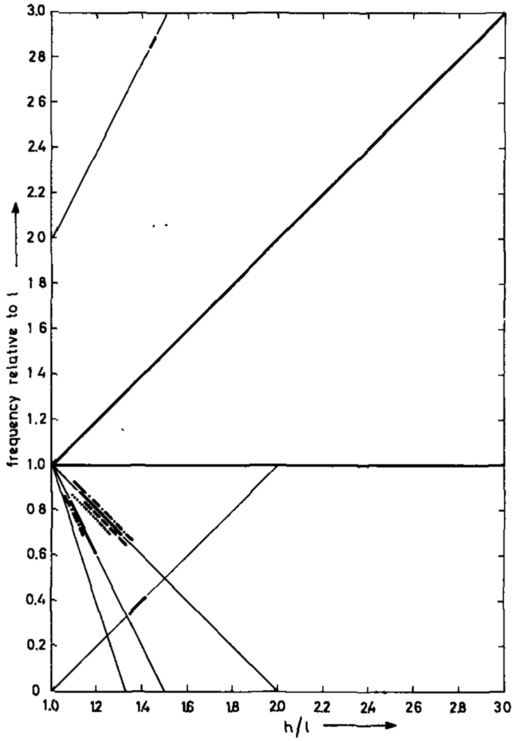


FIGURE 27. Frequency ranges over which combination tones were heard by four subjects for $l=1000$ cps and 40 dB sound-pressure level of the primary tones.

type of line represents each subject in all graphs. Only those lines are copied from Figure 22 for which the corresponding combination tone was distinguished over a certain frequency range.

As the most important conclusions from these experiments we may consider: (1) generally more combination tones were heard for higher sound-pressure levels; (2) the only combination tones heard by all subjects over a shorter or longer frequency range were $h-l$, $2l-h$, and $3l-2h$, always below 1000 cps; (3) with exception of the tones just mentioned, there were large individual differences in the audibility of combination tones; (4) generally more combination tones were heard for $h<\frac{3}{2}l$ than for $h>\frac{3}{2}l$; (5) only one subject heard $h-l$ for $h>2l$; (6) no subject heard $h+l$.

Detectability threshold for the "missing fundamental"

The fact that the pitch of a complex tone does not alter when the fundamental frequency is rather weak or even absent was often explained in the past by stating that the "missing fundamental" is reintroduced by distortion of the ear (see Chapter 6). However, as far as the writer knows, no experimental data on the detectability threshold for this combination tone are available. The following experiments were made for this reason.

Figure 28 represents a block diagram of the apparatus. The complex tone was produced by the same function generator as used in Chapter 2. This generator was adjusted to produce the function

$$\sum_{n=2}^{10} \cos 2\pi nft.$$

The output signal passed through a variable low-pass filter in order to eliminate all fre-

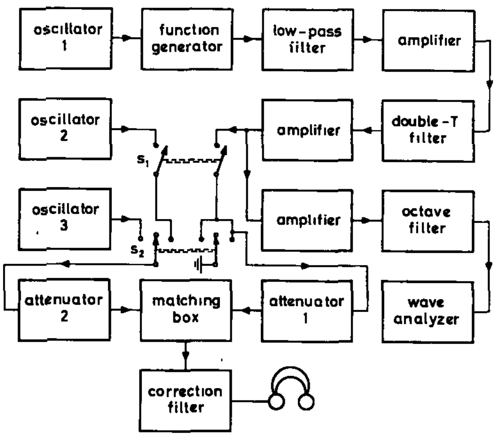


FIGURE 28. Block diagram of the apparatus used for measuring the detectability threshold of the missing fundamental.

quency components above the tenth partial, and a double-T filter (Brüel & Kjaer Frequency and Distortion Measuring Bridge 1607) to eliminate the fundamental frequency, if present. Oscillators 2 and 3 (Hewlett-Packard 200 CD) were incorporated to produce simple tones for comparison purposes. The sound-pressure level of the complex tone and of the simple tones could be adjusted by the attenuators 1 and 2, respectively. The headphone (Beyer DT 48) was provided with a correction filter after the design of Zwicker and Gässler (1952).

Preliminary tests showed that in many cases the fundamental tone was audible only for a sensation level of the complex signal of more than 60 dB. As a consequence, the acoustic signal had to meet stringent requirements, and therefore much care was taken in eliminating the fundamental frequency and testing the output signal of the telephone. It was only possible by using both the function generator and double-T filter to reduce the amplitude of the fundamental frequency of the electric signal to a level sufficiently low: more than 75 dB below the level of the harmonics, controlled continuously during the measurements with a frequency analyzer (Radiometer FRA 1). The Beyer telephone was chosen for its very low nonlinear distortion. The sound-pressure level of the fundamental produced by the telephone was determined as a function of the sound-pressure level of the complex signal by means of an artificial ear. An average-response computer was used to measure the relative level of the fundamental. By comparing this level with the difference between the sound-pressure level of the complex signal for which in the definitive experiments the fundamental became audible and the hearing threshold for a simple tone with the same frequency as this tone, it was checked that the detectability thresholds measured were not the result of distortion of the telephone.

The detectability threshold for the "missing fundamental" was measured in the following way. At first, the experimenter determined the subject's hearing threshold for the complex tone by operating switch S_1 (S_2 in the middle position). Then, the threshold for a simple tone with the same frequency as the fundamental tone, produced by oscillator 2, was determined (S_2 in the right position).

Finally, the sensation level of the complex stimulus was measured at which the fundamental could just be detected. In this case, switch S_1 was in the on-position, and the subject was allowed to operate S_2 for each level of the stimulus so that he could compare this stimulus with simple tones, produced by oscillators 2 and 3 and tuned to the fundamental frequency and a frequency about 10% higher, respectively. These tones were introduced for the same reason as in previous experiments and this proved to be very helpful to obtain reliable responses.

The experiments were carried out for complex tones with (absent) fundamental frequencies of 125, 175, 250, 350, 500, 700, and 1000 cps, respectively. An analysis of the output signal of the telephone showed that for each stimulus the sound-pressure levels of the different harmonics did not differ more than about 10 dB; in general the level of the lower harmonics was somewhat higher than the level of the higher ones.

The same four subjects of the preceding experiments were used. The detectability threshold was determined for each of them monaurally in a random order of the fundamental frequency.

The experimental results are reproduced in Figure 29. The symbols used for each of the subjects are the same as in Figure 20. The upper graph represents the sensation levels of the stimulus at which the fundamental could just be heard, the lower graph gives the corresponding sound-pressure levels.

As the graphs show, also for this combination tone a large spread of the data points

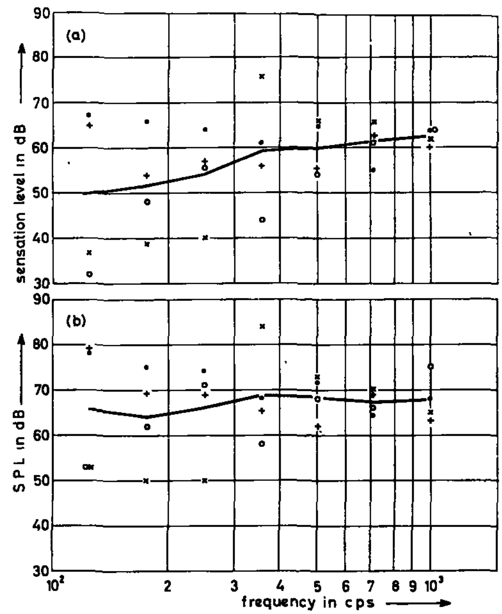


FIGURE 29. Detectability threshold for the missing fundamental of $\sum_{n=2}^{10} \cos 2\pi nft$ as a function of its frequency. The symbols for each of the four subjects are the same as in Figure 20. The solid lines represent the mean values.

exists. The curve in the upper graph gives the impression that for lower fundamental frequencies lower sensation levels are required to hear the corresponding tone than for higher frequencies. However, because of the spread of the data points, this trend is not significant. Averaged over all points, the detectability threshold is 57.0 dB with a standard deviation of 10.5 dB. For the corresponding sound-pressure levels, the mean value is 66.8 dB, with a standard deviation of 8.5 dB.

DISCUSSION

The first question to be asked is whether the experiments support the view that combination tones must be considered as caused by distortion in the ear or the view that they correspond to periodicities in the superimposed sinusoids of the primary tones. The mere fact that nonlinearities occur so frequently in physical systems is not decisive on this point. After the same reasoning, the production in the ear of strong harmonics has been accepted nearly universally but, as we shall see in Chapter 5, experimental evidence does not support this view.

In the author's opinion, there are several reasons to prefer indeed distortion rather than periodicity as the origin of combination tones. The most important objections against the latter view are: (1) combination tones only appear at sensation levels of more than 40-50 dB on the average, whereas beats are even audible at the hearing threshold; (2) from the periodicity view-point, we would expect $h-l$ to be the most distinct combination tone, which is not confirmed by the experiments; (3) it is difficult to understand in which way the various orders of combination tones, with the exception of $h-l$, are related to periodicities in the superimposed sinusoids; (4) combination tones become inaudible when noise in a narrow band around the frequency of these tones is introduced, in the same way as "real" tones do. These objections, briefly formulated, may be sufficient as a justification why the experimental data shall be discussed in terms of nonlinear distortion only.

The most striking conclusion from the experiments is that there are large individual differences in the minimum sensation level for which combination tones appear. This is especially clear from Figure 21; the standard deviation for the various combination tones varies between 6.0 and 16.4 dB. In another way, the same effect is demonstrated by the lines in Figures 23 to 27; only $h-l$, $2l-h$, and $3l-2h$ were heard by all subjects over a shorter or longer frequency range of h.

This fact may be considered as the most important ground for the differences between the results of the investigations treated in the historical review. In particular, the data of Blein (Figure 14), Hällström (Figure 15), Meyer (Figure 17), and Stumpf (Figure 19) agree rather well with our results, with the exception of $h+l$, heard by Stumpf but not by our subjects. As was mentioned, also many other investigators could not detect this combination tone.

However, the results of the experiments of Helmholtz, König, and Krueger are much more different from our data. So, the statement of Helmholtz that for simple primary tones only the combination tones $h-l$ and $h+l$ can be heard, was not confirmed. In this respect, Helmholtz stands alone. The fact that König only communicated the results represented in Figure 16, may be influenced by his preference for the beat-tone theory as an explanation of the existence of combination tones. The data of Krueger (Figure 18) are so complex that they are difficult to evaluate.

As was mentioned, one of the objections against the distortion theory as an explanation of combination tones was that for simple primary tones many investigators did not hear summation tones. In general, the difficulty to detect combination tones with frequencies above the frequency of the lower primary tone was considered as being at variance with this theory. Our experimental data (Figures 23-27) also reflect this difficulty. However, the concept of masking was rather unknown before about 1920. As was shown by Wegel and Lane (1924) and confirmed by several later investigations, lower-frequency tones mask higher-frequency tones much more than the reverse. In this way, it is understandable that mainly combination tones with frequencies below the primary tones are heard.

Concerning the mean values of the minimum sensation level for which combination tones are audible, no simple rule can be given. All mean detectability thresholds investigated in our experiments exceed 40 dB, corresponding with a nonlinear distortion of below 1%. As Janovsky (1929) and von Braunmühl and Weber (1937) have demonstrated, 3 to 5% distortion of speech and music can be introduced without being noticed, so we may conclude that, for usual listening levels, the ear's distortion is sufficiently low to avoid audible combination tones. This fact makes it rather improbable that combination tones represent a constitutive basis for musical consonance, as was stated by Krueger (1906-1910) and Hindemith (1940). In this connection, it is of interest that experiments on the percentage of quadratic and cubic distortion for which distortion of musical intervals within the octave is just perceptible resulted in values of 0.1 to 1.0%, comparable with the distortion percentage of the ear (Weitbrecht, 1950; Haar, 1952; Feldtkeller, 1952; Zwicker and Spindler, 1953).

The conclusion that combination tones are inaudible in practice for usual listening levels of speech and music holds also for the case of the "missing fundamental", as the lower graph of Figure 29 demonstrates. For this reason, the fact that the pitch of a complex tone without fundamental is equal to the pitch of the latter tone cannot be explained by the assumption that the fundamental is reintroduced in the listener's ear. This reasoning, frequently proposed in the past (see Chapter 6), has to be abandoned in favour of the assumption that the periodicity of the complex tone results in a corresponding pitch, as is admitted by many investigators nowadays.

As Figure 21 shows and the analysis of variance (Table 2) demonstrates, the detectability thresholds of the combination tones of 200, 400, and, in particular, 600 cps for the interval 800+1000 cps were significantly lower than for 800+1400 cps. This dependence upon frequency is not compatible with the assumption that for both intervals the distortion is caused by the same nonlinear element. We would expect comparable values in that case because combination tones with the same frequencies are involved.

The results of these experiments suggest also in another way that the combination tones for 800+1000 and 800+1400 cps have different origins. As was mentioned, an analysis of variance with the order instead of the frequency of the combination tones as a source of variance indicated that this order cannot be considered as a significant source. This suggests that, over the subjects, the correlation between the detectability thresholds of combination tones due to quadratic and cubic distortion, respectively, is not larger than between the detectability threshold due to quadratic distortion for one interval and cubic distortion for the other, as was checked explicitly. Also the experiments on the audibility of combination tones as a function of sound-pressure level and h/l demonstrate that for small values of h/l more combination tones are audible than for larger values (Figures 23-27). The same trend is present in the data of Figures 13, 14, 17, and 19. These facts all indicate that the ear's distortion for narrow intervals is larger than for wide intervals, so that it cannot be represented by a frequency-independent nonlinear characteristic. We shall discuss in Chapter 8 the relevance of this conclusion with respect to the question of where combination tones have their origin.

CONCLUSIONS

1. There are large individual differences in the minimum sensation level of primary tones at which combination tones appear. Standard deviations, calculated from experiments with 18 subjects, varied between 6.0 and 16.4 dB for particular combination tones.

2. The most important combination tones are $h-l$, $2l-h$, and $3l-2h$, with h and l representing the frequencies of the higher and lower primary tones, respectively. Only in a few cases, combination tones with frequencies above l were detected. This can be explained by the fact that lower-frequency tones mask higher-frequency tones much more than the reverse.

3. All mean detectability thresholds found exceeded 40 dB, corresponding with a nonlinear distortion of below 1%. As 3 to 5% distortion of speech and music can be introduced without being noticed, we may conclude that, for usual listening levels, the ear's distortion is sufficiently low to avoid audible combination tones. This fact makes it rather improbable that combination tones represent a constitutive basis for musical consonance.

4. Detectability thresholds for combination tones were significantly lower for narrow than for wide tone intervals. This indicates that the distortion of the ear cannot be represented by a frequency-independent nonlinear characteristic.

5. Mean sensation levels of more than 50 dB were required to hear a combination tone corresponding to the missing fundamental of a complex of harmonics. For this reason, the fact that the pitch of a complex tone without fundamental is equal to the pitch of the latter tone cannot be explained by the assumption that the fundamental is reintroduced by the ear's distortion.

REFERENCES

Békésy, G. von: Ueber die nichtlinearen Verzerrungen des Ohres, *Ann. Phys. 20*, 809-827 (1934). [English translation, G. von Békésy: *Experiments in Hearing*, McGraw-Hill Book Company, Inc., New York, pp. 332-344 (1960).]

Bosanquet, R. H. M.: On the Beats of Consonances of the Form h:1, *Phil. Mag. 5th Ser. 11*, 420-436 and 492-506 (1881).

Braunmühl, H. J. von, and W. Weber: Ueber die Störfähigkeit nichtlinearer Verzerrungen, *Akust. Z. 2*, 135-147 (1937).

Chladni, E. F. F.: *Die Akustik*, Verlag Breitkopf und Härtel, Leipzig, pp. 207-209 (1802).

Dennert, H.: Akustisch-physiologische Untersuchungen, *Arch. Ohrenheilk. 24*, 171-184 (1887).

Feldtkeller, R.: Die Hörbarkeit nichtlinearer Verzerrungen bei der Übertragung musikalischer Zweiklänge, *Acustica 2*, *Akust. Beihefte*, 117-124 (1952).

Feldtkeller, R.: Hörbarkeit nichtlinearer Verzerrungen bei der Übertragung von Instrumentenklängen, *Acustica 4*, 70-72 (1954).

Haar, G.: Die Störfähigkeit quadratischer und kubischer Verzerrungen bei der Übertragung von Musik, *Frequenz 6*, 199-206 (1952).

Hällström, G. G.: Von den Combinationstönen, *Ann. Phys. Chem. 24*, 438-466 (1832).

Helmholtz, H. L. F. von: Ueber Combinationstöne, *Ann. Phys. Chem. 99*, 497-540 (1856).

Helmholtz, H. L. F. von: *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*, F. Vieweg & Sohn, Braunschweig, 1st Ed., Chapter 4 (1863).

Hermann, L.: Zur Theorie der Combinationstöne, *Arch. ges. Physiol. 49*, 499-518 (1891).

Hindemith, P.: *Unterweisung im Tonsatz*, I. Teil, Verlag B. Schott's Söhne, Mainz, 2nd Ed., Chapter 3 (1940).

Janovsky, W.: Ueber die Hörbarkeit von Verzerrungen, *Elektr. Nachr.-Techn. 6*, 421-439 (1929).

Jones, A. T.: The Discovery of Difference Tones, *Am. Phys. Teacher 3*, 49-51 (1935).

König, R.: Ueber den Zusammenklang zweier Töne, *Ann. Phys. Chem. 157*, 177-237 (1876). [English translation, R. König: On the Simultaneous Sounding of Two Notes, *Phil. Mag. 5th Ser. 1*, 417-446 and 511-525 (1876).]

Krueger, F.: Beobachtungen an Zweiklängen, *Phil. Studien 16*, 307-379 and 568-664 (1900).

Krueger, F.: Die Theorie der Konsonanz, *Psychol. Studien 1*, 305-387 (1906), *2*, 205-255 (1907), *4*, 201-282 (1909), *5*, 294-411 (1910).

Kuhl, W.: Ueber die Abhängigkeit der Lautstärke des subjektiven Differenztones von der Frequenz der Primärtöne, *Akust. Z. 4*, 43-50 (1939).

Lagrange, J. L.: Recherches sur la nature et la propagation du son, *Miscellanea Taurinensia 1* (1759). [=*Oeuvres de Lagrange 1*, Éditeur Gauthier-Villars, Paris, pp. 39-148 (1867).]

Lummer, O.: Ueber eine empfindliche objective Klanganalyse, *Verhandl. Phys. Gesellsch. Berlin 1886, 5*, 66-69 (1887).

Meyer, M.: Ueber Kombinationstöne und einige hierzu in Beziehung stehende akustische Erscheinungen, *Z. Psychol. Physiol. Sinnesorgane 11*, 177-229 (1896).

Newman, E. B., S. S. Stevens, and H. Davis: Factors in the Production of Aural Harmonics and Combination Tones, *J. Acoust. Soc. Am. 9*, 107-118 (1937).

Ohm, G. S.: Bemerkungen über Combinationstöne und Stösse, *Ann. Phys. Chem. 47*, 463-466 (1839).

Ohm, G. S.: Ueber die Definition des Tones, nebst daran geknüpfter Theorie der Sirene und ähnlicher tonbildender Vorrichtungen, *Ann. Phys. Chem. 59*, 513-565 (1843).

Peterson, J.: Origin of Higher Orders of Combination Tones, *Psychol. Rev. 22*, 512-518 (1915).

Preyer, W.: *Akustische Untersuchungen*, Verlag Gustav Fischer, Jena, pp. 25-29 (1879).

Romieu, J. B.: Nouvelle découverte des sons harmonique graves dont la résonance est très sensible dans les accords des instruments à vent, *Assemblée publique de la Société Royale* *des Sciences tenue dans la grande salle de l'Hôtel de Ville de Montpellier, le 16 décembre 1751*, 77 (1751).

Rücker, A. W., and E. Edser: On the Objective Reality of Combination Tones, *Phil. Mag. 5th Ser. 39*, 341-357 (1895).

Schaefer, K. L.: Über die Erzeugung physikalischer Kombinationstöne mittels des Stentortelephons, *Ann. Phys. 17*, 572-583 (1905).

Sorge, G. A.: *Anweisung zur Stimmung und Temperatur sowohl der Orgelwerke, als auch anderer Instrumente, sonderlich aber des Claviers*, Hamburg (1744).

Stumpf, C.: Beobachtungen über Kombinationstöne, *Z. Psychol. 55*, 1-142 (1910). [= *Beitr. Akust. Musikwiss. 5*, 1-142 (1910).]

Tartini, G.: *Trattato di musica secondo la vera scienza dell' armonia*, Padua, pp. 10-19 (1754).

Vieth, Dir.: Ueber Combinationstöne, in Beziehung auf einige Streitschriften über sie zweier Englischer Physiker, Th. Young und Jo. Gough, *Ann. Phys. 21*, 265-314 (1805).

Voigt, W.: Ueber den Zusammenklang zweier einfacher Töne, *Ann. Phys. Chem. 40*, 652-660 (1890).

Waetzmann, E.: Zur Frage nach der Objektivität der Kombinationstöne, *Ann. Phys. 20*, 837-845 (1906).

Waetzmann, E.: Über die Intensitätsverhältnisse und die Entstehungsweise der Differenztöne höherer Ordnung, *Ann. Phys. 33*, 1299-1304 (1910).

Weber, W.: Ueber die Tartinischen Töne, *Ann. Phys. Chem. 15*, 216-222 (1829).

Wegel, R. L., and C. E. Lane: The Auditory Masking of One Pure Tone by Another and its Probable Relation to the Dynamics of the Inner Ear, *Phys. Rev. 23*, 266-285 (1924).

Weitbrecht, W.: Ueber der Einfluss nichtlinearer Verzerrungen auf die Hörbarkeit von Verstimmungen musikalischer Intervalle, *Fernmeldetechn. Z. 3*, 336-345 (1950).

Wever, E. G., C. W. Bray, and M. Lawrence: The Origin of Combination Tones, *J. Exptl. Psychol. 27*, 217-226 (1940a).

Wever, E. G., C. W. Bray, and M. Lawrence: A Quantitative Study of Combination Tones, *J. Exptl. Psychol. 27*, 469-496 (1940b).

Wundt, W.: *Grundzüge der physiologischen Psychologie*, Vol. 1., Verlag Wilhelm Engelmann, Leipzig, 2nd Ed., pp. 401-410 (1880).

Young, Th.: Outlines of Experiments and Inquiries Respecting Sound and Light, *Phil. Trans. 1800, Part I*, 106-150 (1800).

Zwicker, E.: Der ungewöhnliche Amplitudengang der nichtlinearen Verzerrungen des Ohres, *Acustica 5*, 67-74 (1955).

Zwicker, E., and G. Gässler: Die Eignung des dynamischen Kopfhörers zur Untersuchung frequenzmodulierten Töne, *Acustica 2, Akust. Beihefte*, 134-139 (1952).

Zwicker, E., and W. Spindler: Über den Einfluss nichtlinearer Verzerrungen auf die Hörbarkeit des Frequenzvibratos, *Acustica 3*, 100-104 (1953).

4

TONAL CONSONANCE AND CRITICAL BANDWIDTH

We concluded in Chapter 2 that the ear is able to distinguish partials of a complex of tones if their frequency separation exceeds critical bandwidth. This implies that we may expect interference for frequency differences smaller than this bandwidth. For two simple tones similar in frequency, this interference manifests itself by the long-known phenomenon of beats. The significance of these beats was emphasized strongly by Helmholtz. He used them as a basis for his explanation of why musical consonance is related to simple frequency ratios of the tones involved. Although his conception became well-known, it was criticized severely, in particular by psychologists and musicologists.

In this chapter, the relation between beats and consonance is studied again. To avoid misunderstandings, it may be useful to emphasize in advance that our sole concern is the question why consonance is related to simple frequency ratio. Although the concept of consonance is rather vague and may be different for musicians and laymen, this relationship is always involved. In our opinion, consonance refers to the peculiar sensorial experience associated with isolated tone pairs with simple frequency ratios. We use the term *tonal consonance* to indicate this characteristic experience. As we shall see, experimental results concerning tonal consonance confirm Helmholtz's theory, presenting some quantifications of his conception in which the critical bandwidth will appear to play an important rôle.

HISTORICAL REVIEW

Explanations of consonance

Traditionally, Pythagoras is considered to be the discoverer of the fact that tones produced by a string vibrating in two parts with length ratios of 1:1, 1:2, 2:3, and 3:4, respectively, give much better harmonies than all other ratios. These tone intervals were called *consonances* and on their singular character the harmony of Western music has been developed, especially after, in the Middle Ages, other intervals with ratios of 4:5, 3:5, 5:6, and 5:8 were accepted as imperfect consonances.

The question why consonance is related to simple integral ratios of string lengths has occupied many scholars through the ages. In particular between about 1860 and 1920, numerous studies were devoted to it. The explanations proposed, as far as they are related to hearing theory, are based on one or more of the following data.

1. *Frequency ratio*. One of the first and most important discoveries in acoustics during the rise of modern science in the sixteenth and seventeenth centuries was the dependence of pitch upon frequency. It implied that consonant intervals are characterized by simple frequency ratios and suggested an attractive hypothesis concerning the origin of consonance. So Galilei (1638) stated: "Agreeable consonances are pairs of tones which strike the ear with a certain regularity; this regularity consists in the fact that the pulses delivered by the two tones, in the same interval of time, shall be commensurable in number, so as not to keep the ear drum in perpetual torment, bending in two different directions in order to yield to the ever-discordant impulses." Other scientists as Leibniz and Euler refined this explanation, exchanging the eardrum for the unconsciously counting soul which would prefer intervals the more as the vibrations of the constituting tones concur more frequently. Substantially the same idea was promoted and worked out by Lipps (1885) and Polak (1900), whereas the recent "common long pattern theory" of Boomsliter and Creel (1961) also must be considered as belonging to this group.

2. *Relationship of harmonics*. The discovery in the seventeenth century that the tones of musical instruments are composed of partials (see Chapter 2) gave rise to an alternative explanation of consonance. At first, the mere presence of harmonics with frequency ratios of 1:2, 2:3, etc., in every (complex) tone was considered as a sufficient proof of the consonance of these ratios (Rameau). In the nineteenth century, more-thoroughly formulated implications of the existence of harmonics were presented. Both Helmholtz (1863, Chapters 14 and 15) and Wundt (1880) based the development of melody and harmony on the coinciding harmonics for consonant intervals. The opinion that consonance itself originates in these coincidences was defended more recently by Ogden (1909) and Husmann (1953), although from different points of view. Montani (1947) tried to give this explanation a physiological base.

3. *Beats between harmonics*. The existence of harmonics led also to a quite different hypothesis, in which the phenomenon of consonance was related to beats and roughness, appearing for small frequency differences of simultaneous tones. Although nearly always Helmholtz is mentioned as the originator of this conception, there are much older statements of a quite similar nature (Sorge, 1745-1747). Helmholtz (1863, Chapter 8) noticed that at small frequency differences the beats between two simple tones can be heard individually. At larger differences it becomes impossible to follow the rapid succession of beats, and the sound obtains a rough and unpleasant character. He ascertained that this roughness depends upon both frequency difference and absolute frequency in a compound manner. The roughness has a

frequency-independent maximum for a frequency difference of 30-40 cps, but the maximal roughness increases with frequency. Moreover, the maximal frequency difference for which beats can be noticed increases also with frequency. Beyond this limit, the sound becomes consonant and agreeable. For complex tones as produced by musical instruments, there exist beats between the harmonics of the tones. In this way, Helmholtz (1863, Chapter 10) explained that the smaller the numbers in which the frequency ratio of an interval can be expressed, the more consonant is the interval. The octave, with a frequency ratio of 1:2, is the most consonant interval because all harmonics of the higher tone coincide with harmonics of the lower tone and no beats are introduced. The next most consonant interval is the fifth (2:3), for in this case half of the partials coincides, whereas the other ones lie just midway between partials of the lower tone. He considered it an affirmation of his theory that, in musical practice, thirds (4:5 and 5:6) and sixths (3:5 and 5:8) are avoided in the low-frequency range where partials are nearer to each other than at higher frequencies.

4. *Combination tones.* Although Helmholtz did not deny that beats between combination tones also may contribute to dissonance, this aspect was much more emphasized by Preyer (1879), and in particular by Krueger (1903-1904, 1906-1910). On the basis of detailed experiments on combination tones (see Chapter 3), Krueger concluded that the significance of these tones was strongly underestimated by Helmholtz. Since the total number of combination tones increases with complexity of frequency ratio, these tones could explain the order of consonant intervals, not only for complex but also for simple primary tones. Sandig (1939) compared more recently the character of intervals with both tones presented to the same ear and intervals with one tone presented to the left and the other one to the right ear, respectively, regarding the more neutral character of intervals in the last case as an affirmation of Krueger's theory.

5. *Fusion.* A quite different point of view was developed by Stumpf (1898). In his opinion, neither harmonics nor difference tones are essential to discriminate between consonant and dissonant intervals, whereas he rejected the frequency-ratio theory as mere speculation. Stumpf called attention to the fact, investigated by him before (1890) and confirmed by many others after him, that the degree of fusion ("Verschmelzung") of intervals depends upon simple frequency ratio in the same order as consonance does. By fusion, he meant the tendency of two simultaneous tones to be perceived as a unity. Stumpf understood the close connection to consonance as a causal relation, fusion being the basis of consonance. However, he admitted many years later that this conclusion was not justified and that the relation cannot be considered as a satisfactory explanation of the consonance phenomenon (1926).

Evaluation of these explanations

The existence of these divergent theories suggests that consonance is a complex phenomenon and that conclusive experiments on the value of the explanations mentioned are difficult to find. In contrast with the time before about 1920, modern books on hearing pay only little or no attention to consonance. Is this lack of interest justified and must we agree with those investigators who considered consonance as determined mainly or exclusively by cultural (Cazden, 1945; Lundin, 1947) or even genetic (Moore, 1914; Ogden, 1909) factors?

In answering this question, we have to realize that our consonance perception is indeed profoundly influenced by the development of Western music and musical training. This will be illustrated in two ways.

1. The primary reason why Helmholtz's explanation of consonance by beats was rejected by many investigators was that in their opinion the degree of consonance or dissonance of an interval is not altered by removing the harmonics of the component tones. A study of the observations on which this opinion was based shows that, without exception, the intervals were judged by persons trained musically. This was not considered as a difficulty but, on the contrary, as an essential condition to obtain relevant responses. Stumpf himself, perhaps the most important critic of the beat theory, may be presented as a good illustration. His great interest in the psychology of tone was due to the fact that originally he intended to become a musician (Stumpf, 1883). For him, judgement of a particular tone interval was identical to finding out its musical name, and this knowledge determined entirely the consonance value he attached to the interval. For this reason, he considered intervals like 8:15 and 7:10 as dissonants, also in cases without audible harmonics and combination tones. Apparently, this approach was so self-evident to him (and many others) that he did not consider the possibility that his results had nothing to do with the origin of consonance and dissonance but might be understood only as a demonstration of the success of his musical education and training. The large influence of training was demonstrated by an investigation by Moran and Pratt (1926) in which three subjects, who were able to recognize any presented musical interval, had to adjust the frequency of one of the tones of each of a series of intervals to the correct value for that interval. The results, obtained for simple-tone intervals, indicated that the average frequency adjustments were more in agreement with the interval widths after the equally-tempered scale, as used in music, than after the natural scale, given by simple frequency ratios. These results show that we have to make a clear distinction between interval recognition and consonance judgement. The ability to recognize frequently used intervals does not explain why the singular nature of the impressions produced by particular intervals is related to simple frequency ratios of the component tones.

2. The influence of music on the judgement of tone intervals can be shown also in another way. Originally only 1:1, 1:2, 2:3, and 3:4 were considered as consonant and agreeable intervals. Nowadays, the situation is much more complex. Asking a jury of musicians and psychologists to ascertain the rank order of consonance of all musical intervals within the octave, Malmberg (1917-1918) obtained the order 1:2, 2:3, 3:5, 3:4 and 4:5, 5:8, 5:6, 5:7, 5:9, 8:9, 8:15, and 15:16. Guernsey (1928) has confirmed the well-known fact that musicians make a clear distinction between pleasantness and consonance. It was found in this study that for a group of musicians the ranking of intervals for consonance was about the same as that obtained by Malmberg, but the ordering in terms of pleasantness was quite different: sixths (3:5, 5:8), thirds (4:5, 5:6), fourth (3:4), and minor seventh (5:9) shared the highest rank. For naive subjects, however, consonance and pleasantness are much more similar concepts, as has been investigated in an experiment in which ten subjects had to judge a large number of intervals on ten different semantic scales (van de Geer *et al.*, 1962). A high correlation between consonance and pleasantness scores was found. In fact "consonance" appeared to be used as an evaluation category. For these subjects, too, the sixths, thirds and fourth were the most pleasant intervals, but their evaluation of the octave and fifth was much higher than for musicians, as was also the case in Guernsey's experiments. We may conclude from these results that the original concept of consonance has been split up in two opinions, one held by musicians, the other by naive subjects. This development must be seen as a result of the fact that, in the course of history, preference shifted from intervals given by 1:2, 2:3, and 3:4 to more-complex frequency ratios. For laymen, the meaning of the term *consonance* followed this shift. Musicians, however, maintained the traditional rank order of intervals in terms of consonance, characterized by smoothness and uniformity, independent of evaluation.

After these two digressions on the relation of consonance to music, the question can be asked how to evaluate the various consonance explanations mentioned above. In our attempt to answer this question, we are interested in perception of consonance not so much as a *product* of musical education and training but as a *basis* of it. In our opinion, there exists a typical sensorial phenomenon that is related to simple frequency ratios and is independent of any experience in musical harmony. This particular sensorial phenomenon, which we call *tonal consonance*, may be considered to be basic to the relation between the concept of consonance and simple frequency ratios, as held by musicians and laymen.

With these restrictions in mind, the results of only a few experiments are relevant to decide upon the merits of the five different types of consonance explanation. The most pertinent study is that by Guthrie and Morrill (1928) on the judgement of intervals composed of two simple tones. In this experiment, about 380 subjects were presented with 44 different intervals, with interval widths between 1:1 to beyond 2:3, and the subjects were asked to judge the interval as consonant or dissonant, and as pleasant or unpleasant, respectively. The average results are reproduced in Figure 30. The fact that the two curves are quite similar is in agreement with the conclusion, mentioned above, that the notions consonance and pleasantness are nearly identical for naive subjects.

In this connection, another investigation in which only pleasantness was examined is also relevant. In this study, carried out by Kaestner (1909), pairs of intervals were presented successively to subjects who were

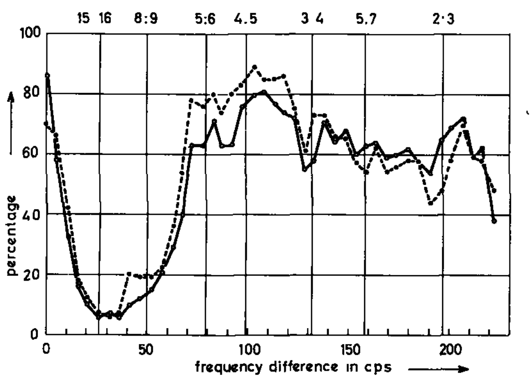


FIGURE 30. Percentage of subjects who judged simple-tone intervals as consonant (solid curve) and pleasant (dashed curve), respectively, plotted as a function of frequency difference between the tones (after Guthrie and Morrill, 1928). The frequency of the lower tone was 395 cps for all intervals.

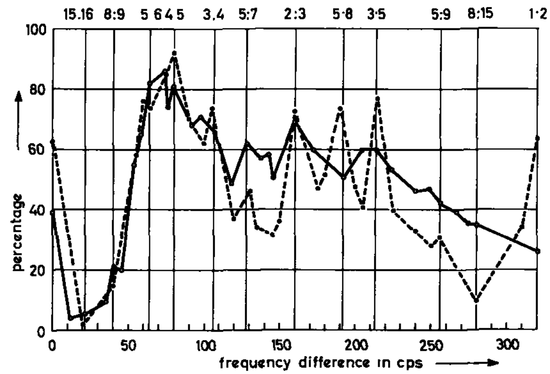


FIGURE 31. Percentage of cases in which a tone interval was judged as more pleasant than the other ones, plotted as a function of frequency difference between the tones (after Kaestner, 1909). The solid curve represents the data for simple, the dashed curve for complex tones. The frequency of the lower tone was 320 cps for all intervals.

asked to indicate which one was more pleasant. The experiments were performed for intervals composed of either simple or complex tones. In both cases, about 30 intervals within the octave were involved and all pairs of intervals were judged. The mean values of the most important results are presented in Figure 31. The simple-tone curve agrees with the curves of Figure 30, whereas the other curve, based on complex tones, shows marked peaks for simple frequency ratios.

These experiments are very useful to evaluate the different explanations of consonance. As we see, for intervals composed of simple tones, simple frequency ratios did not result in singular points of the curves.

On the contrary, the curves suggest that frequency difference rather than frequency ratio is the decisive parameter. For increasing frequency difference, the curves show a marked minimum, followed by a broad maximum.

The only explanation supported by the results of these two experiments is the theory promoted by Helmholtz, according to which the dissonance of an interval is primarily due to rapid beats between the component tones. The minima of the curves correspond in both investigations very well with a frequency difference of 30-40 cps, in accordance with Helmholtz's statement on maximum dissonance. The fact that the curve of Figure 31 based on complex tones shows marked peaks for the intervals corresponding with simple frequency ratios also is in agreement with this explanation.

On the other hand, the experiments do not support the other explanations mentioned above. Against these views, the following objections can be raised:

1. The hypothesis that, anywise, frequency ratio is perceived is contradictory to the finding that the simple-tone curves in Figures 30 and 31 do not have peaks for simple ratios. All evidence in this direction must be due to interval recognition as a result of musical training, the importance of which is demonstrated by the experiments of Moran and Pratt, mentioned above.

2. Insofar as consonance explanations based on relationships of harmonics imply that the presence of harmonics in every complex tone results in a "conditioning" for simple frequency ratios, the objections of (1) again apply. In another view on the influence of harmonics, consonance is considered to be related to the number of coinciding harmonics during actual sounding of two complex tones simultaneously. However, it is not clear how this coincidence may be relevant to consonance other than by the absence of beats or combination tones, because every common partial may be regarded as belonging to only one of the complex tones.

3. The influence of combination tones on consonance perception also is not very probable in view of the data reproduced in Figures 30 and 31. Moreover, the experiments on the audibility of combination tones, treated in Chapter 3, showed that the nonlinear distortion in the hearing organ is so small that it cannot be regarded as a constitutive basis for consonance.

4. The fact that the rank order of consonant intervals is correlated with their degree of fusion cannot be considered as a satisfactory explanation, as Stumpf (1926) himself admitted. This does not mean that the relation has no relevance. However, in this study it will be left out of consideration.

We may conclude from this survey that it is of interest to investigate more thoroughly the hypothesis that tonal consonance, the peculiar character of intervals composed of complex tones with simple frequency ratios, is due to the absence of beats between harmonics of the component tones.

EXPERIMENTS

In the investigation by Guthrie and Morrill, only intervals with a lower tone of 395 cps were used. Kaestner used 256 and 320 cps for the lower frequency. So these studies give no information on the dependence of interval evaluation upon absolute frequency. This information is essential for an evaluation of Helmholtz's theory on the relation between consonance and beats. For this reason, the following experiments were carried out.

Method and procedure

In the experiments, subjects had to judge tone intervals as a function of two parameters: situation of the interval in the frequency range and frequency difference between the component tones. The geometric mean of the frequencies of the two tones was taken as a measure of the first parameter. In order to separate the influence of the parameters as much as possible, this mean frequency has advantages to frequency of the lower tone of the intervals which was used in the earlier studies. For the same reason, different groups of subjects were used for each of the mean frequencies involved.

The subjects judged each tone interval on a 7-point scale, "consonant-dissonant", 1 corresponding with most dissonant, 7 with most consonant. Some subjects asked for the meaning of *consonant*. In that case, the experimenter circumscribed the term by *beautiful* and *euphonious*. This procedure is justified because, as was ascertained earlier (van de Geer *et al.*, 1962), *consonant, beautiful,* and *euphonious* are highly correlated for naive subjects. In fact, they represent one dimension in semantic space: evaluation.

The experimental set-up was very simple. The tones were produced by two sine-wave oscillators and reproduced by a loudspeaker in front of the observer. The sound pressure near the subject's ear was kept at a constant level of about 65 dB *re* 2.10^{-4} dyn/cm^2. The subjects were tested individually in a soundproof room with sound-absorbing walls. The experimenter was seated in another room and presented each interval during about 4 sec. He had to readjust the frequency of the oscillators after each exposure, resulting in a pause of 10 to 20 sec between exposures. An electronic counter was used to adjust the frequencies very accurately.

The experiments were carried out for 5 values of the mean frequency of the intervals: 125, 250, 500, 1000, and 2000 cps. Each subject was used only in one test session, in which he had to judge 12 to 14 different interval-width values around one of these mean frequencies. To avoid the influence of interval recognition, the widths of these intervals were chosen on the basis of frequency difference, not on frequency ratio.

The following procedure was used. Firstly, the subject read written instructions concerning the purpose of the test and the way in which he had to record his responses on a sheet with horizontal lines, each provided with 7 short vertical dashes. After that, a preliminary series of 10 different intervals, chosen at random out of the interval widths

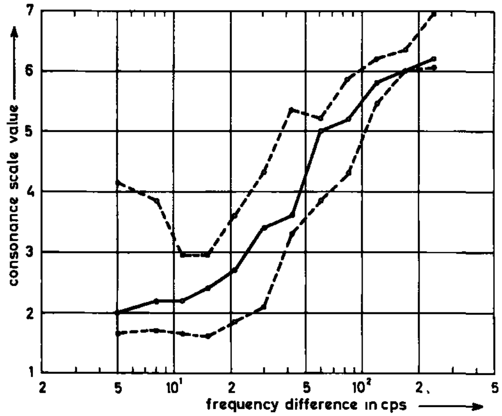


FIGURE 32. Consonance-rating scores of simple-tone intervals with a mean frequency of 125 cps as a function of frequency difference between the tones. The solid curve corresponds with the median, the dashed curves with the lower and upper quartiles of the scores (11 subjects).

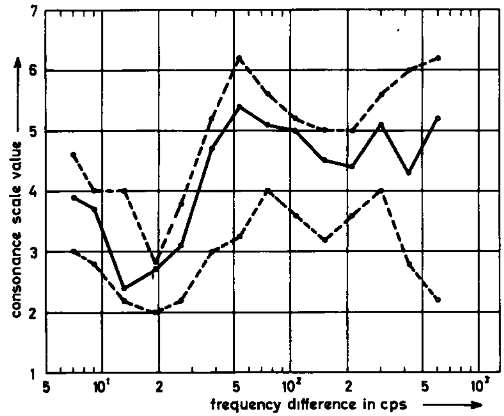


FIGURE 33. As Figure 32, but with mean frequency 250 cps (10 subjects).

used in the experiment, were presented in order to make the subject familiar to the differences between the stimuli and to warrant an adequate use of the 7-point scale. Then, 5 series of 12 to 14 intervals were presented (12 for 125 cps, 14 for the other mean frequencies). Each series contained the same interval widths but in a different (random) order. The first interval of a series was always different from the last one of the preceding series.

The test subjects were young male adults of about 20 years of age and with secondary-school training. For the mean frequencies 125, 250, 500, 1000, and 2000 cps, the number of subjects was 19, 22, 18, 11, and 18, respectively.

Results

To exclude data of subjects who were not able to give consistent responses, for each of them test-retest reliability was determined by computing the correlation coefficient between the scores of the first and the last of the 5 series of interval widths presented to the subjects. Only the data of those subjects were maintained who had a correlation coefficient above 0.5. Their average scores of the 5 series were used for further calculations. In this way, the number of accepted subjects was reduced to 11, 10, 11, 10, and 8, respectively, for the mean frequencies 125 to 2000 cps.

The experimental results for the different mean frequencies are reproduced in Figures 32 to 36 as a function of interval width. In each of these graphs, the solid line connects points representing the median; the other lines

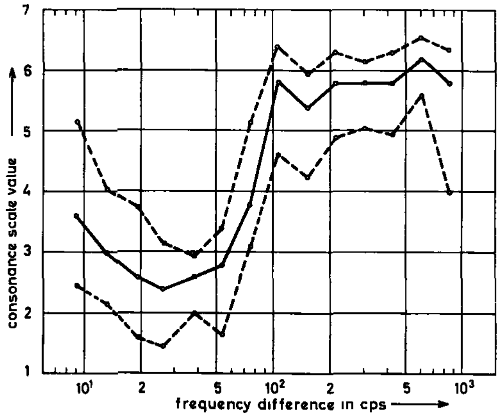


FIGURE 34. As Figure 32, but with mean frequency 500 cps (11 subjects).

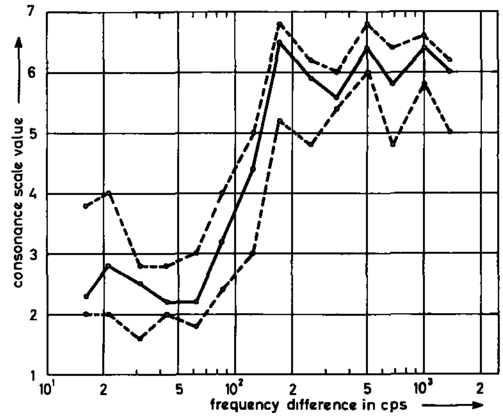


FIGURE 35. As Figure 32, but with mean frequency 1000 cps (10 subjects).

correspond with the lower and upper quartiles of the scores.

Discussion

The curves in Figures 32 to 36 have the same general course as in Figures 30 and 31 (solid line). They show a minimum for small frequency differences, followed by a more or less broad plateau for larger differences. To characterize the curves, two points can be used: the minimum, and the frequency difference for which the plateau is reached. We shall consider each of them.

The interval widths corresponding with the minima of the curves in Figures 32 to 36 are plotted in Figure 37 as a function of mean frequency. Also for the curves of Guthrie and Morrill and of Kaestner, the minima are marked.

The only other data found in literature with which our results can be compared, are from Cross and Goodwin (1893), who published some data concerning the "point at which the harshness of the dissonance produced by the tones of two resonators reaches a maximum". These points, investigated for only one subject, are reproduced in Figure 37.

In comparing and evaluating these data, we have to realize that the minima in the consonance curves are rather broad, so that the points are not very precise. Nevertheless, it will be clear that the experimental results do not confirm Helmholtz's opinion that the frequency difference for maximum roughness is independent of frequency. Although the value of 30-40 cps, given by him, agrees with the data points in the frequency range between 500 and 1000 cps, the general trend of the data indicates that, for increasing fre-

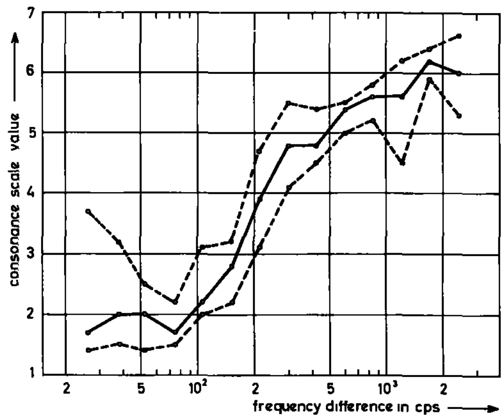


FIGURE 36. As Figure 32, but with mean frequency 2000 cps (8 subjects).

quency, also the interval width for maximum roughness or dissonance increases. The solid curve corresponds with 25% of the critical bandwidth, adopted from a paper of Zwicker *et al.* (1957). The graph suggests that, instead of Helmholtz's hypothesis of a constant frequency difference, a frequency difference proportional to critical bandwidth gives a better fit to the data.

Similar things can be said about the minimum frequency difference of intervals that are judged as consonant. The vertical dashes in Figure 38 represent the interval widths for which the curves in Figures 30 to 36 reach their maximum. Because this value cannot be determined exactly for some curves, dashes instead of points are plotted. In the same graph, relevant data of some other studies are reproduced. The open points correspond with the limit of audible beats as determined by Cross and Goodwin (1893), the crosses correspond with the smallest consonant intervals after an investigation by Mayer (1894). A clear relationship exists between these data, justifying the conclusion that consonance is closely related to the absence of (rapid) beats, as in Helmholtz's theory. The critical-bandwidth curve fits the data rather well.

In conclusion, Helmholtz's theory, stating that the degree of dissonance is determined by the roughness of rapid beats, is confirmed. It appears that maximal and minimal roughness are related to critical bandwidth, with the rule of thumb that maximal tonal dissonance is produced by intervals subtending 25% of the critical band, and that maximal tonal consonance is reached for interval widths of 100% of the critical band-

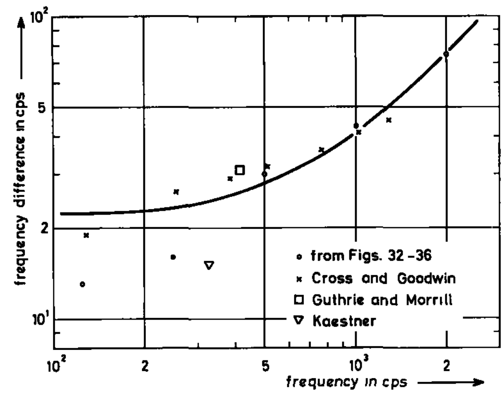


FIGURE 37. Frequency difference of two simple tones for maximum dissonance as a function of the mean frequency of the tones. The solid curve corresponds with 0.25 critical bandwidth as given by Zwicker *et al.* (1957).

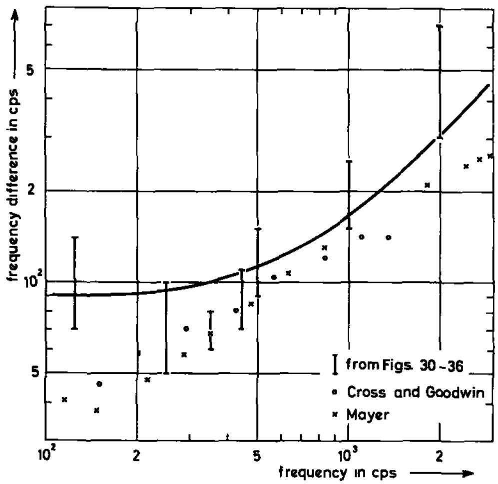


FIGURE 38. Frequency difference of smallest consonant interval of two simple tones as a function of the mean frequency of the tones. The solid curve represents the critical bandwidth.

width. This result confirms the hypothesis, stated in the introduction of this chapter, that interference of simple tones occurs for frequency differences less than the critical bandwidth.

CONSONANCE FOR COMPLEX-TONE INTERVALS

The data of the preceding experiments are used in this section to explain not only why, for complex tones, consonance is related to simple frequency ratio, but also to illustrate some other well-known properties of consonant intervals.

As Figures 32 to 36 show, all curves have quite similar shapes. This means that they all can be substituted by one curve in which consonance score is represented as a function of the interval width in units of critical bandwidth. This standard curve is reproduced in Figure 39. It has been derived by plotting in one graph the data points for each of the mean frequencies as a function of critical bandwidth and drawing the curve that best fits all data. The curve is extended for small frequency differences on the basis of the curves in Figures 30 and 31. By a linear transformation, the evaluation scale is substituted by a "consonance" scale, 1 corresponding with maximum, and 0 with minimum appreciation.

The curve in Figure 39 can be used to get some impression of how consonance varies for complex tones as a function of the frequency difference between the fundamentals.

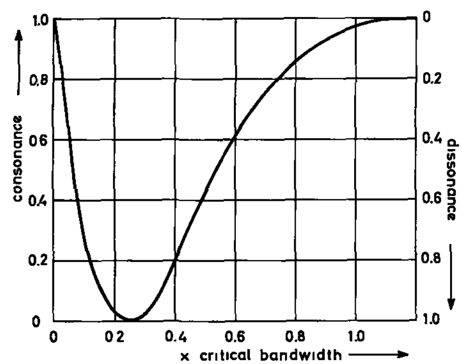


FIGURE 39. Standard curve representing consonance of simple-tone intervals as a function of frequency difference in units of critical bandwidth. The curve is based on the data points in Figures 30-36. The consonance and dissonance scales are arbitrary.

In this case, consonance depends not only upon the frequency difference of the fundamentals, but also of the harmonics.

We assume that the total dissonance of such an interval is equal to the sum of the dissonances of each pair of adjacent partials, using the right-hand scale of Figure 39 to compute the total dissonance. This assumption implies that these dissonance values may be added. Although these presuppositions are rather speculative, they are not unreasonable as a first approximation and may be justified for illustrating how, for complex-tone intervals, consonance depends upon frequency and frequency ratio.

In this way, the curves of Figures 40 and 41 were computed for complex tones consisting of 6 harmonics. Figure 40 illustrates how consonance varies as a function of interval width, and can be compared with similar curves plotted by Helmholtz (1863, Chapter 10) and von Békésy (1935). Figure 41 shows how the consonance of some intervals, given by simple frequency ratios, depends upon frequency.

These curves may be considered as an illustration of the following properties of tone intervals.

1. Singular points of the curve of Figure 40 correspond with simple frequency ratios of the component tones. Since we restricted the number of harmonics to 6, only peaks for frequency ratios that contain the numbers 1 to 6 could appear. If also the seventh and eighth harmonics were included, the curve would have shown extra peaks for 4:7, 5:7, 6:7, 5:8, and 7:8. In this way, it may be

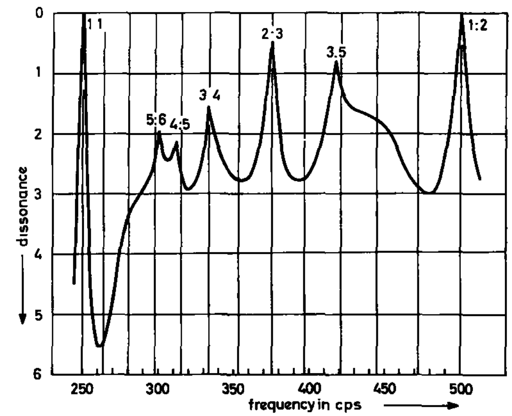


FIGURE 40. Illustration of the way in which consonance of an interval with a lower complex tone of 250 cps and a variable higher one depends upon the frequency of the latter tone. Both complex tones consist of six harmonics. The vertical lines represent interval width after the equally-tempered scale.

clear that, for complex tones as produced by musical instruments, consonance is related to simple frequency ratios.

2. Simpler frequency ratios are represented by sharper peaks. This means that octaves and fifths are much more sensitive to a deviation from their correct frequency ratio than the other consonant intervals are. It explains why the impure thirds are much more tolerable in the equally-tempered scale (vertical lines in Figure 40) than impure octaves and fifths would have been.

3. The rank order of consonant intervals as given by Malmberg (1917-1918) agrees rather well with the relative heights of the peaks in Figure 40 and the curves in Figure 41. Fur-

thermore, Figure 41 suggests that there are only minor differences between the degree of consonance of the fourth and the thirds.

4. As Figure 41 shows, the degree of consonance is nearly independent of frequency over a large range. However, below a critical frequency, the intervals become more and more dissonant, as a consequence of the bend in the critical-bandwidth curve at about 500 cps. The critical frequency is lower for more-consonant intervals. This behaviour reflects the musical practice to avoid thirds at low frequencies and to use mostly octaves or wider intervals.

5. Apart from the range below 100 cps, the dissonance value is 0 for the octave (Figure 41). This means that, for up to 6 harmonics, all frequency differences between adjacent harmonics exceed critical bandwidth. It appears that this does not apply for tones with higher partials. This fact explains why complex tones with strong higher harmonics sound much sharper than tones consisting of only 6 harmonics. This fact was already emphasized by Helmholtz (1863, Chapter 5).

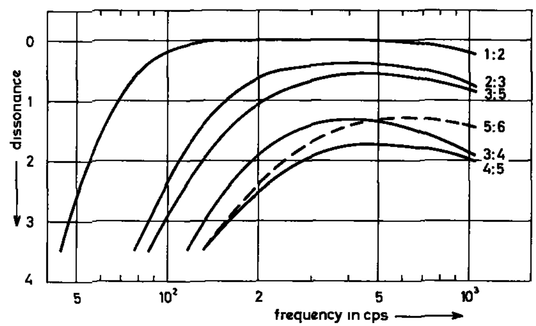


FIGURE 41. Illustration of the way in which consonance of some intervals with simple frequency ratios depends upon the frequency of the lower tone. Both complex tones consist of six harmonics.

This exposition demonstrates that the singular nature of complex-tone intervals with frequency ratios determined by simple integral numbers can be explained by the interference of harmonics. In this connection, recent experiments in which subjects had to compare different tone intervals are of interest. For simple-tone intervals the subjects did not use frequency ratio as a criterion, but for complex-tone intervals they did (Levelt *et al.*, 1966).

STATISTICAL ANALYSIS OF CHORDS IN MUSIC

The preceding section showed that several properties of tone intervals can be explained by interference of partials. This interference occurs, as the experiments indicated, for frequency differences smaller than critical bandwidth. Apparently, this bandwidth plays an important rôle in the sensation of simultaneous tones.

This conclusion raises the interesting question whether in music, too, we may find properties related to critical bandwidth. Some preliminary investigations, in which chords of musical compositions were analyzed (Plomp and Levelt, 1962), were very promising, and for that reason a more detailed study was made.

The basic idea underlying these analyses was the following. During the process of composing, the composer at every moment makes a selection of tones from the total set of tones "available" to him. One of the criteria of selection is that the composed sequence of chords realizes some desired variance in consonance and dissonance. Leaving the time dimension out of consideration, a "vertical" dimension remains: the composition of the chord out of simultaneously present tones. We may get some insight in this vertical dimension by investigating the density distribution of simultaneous tones, partials included, as a function of frequency. This is a statistical approach; it does not give information about the occurrence of specific chords but only about the frequency of occurring of different tone intervals.

An illustration may serve to explain how the analysis was done. Suppose that we are interested in the density distribution of intervals with $c^2 = 523.3$ cps as the lower tone. Firstly, we restrict ourselves to the case where only fundamentals are taken into account. In this case, we take out of a musical composition all chords which contain c^2 and a higher tone simultaneously. We then determine the fraction of time, relative to the total duration of these chords, during which the nearest higher tone is separated from c^2 by a distance of one semitone ($c^2\sharp$ or $d^2\flat$), two semitones (d^2), etc. An example of such a density distribution is given in Figure 42 (solid line). The cumulative distribution, derived from the density distribution by taking the fraction of time the interval does not exceed one semitone, two semitones, etc., also is given (dots and dashes).

The procedure can be repeated by including second harmonics, second and third harmonics, etc. In general, in the case of n harmonics, we take chords that include c^2 as one of the first n harmonics of a lower tone (fundamental included). The density distribution is then calculated for distances between c^2 and the nearest higher tone which may also be one of the first n harmonics of a lower tone (fundamental included). In Figure 42, distributions for $n=6$ are plotted. It is found, as was expected, that the 50% point of the cumulative distribution for $n=6$ gives a smaller interval value than the corresponding point in the cumulative distribution for $n=1$.

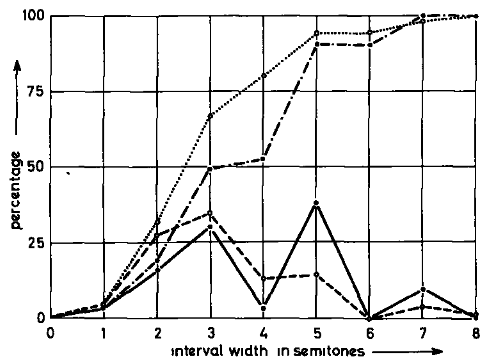


FIGURE 42. Example of interval distributions at $c^2 = 523.3$ cps for $n=1$ (solid curve) and $n=6$ (dashed curve). The other curves represent the cumulative distributions for $n=1$ (dots and dashes) and $n=6$ (dots). The interval distributions were computed from the last movement of J.S. Bach's *Trio Sonata for Organ No. 3* in *c* minor.

fundamental	frequency (cps)	number of harmonic	frequency of harmonic (cps)	deviation from c^2 (cps)
c^2	523.252	1	523.25	0
c^1	261.626	2	523.25	0
f	174.614	3	523.84	+0.59
c	130.813	4	523.25	0
$C\sharp$	103.826	5	519.13	−4.12
F	87.307	6	523.84	+0.59
D	73.416	7	513.91	−9.34
C	65.406	8	523.25	0
$A_1\sharp$	58.270	9	524.43	+1.18
$G_1\sharp$	51.913	10	519.13	−4.12

TABLE 3. Fundamentals with c^2 as the first, second, third, ..., tenth harmonic, respectively. The deviations of the frequency of these harmonics from the frequency of c^2 are indicated in the last column (equally-tempered scale).

Table 3 gives values of frequencies of tones that contain c^2 as their n-th harmonic, with $n=1, 2, \ldots\ldots, 10$. The table also gives the frequencies of the harmonics of these tones on the basis of the equally-tempered scale. As is well-known, these frequencies deviate somewhat from the frequency of c^2 in some cases. These deviations are left out of consideration here.

To facilitate compution of interval distributions for different values of the basic frequency and different numbers of harmonics, special equipment has been developed. It consists of (1) an apparatus to transmute the notes and duration of chords, "played" successively on a key-board, in punch code, using a 8-bit tape, and (2) an apparatus to read out the tape and to compute the interval distribution with both basic frequency and n adjustable.

Two musical compositions were analyzed in this way, namely the last movement of J.S.Bach's *Trio Sonata for Organ No. 3* in c minor, and the third movement (Romanze) of A. Dvořák's *String Quartet Op. 51* in e♭ major. In both cases, interval distributions were computed for $C=65.4$ cps, $G=98.0$ cps, $c=130.8$ cps, $g=196$ cps, $c^1=261.6$ cps, $g^1=392$ cps, etc., and taking into account n harmonics with $n=1, 2, 3, \ldots\ldots, 10$. The interval width which is not exceeded during 25, 50, and 75% of time, respectively, was calculated for each of these distributions (firstly in semitones and from these values in cps).

The results are reproduced in Figures 43 and 44 as a function of frequency, with n as a parameter (solid lines). Since the data for $n=10$ were quite similar to the data for $n=9$, the former case has been left out. The

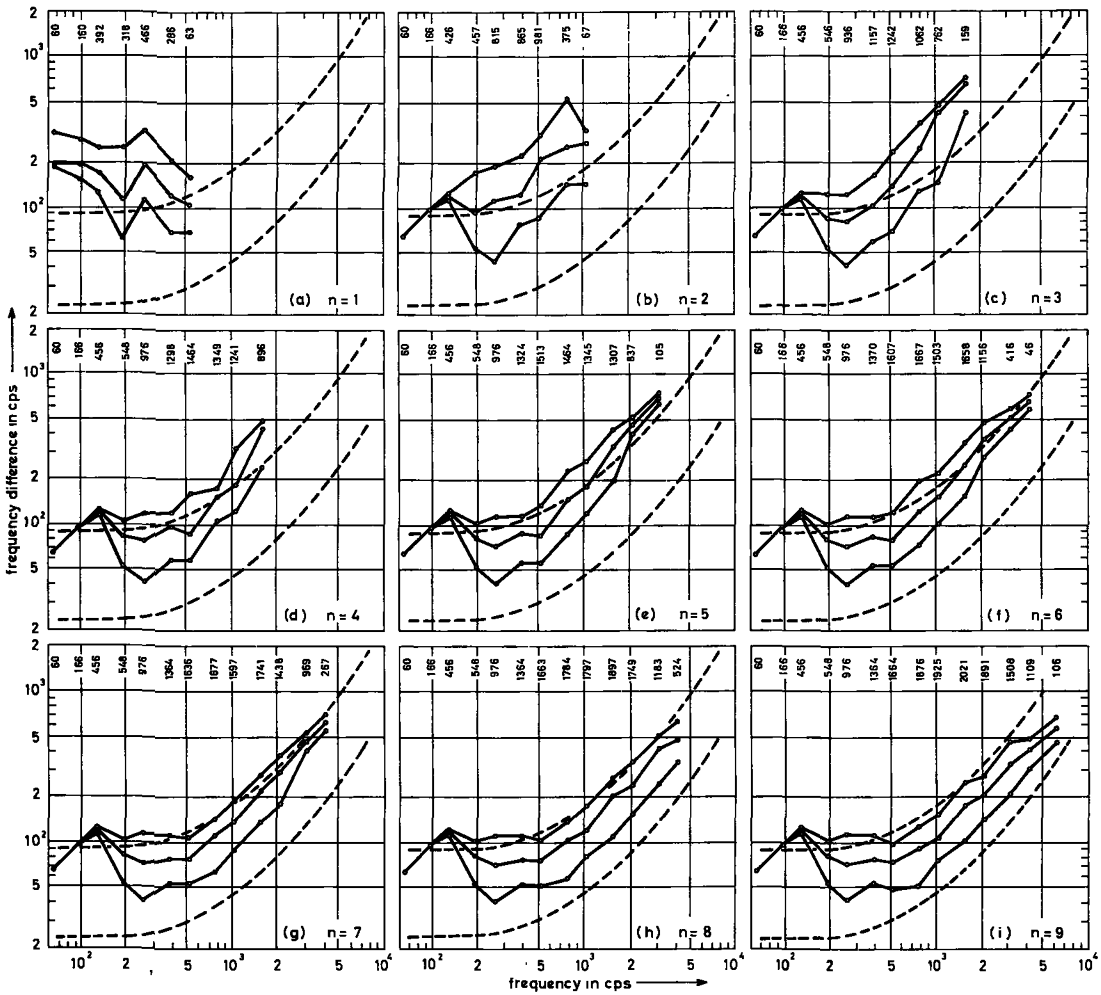


FIGURE 43. Results of a statistical analysis of the chords of the last movement of J. S. Bach's *Trio Sonata for Organ No. 3* in *c* minor with *n* (=number of harmonics taken into account) as a parameter. The solid curves represent the interval width in cps between adjacent partials, plotted as a function of frequency, which is not exceeded in 25, 50 and 75% of time, respectively, computed from curves as reproduced in Figure 42. The dotted curves correspond with critical bandwidth and a quarter of this bandwidth.

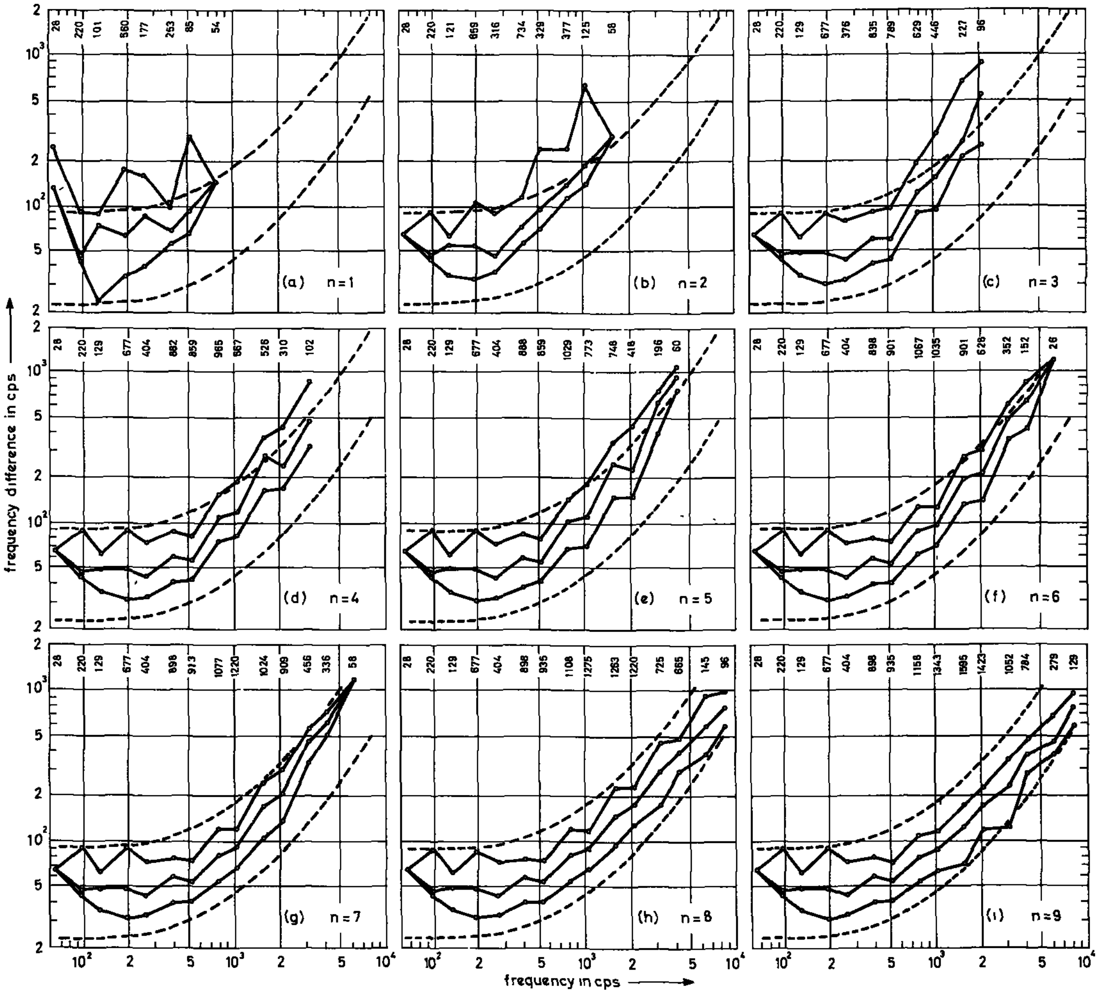


FIGURE 44. Results of a statistical analysis of the chords of the third movement (Romanze) of A. Dvořák's *String Quartet Op. 51* in e♭ major. The curves have the same meaning as in Figure 43.

dashed lines represent the critical bandwidth after Zwicker *et al.* (1957), plotted as a function of the lower cut-off frequency, and a quarter of this bandwidth, corresponding with maximum dissonance (Figure 39). The total duration of time of all chords on which the concerning interval distribution was based is indicated for each frequency, using the duration of the shortest note occurring in the composition as a time unit.

To grasp the significance of the curves, it may be helpful to trace their shift as a function of the number of harmonics. This is done on the basis of the graphs of Figure 43. For the case that only the fundamental tone was taken into account, most of the intervals exceed the critical bandwidth, in particular for the lower frequencies (Figure 43a). It will be clear that, as a function of frequency, all intervals with the same frequency ratio between the component tones correspond with a straight line with a positive slope of 45°. Since the frequency difference of octave intervals is equal to the frequency of the lower tone, we see that for the lower frequencies nearly all intervals of Figure 43a exceed the octave. This implies that, including also the second harmonic, these intervals reduce to octaves, resulting in a line with a slope of 45° through the point $\Delta f=100$ cps at $f=100$ cps (Figure 43b). Above $c=130.8$ cps, however, most intervals are smaller than the octave. Because $n=2$ means that all fundamentals are accompanied by their octaves, the curves of Figure 43b extend to a corresponding higher frequency. The inclusion of the third harmonic manifests itself in the following ways: (1) the points corresponding with the lower frequencies do not shift because the frequencies of the new tones all are above that range; (2) the "density" of tones increases in the middle range, resulting in a shift of the curves to smaller frequency differences; (3) the curves are extended to a 50%-higher frequency, compared with the curves for $n=2$; (4) as most of the intervals for the highest frequencies are fifths, corresponding with the frequency ratio of the second and third harmonics of the highest fundamentals of the composition, this interval determines the course of the curves at the higher frequencies.

Every time when a further harmonic is added, a repetition of this process occurs, with the result that for increasing n (1) the frequency limit below which no new tones are added shifts to higher frequencies; as we saw, this limit is about $c=130.8$ cps for $n=2$, whereas it is about $c^2=523.5$ cps for $n=9$; (2) in the frequency range above this limit, the curves shift to smaller frequency differences; (3) a further extension of the curves to higher frequencies takes place; (4) for the highest frequencies, the course of the curves is mainly determined by the interval $(n-1):n$.

The curves in Figure 44 show the same trends as a function of the number of harmonics. However, in this case, the interval widths between the fundamentals are much smaller than in the former case. Only for $C_1=65.4$ cps do the intervals exceed the octave, as a comparison of the graphs (a) and (b) shows. As a consequence of this fact, the curves in Figure 44 correspond also for $n>1$ with smaller intervals than the curves in Figure 43.

After these more general remarks, we may compare the position of the curves with the critical-bandwidth curves. As we see, the shape of the interval curves agree for increasing n more and more with the dashed curves. In both figures, the agreement is greatest for about 8 harmonics.

These results strongly suggest that critical bandwidth plays an important rôle in music. The significance of this fact can be interpreted in the following way. As we saw above, simple-tone intervals with a frequency difference exceeding the critical band are judged as consonant and do not differentiate in this respect. On the other hand, for smaller frequency differences, consonance evaluation strongly depends upon interval width, with a minimum for about a quarter of critical bandwidth. So it is not surprising that just this range is used for "modulation" between more-consonant and more-dissonant chords. However, it is surprising indeed that, for a number of harmonics representative of musical instruments, this is achieved in about the same measure over a wide frequency range.

We have to realize that this equally deep "penetration" in the borderland between pronounced consonant and dissonant simple-tone intervals, represented by the upper and lower dashed curves in the graphs, respectively, is a result of many factors. As the most important ones we may consider:

1. The fact that in the tone scale as developed in Western music, a lot of intervals agree with simple frequency ratios, so that harmonics of the different component tones of a chord may coincide; otherwise, the shape of the solid curves in Figures 43 and 44 would have been more flat, as a result of more dissonant chords.

2. The fact that the frequencies of the partials of the tones are multiples of the frequency of the fundamental tone. A deviation from this rule would have the same effect as mentioned under (1). This may be regarded as one of the reasons (there are more!) why instruments with inharmonic partials are not used to produce musical chords.

3. The way in which the composer selects his intervals as a function of frequency. We saw above that in Bach's composition the frequency ratio between fundamental tones is larger at lower than at higher frequencies. As a comparison with Figure 43 shows, in this way very dissonant chords are avoided. Although to a smaller degree, this is also the case in Dvořák's string quartet.

4. The number of notes in a chord. It is clear that, generally, the mean frequency difference of adjacent partials decreases for increasing number. The fact that the solid curves in Figure 44 correspond with smaller frequency differences than the curves in Figure 43 may be mainly due to this factor and the third one.

5. The frequency limits between which the fundamentals are chosen and their distribution within this range. So a multiplication of all frequencies by a certain factor shifts all curves both horizontally and vertically to

the same degree. As we see, this would influence their relation to the dashed curves much more for lower than for higher frequencies.

6. The number of harmonics produced by the instruments on which the composition is performed. Only the influence of this factor has been studied here, showing that the frequency range over which a typical harmonic modifies the interval distributions shifts to higher frequencies for increasing n. This implies that the number of harmonics is not very critical. Most musical instruments produce strong harmonics up to a number that may vary from about 6 to 10, although in the latter case the tone has a sharp quality and is more suited for solo parts.

The mere enumeration of these factors does not give us much information about their relative importance. So it would be of interest to know more about the degree to which each factor determines the position of the horizontal and the sloping parts of the curves. Moreover, we should like to have more insight in the way in which their position depends upon both musical style and the instruments for which the composition is written. Further investigations are in preparation to answer these questions.

CONCLUSIONS

1. Over the frequency range considered (125-2000 cps), simple-tone intervals are judged as most dissonant for frequency differences equal to about a quarter of critical bandwidth.

2. Over the same frequency range, simple-tone intervals are judged as consonant for frequency differences exceeding critical bandwidth.

3. These results support the view that tones interfere when their frequency difference is smaller than critical bandwidth, expressing itself in beats for very small frequency differences and roughness or dissonance for larger differences.

4. On the basis of these results, presenting a quantification of Helmholtz's consonance theory, the singular nature of complex-tone intervals with frequency ratios determined by small integral numbers (consonances) can be explained.

5. Interval distributions computed from the chords of some musical compositions strongly suggest that, indeed, critical bandwidth plays an important rôle in music.

REFERENCES

Békésy, G. von: Über akustische Rauhigkeit, *Z. techn. Phys. 16*, 276-282 (1935). [English translation, G. von Békésy: *Experiments in Hearing*, McGraw-Hill Book Company, Inc., New York, pp. 344-354 (1960).]

Boomsliter, P., and W. Creel: The Long Pattern Hypothesis in Harmony and Hearing, *J. Music Theory 5*, No. 2, 3-30 (1961).

Cazden, N.: Musical Consonance and Dissonance: a Cultural Criterion, J. *Aesthet. 4*, 3-11 (1945).

Cross, Ch. R., and H. M. Goodwin: Some Considerations Regarding Helmholtz's Theory of Consonance, *Proc. Am. Acad. Arts and Sciences, New Ser. 19*, 1-12 (1893).

Galilei, G.: *Discorsi e dimostrazioni matematiche interno à due nuoue scienze attenenti alla* *mecanica & i movimenti locali*, Elsevir, Leiden (1638). The quotation is from the English translation: *Dialogues Concerning Two New Sciences*, transl. by H. Crew and A. de Salvio, McGraw-Hill Book Company, Inc., New York, p. 100 (1963).

Geer, J. P. van de, W. J. M. Levelt, and R. Plomp: The Connotation of Musical Consonance, *Acta Psychol. 20*, 308-319 (1962).

Guernsey, M.: The Rôle of Consonance and Dissonance in Music, *Am. J. Psychol. 40*, 173-204 (1928).

Guthrie, E. R., and H. Morrill: The Fusion of Non-Musical Intervals, *Am. J. Psychol. 40*, 624-625 (1928).

Helmholtz, H. L. F. von: *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*, F. Vieweg & Sohn, Braunschweig, 1st Ed. (1863).

Husmann, H.: *Vom Wesen der Konsonanz*, Müller-Thiergarten-Verlag, Heidelberg (1953).

Kaestner, G.: Untersuchungen über den Gefühlseindruck unanalysierter Zweiklänge, *Psychol. Studien 4*, 473-504 (1909).

Krueger, F.: Differenztöne und Konsonanz, *Arch. ges. Psychol. 1*, 205-275 (1903), *2*, 1-80 (1904).

Krueger, F.: Die Theorie der Konsonanz, *Psychol. Studien 1*, 305-387 (1906), *2*, 205-255 (1907), *4*, 201-282 (1909), *5*, 294-411 (1910).

Levelt, W. J. M., J. P. van de Geer, and R. Plomp: Triadic Comparisons of Musical Intervals, *Brit. J. Mathem. Statist. Psychol.*, in press (1966).

Lipps, Th.: *Psychologische Studien*, Verlag G. Weiss, Heidelberg, pp. 92-161 (1885).

Lundin, R. W.: Toward a Cultural Theory of Consonance, *J. Psychol. 23*, 45-49 (1947).

Malmberg, C. F.: The Perception of Consonance and Dissonance, *Psychol. Monogr. 25*, No. 2, 93-133 (1917-1918).

Mayer, A. M.: Researches in Acoustics. No. IX, *Phil. Mag. 5th Ser. 37*, 259-288 (1894).

Montani, A.: Outline of a Physiological Theory of Musical Consonance, *Rivista Musicale Italiana 49*, 168-176 (1947).

Moore, H. T.: The Genetic Aspects of Consonance and Dissonance, *Psychol. Monogr. 17*, No. 2, 1-68 (1914).

Moran, H., and C. C. Pratt: Variability of Judgments on Musical Intervals, *J. Exptl. Psychol. 9*, 492-500 (1926).

Ogden, R. M.: A Contribution to the Theory of Tonal Consonance, *Psychol. Bull. 6*, 297-303 (1909).

Plomp, R., and W. J. M. Levelt: Musical Consonance and Critical Bandwidth, in: *Proceedings of the Fourth International Congress on Acoustics*, Copenhagen, Paper P 55 (1962).

Polak, A. J.: *Über Zeiteinheit in Bezug auf Konsonanz, Harmonie und Tonalität*, Verlag Breitkopf & Härtel, Leipzig (1900).

Preyer, W.: *Akustische Untersuchungen*, Verlag G. Fischer, Jena, pp. 44-61 (1879).

Sandig, H.: Beobachtungen an Zweiklängen in getrenntohriger und beidohriger Darbietung. Ein Beitrag zur Theorie der Konsonanz, *Neue Psychol. Studien 14*, 25-131 (1939).

Sorge, G. A.: *Vorgemach der musicalischen Composition*, Verlag des Autoris, Lobenstein, pp. 333, 334 (1745-1747).

Stumpf, C.: *Tonpsychologie*, Vol. 1, Verlag S. Hirzel, Leipzig, Preface (1883).

Stumpf, C.: *Tonpsychologie*, Vol. 2, Verlag S. Hirzel, Leipzig, pp. 127-218 (1890).

Stumpf, C.: Konsonanz und Dissonanz, *Beitr. Akust. Musikwiss. 1*, 1-108 (1898).

Stumpf, C.: *Die Sprachlaute*, Verlag J. Springer, Berlin, p. 281 (1926).

Wundt, W.: *Grundzüge der physiologischen Psychologie*, Verlag W. Engelmann, Leipzig, 2nd Ed., Vol. 1, pp. 402-408, and Vol. 2, pp. 35-48 (1880).

Zwicker, E., G. Flottorp, and S. S. Stevens: Critical Band Width in Loudness Summation, *J. Acoust. Soc. Am. 29*, 548-557 (1957).

5

BEATS OF MISTUNED CONSONANCES

The experiments in the preceding chapter showed that interference between simultaneous tones occurs essentially when the frequency difference is less than a critical band. Similarly, the experimental results discussed in Chapter 2 indicated that the ear's frequency-analyzing power is determined by the same bandwidth. These results suggest that the ear is provided with a set of band-pass filters with bandwidth equal to the critical band. The existence of combination tones appeared to be compatible with this conception, if we accept nonlinear distortion as an additional mechanism (Chapter 3).

There is, however, another phenomenon which seems to contradict this rôle of the critical band in tone perception. Two simple tones may also give rise to a beat sensation for frequency differences much larger than critical bandwidth, namely when their frequency ratio is slightly different from $m:n$, with m and n both small integral numbers. These so-called *beats of mistuned consonances* are usually considered as resulting from nonlinear distortion too, and many investigators, following Wegel and Lane (1924), have used the beats for the case $m=1$ to determine the ear's nonlinearity.

In this chapter, the traditional explanation of the beats of mistuned consonances will be questioned and experimental evidence shall be adduced indicating that the beats have to be explained in quite a different way.

HISTORICAL REVIEW

The phenomenon

The occurrence of beats when consonant tone intervals, produced by musical instruments, are slightly mistuned must have been well-known for a long time. Perhaps the first attempt to explain this phenomenon was made in 1628 by the Dutch scientist Beeckman, who discovered the relation between the length of a string and its frequency of vibration several years before the law was published by Mersenne and Galilei (de Waard, 1945). Beeckman based his explanation on the deviation from the periodic concurrence of vibrations for perfect consonances. In his opinion, when the higher tone of a fifth (2:3) has one vibration per unit of time more than for perfect tuning, one beat per unit of time is heard. As we shall see below, this reasoning is incorrect; not one but two beats per unit of time are heard.

Only a few years later, Mersenne called attention to the fact that a vibrating string produces a series of harmonically related tones (see Chapter 2). The existence of harmonics offered a simple explanation of beats of mistuned consonances. For example, consider the mistuned fifth 200+301 cps, each tone consisting of a series of harmonics; the third harmonic of the lower tone (= 600 cps) and the second harmonic of the upper tone (=602 cps) produce interference beats at the rate of two per second. In this way, the beats of mistuned consonances were reduced to beats of mistuned unisons.

For *complex* tones, it was not difficult to explain the beats. But how must they be interpreted for intervals composed of *simple* tones? This case was first studied by Scheibler, a merchant in Crefeld (Germany), who was interested in more-accurate methods for tuning musical instruments. A description of his experiments and the way in which they were explained was given by Röber (1834). Scheibler used tuning forks and examined the relation between the frequencies of the primary tones and the number of beats observed. His most important conclusions were: (1) two tones of p and $np+a$ cps, respectively, give rise to a beats per second, with decreasing strength for increasing n; (2) two tones of $2p+a$ and $3p+b$ cps (mistuned fifth) result in $3a-2b$ beats per second; (3) the frequencies $3p+a$ and $4p+b$ cps (mistuned fourth) give $4a-3b$ beats per second. With tuning forks, Scheibler was not able to determine the rate of beats for other mistuned consonances, as the thirds 4:5 and 5:6, because in those cases the beats were too weak to be counted.

Ohm (1839) published some years later a general formula for the number of beats: two tones of M and N cps, respectively, with $M:N$ only slightly different from $m:n$, give rise to

$$A = mN - nM$$

beats per second. With $M = mp$ and $N = np + a$, this equation can also be written as

$$A = ma.$$

As is seen easily, all cases studied by Scheibler agree with this formula.

Later investigators showed that, for loud tones, the beats are audible over a larger range of m- and n-values than had been found by Scheibler. So, for low frequencies, König (1876) was able to hear beats produced by p and $np+a$ cps up to $n=9$, whereas he also detected beats for the mistuned consonances 2:5, 3:5, 4:5, 5:6, 3:7, 4:7, and other ratios. Cotton (1935), with a steady higher tone of 1000 cps and a variable lower one between 500 and 1000 cps, found no less than 19 frequency ratios around which beats could be observed.

Explanations proposed

Since these beats were discovered for simple tones, three different origins of their existence has been proposed.

1. *Combination tones*. This point of view was held by Scheibler and expounded by Röber

in the article mentioned above. In the same way as Hällström had done some years before (see Chapter 3), Scheibler admitted that first-order combination tones may lead to second-order ones, etc. For example, his explanation of why tones of p and $4p+a$ cps, respectively, produce a beats per second is the following: the tones p and $4p+a$ give rise to the combination tone $(4p+a)-p=3p+a$; in the same way, the second-order combination tone $(3p+a)-p=2p+a$ and the third-order one $(2p+a)-a=p+a$ are created; the last combination tone interferes with p, resulting in $(p+a)-p=a$ beats per second. For the mistuned fifth $(2p+a):(3p+b)$, the steps are: $(3p+b)-(2p+a)\rightarrow p+b-a$; $(2p+a)-(p+b-a)\rightarrow p+2a-b$; $(p+2a-b)-(p+b-a)\rightarrow 3a-2b$ beats per second. This way of explaining beats was also adopted by Helmholtz (1856, 1863), although he was only able to hear first-order combination tones directly. In fact, he considered the existence of beats of mistuned consonances as a proof that also higher-order combination tones are produced.

The explanation by means of various orders of combination tones looks rather artificial; it presumes in some cases the presence of a number of combination tones that are never heard individually. Therefore, other investigators modified the theory, assuming that the primary tones produce different combination tones directly. So Bosanquet (1881a) explained the beats of the interval $p:(3p+a)$ as caused by interference between the third-power combination tone $(3p+a)-2.p=p+a$ and the lower primary tone p. Stumpf (1910) thought that beats for the mistuned fifth $2p:(3p+a)$ resulted from interference between the combination tones $(3p+a)-2p=p+a$ and $2.2p-(3p+a)=p-a$. For the last case, the same interpretation was given more recently by Haar (1951).

2. *Aural harmonics.* This explanation was promoted by Müller (1871) and Hermann (1896). Helmholtz, although preferring the explanation by combination tones, left open the possibility that, for strong primary tones, aural harmonics produced by the ear's distortion also may contribute to the appearance of beats. This approach also accounts very easily for the number of beats given by Ohm's formula: the m-th harmonic of the higher tone of N cps interferes with the n-th harmonic of the lower tone of M cps, resulting in $mN-nM$ beats per second.

The explanation of beats by means of aural harmonics was adopted by Wegel and Lane (1924) in their pioneer study on masking. They interpreted the beats of mistuned intervals with frequency ratio $1:n$ as a demonstration of the ear's nonlinearity and introduced the "method of best beats" for investigating it. In this method, the intensity of an aural harmonic is approximately equal to the intensity of an auxiliary tone of a slightly different frequency that gives the most pronounced beats. The method was accepted and used to study the intensity of aural harmonics by Fletcher (1930), von Békésy (1934), Moe (1942), Egan and Klumpp (1951), Opheim and Flottorp (1955), and Lawrence and Yantis (1956a, 1956b).

3. *Variations in waveform.* Both previous explanations are in harmony with the conception of the ear as a frequency analyzer; they have in common that the additional mechanism of nonlinear distortion is regarded as the origin of beats of mistuned consonances. In contrast with this, the third interpretation appears to attack the frequency-analyzer conception by stating that beats are related to variations in the waveform of the superimposed sinusoids. This implies that simple tones interfere for much larger frequency differences than the critical bandwidth. Because two simple tones with frequency ratio $m:n$, slightly mistuned, can be understood as a perfect consonance with one tone shifting continuously in phase, the ear's ability to perceive waveform variations may be interpreted also as a demonstration that the ear is sensitive to phase. This ability agrees with the conception of the ear as a periodicity detector.

As far as the writer knows, De Morgan (1864) proved for the first time that the number of waveform variations of the superimposed sinusoids agrees with Ohm's formula[1]. Essentially, his reasoning runs as follows. Suppose that the constituent tones of a consonant interval are given by the sound pressures $x=\cos 2\pi mpt$ and $y=\cos 2\pi npt$, respectively ($n>m$). It is clear that $x=1$ at all multiples of $1/mp$ sec and $y=1$ at all multiples of $1/np$ sec, so we conclude that all these maxima of x and y occur at particular multiples of $1/mnp$ sec. There are other multiples of $1/mnp$ at which $x=1$, whereas $y=1$ at the next-higher multiples. This implies that a time shift of $-1/mnp$ sec of y results in a coincidence of that maxima of x and y, and in that case the combined waveform is identical to the original one for which these maxima coincide at $t=0$. The only difference is a shift in time. Now we take the case that the frequency of y is not np but $np+a$ cps ($0<a\ll np$). This means that a period of y does not last $1/np$ but $1/(np+a)$ sec, so it is shortened by

$$1/np-1/(np+a)=a/np(np+a)$$

sec. After a certain number of periods of y, the sum of successive time lags is equal to $1/mnp$ sec and in that case the waveform is identical to the original one. This occurs after a number of periods of y equal to

$$\frac{1}{mnp}\Big/\frac{a}{np(np+a)}=\frac{np+a}{ma}$$

periods. Since one vibration takes $1/(np+a)$ sec, we may also say that the original waveform is repeated after $1/ma$ sec, so we have ma times in a second that the same waveform occurs. In this way, De Morgan showed that the number of periodic changes in waveform agrees with the number of beats after Ohm's formula. Although this derivation can be formulated more carefully and, moreover, holds only strictly if $1/mnp$ is a multiple of $a/np(np+a)$, the reasoning is correct in principle, as can be demonstrated by count-

[1] De Morgan refers in his paper at large to a book of R. Smith called *Harmonics* and published in 1749 (2nd Ed. 1759). Smith may be regarded as the originator of the view that beats are related to variations in waveform. See also Bosanquet (1881b).

ing the number of waveform variations with the aid of a cathode-ray oscilloscope.

The opinion that beats are related to waveform was also promoted by König (1876) in an extensive study on beats and secondary tones. He ascertained by means of Lissajous' figures that, for simple tones of p and $np+a$ cps, respectively, maxima and minima of the amplitude occur at the rate of a per second. The combination tones, considered by Helmholtz as products of distortion, were interpreted by König as "beat tones", into which the beats change for increasing rate. He did not explain the beats produced by two tones of $2p$ and $(2n+1)p+a$ cps, respectively, directly by waveform variations but by interference between the beat tones $\{(2n+1)p+a\}-2np=p+a$ and $(2n+2)p-\{(2n+1)p+a\}=p-a$, resulting in $(p+a)-(p-a)=2a$ beats per second. He could observe these beats up to $n=4$ for $2p=128$ cps; the maximum n-value for which the beats were audible decreased for higher frequencies. He tried to explain the beats of mistuned consonances with more-complex ratios as 3:4, 3:5, 3:7, 3:8, 4:5, and 4:7 in a similar way.

Although the hypothesis that beats are caused by nonlinear distortion was accepted by most investigators during the last century, there were others who continued, for reasons we discuss below, and mostly with more or less hesitation, to relate beats to waveform (Thomson, 1878; ter Kuile, 1904; Beasly, 1931; Cotton, 1935; Lottermoser, 1937; Meyer, 1949, 1953, 1954, 1957; von Békésy, 1957; Chocholle and Legouix, 1957b). A recent study of monaural phase perception by Goldstein (1965) takes a more definite stand in support of the waveform hypothesis.

Experimental evidence concerning these theories

We now may ask to what extent each of the explanations mentioned is supported by experimental evidence.

It is obvious that the first question to be asked must be: how do these beats sound? In other words: is it possible to discriminate between the three hypotheses on the basis of the sensation of beats?

The diverging descriptions given to this sensation indicate that it is not easy to express by words what one hears in this case. For intervals as 2:3, 3:4, and 4:5, slightly mistuned, some investigators (ter Kuile, 1904; Beasly, 1931; Cotton, 1935) localized the beats on the primary tones as well as on the difference tones, but more frequently only the primary tones are mentioned. The beats were described in various ways: as an alternate perception of the tones (De Morgan, 1864; König, 1876), as periodic variations in timbre (De Morgan, 1864; Thomson, 1878), in loudness (De Morgan, 1864; Beasly, 1931), or in pitch (Beasly, 1931). Although Helmholtz (1856) was very reserved in answering the question which tones beat, in his opinion, the combination tones predicted by Scheibler can be heard as beating. Peterson (1908) stated that "one invariably locates the beating in the *primary* tones. Even when the first difference tone of the imperfect fifth is plainly audible I can never locate the beating solely in this tone. The primaries themselves always beat for me." After Haar (1951), for mistuned fifths up to 500 cps, one has the impression that the interval itself beats, but above this frequency one hears beats of the combination tones.

Of the mistuned intervals with frequency ratio $1{:}n$, in particular the imperfect octave was studied. The beats of these intervals were described as periodic variations in timbre (Helmholtz, 1856; Trimmer and Firestone, 1937) and in loudness of the lower tone (König, 1881; Trimmer and Firestone, 1937). Bosanquet (1879, 1881a), Stumpf (1896), and Schouten (1938) made a distinction between two types of beats, manifesting themselves in variations in loudness of the lower and of the higher tone, respectively, both with the same rate. In their opinion, the higher beats are caused by harmonics of the sound source (after Schouten also by the ear's nonlinearity) and can be eliminated by acoustical measures, whereas the lower beats are always audible.

This enumeration does not point to one of the explanations on the origin of beats exclusively. Although it is difficult indeed to draw any conclusion from the circumscriptions of the beating phenomenon for mistuned intervals of the form $1{:}n$, a somewhat more positive result can be obtained for the intervals given by 2:3, 3:4, etc. The combination tones for these intervals are below the lower primary tone and the harmonics are above the higher tone (for example: 200+301 cps may give beats between the combination tones of 99 and 101 cps, or between the harmonics of 600 and 602 cps). It is of interest that, although in some cases the beats were localized also on the combination tones, in most cases the primary tones themselves were considered as beating and, moreover, that in no case the beats were localized on harmonics of the tones. This is the only result of the introspective method of examining beats that discriminates between the hypotheses concerning their origin. It may be considered as a support of the proposition that neither combination tones nor aural harmonics are always essential in producing beats.

Beside the question of how beats sound, we may consider other arguments put forward in favour of the different opinions concerning their origin. A survey shows that the number of experimentally supported arguments is very small. Most investigators were completely satisfied by the fact that the beat rate agreed with their calculations. Their preference to one of the explanations was in nearly all cases only a result of theoretical considerations on the process of hearing. So, for the explanation of beats by combination tones no other motives were put forward than the existence of these tones. Similarly, the view that the beats must be due to aural harmonics was argued by pointing to the nonlinearity of the ear. The assumption of several investigators that beats of mistuned consonances are related to periodic waveform variations has to be regarded as a consequence of their denial of the other two ways of explanation rather than as a result of positive evidence for it.

For this reason, it is of interest to mention the arguments brought forward against the hypotheses that beats are due to combination tones or to aural harmonics. Objections against the first hypothesis were: (1) it is very improbable that the beats of, for instance, $p:(8p+a)$ must be explained by a series of combination tones with frequencies $7p+a$, $6p+a$, etc., since already $7p+a$ is inaudible (König, 1876); (2) for the mistuned octave, the most pronounced beats are obtained when the lower tone is much louder than the higher one, which does not seem to be the optimal situation for a loud combination tone (Thomson, 1878; König, 1876; Stumpf, 1910); (3) although both investigators accepted the explanation of beats by combination tones, Bosanquet (1879) considered it as a difficulty that he did not hear the combination tone itself when, for a mistuned octave, the frequency of the lower tone was raised, whereas Stumpf (1910) pointed to the fact that also for increasing higher tone the combination tone does not become audible.

The view that beats of mistuned consonances are due to aural harmonics was criticized by several investigators on the ground that these harmonics themselves are never heard (König, 1881; Lottermoser, 1937; Meyer, 1949; Chocholle and Legouix, 1957a; contradicted by Lawrence and Yantis, 1957). However, this objection could not prevent the present general acceptance of the method of best beats as a means to determine the ear's nonlinearity, as an examination of books on hearing shows.

In conclusion, after considering the psychophysical evidence for and against the three hypotheses about the origin of beats of mistuned consonances, we may ascertain that more research is required to solve the problem.

Beats of chords consisting of three tones

Before passing on to some experiments about mistuned consonances, it may be of interest to draw the attention to the fact that beats are produced also by chords consisting of three simultaneous tones of $(n-1)p$, np, and $(n+1)p$ cps, respectively, with one tone slightly mistuned. These beats too were already studied extensively by Scheibler (Röber, 1834). He discovered that a mistuning of a cps of the lower or the higher tone resulted in a beats per second but that for the same mistuning of the middle tone not a but $2a$ beats per second are produced. His explanation was, just as for tone pairs, based on combination tones: in the case that the lower tone is $(n-1)p+a$ cps, the beats are produced by the combination tones $np-\{(n-1)p+a\}=p-a$ and $(n+1)p-np=$ $=p$, whereas for a middle tone of $np+a$ cps, the beats are due to the combination tones $(np+a)-(n-1)p=p+a$ and $(n+1)p-$ $(np+a)=p-a$. In this way, he could explain the number of beats.

Helmholtz (1863) interpreted also the beats in this way, but this view later met strong opposition. Thomson (1878) called attention to the fact that these beats are much more distinct than in the case of two tones and that already they occur at very low sound levels. Using tuning forks, he ascertained: "The sound dies beating, the beats being distinctly heard all through a large room as long as the faintest breath of the sound is perceptible, ... being, in fact, sometimes the very last sound heard." This makes it rather improbable that these beats are caused by combination tones. On the other hand, it is also very unlikely that the beats are due to aural harmonics.

In 1902, ter Kuile published a detailed study on the beats of chords consisting of three tones. He ascertained that the whole chord is beating periodically and that the beats have to be considered as periodic changes in timbre. Since the mistuned tone may be regarded as a tone with a correct frequency but shifting phase, he plotted the superimposed waveform of the three tones for various values of phase and showed that the beat rate agrees with the number of waveform variations per second. The correctness of this reasoning can be demonstrated more conveniently with the aid of an oscilloscope. On the ground of this agreement and the fact that the beats are audible at very low sensation levels, ter Kuile concluded that the beats are not due to distortion of the ear but are indeed related to waveform variations. This conclusion is supported by more-recent studies about this phenomenon (Mathes and Miller, 1947; Zwicker, 1952; Goldstein, 1965). This interpretation makes it more probable that also in the case of two tones, the beats are related to waveform variations.

Audibility of aural harmonics

It was mentioned above that several investigators criticized the explanation of beats of mistuned consonances by aural harmonics on the ground that these harmonics themselves are never heard. The following experiment was carried out to test this statement.

Test subjects were presented with simple tones and the sensation level was determined at which aural harmonics became audible. The experimental procedure shall be explained on the basis of a block diagram of the apparatus, represented in Figure 45. Firstly, by operating switch S_2 (switch S_1 in the middle position), the experimenter interrupted a tone of f cps produced by sine-wave generator 2 (Hewlett-Packard 200 CD) and determined, with the aid of attenuator 2, the hearing threshold of the subject for that tone. Then, he adjusted the sensation level of f to 40 dB and asked the subject to operate freely S_1 (S_2 in the on-position), so that he

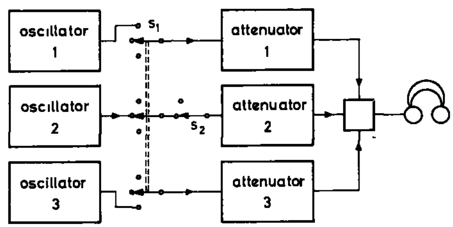


FIGURE 45. Block diagram of the apparatus used for investigating the audibility of aural harmonics.

could listen successively to f and to two auxiliary tones of nf and $(n+\frac{1}{2})f$ cps, respectively, produced by sine-wave generators 1 and 3 (Hewlett-Packard 200 AB). The subject was free to adjust attenuators 1 and 3 to convenient values and had to decide whether an aural harmonic of nf cps was audible in the middle position of S_2. If this was indeed the case, the experimenter determined by means of attenuator 2 the minimum sensation level of f at which the harmonic could be heard by the subject. In the opposite case, the procedure was repeated at a sensation level of 50 dB, and so on. The use of two auxiliary tones proved to be very helpful for the subject to make a decision whether a particular aural harmonic was audible or not. In this way, the audibility was examined for $n=2, 3, 4, \ldots, 8$.

The tones were presented monaurally by means of a high-quality headphone (Beyer DT 48) to the subject who was seated in a soundproof room; the experimenter was seated in an adjacent room. Ten subjects participated in the experiment. Their ability to hear aural harmonics was investigated for $f=125$, 250, 500, 1000, and 2000 cps. For these frequencies, the maximum sensation level applied was about 65, 90, 100, 100, and 100 dB, respectively.

In Table 4, the percentage of subjects who heard aural harmonics is given as a function of f and n, independent of sensation level. The mean values indicate that fewer harmonics were heard for lower test frequencies than for higher ones. The audibility of aural harmonics varied largely over the subjects (two of them did not hear any aural harmonic). The sensation level at which the harmonics became audible was, averaged over all n-values and subjects, 45.5, 82.4, 84.5, 91.5, and 87.8 dB for $f=125$, 250, 500, 1000, and 2000 cps, respectively.

These data show that the statement that aural harmonics are inaudible, is not true. There appear to be, however, large differences in the individual ability to hear these harmonics. Moreover, the sensation levels at which they became audible were rather high. These findings do not eliminate aural harmonics as a possible source of the beats of mistuned consonances with frequency ratio $1:n$. Nevertheless, they do not support this explanation, because these beats were observed very easily by subjects who did not hear any aural harmonic.

Masking of combination tones and aural harmonics

It is possible to ascertain whether the beats of mistuned consonances with frequency ratio $m:n$, with $m\neq 1$, are due to distortion or not. This may be illustrated by an example. Two tones with frequencies 200 and 301 cps, respectively, give rise to two beats per second. If these beats are caused by combination tones, they imply that the quadratic-distortion product $301-200=101$ cps beats with the cubic-distortion product $2.200-301=99$ cps. If, on the other hand, the beats are caused by aural harmonics, they imply that the quadratic-distortion product $2.301= =602$ cps beats with the cubic-distortion

f (cps)	number of aural harmonic n 2	3	4	5	6	7	8	mean value
125	0	20	0	50	10	0	0	11
250	10	10	10	20	0	10	20	11
500	0	10	20	30	20	10	20	16
1000	10	20	30	30	40	30	0	23
2000	30	30	20	20	30	40	—	28
mean value	10	16	16	30	20	18	10	17

TABLE 4. Percentage of subjects that heard aural harmonics.

product 3.200=600 cps. In the first case, the frequencies of the beating tones are lower than the primary tones, whereas in the second case they are higher.

This suggests that, by introducing noise in the frequency ranges below 200 cps and beyond 301 cps, it must be possible to mask the secondary tones. This would result in a disappearance of the beats if they are due to distortion. Experiments showed that the beats do not disappear by introduction of noise in the frequency bands mentioned.

This test was repeated for many other tone intervals. The results indicated that at low frequencies (below about 500 cps), where the beats are much more distinct than at high frequencies, as the experiments of the next section will show, the beats of mistuned consonances are neither caused by combination tones nor by aural harmonics. At high frequencies (above about 1000 cps), however, in the few cases that beats were audible, these beats appeared to be due to combination tones. In the intermediate frequency range, sometimes two types of beats were distinguished; one type was caused by combination tones and could be masked by noise, whereas the other type could not. In no case, aural harmonics seemed to play a rôle in the production of beats.

These experiments do not support the view that the beats of mistuned consonances with frequency ratio $m:n$ ($m \neq 1$) can be explained solely by distortion. In the low-frequency range, where the beats are most distinct, they have apparently another origin. Nearly all observations summarized in the historical review have been made in this frequency range.

Unfortunately, this masking procedure cannot be applied directly for studying the beats of tone intervals $1:n$, slightly mistuned. In that case, the secondary tones beat with the primary tones themselves, so the former tones cannot be masked without masking the latter tones too.

Sound-pressure levels for best beats

Since the previous method was not applicable to the case of slightly mistuned tone intervals of the type $1:n$, other methods were sought to study these intervals more explicitly. These experiments are treated in this and the next section.

The first method concerned the sound-pressure level at which the beats are most distinct. As was mentioned above, this "method of best beats" has been used by several investigators to determine the intensity of the hypothetical aural harmonics. In their experiments, only beats for tone intervals with frequency ratio $1:n$ were taken into account. For this reason, it seemed desirable to repeat these experiments including also the case $m:n$ with $m \neq 1$.

Figure 46 represents a block diagram of the apparatus. The (constant) lower and (variable) higher tones were produced by sine-wave oscillators 1 and 2 (Hewlett-Packard 200 CD), respectively. The frequency ratio of these tones was controlled by an electronic counter, whereas the Lissajous' figure was made visible with a cathode-ray oscilloscope. The signals passed through attenuators 1 and 2, respectively, and were presented to the headphone (Beyer DT 48) via the matching box. The amplitude of the variable tone (oscillator 2) was controlled during the experiments by an electronic voltmeter; switch S_1 was included to switch off the constant tone (oscillator 1) every time that this amplitude was measured. An auxiliary

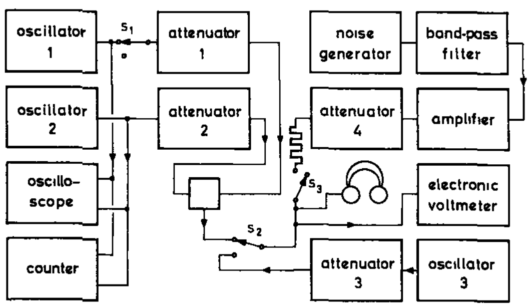


FIGURE 46. Block diagram of the apparatus used for investigating the sound-pressure level for best beats.

tone, produced by sine-wave oscillator 3, could be connected to the headphone by means of switch S_2. Possible combination tones could be masked by a band of noise produced by a white-noise generator and a 1/3-octave band-pass filter (Brüel & Kjaer Audio Frequency Spectrometer 2109). This noise was introduced by switch S_3; the resistor was included in order to prevent any influence of the position of S_3 on the amplitude of the tones.

The following procedure was used. At first, both the frequency and the sound-pressure level of the lower tone were adjusted to the values wanted. Then the experimenter slowly increased the frequency of the variable tone starting from the frequency of the constant tone, and at every frequency for which slow variations of the Lissajous' figure were observed, he asked the test subject whether slow beats, coinciding in phase with the movements of this figure, were audible (the figure was also visible to the subject). If beats were audible indeed, the subject was asked to vary the amplitude of the higher tone by means of attenuator 2 and to determine the attenuator's position for which the beats were most distinct. The subject was always allowed to adjust the beat rate to the value he preferred. After that, the experimenter noted the frequency ratio of the consonant tone interval around which the beats were heard and the amplitude of the signal for best beats. The latter value was converted afterwards into sound-pressure level by means of the real-ear frequency-response characteristic of the telephone. In this way, the whole frequency range above the constant lower tone was searched for audible beats.

Care was taken during the experiments that the beats observed were not caused by combination tones. For this reason, the subject could listen successively, by operating S_2, to the test tones and an auxiliary tone. If a combination tone was audible, he made the frequency and the amplitude of the auxiliary tone equal to the corresponding parameters of the combination tone. Then, with the aid of S_3, noise in the same frequency band was introduced at a sound-pressure level that the auxiliary tone was just masked. This implies that the combination tone was masked, too, by that noise. If the beats did not disappear by introducing the noise, these beats were considered as not resulting from combination tones. Mostly, it was rather easy to determine whether the beats were caused by combination tones, for in that case they were localized on that tones. In the experimental results, only beats are included that were not due to combination tones. As was discussed in the preceding section, this procedure for identifying the beats was only applicable for frequency ratios $m:n$ with $m \neq 1$.

The experiments were carried out for constant lower tones of 125, 250, 500, 1000, and 2000 cps, respectively. They were started at a sound-pressure level of 100 dB for each of these frequencies and repeated at lower levels, in steps of 10 dB, as long as beats were audible. Four subjects participated in the experiments, the two subjects who did not hear any aural harmonic in the previous experiment included; they listened to the tones monaurally. The subjects were tested twice; in the first session, the procedure communicated above was followed entirely, whereas in the second session the experimenter tuned the variable frequency in random order to the frequencies for which in the first session beats had been noticed. The results of both sessions were averaged and considered as the definitive data.

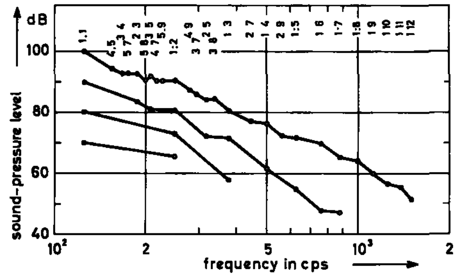


FIGURE 47. Sound-pressure level, averaged over four subjects, for best beats of mistuned consonances as a function of the frequency of the higher tone. The sound-pressure level of the constant lower tone of 125 cps is given for each curve by the most-left point. The frequency ratios of the tone intervals around which beats were heard are indicated.

The experimental results, averaged over the subjects, are reproduced in Figures 47-51. Only those intervals are included for which all subjects heard beats; in this respect, however, the differences between the subjects were not large.

Since in these graphs all beats for $m \neq 1$ caused by combination tones were excluded (aural harmonics appeared to be of no

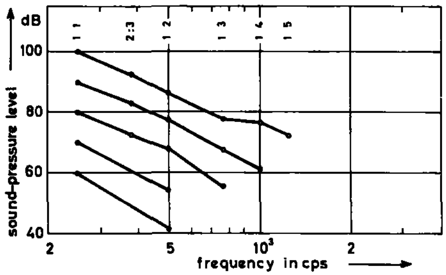


FIGURE 48. As Figure 47, but with a lower tone of 250 cps.

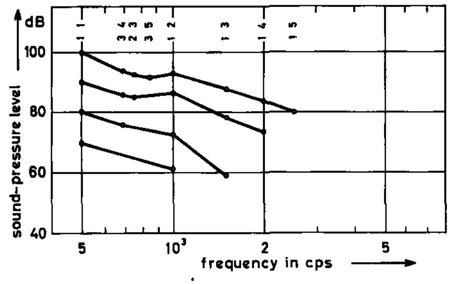


FIGURE 49. As Figure 47, but with a lower tone of 500 cps.

significance, as we saw in the preceding section), it is of interest to know in which cases they were noticed. For a lower tone of 125 cps, no beats of combination tones were ever heard, whereas for 2000 cps, however, all beats around frequency ratios $m:n$ ($m \neq 1$) appeared to be due to combination tones; in the intermediate range, both types of beats were observed.

The graphs show that the data points for best beats can be fitted very well by smooth curves. This was also the case for the data for the subjects individually, although the slopes of their curves differed somewhat. It implies that the sound-pressure level for best beats depends upon frequency difference rather than on frequency ratio of the tones. Since the beats around $m:n$ with $m \neq 1$ could not be reduced to beats of combination tones or aural harmonics, this result strengthens the evidence that also the beats for $m=1$ are not caused by nonlinear distortion. For example, it seems to be rather improbable that, in Figure 47, the beats for 1:2, slightly mistuned, are caused by distortion, whereas the beats for 5:9 and 4:9, the adjacent data points, are not; if the beats result from different mechanisms, we should expect that best heats occur at much more-different sound-pressure levels than has been found experimentally.

The diagrams also indicate that the number of consonances for which beats were audible diminished considerably for increasing frequency of the lower tone. This fact, too, does not support the view that the beats for the case $m=1$ result from distortion because it is not clear why we are able to hear beats up to 1:13 for a lower tone of 125 cps and only up to 1:2 for 2000 cps.

The sweep-tone effect

In the experiments treated thus far, the audibility of beats of mistuned consonances was investigated, passing over the question how these beats sound. In view of the fact that in the literature various answers to this question were found, it was decided at the outset to study the origin of beats

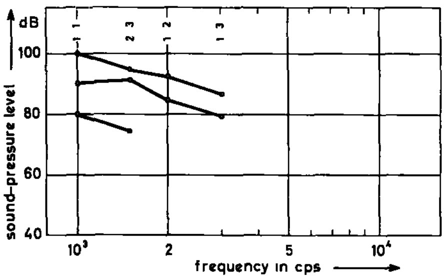


FIGURE 50. As Figure 47, but with a lower tone of 1000 cps.

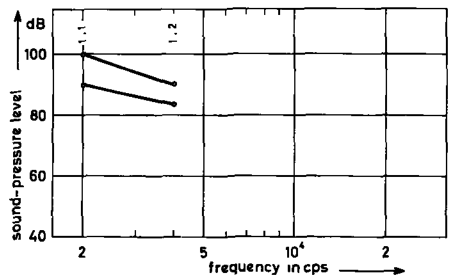


FIGURE 51. As Figure 47, but with a lower tone of 2000 cps.

along other lines. During the experiments, however, the experimenter got the impression that in many cases periodic variations in pitch can be heard in the beats. These pitch variations seemed to be most distinct for slightly mistuned intervals given by 1:3, 1:5, etc. It appeared of interest to pay some attention to this phenomenon.

Although it is difficult to describe the beat sensation, it can be illustrated by a particular example. It was observed that the beats of 200+599 cps sound somewhat different from the beats of 200+601 cps. The time function of the pitch variation can be compared in both cases with a saw tooth, but in the first case the slow phase corresponds with an increase and in the latter case with a decrease in pitch. For the tone intervals mentioned, this difference was so distinct that no errors were made in the identification of the intervals. This capability was also investigated for other subjects. Out of a group of ten subjects 50% heard this difference. We call this phenomenon the *sweep-tone effect*.

In consequence of this result, the question was asked whether it is possible to investigate more quantitatively how the pitch of the sweep tone varies in time. The tone interval 200+600 cps, with the higher tone slightly mistuned, can be interpreted as

$$a \sin 2\pi.200t + b \sin (2\pi.600t + \varphi)$$

with continuously shifting φ. So the problem was stated in this way: is it possible to determine the relation between φ and the pitch of the sweep tone by testing different constant values of φ successively?

For investigating this, an apparatus was designed that produced a stimulus corresponding to the equation, with φ variable in steps of 18°. Figure 52 represents a block diagram of the equipment. Sine-wave oscillators 1 and 2 produced 12,000 and 12,020 cps, respectively. Function generator 1, already described in Chapter 2, was adjusted to produce a sinusoidal waveform; since this generator consisted of 60 shift-register elements, the frequency of this sinusoid was 200 cps. Function generator 2 was of the same design, but it consisted of 20 elements;

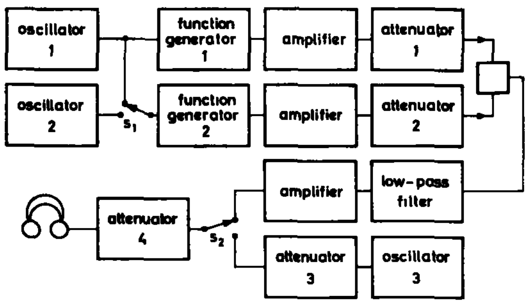


FIGURE 52. Block diagram of the apparatus used for measuring the pitch of the sweep tone.

with switch S_1 in the lower position, it produced 601 cps, and with S_1 in the upper position 600 cps. Provisions were made that φ could be varied in steps of 18° by applying a reset pulse to any one of the 20 elements of generator 2. Amplifiers and attenuators were inserted in both channels, so the output signal of the matching box was equal to the equation, with a constant φ for S_1 in the upper position and a continuously shifting φ for S_1 in the lower position. Higher-frequency components were eliminated by means of a low-pass filter with a cut-off frequency of 870 cps. The subject could listen successively to this stimulus and a simple tone produced by oscillator 3. The stimuli were presented monaurally by headphone (Beyer DT 48).

The experiment was carried out at a sound-pressure level of the 200-cps tone of 100 dB. With S_1 in the lower position, attenuator 2 was adjusted to the value for which the beats were most distinct. Then, S_1 was switched to the other position and the subject was asked to listen successively, by operating S_2, to the tone interval 200+600 cps and the auxiliary tone, and to make the pitch of the latter tone equal to the pitch of the "frozen" sweep tone. This determination was made for the 20 different values of φ in random order. Since the task appeared to be very difficult, the experiment was repeated on 10 different days, each time with a different order of φ-values.

In Figure 53, the result is reproduced as a function of φ. The value of φ relates to the waveform of the stimulus as measured with the aid of an artificial ear.

Although there is a large spread of the data points, the general trend is clear. Sometimes, two pitches were distinguished simultaneously, one of about 610 cps and the other of about 730 cps. The solid curve, drawn

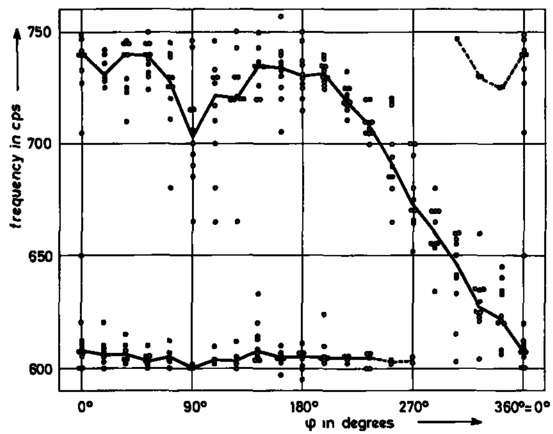


FIGURE 53. Frequency of the comparison tone with a pitch equal to the pitch of the sweep tone, plotted as a function of φ.

through the mean values of the data points, indicates that, as a function of φ, the sweep tone indeed varied in a way that can be compared with a saw tooth: for increasing φ, there is a region over which the pitch lowers gradually; this situation corresponds with the sweep tone heard for 200+601 cps.

After the establishment of the sweep-tone effect, the question was asked whether this phenomenon had been observed earlier by other investigators. It appeared that Stumpf (1910) has noticed "inharmonic" combination tones which may be identified with the sweep tone. Recently, Nordmark brought my attention to an article of him in which he had described the effect very briefly (Nordmark, 1960).

It seems rather difficult to explain the sweep-tone effect in terms of nonlinear distortion by the hearing mechanism. The fact that the effect manifests itself most clearly for mistuned consonances with frequency ratio $1:n$ supports the view that these beats are not caused by aural harmonics.

CONCLUSIONS

1. The statement that aural harmonics are never audible is not correct, although the ability of a subject to hear aural harmonics does not seem to be related to his ability to hear beats of mistuned consonances.

2. The beats of mistuned intervals with frequency ratios $m:n$, with $m \neq 1$, cannot be explained solely by nonlinear distortion; although combination tones appear to play some rôle, larger for high than for low frequencies, the beats at low frequencies, which are the most distinct ones, must have another origin.

3. The experiments on the sound-pressure level for best beats and the sweep-tone effect both strongly suggest that this conclusion also holds for the case $m=1$; this means that the method of best beats cannot be used for investigating the ear's distortion.

4. The results support the view that the beats of mistuned consonances are related to periodic variations of the waveform of the superimposed sinusoids, indicating that simple tones also interfere for frequency distances much larger than the critical bandwidth; the implications of this conclusion are discussed in Chapter 8.

REFERENCES

Beasly, W.: Differential Responses to Cyclic Phase Variations in Compound Sound, *J. Gen. Psychol. 5*, 329-349 (1931).

Békésy, G. von: Ueber die nichtlinearen Verzerrungen des Ohres, *Ann. Phys. 20*, 809-827 (1934). [English translation, G. von Békésy: *Experiments in Hearing*, McGraw-Hill Book Company, Inc., New York, pp. 332-344 (1960).]

Békésy, G. von: Sensations on the Skin Similar to Directional Hearing, Beats, and Harmonics of the Ear, *J. Acoust. Soc. Am. 29*, 489-501 (1957). [= *Experiments in Hearing*, pp. 568-590.]

Bosanquet, R. H. M.: On the Present State of Experimental Acoustics, with Suggestions for the Arrangement of an Acoustic Laboratory and a Sketch of Research, *Phil. Mag. 5th Ser. 8*, 290-305 (1879).

Bosanquet, R. H. M.: On the Beats of Consonances of the Form h:1, *Phil. Mag. 5th Ser. 11*, 420-436 and 492-506 (1881a).

Bosanquet, R.H.M.: On the History of the Theory of the Beats of Mistuned Consonances, *Phil. Mag. 5th Ser. 12*, 270-282 (1881b).

Chocholle, R., and J.P.Legouix: On the Inadequacy of the Method of Beats as a Measure of Aural Harmonics, *J. Acoust. Soc. Am. 29*, 749-750 (1957a).

Chocholle, R., and J.P.Legouix: About the Sensation of Beats between Two Tones whose Frequencies are Nearly in a Simple Ratio, *J. Acoust. Soc. Am. 29*, 750 (1957b).

Cotton, J.C.: Beats and Combination Tones at Intervals between the Unison and the Octave, *J. Acoust. Soc. Am. 7*, 44-50 (1935).

De Morgan, A.: On the Beats of Imperfect Consonances, *Trans. Cambr. Phil. Soc. 10*, 129-141 (1864).

Egan, J.P., and R.G.Klumpp: The Error Due to Masking in the Measurement of Aural Harmonics by the Method of Best Beats, *J. Acoust. Soc. Am. 23*, 275-286 (1951).

Fletcher, H.: A Space-Time Pattern Theory of Hearing, *J. Acoust. Soc. Am. 1*, 311-343 (1930).

Goldstein, J.L.: *An Investigation of Monaural Phase Perception*, Doctoral Dissertation, University of Rochester, Rochester, New York (1965).

Haar, G.: Über den Charakter der Nichtlinearität des Ohres, *Funk und Ton 5*, 248-257 (1951).

Helmholtz, H.L.F.von: Ueber Combinationstöne, *Ann. Phys. Chem. 99*, 497-540 (1856).

Helmholtz, H.L.F. von: *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*, F.Vieweg & Sohn, Braunschweig, 1st. Ed., Chapter 11 (1863).

Hermann, L.: Zur Frage betreffend den Einfluss der Phasen auf die Klangfarbe, *Ann. Phys. Chem. 58*, 391-401 (1896).

König, R.: Ueber den Zusammenklang zweier Töne, *Ann. Phys. Chem. 157*, 177-237 (1876). [English translation, R.König: On the Simultaneous Sounding of Two Notes, *Phil. Mag. 5th Ser. 1*, 417-446 and 511-525 (1876).]

König, R.: Ueber den Ursprung der Stösse und Stosstöne bei harmonischen Intervallen, *Ann. Phys. Chem. 12*, 335-349 (1881).

Kuile, Th.E.ter: Einfluss der Phasen auf die Klangfarbe, *Arch. ges. Physiol. 89*, 333-426 (1902).

Kuile, Th.E.ter: *Over phasen, zwevingen en klankaard*, Doctoral Dissertation. University of Amsterdam (1904).

Lawrence, M., and P.A.Yantis: Thresholds of Overload in Normal and Pathological Ears, *A.M.A. Arch. Otol. 63*, 67-77 (1956a).

Lawrence, M., and P.A.Yantis: Onset and Growth of Aural Harmonics in the Overloaded Ear, *J. Acoust. Soc. Am. 28*, 852-858 (1956b).

Lawrence, M., and P.A.Yantis: In Support of an "Inadequate" Method for Detecting "Fictitious" Aural Harmonics, *J. Acoust. Soc. Am. 29*, 750-751 (1957).

Lottermoser, W.: Bemerkungen zu den subjektiven, harmonischen Teiltönen nach G.v.Békésy, *Akust. Z. 2*, 148 (1937).

Mathes, R.C., and R.L. Miller, Phase Effects in Monaural Perception, *J. Acoust. Soc. Am. 19*, 780-797 (1947).

Meyer, M.F.: Beats from Combining a Unit Frequency with a Mistuned Multiple, *Am. J. Psychol. 62*, 424-430 and 592-593 (1949).

Meyer, M.F.: Crucial Experiments in Cochlear Mechanics, *Am. J. Psychol. 66*, 261-269 (1953).

Meyer, M.F.: A Test of the Hydraulic Theory of Cochlear Mechanics through Multiple Frequency Beats, *Am. J. Psychol. 67*, 39-55 (1954).

Meyer, M.F.: Aural Harmonics are Fictitious, *J. Acoust. Soc. Am. 29*, 749 (1957).

Moe, C.R.: Experimental Study of Subjective Tones Produced within the Human Ear, *J. Acoust. Soc. Am. 14*, 159-166 (1942).

Müller, J.J.: Ueber die Tonempfindungen, *Ber. Verhandl. Kön. Sächs. Gesellsch. Wissensch. Leipzig, Math.-Phys. Classe 23*, 115-124 (1871).

Nordmark, J.: Perception of Distance in Animal Echo-Location, *Nature 188*, 1009-1010 (1960).

Ohm, G.S.: Bemerkungen über Combinationstöne und Stösse, *Ann. Phys. Chem. 47*, 463-466 (1839).

Opheim, O., and G.Flottorp: The Aural Harmonics in Normal and Pathological Hearing, *Acta Oto-Laryngol. 45,* 513-531 (1955).

Peterson, J.: Combination Tones and Other Related Auditory Phenomena, *Psychol. Rev. 9, Psychol. Monogr. No. 39* (1908).

Röber, A.: Untersuchungen des Hrn. Scheibler in Crefeld über die sogenannten Schläge, Schwebungen oder Stösse, *Ann. Phys. Chem. 32,* 333-362 and 492-520 (1834).

Schouten, J.F.: The Perception of Subjective Tones, *Proc. Kon. Nederl. Akad. Wetensch. 41,* 1086-1093 (1938).

Stumpf, C.: Ueber die Ermittelung von Obertönen, *Ann. Phys. Chem. 57,* 660-681 (1896).

Stumpf, C.: Beobachtungen über Kombinationstöne, *Z. Psychol. 55,* 1-142 (1910). [=*Beitr. Akust. Musikwissensch. 5,* 1-142 (1910).]

Thomson, W.: On Beats of Imperfect Harmonies, *Proc. Roy. Soc. Edinb. 9,* 602-612 (1878).

Trimmer, J.D., and F.A. Firestone: An Investigation of Subjective Tones by Means of the Steady Tone Phase Effect, *J. Acoust. Soc. Am. 9,* 24-29 (1937).

Waard, C. de: *Journal tenu par Isaac Beeckman de 1604 à 1634, publié avec une introduction et des notes par C. de Waard,* Vol. 3, Martinus Nijhoff, La Haye, p. 51 (1945).

Wegel, R.L., and C.E. Lane: The Auditory Masking of One Pure Tone by Another and its Probable Relation to the Dynamics of the Inner Ear, *Phys. Rev. 23,* 266-285 (1924).

Zwicker, E.: Die Grenzen der Hörbarkeit der Amplitudenmodulation und der Frequenzmodulation eines Tones, *Acustica 2, Akust. Beihefte,* 125-133 (1952).

6

PITCH OF COMPLEX TONES

It was pointed out in Chapter 1 that the periodic sound waves produced by musical instruments and the human voice usually consist of a large number of harmonics. As we saw in Chapter 2, the organ of hearing is able to distinguish the first 5 to 8 harmonics of a complex tone individually, indicating that the auditory mechanism may be compared with a (continuous) frequency analyzer with a selectivity corresponding to the critical bandwidth.

In normal listening, however, we are not aware of the individual harmonics, but the complex tone is always perceived as a single tone with one definite pitch equal to the pitch of the fundamental. The "classical" explanation attributes this to the high relative loudness of the fundamental, which either exists in the sound itself or is introduced by the ear's nonlinearity. There are indications, however, that the presence of the fundamental is not essential for hearing a corresponding pitch but that pitch is related to periodicity. For that reason, it appeared to be of interest to investigate more extensively on which physical parameter the pitch of complex tones is based: on the frequency of the fundamental or on the periodicity of the soundwave as a whole. If the latter should be the case indeed, then one is confronted with the question of how this fact can be reconciled with the ear's frequency-analyzing power.

HISTORICAL REVIEW

Not until the publication of a paper by Seebeck in 1841 was it realized that the pitch of complex tones might be a controversial question. Seebeck reported the results of some investigations on the conditions for hearing tones. His sound sources were rotating discs provided with a concentric ring of holes; the interspaces between the holes were in most cases larger than the diameters of the holes. Air was blown through a tube perpendicular to the rotating disc so that the air stream was interrupted periodically by the interspaces between the holes. With equidistant holes, these siren discs produced a tone with a pitch corresponding to the number of holes per second that passes the air tube. The use of these devices demonstrated that short periodic pressure pulses give rise to tone sensation, so it is not surprising that Seebeck interpreted his findings in terms of periodicities perceived as tones.

Seebeck used this siren to investigate the pitch for discs with different distances between the holes, namely with distances alternately equal to a, b, a, b, etc. and a, b, c, a, b, c, etc., respectively. His most important conclusions were that (1) the hearing organ is able to analyze such a sequence of irregular pulses into two and three sequences of periodic pulses, respectively, for only one pitch is heard equal to the pitch of a disc with equidistant holes at a distance of $a+b$ in the first example, and $a+b+c$ in the second one; (2) for small differences between the distances a and b, and a, b, and c, respectively, the ear overlooks the irregularities and a pitch is heard corresponding to $\frac{1}{2}(a+b)$ and $\frac{1}{3}(a+b+c)$, respectively.

Ohm criticized two years later Seebeck's interpretation of these experiments (Ohm, 1843). Up to now, so he wrote, he had accepted as an established truth that a sinusoidal wave given by $\sin(2\pi ft+\varphi)$ is required to hear a pitch corresponding to the frequency f. This is Ohm's famous definition of tone, mentioned previously in Chapter 2. In his opinion, Seebeck's conclusions were at variance with this definition for they suggested that the pitch of tones is based on periodicities of acoustic pulses. Introducing Fourier's theorem on the resolution of a periodic function into sinusoids, Ohm demonstrated that the sounds produced by the siren discs contained sinusoidal components corresponding to the pitches heard. He tried in this way to reconcile Seebeck's experimental results with the "traditional" conception of the origin of tone pitch.

In the continuation of this interesting discussion, the difference between the views of Seebeck and Ohm about the physical correlate of pitch becomes more clear. In view of the fact that the pitch corresponding to the period of acoustic pulses of a siren disc with equidistant holes is heard much stronger than the harmonics, Seebeck (1843) concluded that the presence of a sinusoidal component with frequency f is not essential for hearing a pitch corresponding to this frequency. Ohm (1844) replied that this phenomenon might be due to "acoustical illusion" of the ear. Rightly, Seebeck (1844a) took this answer as a confirmation of his point of view because, as he said, only the ear can decide in which way tones are perceived. Moreover, he indicated that only by assuming that the pitch of a periodic sound wave is based on the total of the sinusoidal components, the existence of different timbres of tones with the same pitch can be understood. In a closing paper (Seebeck, 1844b), he tried to prove quantitatively that periodic sound waves with very faint fundamental can be produced. He also found in those cases that it was difficult to distinguish the higher harmonics, but that the sound as a whole had a distinct pitch corresponding to the periodicity of the pulses.

The fact that Helmholtz (1863) stronglv promoted Ohm's view did not mean the end of the controversy. König published in 1876 a detailed study on beats and combination tones, including also some investigations with siren discs. In one of these experiments, he rotated such a disc in front of a vibrating tuning fork. He found for an interruption rate of 128 per second and a tuning fork of

n.128 cps (n=integral number) that a pitch corresponding to 128 cps could be heard. The strength and clearness of the pitch of this "interruption tone" increased for increasing n. This stimulus, consisting of periodic tone pulses, can be interpreted as a complex tone of 128 cps (see Figure 1, in which a similar signal, consisting of tone pulses with a rectangular envelope, is presented). In another experiment, König blew siren discs provided with equidistant holes varying periodically in diameter, and also in this case the variations resulted in a definite pitch. Although König did not touch the question of the Fourier spectrum of these sounds, it is evident that the stimulus in both cases consisted of a complex tone with harmonics much stronger than the fundamental.

König's experiments directed the attention of several other investigators to the phenomenon of interruption tones. Dennert (1887) repeated and extended these experiments and came to the same conclusion as König did. In his opinion, the existence of interruption tones is incompatible with Ohm's acoustical law. Hermann (1890), in a study on the frequency spectrum of sung vowels, determined that in many cases the fundamental is nearly absent; nevertheless, always a pitch equal to the pitch of this tone is heard. Pursuing his interest in this question, he carried out experiments similar to those of König and Dennert and found that when a tone is periodically interrupted, the tone disappears as the interruption rate increased and a tone with a pitch corresponding to this rate becomes audible. Apparently, so he concluded, the ear is capable of perceiving periodic variations in amplitude as a tone. Pipping (1895) made an interesting attempt to reconcile this phenomenon with Helmholtz's hearing theory by introducing a distinction between the pitch of a tone and the pitch of a clang ("Klang" in German): for a complex tone, we may direct our attention to the partials (tone pitch), but it is also possible to pay attention to the total impression of the sound (clang pitch). The latter pitch is not influenced by the absence of the fundamental, a group of partials being *sufficient to perceive this pitch.*

Pipping also pointed to the fact that in cases that the stimulus does not contain the fundamental, this tone will be reintroduced by nonlinear distortion in the hearing organ. Everett (1896) and Schaefer (1899) even suggested that the interruption tone itself may be due to nonlinearity. They tried in this way to obviate the difficulty how the existence of this tone could be squared with the frequency principle in hearing. For that reason, the origin of interruption tones was studied more extensively by Schaefer and Abraham (1901, 1902). On the basis of their experiments with siren discs, they found that in many cases the objective existence in the sound of a partial with a pitch equal to the pitch of the interruption tone could be demonstrated with the aid of an acoustic resonator, whereas they explained the latter tone in the other cases as a difference tone produced in the ear. Thus in all cases the interruption-tone perception was attributed to a simple tone in accordance with Ohm's law.

Meanwhile, more investigations on interruption tones were published. Zwaardemaker (1900) placed the tone source in front of a microphone and, by means of an electric interrupter, a strong interruption tone could be perceived. These experiments were repeated by Schaefer and Abraham (1904), who obtained a somewhat different result; they only heard a distinct tone corresponding to the interruption rate if the frequency of the tone source was a multiple of the number of interruptions per second. Schulze (1908) explained these results in terms of Helmholtz's hearing theory. Hermann (1912), however, pointed to the fact that the results contradicted Schaefer and Abraham's opinion that the interruption tone has to be considered as a difference tone. Interrupting a tone of f cps g times per second gives rise to other tones of $f-g$, $f+g$, $f-2g$, $f+2g$ cps, etc; every pair of adjacent partials will contribute, in the case of nonlinear distortion, to the production of a difference tone of g cps, independent of the fact whether $f=ng$ (n=integral number) or not. So the fact that a distinct tone corresponding to g was heard only for $f=ng$ does not support Schaefer and Abraham's own interpretation that this tone was a difference tone.

Using similar equipment, Hermann performed new experiments and came to the interesting conclusion that sometimes for $f \neq ng$ the pitch that is heard does not correspond to g, but to a value differing by 10 to 20%. He was not able to explain this effect. For $f \neq ng$ the stimuli used by Hermann consisted of a set of inharmonic partials, separated by g cps. Therefore, these experiments anticipated some much more recent experiments wherein similar pitch shifts were found. Without his realizing it, Hermann's findings represented one of the best demonstrations that the interruption tone is not a difference tone, but has to be understood as the result of periodicity perception.

Summarizing our historical review on the pitch of a complex tone up to about 1920, we may conclude that at that time the problem was still just as controversal as it was 80 years before. Most workers followed Schaefer and others in their attempt to reconcile the experimental evidence on the pitch of complex tones with the frequency principle by pointing to the fact that mostly the fundamental is not completely absent and that the ear's distortion may reintroduce this tone. However, there were some investigators who did not accept this point of view, because it could not explain why the fundamental pitch of complex tones is so dominant. Perhaps it was not possible to solve this problem with the equipment available at that time.

Even after the introduction of electronic equipment in hearing research, it took many years to surmount the impasse. Fletcher (1924) used electric filters to eliminate the lower harmonics of complex tones and noticed that the pitch did not change. He studied synthesized complex tones consisting of multiples of 100 cps and found that three consecutive harmonics were sufficient to give a clear musical tone with a pitch equal to the pitch of a simple tone of 100 cps. Fletcher explained these results by assuming that in all these cases the missing fundamental is reintroduced by the ear's nonlinear distortion. Jeffress (1940) investigated the pitch of complex tones in which the fundamental was removed and obtained different results. His subjects found great difficulty in comparing this pitch with the pitch of the fundamental, and in most cases they preferred the octave of the latter tone for best match.

We now arrive at some pioneer investigations by Schouten, carried out in 1938 to 1940, the importance of which was not fully realized before about 1955. The experiments were carried out with the aid of an optical siren with which it was possible to produce periodic sounds with any desired waveform. In a first study, Schouten (1938) investigated the pitch of periodic pulses of finite width (one twentieth of the repetition time of 1/200 sec) of which the fundamental of 200 cps was cancelled completely. This was verified by the absence of beats when an additional tone of 206 cps was introduced. The pitch of the complex tone, however, was exactly the same as the pitch of the periodic pulse with the fundamental included. In a second paper, Schouten (1940a) described more extensively the sensation corresponding to a periodic sound wave with many partials and he stated: "The lower harmonics can be perceived individually and have almost the same pitch as when sounded separately. The higher harmonics, however, cannot be perceived separately but are perceived collectively as one component (the residue) with a pitch determined by the periodicity of the collective wave form, which is equal to that of the fundamental tone." A third paper (1940b), dealing with the consequences of these investigations for hearing theory, reconciled periodicity perception with auditory frequency analysis. This paper will be discussed further in Chapter 8.

Schouten's conclusion, in agreement with the views of some scientists mentioned above, that periodicity of waveform may lead to a definite pitch sensation, was strongly opposed by Hoogland (1953) as a representative of the "traditional" view that a sinusoidal sound wave is necessary to hear a corresponding pitch. Although Hoogland's observations, indicating that the ear's nonlinearity does introduce the fundamental of a series of higher harmonics, are correct, his opinion that also Schouten had perceived a difference tone shows that he missed the essence of Schouten's exposition.

Other experiments, however, supported Schouten's results. Davis *et al.* (1951) led brief periodic pulses through a band-pass filter with a centre frequency of 2000 cps and asked subjects to match the pitch of this signal to the pitch of a simple tone. The listeners varied greatly in the accuracy with which the latter pitch was matched to either the period of the pulses (about 130 pps) or the band-pass frequency; errors of exactly one octave were particularly common.

A much more conclusive experiment was carried out by Licklider. He demonstrated in 1954 that a pitch corresponding to the periodicity of a series of high-frequency harmonics of a low-frequency (missing) fundamental was audible even when low-frequency noise, sufficiently loud to mask a possible fundamental created in the ear, was introduced (see his paper of 1956). He made a tape-recording of pairs of tone pulses; the first pulse in each pair consisted of low-frequency sinusoids and the second pulse of high-frequency harmonics of the same low frequency. The first tone pulse was definitely louder than the complex of harmonics. The successive pairs of tone pulses progressed up and down in the scale of frequency. The tone pulses consisting of sinusoids disappeared by the introduction of the masking noise, but the low pitch of the other tone pulses could be heard right through the noise, going up and down. During Licklider's visit to Utrecht, in 1955, the writer had the opportunity to hear this tape-recording; the demonstration was very impressive.

Thurlow and Small (1955) repeated nearly at the same time the experiments of Davis *et al.*, mentioned above, but they used band-pass filters with different centre frequencies. For a 1000-cps filter, the pitch corresponded most closely to the repetition rate of the pulses (100 pps) and no pitch corresponding to 1000 cps was heard. For a 5000-cps filter, both low and high pitches were heard. The stimuli were presented at low sensation levels, and spectral analysis showed that no low frequency corresponding to pulse rate was present.

In another experiment, Small (1955) used periodic tone pulses by interrupting tones of 1000, 2000, 4000, and 8000 cps, respectively, with an interruption rate of 100 per second. Varying also pulse duration and rise-fall time, he asked listeners to match the pitch of the periodic tone pulses to the pitch of a sinusoidal comparison tone. The stimuli were presented at a very low sensation level of 20 dB in order to avoid audible tones at 100 cps, corresponding with repetition rate. Two pitches were identified, one low pitch corresponding to periodicity, and a high pitch corresponding to the frequency of the interrupted tone. The periodicity pitch decreased in audibility with the increase of any of the following variables: frequency, pulse duration and rise-fall time. These variables had the opposite effect on the perception of the high pitch, which generally was audible only for the longer pulse durations.

The stimuli used in this and the foregoing experiment can be regarded as consisting of a series of high harmonics of a (missing) fundamental. Apparently, the audibility of periodicity pitch for 100 cps decreases when the frequency of the strongest harmonic shift from 1000 to 8000 cps and increases when the stimulus contains more harmonics. Small found that periodicity-pitch perception is not limited to repetition rates that are submultiples of the frequency of the interrupted tone; varying the repetition rate of the tone pulses resulted in a corresponding pitch shift. In this case, the partials of the complex signal are no multiples of the repetition rate.

The perception of pitch of inharmonic

complexes of tones was studied more extensively by de Boer (1956). In his experiments, based on an important observation by Schouten (1940c), he made use of amplitude modulation as described in the following example. A simple tone of 2000 cps is modulated in amplitude by a signal composed of tones of 200, 400, and 600 cps, resulting in a complex of tones with frequencies of 1400, 1600, 1800, 2000, 2200, 2400, and 2600 cps. This stimulus has a clear pitch equal to the pitch of a 200-cps tone, even for low sensation levels (20 to 40 dB was used). By shifting the carrier frequency from 2000 cps up to, for instance, 2030 cps, the originally harmonic complex of tones is transformed in an inharmonic complex with frequencies of 1430, 1630, 1830, 2030, 2230, 2430, and 2630 cps. De Boer observed that this transformation results in a pitch shift from a value corresponding to a tone of 200 cps to a value corresponding to about 203 cps. This effect cannot be attributed to nonlinear distortion for in both stimuli the frequency difference between adjacent partials is 200 cps. De Boer demonstrated that the phenomenon can be explained by assuming that pitch sensation is caused by periodicities in waveform. It is interesting to notice that similar pitch shifts were already observed by Hermann (1912), as we saw above.

Later investigations showed that the audibility of periodicity depends upon both the frequency of the harmonics presented and the frequency of the missing fundamental. Flanagan and Guttman (1960) carried out experiments in which the pitch of a periodic train of very short pulses was matched to that of another train whose fundamental tone was removed. They found that, for pulse rates less than 100 pps, the pitches are matched to equal pulse rate, regardless of the polarity pattern of the pulses (positive and negative pulses alternated). For fundamental frequencies of 1000 cps and higher, subjects tend to equate the fundamental of the matching tone to the lowest spectral component present in the matched stimulus. In between these ranges, in particular between fundamental frequencies of 200 and 500 cps, there is a decided tendency to equate fundamental frequencies, so in this range the polarity pattern of the pulses is indeed taken into account.

The limits of periodicity pitch were studied in another way by Ritsma (1962) with a harmonic complex of 3 tones produced by modulating the amplitude of a simple tone with frequency f with another simple tone with frequency g (see Figure 1). The minimal modulation depth required to hear periodicity pitch was determined for $f/g=4, 5, 6, 7$, etc., respectively, as a function of g. The results indicate that (1) periodicity is heard more clearly as modulation depth increases; (2) no periodicity pitch is audible for $f>$ about 5000 cps nor for $g>$ about 800 cps either. The latter result agrees with the finding of Flanagan and Guttman, mentioned above.

From this review on the pitch of complex tones, we may conclude that experimental evidence has shown clearly that, over a wide range, the presence of the fundamental is not required to perceive a definite pitch equal to the pitch of the fundamental. Apparently, the ear is able to use periodicity as a basis of pitch perception.

EXPERIMENTS

The conclusion that the ear appears to be able to use the period of a sound wave as a basis of pitch perception deviates from the "classical" opinion that pitch is related to frequency. It seems rather unlikely, however, that the ear is provided with two pitch-detecting mechanisms, one using frequency and the other periodicity.

For a better insight into this question, it is of interest to know more about the relative contribution of the fundamental to the pitch of a complex tone. If experimental evidence should indicate that this pitch is mainly or exclusively based on the tone's periodicity, this would support the view that frequency is not a relevant parameter for pitch. In order to answer this question, the following experiments were carried out.

Method

If the pitch of a complex tone is mainly based on the frequency of the fundamental, we may expect that the pitch is much more influenced by a 10% frequency decrease of the fundamental than by a simultaneous 10% frequency increase of the second and higher harmonics. This was studied by presenting the ear, successively, with two stimuli:

stimulus 1: $$\sum_{n=1}^{12} a_n \cos 2\pi n f t,$$

stimulus 2:

$$a_1 \cos 2\pi(0.9f)t + \sum_{n=2}^{12} a_n \cos 2\pi n(1.1f)t.$$

The subject's task was to indicate whether the pitch of stimulus 2 was higher or lower than of stimulus 1.

Since it appeared that, over a large frequency range, the pitch of stimulus 2 was definitely judged as higher than of stimulus 1, demonstrating that pitch depends upon the periodicity of the second and higher harmonics together rather than on the frequency (or periodicity) of the fundamental, additional experiments were carried out in which stimulus 2 was of the more general form

$$\sum_{n=1}^{m} a_n \cos 2\pi n(0.9f)t + \sum_{n=m+1}^{12} a_n \cos 2\pi n(1.1f)t,$$

with $m=1, 2, 3, 4$, respectively. It was possible in this way to investigate to what degree pitch is related to the lower or the higher harmonics of a complex tone.

In one experiment the amplitudes of all harmonics were equal to each other ($a_n=1$). Since this situation may not be representative for complex tones as used in practice, a circuit with a frequency-response characteristic of -6 dB per octave was introduced in a second experiment ($a_n=1/n$). The relative amplitudes of the harmonics agree in the latter case rather well with the mean situation for speech vowels and musical tones.

Apparatus and procedure

Figure 54 represents a block diagram of the apparatus. Function generator 1 was the same as used in Chapter 2. It was adjusted to produce the function

$$\sum_{n=m+1}^{12} \cos 2\pi ngt.$$

Function generator 2 consists of a circuit that divides the frequency of the input signal by a factor 3, followed by a circuit similar to function generator 1 but with 20 shift-register elements (also used in the sweep-tone experiments, Chapter 5). This generator produced the function

$$\sum_{n=1}^{m} \cos 2\pi ngt.$$

Periodic input signals of frequency $60g$ are required to produce these functions. These signals could be taken from three sine-wave oscillators (Hewlett-Packard 200 CD) by means of the electronic switch S_1. Oscillator 1 was adjusted to $60f$, oscillator 2 to $66f$, and oscillator 3 to $54f$. In position 1 of S_1, both function generators are connected to oscillator 1, so the superimposed output signals of the generators result in a signal equal to stimulus 1 with $a_n=1$. In position 2, the generators are connected to oscillators 2 and 3, respectively, resulting in an output signal equal to stimulus 2, also with $a_n=1$. Switch S_1 was controlled by an electronic circuit with two switching programs represented by A and B in Figure 55. In both conditions, the stimuli were presented two times, with a stimulus duration of 200 msec. Possible frequency components above $12f$ were removed by means of a low-pass filter. A -6 dB per octave filter was introduced in the upper position of switch S_2, corresponding with $a_n=1/n$. The headphone (Beyer DT 48) was provided with a correction filter after the design of Zwicker and Gässler (1952).

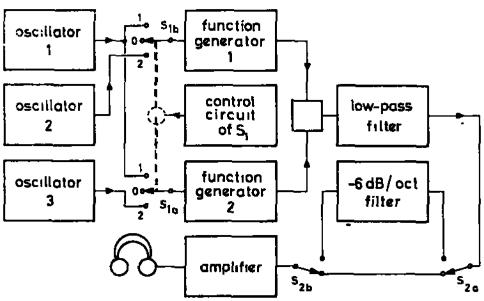


FIGURE 54. Block diagram of the apparatus.

The equipment used for adjustment and control of the signals is not represented in Figure 54. The periodic functions were adjusted in the same way as treated in Chapter 2. The amplitude of the fundamental of the signal produced by generator 1 was in all cases about 40 dB or more lower than of the partials wanted. Every time that S_1 was switched to position 1, a resetting pulse was applied automatically to both function generators, which assured that the phase relation between their output signals was always the same. These signals were controlled continuously during the experiments with the aid of a double-beam oscilloscope. In addition, the experimenter could check

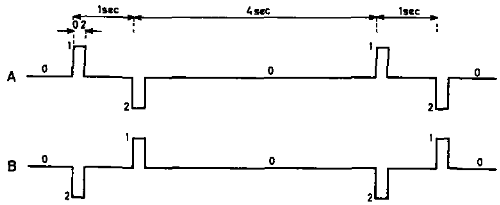


FIGURE 55. Position of switch S_1 for conditions A and B. The positions 1 and 2 correspond with the presentation of stimuli 1 and 2, respectively.

the stimuli presented to the subject by means of a loudspeaker.

The experiments were carried out for $f=125$, 175, 250, 350, 500, 700, 1000, 1400, and 2000 cps, respectively. For each value of f, both the cut-off frequency of the low-pass filter and the tilting frequency of the -6 dB per octave filter, if used, were adapted to f. The stimuli were presented monaurally at a sound-pressure level of 60 dB.

The following procedure was used. The subject was presented with two identical pairs of successive tone pulses after the scheme reproduced in Figure 55 (condition A or B). After that, he had to indicate whether the first or the second tone pulse of the pair was higher in pitch. In order to avoid the influence of timbre, no comparison tone pulses were used, but the subject was asked to reproduce the pitch difference between the stimuli vocally by singing, humming or whistling. He was told before that not the absolute pitch but only the shift of it was important.

In a typical test session, the subject had to judge the pitch shift two times for all values of f mentioned above, since the stimuli were presented once in condition A and once in condition B. These 18 presentations were given in random order, so the subject did not know *a priori* whether stimulus 1 or stimulus 2 preceded. In 8 test sessions, the experiment was repeated for $m=1$, 2, 3, 4, with $a_n=1$ and $a_n=1/n$, respectively. Fifteen subjects participated in the experiments and care was taken that no one knew the physical properties of the stimuli and the feature of the experiments. The subjects were tested individually.

Results

Figures 56 and 57 present the results of the experiments. The percentage of responses, averaged over fourteen subjects, in which the pitch of stimulus 2 was judged as higher than of stimulus 1 is plotted as a function of f, with m as a parameter. The data points in Figure 56 were obtained for $a_n=1$, and the points in Figure 57 for $a_n=1/n$. Stated differently, the percentages indicate how often the shift in pitch corresponded to the shift in frequency of the higher partials of the stimuli. From the original group of fifteen subjects, one had to be omitted because he did not give reliable results even in the case of judging the direction of pitch shift for tones produced by a piano.

The curves show that, for $m=1$ and $m=2$, there is a distinct transition between a fundamental-frequency range over which the pitch shift follows the higher harmonics and a range over which it follows the fundamental.

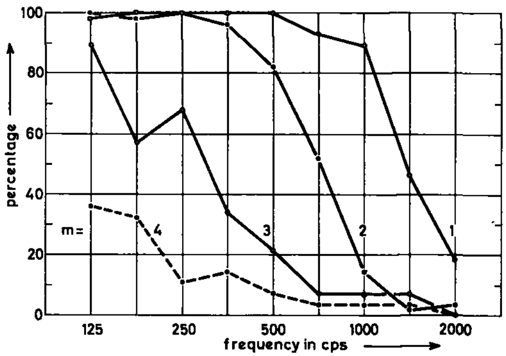


FIGURE 56. Percentage of responses, averaged over 14 subjects, in which the pitch of the stimulus $\sum_{n=1}^{m} \cos 2\pi n(0.9f)t + \sum_{n=m+1}^{12} \cos 2\pi n(1.1f)t$ was judged as higher than of $\sum_{n=1}^{12} \cos 2\pi nft$ as a function of f, with m as a parameter.

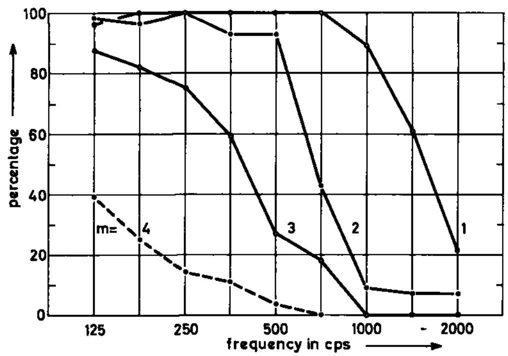


FIGURE 57. Percentage of responses, averaged over 14 subjects, in which the pitch of the stimulus $\sum_{n=1}^{m} \frac{1}{n} \cos 2\pi n(0.9f)t + \sum_{n=m+1}^{12} \frac{1}{n} \cos 2\pi n(1.1f)t$ was judged as higher than of $\sum_{n=1}^{12} \frac{1}{n} \cos 2\pi nft$ as a function of f, with m as a parameter.

It is of interest to point out that, for frequencies within these ranges, the subjects considered their task as a very easy one; they did not notice that the harmonic complex of tones was split up in two parts with opposite frequency shifts but had the impression that the pitch of the whole stimulus was varied. As to the transition range, some subjects remarked that they had noticed both shifts indeed. In this respect, the experiments for $m=3$ and $m=4$ were judged as much more difficult than for the lower m-values.

Discussion

The curves in Figures 56 and 57 for $m=1$ show that, for fundamental frequencies below about 700 cps, the responses agree with the frequency shift of the higher harmonics. This implies that the pitch of complex tones is determined in this frequency range by the harmonics and not by the fundamental. For higher frequencies, however, the situation is different. Beyond about 1400 cps, the pitch followed the fundamental in the majority of cases. This result can be compared with the finding by Flanagan and Guttman (1960) that, for 1000 cps and higher, pitch matches of complex tones are determined by the lowest partial of the stimuli. Furthermore, it is of interest to remember that, for harmonic complexes consisting of only 3 adjacent partials, Ritsma (1962) could observe periodicity pitch for (missing) fundamentals up to about 800 cps; it is obvious that for more partials, as in our experiments, this limit shifts to a higher frequency.

For higher m-values, the curves shift to lower frequencies. It is remarkable that there are only minor differences between the curves of Figure 56 and Figure 57, with $a_n=1$ and $a_n=1/n$, respectively, indicating that the relative amplitudes of the partials do not influence the pitch of complex tones over this range. It is also remarkable that even in the case that the frequency of the first 3 harmonics of 125 cps were lowered, still about 90% of the responses was determined by the higher harmonics.

In the writer's opinion, the experimental data are significant in answering the question whether pitch is based on frequency or on periodicity. Although this point is discussed more extensively in Chapter 8, some conclusions seem to be justified already. The finding that the pitch of complex tones below about 1400 cps is determined by the harmonics rather than by the fundamental strongly suggests that pitch is not derived from frequency but from periodicity. It is not very probable that the situation should be different in the frequency range above 1400 cps. The highest pitch that can be produced by the human voice does not exceed this frequency, so it is not clear why we should have a second pitch mechanism that is never used in social life. The consequence of this reasoning is that also the pitch of simple tones is based on periodicity. The fact that beyond 1400 cps pitch is related to the fundamental can be explained, as is discussed in Chapter 8, by the limit of the ear's ability to detect periodicities.

In the light of these results, we understand why in practice, complex tones are always characterized by one definite pitch. If pitch is based on frequency, we should expect a number of different pitches to be audible, just as in the case of an inharmonic complex of tones. If, however, pitch is based on periodicity, the existence of one pitch, related to the period of the stimulus, is obvious.

The question can be asked how this view is compatible with our earlier conception of the ear as a frequency analyzer (Chapter 2). *At first sight, it seems difficult indeed to* reconcile preservation of periodicity with this analyzing process. We must realize, however, that the frequency-analyzing power is determined by the critical bandwidth, so it is rather limited. Moreover, the experiments of Chapter 5 showed that simple tones do interfere over much larger frequency distances. We will see in Chapter 8 how, owing to these facts, periodicity is preserved.

CONCLUSIONS

1. For fundamental frequencies up to about 1400 cps, the pitch of a complex tone is determined by the frequency of the second and higher harmonics and not by the frequency of the fundamental, whereas beyond this frequency the opposite holds; this is the case both for tones with harmonics of equal amplitude as for tones with harmonics of which the amplitudes fall down with 6 dB

per octave.

2. For fundamental frequencies up to about 700 cps, the pitch is determined by the third and higher harmonics, for frequencies up to about 350 cps by the fourth and higher harmonics.

3. The experimental results strongly suggest that the pitch of complex tones is based on periodicity rather than on frequency; it is reasonable that this also holds for simple tones.

REFERENCES

Boer, E. de: *On the "Residue" in Hearing*, Doctoral Dissertation, University of Amsterdam (1956).

Davis, H., S. R. Silverman, and D. R. McAuliffe: Some Observations on Pitch and Frequency, *J. Acoust. Soc. Am. 23*, 40-42 (1951).

Dennert, H.: Akustisch-physiologische Untersuchungen, *Arch. Ohrenheilk. 24*, 171-184 (1887).

Everett, J. D.: On Resultant Tones, *Phil. Mag. 5th Ser. 41*, 199-207 (1896).

Flanagan, J. L., and N. Guttman: Pitch of Periodic Pulses without Fundamental Component, *J. Acoust. Soc. Am. 32*, 1319-1328 (1960).

Fletcher, H.: The Physical Criterion for Determining the Pitch of a Musical Tone, *Phys. Rev. 23*, 427-437 (1924).

Helmholtz, H. L. F. von: *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*, F. Vieweg & Sohn, Braunschweig, 1st Ed., Chapter 4 (1863).

Hermann, L.: Phonophotographische Untersuchungen. III, *Arch. ges. Physiol. 47*, 347-391 (1890).

Hermann, L.: Neue Versuche zur Frage der Unterbrechungstöne, *Arch. ges. Physiol. 146*, 249-294 (1912).

Hoogland, G. A.: *The Missing Fundamental*, Doctoral Dissertation, University of Utrecht (1953).

Jeffress, L. A.: The Pitch of Complex Tones, *Am. J. Psychol. 53*, 240-250 (1940).

König, R.: Ueber den Zusammenklang zweier Töne, *Ann. Phys. Chem. 157*, 177-237 (1876). [English translation, R. König: On the Simultaneous Sounding of Two Notes, *Phil. Mag. 5th Ser. 1*, 417-446 and 511-525 (1876).]

Licklider, J. C. R.: Auditory Frequency Analysis, in C. Cherry (Ed.): *Information Theory*, Butterworths Scientific Publications, London, pp. 253-268 (1956).

Ohm, G. S.: Ueber die Definition des Tones, nebst daran geknüpfter Theorie der Sirene und ähnlicher tonbildender Vorrichtungen, *Ann. Phys. Chem. 59*, 513-565 (1843).

Ohm, G. S.: Noch ein Paar Worte über die Definition des Tones, *Ann. Phys. Chem. 62*, 1-18 (1844).

Pipping, H.: Zur Lehre von den Vocalklängen, *Z. Biol. 13*, 524-583 (1895).

Ritsma, R. J.: Existence Region of the Tonal Residue. I, *J. Acoust. Soc. Am. 34*, 1224-1229 (1962).

Schaefer, K. L.: Eine neue Erklärung der subjectiven Combinationstöne auf Grund der Helmholtz'schen Resonanzhypothese, *Arch. ges. Physiol. 78*, 505-526 (1899).

Schaefer, K. L., and O. Abraham: Studien über Unterbrechungstöne, *Arch. ges. Physiol. 83*, 207-211 (1901), *85*, 536-542 (1901), *88*, 475-491 (1902).

Schaefer, K. L., and O. Abraham: Zur Lehre von den sogenannten Unterbrechungstönen, *Ann. Phys. 13*, 996-1009 (1904).

Schouten, J. F.: The Perception of Subjective Tones, *Proc. Kon. Nederl. Akad. Wetensch. 41*, 1086-1093 (1938).

Schouten, J. F.: The Residue, a New Component in Subjective Sound Analysis, *Proc. Kon. Nederl. Akad. Wetensch. 43*, 356-365 (1940a).

Schouten, J. F.: The Residue and the Mechanism of Hearing, *Proc. Kon. Nederl. Akad. Wetensch. 43*, 991-999 (1940b).

Schouten, J. F.: The Perception of Pitch, *Philips Techn. Rev. 5*, 286-294 (1940c).

Schulze, F.A.: Die Übereinstimmung der als Unterbrechungstöne bezeichneten Klangerscheinungen mit der Helmholtzschen Resonanztheorie, *Ann. Phys. 26*, 217-234 (1908).

Seebeck, A.: Beobachtungen über einige Bedingungen der Entstehung von Tönen, *Ann. Phys. Chem. 53*, 417-436 (1841).

Seebeck, A.: Ueber die Sirene, *Ann. Phys. Chem. 60*, 449-481 (1843).

Seebeck, A.: Ueber die Definition des Tones, *Ann. Phys. Chem.* 63, 353-368 (1844a).

Seebeck, A.: Ueber die Erzeugung von Tönen durch getrennte Eindrücke, mit Beziehung auf die Definition des Tones, *Ann. Phys. Chem. 63*, 368-380 (1844b).

Small, A.M., Jr.: Some Parameters Influencing the Pitch of Amplitude Modulated Signals, *J. Acoust. Soc. Am. 27*, 751-760 (1955).

Thurlow, W.R., and A.M. Small, Jr.: Pitch Perception for Certain Periodic Auditory Stimuli, *J. Acoust. Soc. Am. 27*, 132-137 (1955).

Zwaardemaker, H.: Ueber Intermittenztöne, *Arch. Anat. Physiol. Physiol. Abt., Suppl. Band*, 60-67 (1900).

Zwicker, E., and G. Gässler: Die Eignung des dynamischen Kopfhörers zur Untersuchung frequenzmodulierter Töne, *Acustica 2, Akust. Beihefte*, 134-139 (1952).

7

PHYSIOLOGICAL CORRELATES OF TONE PERCEPTION

The experiments treated in the preceding chapters gave us some information about the tone-perception properties of the hearing mechanism. The conclusions drawn from these experiments were all formulated in psychophysical terms. This means that they inform us about the over-all characteristics of the hearing organ, considered as a "black box", but not about the way in which this organ is actually organized. The organization can only be elucidated by comparing the experimental data with the physiological knowledge of the ear. In this chapter, an exposition of the physiology of hearing is presented, and in Chapter 8 our experimental data are discussed in the light of the outcome of this survey.

The physiological knowledge relevant to tone perception is presented in the form of a historical review. This presentation has the advantage that one is able to compare the status of this knowledge at a certain moment with the contemporary status of psychophysical research treated in preceding chapters. Since the history of the progress in the knowledge of the ear's physiology is tightly interwoven with the development in auditory theory, attention is given also to this development. Only those theories are discussed which, in the writer's opinion, are relevant for explaining the experimental results presented in this study.

FOUR BASIC HYPOTHESES

Helmholtz's resonance theory

Undoubtedly, Helmholtz must be considered as the founder of modern hearing theory in his admirable attempt to explain the ear's capability to discriminate tones sounding simultaneously. As we saw in Chapter 2, Ohm introduced the view that the analyzing power of the hearing organ may be compared with the way in which periodic functions can be analyzed mathematically by applying Fourier's theorem (Ohm, 1843). Helmholtz fully recognized the significance of this hypothesis and based his theory on it in his famous book "Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik" (1863, Chapter 6).

Helmholtz's hearing theory can be considered as an elaboration of three hypotheses. In general terms, the first one is:

Hypothesis I. The analysis of sound is accomplished in the inner ear by means of a large number of resonators tuned to different *frequencies from low to high.*

Although the view that resonators may play an important rôle in hearing was suggested many times before him, Helmholtz for the first time worked it out in an attractive and cnvioncing way. Originally, he identified his resonators with the so-called arches of Corti, described by Corti in an important anatomical study on the inner-ear structures (1851). However, in the later editions of his book, Helmholtz revised his opinion and assumed that the transverse fibres of the basilar membrane act as resonators, the high frequencies corresponding with the basal and the low frequencies with the apical end of the cochlea. His arguments were: (1) in the cochlea of birds, no arches of Corti are found, as Hasse (1867) had shown; (2) the width of the basilar membrane varies from about 0.04 mm at its basis up to about 0.5 mm at the helicotrema (observation by Hensen, 1863); (3) the membrane is much more tightly stretched in its transverse than in its longitudinal direction. On the basis of psychophysical evidence, Helmholtz estimated the selectivity of the resonators; the values given by him correspond with a resonance bandwidth of 4% of the resonance frequency, so the bandwidth is proportional to frequency.

Furthermore, Helmholtz assumed that the vibrations of the transverse fibres of the basilar membrane give rise to activity of corresponding nerve fibres in the organ of Corti. This led to the second hypothesis:

Hypothesis II. A particular tone-pitch corresponds to each of the numerous nerve fibres in such a way that pitch decreases gradually from the basal to the apical end of the organ of Corti.

This assumption was proposed by Helmholtz as an extension of Müller's doctrine of "specific energies" of the different senses. Müller (1840) had suggested that stimulating a fibre of the optical nerve always results in a visual sensation, stimulating a fibre of the acoustical nerve in an auditory sensation, etc.

The two hypotheses, which were introduced to explain the frequency-analyzing power of the ear, also accounted for the interference of tones with a small frequency difference, but left unanswered the question how combination tones and beats of mistuned consonances come into existence. For that reason Helmholtz proposed an additional auditory mechanism:

Hypothesis III. The sound transmission of the ear is characterized by nonlinear distortion.

The implications of this hypothesis were treated in Chapter 3, whereas we saw in Chapter 5 how Helmholtz used combination tones to explain the beats of mistuned consonances. As sources of these tones, he considered the tympanic membrane and the joint between maleus and incus.

In this way, Helmholtz was able to present a rather convincing explanation of most experimental data available a hundred years ago. The Achilles' heel of his conception was why periodic sound waves are always characterized by a pitch corresponding to the fundamental. Helmholtz was not concerned about this phenomenon, because in his opinion the fundamental in practical sounds is always much stronger than the higher harmonics.

Helmholtz's theory became widely accepted soon after its publication under the names of *resonance theory* and *place theory*. The fact that it was based, apart from distortion, on two independent hypotheses, and not on one, was mostly overlooked. This has hampered a clear view on the merits of his theory, as we shall see below.

Periodicity of nerve impulses: an alternative hypothesis for pitch perception

The success of Helmholtz's resonance theory does not mean, however, that no valid criticism was put forward against it. Some objections shall be discussed in the next section, after first reviewing the way in which new insights in nerve physiology were used by Wundt and others to explain pitch. Since these ideas appear to be rather unknown nowadays, in contrast with Helmholtz's opinions, they are treated extensively below.

In 1866, Helmholtz announced in a lecture that stimulating muscles via their nerves with periodic electric pulses gives rise to synchronous mechanical vibrations of the muscle. Up to about 240 pps, these vibrations manifested themselves by a clear tone (Helmholtz, 1868). Bernstein, his former assistent and successor as professor in physiology in Heidelberg, published some years later a book entirely devoted to the (over-all) responses of nerves and muscles on electric stimuli (Bernstein, 1871). Applying an ingenious technique, he discovered that electric stimulation initiates a negative nerve impulse with a duration of 0.6-0.7 msec, and a propagation speed of about 29 m per second; the corresponding values for muscles were 4 msec and 2.9 m per second, respectively. In analogy with his finding that the reaction of the muscle decreases when the muscle impulses begin to overlap (about 250 stimuli per second), Bernstein suggested that the nerve is able to produce synchronous impulses up to a stimulation frequency of about 1600 cps, the reciprocal of the impulse duration of 0.6 msec. (Some years later, Bernstein (1875) found that the "muscle tone" could be heard up to about 1000 cps.)

It is very interesting to read how Bernstein related this conclusion to hearing. Following Helmholtz, he accepted the existence of a large set of resonators in the inner ear. Supported by the fact that mechanical vibrations are perceived as such by touch, he assumed that, for frequencies up to about 1600 cps, synchronous impulses are evoked in the nerve fibres, and then continues: "It is very remarkable indeed that the most used tones in music, which are the most agreeable ones to our ear and are discriminated well by it, lie below 1600 cps." From this, Bernstein suggested: "The intensity of stimulation will, up to this moment, increase with the number of impulses passing a cross section of the nerve in a unit of time." However, he did not consider the possibility that pitch might be related to the rate of nerve impulses, but adopted Helmholtz's conception of the specific energies as the basis of pitch.

This possibility was not overlooked, however, by Wundt. In the second edition of his work "Grundzüge der physiologischen Psychologie" (1880), he criticized Helmholtz's view that with any simple tone a particular nerve fibre corresponds, for this would imply a nearly infinite number of fibres. Helmholtz's attempt to escape this consequence by assuming that a tone which stimulates two nerve fibres produces an intermediate pitch was rejected by Wundt, since in that case not one but two pitches should be heard. Accepting that the different locations along the basilar membrane are tuned to different frequencies (Hypothesis I), Wundt proposed an alternative for Hypothesis II:

Hypothesis IV. Tones give rise to synchronous nerve impulses whose rate determines pitch.

Wundt tried to evade the difficulty that according to this hypothesis Bernstein's findings would suggest a pitch limit at about 1600 cps. He explained that not the total duration of the nerve impulses but the much shorter duration of their peaks might determine the highest pitch audible, and concluded: "Accepting this, there are no objections against the assumption that sound stimulation has to be considered as a special form of intermittent nerve stimulation and that in particular the tone sensation is based on a regular periodic progress of stimulation processes in the acoustic fibres themselves."

In particular, two arguments were put forward in favour of Hypothesis IV, namely the phenomenon of binaural beats and the claim that the auditory nerve can be stimulated directly by sound.

1. *Binaural beats.* Dove (1839) had pointed for the first time to the fact that stimulating the ears separately with tones of slightly different frequencies gives rise to slow "binaural beats". Usually, they were explained as resulting from bone conduction between the ears (Seebeck, 1846; Mach, 1875; Stumpf, 1890, p. 458; Schaefer, 1891). Thompson, who discovered the beats independently, found that they do not change over into a difference tone when the frequency difference is increased (1877, 1878, 1881). Therefore, he suggested that binaural beats are caused by interference in a higher centre of the auditory pathway.

A short paper by Scripture in 1892, in which he came to the same conclusion as Thompson, brought the interpretation of binaural beats into the centre of interest and led to a violent discussion about their origin. Nowadays it is universally accepted that the beats cannot be explained by bone conduction from ear to ear but with the simple equipment available at that time a conclusive answer was difficult to reach. The theoretical consequence of a "central" origin was very important, however, for it would imply that phase and thus periodicity is preserved in the auditory nerve. This may have been the reason that Wundt (1893) and Ewald (1894) joined in the discussion, started by Scripture (1892, 1893) and Schaefer (1893a). At the same time, the same problem was treated in an independent paper of Cross and Goodwin (1893). Against Schaefer, they all supported the view that the existence of binaural beats had to be considered as a clear evidence that the periodicity of sound waves is indeed preserved in the auditory nerve.

Beside the repeated criticism by Schaefer (1893b, 1895, 1901; also Rostosky, 1902) of the explanation of binaural beats by central interaction, the reaction of Bernstein, the originator of the view that synchronous nerve impulses are evoked, must be mentioned. He, too, sharply rejected Wundt's opinion that pitch might be based on the repetition rate of nerve impulses rather than on the specific energies of the nerve fibres (Bernstein, 1894). He supposed that Wundt's interpretation of the existence of binaural beats was not correct and for that reason he published the results of some experiments carried out already in 1887.

Bernstein's method for investigating whether or not the beats are caused by bone conduction is so amusing that even for this reason a description may be justified. By means of a short double biting board he achieved a close mechanical contact between the bones of his head and of the head of his wife. From another room, two tones of slightly different frequencies were supplied by rubber tubes, one to one of his ears and the other one to one of his wife's ears. Bernstein's expectation that, by bone conduction, also in this case beats would be audible, was not confirmed. Neither he nor his wife heard beats, in contrast with the result when the tones were presented separately to the ears of the same subject. Since he would not exclude the possibility that this negative result was caused by attenuation of the sound by the biting board, he did not consider the experiment as conclusive. Although he admitted that the explanation by central interaction might be correct, supporting his assumption that the periodicity of the stimulus is preserved in the auditory nerve, he was not prepared to admit that pitch might be based on this periodicity.

2. *Direct stimulation of the auditory nerve.* The sensational conclusion that the cochlea is not essential for obtaining an auditory sensation was drawn independently by Fano and Massini (1891) and by Ewald (1892). They based their opinion on the positive reactions on sound by pigeons with removed hearing organs. The conclusion was severely criticized by Matte (1894), Bernstein (1895), Strehl (1895), and Kuttner (1896), and defended by Ewald (1895) and Wundt (1895). Apparently, the reactions were due to stimulation of the tactile sense by the strong sounds used.

On the basis of these two findings, Wundt (1893) proposed a description of the hearing process by which he tried to explain most of the tone-perception data known to him. On the one hand, he accepted, just as in 1880, Hypothesis I as an explanation of the ear's frequency-analyzing power, whereas he also maintained Helmholtz's interpretation of combination tones (Hypothesis III). On the other hand, he assumed that bone conduction, which bypasses the cochlea, provides a direct, perhaps much weaker, stimulation of the auditory nerve. The second mode of stimulation could account for König's beat tones and the beats of mistuned consonances. As we saw already, he assumed that the periodicity of sound waves is preserved in the auditory nerve as a basis of pitch (Hypothesis IV).

Although Wundt's hearing theory could explain interruption tones and beats of mistuned consonances much better than the theory of Helmholtz did, its influence was small. This lack of acceptance may have been due to Wundt's emphasis on the dubious experiments on the ability of the auditory nerve to react on sound directly. Also the fact that Wundt did not work out his theory in detail must have been an important factor.

Only the assumption that periodic sound waves give rise to synchronous nerve impulses as a basis for pitch perception (Hypothesis IV) was accepted by some other investigators. The most famous promotor of this view was Rutherford in a lecture presented in 1886. He compared the organ of hearing with a telephone, neglecting frequency analysis. Apparently independent of Wundt, he too pointed to the "muscle tone", investigated by himself, as an argument in favour of the periodicity hypothesis. He published some years later a more extensive survey of his theory in which he no longer excluded the possibility of frequency analysis in the cochlea before the auditory nerves are stimulated (Rutherford, 1898). As he emphasized, by assuming that periodicity is preserved, tonal consonance can be explained by the perception of frequency ratio in the higher auditory centres. Several other investigators who considered frequency ratio as the origin of consonance by implication shared the view that synchronous nerve impulses occur (Lipps, 1885; Polak, 1900; Adler, 1902).

LATER DEVELOPMENTS

We saw in the preceding sections that the beginning of modern tone-perception theory was mainly governed by four hypotheses, two of which were considered as contradictory (II and IV). In fact, the whole later development in physiological knowledge and theoretical insight can be regarded as an elaboration of these hypotheses. A brief description of this development is given in the following sections.

The resonance principle (Hypothesis I)

Helmholtz stated in the first hypothesis that the inner ear must be considered as a large set of resonators tuned to different frequencies. The discussion about the value of his hearing theory concerned mainly the question whether or not this assumption was correct. The view that fibres of 0.5 mm length should be tuned to low frequencies did not sound very credible and we may suppose that many agreed with Stumpf's statement: "It remains wonderful, however, that so small particles can resonate even on the lowest tones that we produce by strings of enormous size and by which we can bring into resonance only strings of the same size." (Stumpf, 1890, p. 92). So it is understandable that it was criticized severely, both on anatomical and physical grounds. Some investigators tried to save the resonance hypothesis by supposing that the resonators must be sought in other structures of the cochlea: the hair cells (Baer, 1872; Hermann, 1894; Myers, 1904; Specht, 1926) or the tectorial membrane (Kishi, 1907; Shambaugh, 1907, 1909, 1911; Leiri, 1932). Others, however, rejected the resonance hypothesis entirely, proposing new hearing theories in which the frequency-analyzing power of the hearing organ was approached in quite a different way (Meyer, 1896, 1898, 1899, 1907; Ewald, 1899, 1903; Wrightson, 1918; and many others).

In the meantime, attempts were made at verifying Helmholtz's assumption that high tones are perceived at the basal and low tones at the apical end of the cochlea. Examination of the hearing organs of some people with hearing defects did not give conclusive results (Moos and Steinbrügge, 1881; Stepanow, 1886), but experiments with animals were more successful. Baginsky (1883) observed that destruction of the apical part of the cochlea of dogs resulted in a hearing loss for low tones; using guinea pigs, Corradi (1891) came to the same conclusion. Yoshii (1909) found that prolonged auditory exposure of some guinea pigs to intensive tones of different frequencies produced local damage of the organ of Corti; the damaged place shifted to the stapes for increasing frequency. Held and Kleinknecht (1927) succeeded in producing a local cochlear lesion in the basal turn resulting in a loss of hearing around 5300 cps.

Since that time more-systematical investigations about the localization of high and low tones in the cochlea have been carried out. These studies include: post-mortem examinations of the hearing organs of persons with hearing defects (*e.g.* Ciocco, 1934; Crowe *et al.*, 1934; Oda, 1938), the effect of local injuries on the electric responses of the cochlea of cats and guinea pigs (*e.g.* Stevens *et al.*, 1935; Walzl and Bordley, 1942; Schuknecht, 1960), the effect of stimulation deafness on the cochlea (*e.g.* Kemp, 1935, 1936; Smith, 1947; Smith and Wever, 1949; Davis *et al.*, 1953), and electric responses of the cochlea during acoustic stimulation (*e.g.* Culler, 1935; Culler *et al.*, 1937, 1943). On the basis of these different methods, the fact is now well established that the stimulated region of the basilar membrane shifts for decreasing frequency from the basal to the apical end. In one respect, however, the experimental evidence deviates from Helmholtz's conception: much broader regions are stimulated by a simple tone, particularly at low frequencies, than he had presumed.

Psychophysical experiments also did not support the view that the inner ear acts as a large set of sharply tuned resonators. Wegel and Lane published in 1924 their study on masking, which is defined as the threshold shift of one simple tone that is caused by another simultaneously presented tone. It appeared, and was confirmed by many later measurements, that the masking pattern extends over a much larger frequency range than would be expected on the basis of sharp resonators. Therefore, following Roaf (1922) and Fletcher (1923), Wegel and Lane suggested that the resonators are coupled to a much higher degree than was assumed by Helmholtz.

The only way, however, to obtain a conclusive answer to the question of how the basilar membrane reacts to tones is to open the cochlea and to observe the vibration pattern along the membrane. Ewald (1903) and Wilkinson (Wilkinson and Gray, 1924) constructed mechanical models of the inner ear, but these models could be considered as illustrations of the theoretical presuppositions of their makers rather than as demonstrations of the actual situation in the cochlea. In this respect, the cochlea-model experiments by von Békésy (1928) were an important step forward, because he adapted the properties of his model carefully to the hearing organ. He concluded on the basis of these experiments that the vibrations of the basilar membrane are highly damped and may be regarded as traveling waves with an envelope having a rather flat maximum that moves from the basal to the apical region with decreasing frequency. Moreover, he was able to support these conclusions with a few direct observations of the vibration pattern of the basilar membrane.

Von Békésy pursued this important study with several others on the same subject (*e.g.* von Békésy, 1941, 1942, 1943, 1944a, 1947). As the most significant result of these investigations we may consider his success in measuring the resonance curves at six positions along the basilar membrane. These curves showed definitely that the conception of a large set of sharply tuned resonators must be abandoned. Instead of a sharply defined region, a rather broad region of the basilar membrane was put into vibration by simple tones, in which the position of the maximum along the membrane was a function of frequency.

On the basis of von Békésy's experimental data, several investigators developed mathematical theories of the mechanics of the cochlea (Zwislocki, 1948, 1950, 1953; Peterson and Bogert, 1950; Ranke, 1950; Fletcher, 1951). Both mechanical (Diestel, 1954; Tonndorf, 1958, 1959, 1960, 1962) and electric (Bogert, 1951; Oetinger and Hanser, 1961; Wansdronk, 1961) models were designed to study more extensively the operation of the cochlea as a function of the physical parameters of the sound.

Von Békésy's direct observations of the vibration pattern of the basilar membrane were carried out at amplitudes corresponding to very high sound-pressure levels of about 140 dB. His arguments advanced in favour of the point of view that the resonance curves measured at this level are representative for normal listening levels may be criticized. This reservation concerns also the mathematical theories based on von Békésy's data.

Investigations on the responses of single nerve fibres to different frequencies confirm the resonance principle, but they don't give conclusive results about the steepness of the slopes of the resonance curves either. Galambos and Davis (1943) found for secondary neurons steep slopes of the threshold-intensity curve as a function of frequency. Tasaki (1954) could confirm for primary neurons the extreme steepness of the curves on the high-frequency side but the low-frequency side was much less steep (see also Kiang, 1965).

Place pitch (*Hypothesis II*)

Helmholtz's assumption that a particular pitch corresponds to each nerve fibre, implying that pitch is correlated to place along the basilar membrane, was criticized by Wundt (1880), as previously discussed. However, this assumption was accepted widely, although it remained difficult to understand that only one or two nerve fibres are stimulated by a simple tone. Stumpf (1890, p. 115) tried to avoid this consequence by supposing that stimulating a nerve fibre does not always give rise to exactly the same pitch, but that this pitch is variable within narrow limits. In his view, a simple tone activates several resonators and nerve fibres at the same time but, owing to the small variability, they all produce the same pitch.

This problem was confronted much more carefully by Gray (1900). He proposed a modification of Helmholtz's conception that was based upon an analogy between the basilar membrane and the skin. From the fact that one experiences a relatively localized sensation upon pressing a small object firmly on the skin, he concluded that in hearing only the place of maximal amplitude of vibration will lead to pitch sensation. This modification was widely adopted by other investigators in the field of audition.

The experiments by von Békésy, mentioned in the preceding section, led to a new discussion on the subject. His finding that the maxima of the vibration patterns along the basilar membrane are not sharp, particularly for low frequencies, raised the question of how this fact may be compatible with the ear's ability to detect very small frequency differences. Von Békésy himself gave much attention to this problem and sought to explain it by sharpening processes in the neural pathway (1944b). He constructed a large and simplified model of the cochlea which used the tactile sense of the arm to simulate the organ of Corti and the auditory pathway (von Békésy, 1955, 1957). It appeared that even in the case of a very broad maximum of the pattern of vibration, only a small section was felt subjectively to vibrate. The correspondence between hearing and the observations with his model was considered by von Békésy as a strong support of the view that the place of maximal stimulation along the basilar membrane corresponds to pitch. The problem of how the high sensitivity for frequency differences has to be reconciled with the broad tuning of the cochlear analyzing mechanism also was discussed by some other investigators (Huggins and Licklider, 1951; Huggins, 1952; Pimonow, 1959).

The origin of combination tones (*Hypothesis III*)

We saw above that Helmholtz explained combination tones by nonlinear distortion in the hearing organ. Although this explanation was criticized by several investigators at that time (see Chapter 3), it was ultimately generally accepted.

The only point in question was where the distortion had to be localized. Helmholtz himself considered the tympanic membrane and the joint between malleus and incus as sources of combination tones. Dennert (1887), however, found that patients without tympanic membrane, malleus and incus also heard combination tones. Preyer (1889) came to the opposite conclusion, but Dennert's observation was confirmed by Bingham (1907) and Schaefer (1913). This suggests that combination tones have their origin in the inner ear, and this view was accepted by others (Schaefer, 1899; Waetzmann, 1907; Peterson, 1910).

More recently, the subject was studied by von Békésy (1934). He examined the sound radiated from the tympanic membrane outwards during stimulation with two simultaneous tones. It appeared that combination tones are not due to nonlinear vibration of this membrane. Furthermore, he discovered that the introduction of a negative or positive static pressure into the external meatus changed the loudness of difference tones. This would imply that these tones are produced in the middle ear.

Another way to investigate the origin of combination tones is by measuring cochlear microphonics. At low sound-pressure levels this electric potential is a true replica of the sound wave, but at higher levels the potential is distorted. Wever *et al.* (1940) analyzed the cochlear microphonics of guinea pigs, evoked by two simultaneous tones. In one condition, the sound was presented in a normal way, whereas in a second condition the stapes was stimulated mechanically after removal of the tympanic membrane, malleus, and incus. From the fact that in both cases the frequency analysis of the cochlear microphonics resulted in quite similar curves, they concluded that the combination tones arise mainly, if not wholly, in structures beyond the stapes. This conclusion was supported by experiments in which the static pressure in the middle ear was changed (Wever *et al.*, 1941), a result which contradicted von Békésy's findings just mentioned. Further investigations, in which the cochlear microphonics were compared with the sound radiated from the cochlea at the round window, suggested that the main source of combination tones must be sought in the sensory processes, where the microphonic potential is evoked, and not in the mechanical part of the inner ear (Wever and Lawrence, 1954).

Periodicity pitch (Hypothesis IV)

The last assumption about the way in which tones are perceived concerns the synchronism of nerve impulses as a basis for pitch. This view, promoted by Wundt, Rutherford and some others, was strongly opposed by the adherents of Helmholtz's resonance theory. One reason for this opposition was that periodicity pitch was mostly set against both the resonance principle (Hypothesis I) and place pitch (Hypothesis II) instead of against the latter assumption alone (so by Boring, 1926). Since most workers in the field of audition were not prepared to give up the frequency-analyzing power of the cochlea, they also rejected the possibility that pitch may be related to periodicity. In this respect, Köhler (1915) has to be considered as an exception.

The second reason was that there was no experimental evidence available indicating that the auditory nerve is able, indeed, to transmit nerve impulses at a rate corresponding to high frequencies. Fletcher's opinion (1923) that this rate cannot exceed about 50 pps was shared by many of his contemporaries.

This situation altered after about 1930. In that year, Wever and Bray (1930a, 1930b) published the results of some experiments on the limit of synchronous nerve impulses in the auditory nerve of the cat. They found with the aid of electronic amplifiers an upper limit of 5200 cps. Later observations (*e.g.* Davis *et al.*, 1934; Davis, 1935; Derbyshire and Davis, 1935) strongly suggested, however, that Wever and Bray had measured a mixture of cochlear microphonics and nerve impulses, and that synchronism of the impulses is not preserved beyond 3000-4000 cps.

This does not mean, however, that a single nerve fibre should be able to transmit this large number of impulses per second. The shortest refractory periods measured are about 0.5 msec, implying an absolute upper limit of about 2000 pps for the single fibre. (This value of 0.5 msec agrees remarkably well with the 0.6-0.7 msec found by Bernstein a century ago.) The short refractory period only holds immediately after the beginning of stimulation, but increases for longer duration times. Moreover, the excitability of a nerve fibre decreases considerably for continuing stimulation.

The observation that synchronism is maintained up to high frequencies can be explained by the fact that the response of a group of nerve fibres is involved. The reasoning is as follows. Suppose that in a particular nerve fibre not every cycle of a continuous tone initiates an impulse, but only every n-th cycle. If different fibres react on different cycles, the frequency of the tone will be preserved up to a much higher limit by a group than by a single fibre. This plausible explanation, presented here in a simplified form, was proposed by Wever and Bray (1930c) under the name "volley principle", although the idea had already been suggested by Troland (1929, 1930). The underlying condition that nerve impulses are evoked at a particular point in the sinusoidal wave has been proved experimentally by Galambos and Davis (1943). They found in some cases a maximum variability of about 0.25 msec in the point where the nerve impulse arises, pointing to about 4000 cps as the upper limit of synchronism.

It will be evident that these findings led to a re-birth of the conception that pitch is based on the periodicity of nerve impulses. In particular, Wever has to be considered as the promotor of this view (*e.g.* Wever, 1933, 1949, 1951). He tried to reconcile place pitch with periodicity pitch by assuming that the pitch of low tones (below about 400 cps), leading to broad vibration patterns, is solely based on the repetition rate of nerve impulses, and the pitch of high tones (above about 5000 cps) solely on the maximum of the vibration pattern along the basilar membrane. In the frequency range between 400 and 5000 cps both principles were regarded as complementary.

Although this hearing theory has been accepted by many investigators nowadays (see *e.g.* Cremer, 1951; Fletcher, 1953), it was criticized recently by von Békésy, mainly on the basis of his analogy between the ear and the skin. Some of his arguments against the view that periodic nerve impulses may be a cue for pitch are discussed in the next chapter.

REFERENCES

Adler: Eine Rhythmus-Theorie des Hörens, *Z. Ohrenheilk. 41*, 143-152 (1902).

Baer, O.: *Ueber das Verhältniss des heutigen Standpunktes der Anatomie des Corti'schen Organs zur Theorie der Tonempfindungen*, Dissertation. University of Breslau (1872).

Baginsky, B.: Die Function der Gehörschnecke, *Arch. pathol. Anat. Physiol. klin. Med. 94*, 65-85 (1883).

Békésy, G. von: Zur Theorie des Hörens; Die Schwingungsform der Basilarmembran, *Phys. Z. 29*, 793-810 (1928). [English translation, G. von Békésy: *Experiments in Hearing*, McGraw-Hill Book Company, Inc., New York, pp. 404-429 (1960).]

Békésy, G. von: Über die nichtlinearen Verzerrungen des Ohres, *Ann. Phys. 20*, 809-827 (1934). [*Experiments in Hearing*, pp. 332-344.]

Békésy, G. von: Über die Elastizität der Schneckentrennwand des Ohres, *Akust. Z. 6*, 265-278 (1941). [*Experiments in Hearing*, pp. 469-480.]

Békésy, G. von: Über die Schwingungen der Schneckentrennwand beim Präparat und Ohrenmodell, *Akust. Z. 7*, 173-186 (1942). [*Experiments in Hearing*, pp. 429-446).]

Békésy, G. von: Über die Resonanzkurve und die Abklingzeit der verschiedenen Stellen der Schneckentrennwand, *Akust. Z. 8*, 66-76 (1943). [*Experiments in Hearing*, pp. 446-460.]

Békésy, G. von: Über die mechanische Frequenzanalyse in der Schnecke verschiedener Tiere, *Akust. Z. 9*, 3-11 (1944a). [*Experiments in Hearing*, pp. 500-510.]

Békésy, G. von: Über die Frequenzauflösung in der menschlichen Schnecke, *Acta Oto-Laryngol. 32*, 60-84 (1944b). [*Experiments in Hearing*, pp. 535-539.]

Békésy, G. von: The Variation of Phase along the Basilar Membrane with Sinusoidal Vibrations, *J. Acoust. Soc. Am. 19*, 452-460 (1947). [*Experiments in Hearing*, pp. 460-469.]

Békésy, G. von: Human Skin Perception of Traveling Waves Similar to those on the Cochlea, *J. Acoust. Soc. Am. 27*, 830-841 (1955). [*Experiments in Hearing*, pp. 547-568.]

Békésy, G. von: Sensations on the Skin Similar to Directional Hearing, Beats, and Harmonics of the Ear, *J. Acoust. Soc. Am. 29*, 489-501 (1957). [*Experiments in Hearing*, pp. 568-590.]

Bernstein, J.: *Untersuchungen über den Erregungsvorgang im Nerven- und Muskelsysteme*, Carl Winter's Universitätsbuchhandlung, Heidelberg (1871).

Bernstein, J.: Ueber die Höhe des Muskeltones bei elektrischer und chemischer Reizung, *Arch. ges. Physiol. 11*, 191-196 (1875).

Bernstein, J.: Ueber die specifische Energie des Hörnerven, die Wahrnehmung binauraler (diotischer) Schwebungen und die Beziehungen der Hörfunktion zur statischen Funktion des Ohrlabyrinths, *Arch. ges. Physiol. 57*, 475-494 (1894).

Bernstein, J.: Ueber das angebliche Hören laby-

rinthloser Tauben, *Arch. ges. Physiol. 61,* 113-122 (1895).
Bingham, W.V.D.: The Rôle of the Tympanic Mechanism in Audition, *Psychol. Rev. 14,* 229-243 (1907).
Bogert, B.P.: Determination of the Effects of Dissipation in the Cochlear Partition by Means of a Network Representing the Basilar Membrane, *J. Acoust. Soc. Am. 23,* 151-154 (1951).
Boring, E.G.: Auditory Theory with Special Reference to Intensity, Volume, and Localization, *Am. J. Psychol. 37,* 157-188 (1926).
Ciocco, A.: A Statistical Approach to the Problem of Tone Localization in the Human Cochlea, *Hum. Biol. 6,* 714-721 (1934).
Corradi, C.: Ueber die functionelle Wichtigkeit der Schnecke, *Arch. Ohrenheilk. 32,* 1-14 (1891).
Corti, A.: Recherches sur l'organe de l'ouïe des mammifères, *Z. wiss. Zool. 3,* 109-169 (1851).
Cremer, L.: Über die ungelösten Probleme in der Theorie der Tonempfindungen, *Acustica 1,* 83-96 (1951).
Cross, Ch. R., and H.M. Goodwin: Some Considerations Regarding Helmholtz's Theory of Consonance, *Proc. Am. Acad. Arts and Sciences 19, whole Ser. 27,* 1-12 (1893).
Crowe, S.J., S.R. Guild, and L.M. Polvogt: Observations on the Pathology of High-Tone Deafness, *Bull. Johns Hopkins Hosp. 54,* 315-379 (1934).
Culler, E.A.: In Symposium: Is There Localization in the Cochlea for Low Tones?, *Ann. Otol. Rhinol. Laryngol. 44,* 807-813 (1935).
Culler, E.A., J.D. Coakley, K. Lowy, and N. Gross: A Revised Frequency-Map of the Guinea-Pig Cochlea, *Am. J. Psychol. 56,* 475-500 (1943).
Culler, E.A., J. Willmann, and F.A. Mettler: Mapping the Cochlea, *Am. J. Physiol. 119,* 292 (1937).
Davis, H.: The Electrical Phenomena of the Cochlea and the Auditory Nerve, *J. Acoust. Soc. Am. 6,* 205-215 (1935).
Davis, H., A.J. Derbyshire, M.H. Lurie, and L.J. Saul: The Electric Response of the Cochlea, *Am. J. Physiol. 107,* 311-332 (1934).
Davis, H., and Associates: Acoustic Trauma in the Guinea Pig, *J. Acoust. Soc. Am. 25,* 1180-1189 (1953).
Dennert, H.: Akustisch-physiologische Untersuchungen, *Arch. Ohrenheilk. 24,* 171-184 (1887).
Derbyshire, A.J., and H. Davis: The Action Potentials of the Auditory Nerve, *Am. J. Physiol. 113,* 476-504 (1935).
Diestel, H.G.: Akustische Messungen an einem mechanischen Modell des Innenohres, *Acustica 4,* 489-499 (1954).
Dove, H.W.: *Repertorium der Physik,* Vol. 3, Verlag Veit & Comp., Berlin, pp. 404, 405 (1839).
Ewald, J.R.: *Physiologische Untersuchungen ueber das Endorgan des Nervus Octavus,* Verlag J.F. Bergmann, Wiesbaden (1892).
Ewald, J.R.: Die centrale Entstehung von Schwebungen zweier monotisch gehörten Töne, *Arch. ges. Physiol. 57,* 80-88 (1894).
Ewald, J.R.: Zur Physiologie des Labyrinths. 3. Mittheilung. Das Hören der labyrinthlosen Tauben, *Arch. ges. Physiol. 59,* 258-275 (1895).
Ewald, J.R.: Zur Physiologie des Labyrinths. VI. Eine neue Hörtheorie, *Arch. ges. Physiol. 76,* 147-188 (1899).
Ewald, J.R.: Zur Physiologie des Labyrinths. VII. Die Erzeugung von Schallbildern in der Camera acustica, *Arch. ges. Physiol. 93,* 485-500 (1903).
Fano, G., and G. Masini: Beitrag zur Physiologie des inneren Ohres, *Centralblatt für Physiol. 4,* 787-788 (1891).
Fletcher, H.: Physical Measurements of Audition and their Bearing on the Theory of Hearing, *Bell Syst. Techn. J. 2,* 145-180 (1923).
Fletcher, H.: On the Dynamics of the Cochlea, *J. Acoust. Soc. Am. 23,* 637-645 (1951).
Fletcher, H.: *Speech and Hearing in Communication,* D. Van Nostrand Company, Inc., New York (1953).
Galambos, R., and H. Davis: The Response of Single Auditory-Nerve Fibers to Acoustic

Stimulation, *J. Neurophysiol. 6*, 39-57 (1943).
Gray, A.A.: On a Modification of the Helmholtz Theory of Hearing, *J. Anat. Physiol. 34*, 324-350 (1900).
Hasse, C.: Die Schnecke der Vögel, *Z. wiss. Zool. 17*, 56-104 (1867).
Held, H., and F. Kleinknecht: Die lokale Entspannung der Basilarmembran und ihre Hörlücken, *Arch. ges. Physiol. 216*, 1-31 (1927).
Helmholtz, H.L.F. von: *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*, F. Vieweg & Sohn, Braunschweig, 1st Ed. (1863).
Helmholtz, H.L.F. von: Ueber den Muskelton, *Verhandl. Naturhist.-Medicin. Vereins Heidelberg 4*, 88-90 (1868).
Hensen, V.: Zur Morphologie der Schnecke des Menschen und der Säugethiere, *Z. wiss. Zool. 13*, 481-512 (1863).
Hermann, L.: Beiträge zur Lehre von der Klangwahrnehmung, *Arch. ges. Physiol. 56*, 467-499 (1894).
Huggins, W.H.: A Phase Principle for Complex-Frequency Analysis and its Implications in Auditory Theory, *J. Acoust. Soc. Am. 24*, 582-589 (1952).
Huggins, W.H., and J.C.R. Licklider: Place Mechanisms of Auditory Frequency Analysis, *J. Acoust. Soc. Am. 23*, 290-299 (1951).
Kemp, E.H.: A Critical Review of Experiments on the Problem of Stimulation Deafness, *Psychol. Bull. 32*, 325-342 (1935).
Kemp, E.H.: An Experimental Investigation of the Problem of Stimulation Deafness, *J. Exptl. Psychol. 19*, 159-171 (1936).
Kiang, N.Y.: Stimulus Coding in the Auditory Nerve and Cochlear Nucleus, *Acta Oto-Laryngol. 59*, 186-200 (1965).
Kishi, K.: Cortische Membran und Tonempfindungstheorie, *Arch. ges. Physiol. 116*, 112-123 (1907).
Köhler, W.: Akustische Untersuchungen. III. *Z. Psychol. 72*, 1-192 (1915).
Kuttner, A.: Die Hörfähigkeit labyrinthloser Tauben, *Arch. ges. Physiol. 64*, 249-261 (1896).
Leiri, F.: Über die Perzeption der musikalischen Töne, *Z. Sinnesphysiol. 63*, 151-184 (1932).
Lipps, Th.: *Psychologische Studien*, Verlag Georg Weiss, Heidelberg, pp. 92-161 (1885).
Mach, E.: Bemerkungen über die Function der Ohrmuschel, *Arch. Ohrenheilk. 9*, 72-76 (1875).
Matte, F.: Experimenteller Beitrag zur Physiologie des Ohrlabyrinthes, *Arch. ges. Physiol. 57*, 437-475 (1894).
Meyer, M.: Über Kombinationstöne und einige hierzu in Beziehung stehende akustische Erscheinungen, *Z. Psychol. Physiol. Sinnesorgane 11*, 177-229 (1896).
Meyer, M.: Zur Theorie der Differenztöne und der Gehörsempfindungen überhaupt, *Z. Psychol. Physiol. Sinnesorgane 16*, 1-34, and *17*, 8-14 (1898). [= *Beitr. Akust. Musikwiss. 2*, 25-65 (1898).]
Meyer, M.: Zur Theorie des Hörens, *Arch. ges. Physiol. 78*, 346-362 (1899).
Meyer, M.: *An Introduction to the Mechanics of the Inner Ear*, Univ. Missouri Stud., Sci. Ser. 2, No. 1 (1907).
Moos, S., and H. Steinbrügge: Ueber Nervenatrophie in der ersten Schneckenwindung. Physiologische und pathologische Bedeutung derselben, *Z. Ohrenheilk. 10*, 1-15 (1881).
Müller, J.: *Handbuch der Physiologie des Menschen*, Vol. 2, Verlag J. Hölscher, Coblenz, Chapter 2 (1840).
Myers, C.S.: Recent Developments of Helmholtz's Theory of Hearing, *Rep. 74th Meeting Brit. Ass. Adv. Science, Cambridge 1904*, 750 (1905).
Oda, D.: Observations on the Pathology of Impaired Hearing for Low Tones, *Laryngoscope 48*, 765-792 (1938).
Oetinger, R., and H. Hauser: Ein elektrischer Kettenleiter zur Untersuchung der mechanischen Schwingungsvorgänge im Innenohr, *Acustica 11*, 161-177 (1961).
Ohm, G.S.: Ueber die Definition des Tones, nebst daran geknüpfter Theorie der Sirene und ähnlicher tonbildender Vorrichtungen, *Ann. Phys. Chem. 59*, 513-565 (1843).
Peterson, J.: Do Combination Tones Originate in

the Middle Ear?, *Psychol. Bull. 7*, 169-173 (1910).
Peterson, L.C., and B.P.Bogert: A Dynamical Theory of the Cochlea, *J. Acoust. Soc. Am. 22*, 369-381 (1950).
Pimonow, L.: Contribution a l'étude du problème de la sélectivité tonale de l'oreille, *Acustica 9*, 345-352 (1959).
Polak, A.J.: *Über Zeiteinheit in Bezug auf Konsonanz, Harmonie und Tonalität*, Verlag Breitkopf & Härtel, Leipzig (1900).
Preyer, W.: Über Combinationstöne, *Verhandl. Physikal. Gesellsch. Berlin 1886, 8*, 15-20 (1889).
Ranke, O.F.: Theory of Operation of the Cochlea: A Contribution to the Hydrodynamics of the Cochlea, *J. Acoust. Soc. Am. 22*, 772-777 (1950).
Roaf, H.E.: The Analysis of Sound Waves by the Cochlea, *Phil. Mag. 6th Ser. 43*, 349-354 (1922).
Rostosky, P.: Ueber binaurale Schwebungen, *Phil. Studien 19*, 557-598 (1902).
Rutherford, W.: A New Theory of Hearing, *J. Anat. Physiol. 21*, 166-168 (1886).
Rutherford, W.: Tone Sensation with Reference to the Function of the Cochlea, *Lancet 76, Vol. 2*, 389-394 (1898).
Schaefer, K.L.: Ein Versuch über die intrakranielle Leitung leisester Töne von Ohr zu Ohr, *Z. Psychol. Physiol. Sinnesorgane 2*, 111-114 (1891).
Schaefer, K.L.: Ist eine cerebrale Entstehung von Schwebungen möglich?, *Z. Psychol. Physiol. Sinnesorgane 4*, 348-350 (1893a).
Schaefer, K.L.: Nochmalige Ablehnung der cerebralen Entstehung von Schwebungen, *Z. Psychol. Physiol. Sinnesorgane 5*, 397-401 (1893b).
Schaefer, K.L.: Beweise gegen Wundt's Theorie von der Interferenz akustischer Erregungen im Centralorgan, *Arch. ges. Physiol. 61*, 544-550 (1895).
Schaefer, K.L.: Eine neue Erklärung der subjectiven Combinationstöne auf Grund der Helmholtz'schen Resonanzhypothese, *Arch. ges. Physiol. 78*, 505-526 (1899).
Schaefer, K.L.: Ueber die intracranielle Fortpflanzung der Töne, insbesondere der tiefen Töne, von Ohr zu Ohr, *Arch. Ohrenheilk. 52*, 151-155 (1901).
Schaefer, K.L.: Über die Wahrnehmbarkeit von Kombinationstönen bei partiellem oder totalem Defekt des Trommelfelles, *Beitr. Anat. Physiol. Pathol. Therapie Ohres, Nase, Halses 6*, 207-218 (1913).
Schuknecht, H.F.: Neuroanatomical Correlates of Auditory Sensitivity and Pitch Discrimination in the Cat, in G.L. Rasmussen and W.F. Windle (Eds.): *Neural Mechanisms of the Auditory and Vestibular Systems*, Charles C. Thomas, Springfield, Chapter 6 (1960).
Scripture, E.W.: Einige Beobachtungen über Schwebungen und Differenztöne, *Phil. Studien 7*, 630-632 (1892).
Scripture, E.W.: Ist eine cerebrale Entstehung von Schwebungen möglich?, *Phil. Studien 8*, 638-640 (1893).
Seebeck, A.: Beiträge zur Physiologie des Gehör- und Gesichtssinnes, *Ann. Phys. Chem. 68*, 449-465 (1846).
Shambaugh, G.E.: A Restudy of the Minute Anatomy of Structures in the Cochlea with Conclusions Bearing on the Solution of the Problem of Tone Perception, *Am. J. Anat. 7*, 245-257 (1907).
Shambaugh, G.E.: Die Membrana tectoria und die Theorie der Tonempfindung, *Z. Ohrenheilk. 59*, 159-168 (1909).
Shambaugh, G.E.: Die Frage der Tonempfindung, *Arch. ges. Physiol. 138*, 155-158 (1911).
Smith, K.R.: The Problem of Stimulation Deafness. II. Histological Changes in the Cochlea as a Function of Tonal Frequency, *J. Exptl. Psychol. 37*, 304-317 (1947).
Smith, K.R., and E.G. Wever: The Problem of Stimulation Deafness. III. The Functional and Histological Effects of a High-Frequency Stimulus, *J. Exptl. Psychol. 39*, 238-241 (1949).
Specht, F.: Zur Physiologie des Hörens, *Arch. Ohren-, Nasen- und Kehlkopfheilk. 114*, 1-80

(1926).
Stepanow, E.M.: Zur Frage über die Function der Cochlea, *Monatsschr. Ohrenheilk. Kehlkopf-, Nasen-, Rachen-Krankh. 20*, 116-124 (1886).
Stevens, S.S., H.Davis, and M.H.Lurie: The Localization of Pitch Perception on the Basilar Membrane, *J. Gen. Psychol. 13*, 297-315 (1935).
Strehl, H.: Beiträge zur Physiologie des inneren Ohres, *Arch. ges. Physiol. 61*, 205-234 (1895).
Stumpf, C.: *Tonpsychologie*, Vol. 2, Verlag S. Hirzel, Leipzig (1890).
Tasaki, I.: Nerve Impulses in Individual Auditory Nerve Fibers of Guinea Pig, *J. Neurophysiol. 17*, 97-122 (1954).
Thompson, S.P.: On Binaural Audition, *Phil. Mag. 5th Ser. 4*, 274-276 (1877).
Thompson, S.P.: Phenomena of Binaural Audition – Part II, *Phil. Mag. 5th Ser. 6*, 383-391 (1878).
Thompson, S.P.: Phenomena of Binaural Audition – Part III, *Phil. Mag. 5th Ser. 12*, 351-355 (1881).
Tonndorf, J.: Harmonic Distortion in Cochlear Models, *J. Acoust. Soc. Am. 30*, 929-937 (1958).
Tonndorf, J.: Beats in Cochlear Models, *J. Acoust. Soc. Am. 31*, 608-619 (1959).
Tonndorf, J.: Dimensional Analysis of Cochlear Models, *J. Acoust. Soc. Am. 32*, 493-497 (1960).
Tonndorf, J.: Time/Frequency Analysis along the Partition of Cochlear Models: A Modified Place Concept, *J. Acoust. Soc. Am. 34*, 1337-1350 (1962).
Troland, L.T.: The Psychophysiology of Auditory Qualities and Attributes, *J. Gen. Psychol. 2*, 28-58 (1929).
Troland, L.T.: Psychophysiological Considerations Relating to the Theory of Hearing, *J. Acoust. Soc. Am. 1*, 301-310 (1930).
Waetzmann, E.: Zur Theorie der Kombinationstöne, *Ann. Phys. 24*, 68-78 (1907).
Walzl, E.M., and J.E. Bordley: The Effect of Small Lesions of the Organ of Corti on Cochlear Potentials, *Am. J. Physiol. 135*, 351-360 (1942).
Wansdronk, C.: *On the Mechanism of Hearing*, Doctoral Dissertation, University of Leiden (1961). [=*Philips Res. Repts Suppl. No. 1* (1962).]
Wegel, R.L., and C.E.Lane: The Auditory Masking of One Pure Tone by Another and its Probable Relation to the Dynamics of the Inner Ear, *Phys. Rev. 23*, 266-285 (1924).
Wever, E.G.: The Physiology of Hearing: The Nature of Response in the Cochlea, *Physiol. Rev. 13*, 400-425 (1933).
Wever, E.G.: *Theory of Hearing*, John Wiley & Sons, Inc., New York (1949).
Wever, E.G.: Some Remarks on the Modern Status of Auditory Theory, *J. Acoust. Soc. Am. 23*, 287-289 (1951).
Wever, E.G., and C.W.Bray: Action Currents in the Auditory Nerve in Response to Acoustical Stimulation, *Proc. Nat. Acad. Science 16*, 344-350 (1930a).
Wever, E.G., and C.W.Bray: The Nature of Acoustic Response: The Relation between Sound Frequency and Frequency of Impulses in the Auditory Nerve, *J. Exptl. Psychol. 13*, 373-387 (1930b).
Wever, E.G., and C.W.Bray; Present Possibilities for Auditory Theory, *Psychol. Rev. 37*, 365-380 (1930c).
Wever, E.G., C.W.Bray, and M.Lawrence: The Origin of Combination Tones, *J. Explt. Psychol. 27*, 217-226 (1940).
Wever, E.G., C.W.Bray, and M.Lawrence: Effect of Middle Ear Pressure upon Distortion, *J. Acoust. Soc. Am. 13*, 182-187 (1941).
Wever, E.G., and M.Lawrence: *Physiological Acoustics*, Princeton University Press, Princeton, N.J., pp. 163-172 (1954).
Wilkinson, G., and A.A.Gray: *The Mechanism of the Cochlea. A Restatement of the Resonance Theory of Hearing*, Macmillan and Co., Ltd., London (1924).
Wrightson, T.: *An Enquiry into the Analytical Mechanism of the Internal Ear*, Macmillan and Co., Ltd., London (1918).
Wundt, W.: *Grundzüge der physiologischen Psychologie*, Vol. 1, Verlag W.Engelmann, Leipzig, 2nd Ed., pp. 315-317 (1880).
Wundt, W.: Ist der Hörnerv direct durch Ton-

schwingungen erregbar?, *Phil. Studien 8*, 641-652 (1893).

Wundt, W.: Zur Frage der Hörfähigkeit labyrinthloser Tauben, *Arch. ges. Physiol. 61*, 339-341 (1895).

Yoshii, U.: Experimentelle Untersuchungen über die Schädigung des Gehörorgans durch Schalleinwirkung, *Z. Ohrenheilk. 58*, 201-251 (1909).

Zwislocki, J.: *Theorie der Schneckenmechanik*, Doctoral Dissertation, University of Zürich (1948). [=*Acta Oto-Laryngol. Suppl. 72* (1948).]

Zwislocki, J.: Theory of the Acoustical Action of the Cochlea, *J. Acoust. Soc. Am. 22*, 778-784 (1950).

Zwislocki, J.: Review of Recent Mathematical Theories of Cochlear Dynamics, *J. Acoust. Soc. Am. 25*, 743-751 (1953).

8

IMPLICATIONS OF THE EXPERIMENTS FOR HEARING THEORY

In this final chapter, the results of the experiments described in Chapters 2 to 6 are discussed in succession in the light of the physiological data presented in the preceding chapter. For convenience, a summary of the most important experimental results is given first.

1. The partials of a complex of tones can be discriminated only when their frequency separation exceeds critical bandwidth.

2. For sound-pressure levels above about 50 dB, two tones give rise to combination tones, predominantly with frequencies below the lower primary tone; these combination tones are apparently due to nonlinear distortion and their detectability threshold is lowest for small frequency differences between the primary tones.

3. Two simple tones within a critical band interfere, resulting in dissonance which is maximal for a frequency difference equal to about a quarter of the critical band; on the basis of this interference, the singular character of complex tones with simple frequency ratios can be explained.

4. For frequency differences exceeding the critical bandwidth, two simple tones with frequency ratio slightly different from a simple ratio may give rise to a beat sensation; these beats cannot be explained solely by nonlinear distortion but seem to be related mainly to periodic variations in waveform.

5. The pitch of complex tones with a fundamental frequency up to about 1400 cps is not based on the frequency of the fundamental but, apparently, on the periodicity of the waveform.

DISCUSSION OF THE EXPERIMENTAL RESULTS

Limit of the ear's frequency-analyzing power

The finding that the partials of a complex of tones can be distinguished only when their frequency separation exceeds critical bandwidth presents new evidence for the basic significance of the latter measure. Moreover, it may contribute to a better understanding of the physiological background of the critical band.

We saw in Chapter 7 that a simple tone gives rise to a vibration pattern along the basilar membrane with a maximum whose location shifts as a function of frequency. We may conclude from this that the cochlea is able to accomplish a frequency analysis. It is attractive to identify the critical bandwidth with the width of the vibration pattern.

In the writer's opinion, this supposition is justified irrespective of the answer to the question whether pitch is determined by the place of maximal displacement or by the rate of nerve impulses. If pitch is related to place, we may suppose that simultaneous tones can be discriminated only when each simple tone produces a separate vibration maximum. If pitch is related to periodicity, the same supposition holds, because it seems likely that periodic impulses synchronous to the frequency of a particular partial of a multitone stimulus are evoked only if cochlear frequency analysis separates this partial sufficiently from adjacent ones. It is difficult to imagine how, when a group of nerve fibres is stimulated by two or more superimposed sinusoidal waves, periodic impulse trains corresponding to each of these waves are derived.

The investigation of the masking pattern of a complex tone of 500 cps (Figure 13) supports the view that critical bandwidth must be identified with the width of the vibration pattern along the basilar membrane, since only the partials with frequency differences exceeding the critical band led to peaks in the masking pattern. The sharpness of the peaks may be attributed to the fact that the probe tone does not stimulate an infinitesimally small part of the basilar membrane but also an area equal to the critical band. For that reason, the threshold of audibility of the probe tone depends upon many factors including the steepness of the slopes of the vibration patterns.

Finally, we may wonder why two simple tones sounding simultaneously are distinguished individually for smaller frequency differences than the partials of a multitone stimulus (Figure 9). On the basis of place pitch this is difficult to understand. Greenwood (1961a, 1961b) has demonstrated that the masking pattern of two-tone stimuli shows two peaks only for frequency differences exceeding the critical band, and we may expect that the same holds for the superimposed vibration patterns along the basilar membrane. The following explanation can be proposed on the basis of periodicity pitch. The difference between two and more tones, respectively, is that in the first case the part of the higher tone's vibration pattern on the high-frequency side of its maximum and the part of the lower tone's vibration pattern on the low-frequency side of its maximum are not "masked" by other tones. It is not improbable that these parts contribute to pitch discrimination by facilitating the detection of the periodicities of each of the two tones, resulting in a somewhat smaller minimum frequency difference than for multitone stimuli. The fact that this improvement is larger at low than at high frequencies supports this explanation, for in the latter case periodicity is no longer preserved.

The origin of combination tones

One of the experimental results in Chapter 3 was that the detectability threshold for combination tones is significantly lower for small than for large frequency differences between the primary tones. From this, the conclusion was drawn that the ear's distortion cannot be represented by a frequency-independent nonlinear characteristic.

This fact throws some light on the origin of combination tones. If they are created in the middle ear, we would expect a much smaller dependence upon frequency difference than actually was found, because the middle-ear's frequency-response curve is rather flat. This suggests that combination tones are created in the inner ear rather than in the middle ear. The frequency-analyzing power of the cochlea explains why tones with a small frequency difference produce much more distortion than tones with a large difference, since in the first case the stimulation patterns overlap much more than in the latter case. So our conclusion is that the main distortion source has to be localized in or beyond the frequency-analyzing mechanism.

As was mentioned in Chapter 7, the same conclusion was drawn by Wever *et al.* on the basis of quite different experiments in which the cochlear microphonics of animals were measured. Wever and Lawrence (1954) even went a step further in their statement that the sensory processes have to be considered as the main source of combination tones. In the writer's opinion, however, we have to be very cautious in the interpretation of the distortion of cochlear microphonics. Combination tones produced by nonlinearities of the mechanics of the cochlea will manifest themselves in this electric potential but, reversely, distortion of this signal does not imply that audible combination tones are produced.

The assumption that combination tones as investigated in psychophysical experiments are caused by nonlinearities in the transition of mechanical vibrations into cochlear microphonics raises the question of how this distortion may lead to corresponding tone sensations. This problem applies both to place pitch and periodicity pitch. If pitch is based on the place of maximal vibration, it is essential for hearing a combination tone that the corresponding place of the basilar membrane is stimulated. Then the question may be asked of how this can be accomplished by sensory processes of hair cells at a distant place of the cochlea. The ascertainment by Six (1956) that cochlear microphonics corresponding to combination tones have their maximum at the same place as the primary tones, contradicts this possibility. If, on the other hand, pitch is based on the periodicity of nerve impulses, the problem arises how impulses that are synchronous with the frequency of combination tones can be initiated when the waveform of the cochlear microphonics is flattened (the common type of distortion). Indeed, we may expect that in this case, since the impulses are evoked at a particular moment of the cycle, peak limiting does not influence this moment. For aural harmonics, the same objection has been formulated by Newman (1951). Moreover, Wever *et al.* (1940) have investigated that the amplitude of combination tones in the cochlear microphonics due to cubic distortion is proportional to the third power of the amplitude of the primary tones, whereas Zwicker (1955) has shown that the increase of sensation level of these combination tones does not follow this rule.

We conclude from this that, although the nonlinearity of the transition of mechanical vibrations into cochlear microphonics may be larger than the nonlinearity in the mechanical part of the cochlea itself, audible combination tones have their origin in the mechanical part.

Tonal consonance and critical bandwidth

The experiments treated in Chapter 4 showed that critical bandwidth plays an important rôle in our perception of tone intervals. Apparently, simple tones with a frequency difference larger than this bandwidth do not interfere to such a degree as for smaller frequency differences.

This result supports the view that critical bandwidth has to be understood as the width of the vibration pattern along the basilar membrane which is produced by simple tones. Independent of the question whether pitch is related to place or to periodicity, we have to expect that tones within the same critical band interfere strongly.

The fact that maximal dissonance occurs at about a quarter of critical bandwidth does not contradict this conclusion. This interval width must be regarded as a result of various factors of which the inability of the ear to follow rapid amplitude variations is an important one.

The origin of beats of mistuned consonances

We suggested in Chapter 5 that, although combination tones do play a rôle, the beats of mistuned consonances are related mainly to periodic variations in waveform. The consequence of this view is that two simple tones do not only interfere for frequency differences less than the critical bandwidth but also for much larger differences. Our discussion about possible physiological mechanisms underlying these conclusions starts with this consequence.

We saw in Chapter 7 that simple tones stimulate a rather broad area of the basilar membrane. Above, it was suggested that critical bandwidth might be identified with the width of this vibration pattern. Apparently, we have to make a distinction between the width of the vibration pattern (comparable with the bandwidth of an electric band-pass filter) and the steepness of the slopes of the vibration pattern outside the critical band. The interference of tones with a large frequency difference can be explained by assuming that these slopes are rather flat. This idea has been discussed by Goldstein (1965) in a study in which he has given extensive consideration to the slope of the stimulation pattern.

Figure 58 illustrates how tones may interfere. The vibration pattern of the lower tone is represented by curve A and of the higher tone by curve B. It is tempting to suppose that maximal interference or maximal beat sensation occurs for a particular interval when the peak of the vibration pattern

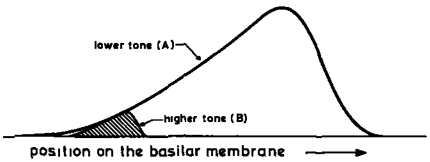


FIGURE 58. Illustration of the interference of the vibration patterns along the basilar membrane produced by two simple tones.

corresponding to the higher tone equals the amplitude of vibration at the same position resulting from the lower tone (condition as represented in the graph). This supposition is corroborated by the observation that, in the low-frequency range, best beats seem to occur when the sound-pressure level of the higher tone is somewhat lower than required for hearing this tone individually. For higher frequencies, however, this rule does not hold always, so it appears to be unjustified to consider the sensation level of the higher tone for best beats as a direct measure of the vibration pattern of the lower tone. The least we can say is that this level, in view of its monotonic decline for increasing frequency difference, is related to that vibration pattern, although probably in a more complex way.

This exposition shows that two tones with a frequency difference exceeding critical bandwidth may give rise to interference of their corresponding vibration patterns along the basilar membrane, but it does not explain what property of the interfering vibration patterns produces beats for mistuned consonances. It is reasonable to suggest that the beats have their origin in the neighbourhood of the peak of the vibration pattern of the higher tone. Recent experiments by Goldstein (1965), indicating that the beats are more easily masked by a band of noise with a centre frequency equal to the frequency of the higher tone than for other centre frequencies, support this view. The question to be answered now is how slow waveform variations become audible as beats.

In Figure 59, the waveforms of two superimposed sinusoids with frequency ratio of 1:2, 1:3, and 2:3, respectively, are reproduced for 4 different phase relations. We may suppose that the interference of tones in the cochlea results in similar displacements of the basilar membrane as a function of time. In the case of slightly mistuned consonances, the waveform changes continuously, corresponding with a gradual increase or decrease of φ.

In a study on similarities between the skin and the organ of hearing, von Békésy (1957) considered the question whether the beats might be related to the periodic variation in the maximum of the waveform's envelope. Although he judged this explanation as inadequate, it was promoted more recently by Schügerl (1964). On the basis of a mathematical treatment of the relative magnitude of the envelope variation as a function of φ for different values of the amplitude ratio and of the frequency ratio $m:n$ of the individual tones, he showed that this variation decreases for increasing $m+n$ and concluded that the limit of $m+n$ for which beats are audible is determined by the just noticeable amplitude variation of a simple tone.

There are, however, some objections against

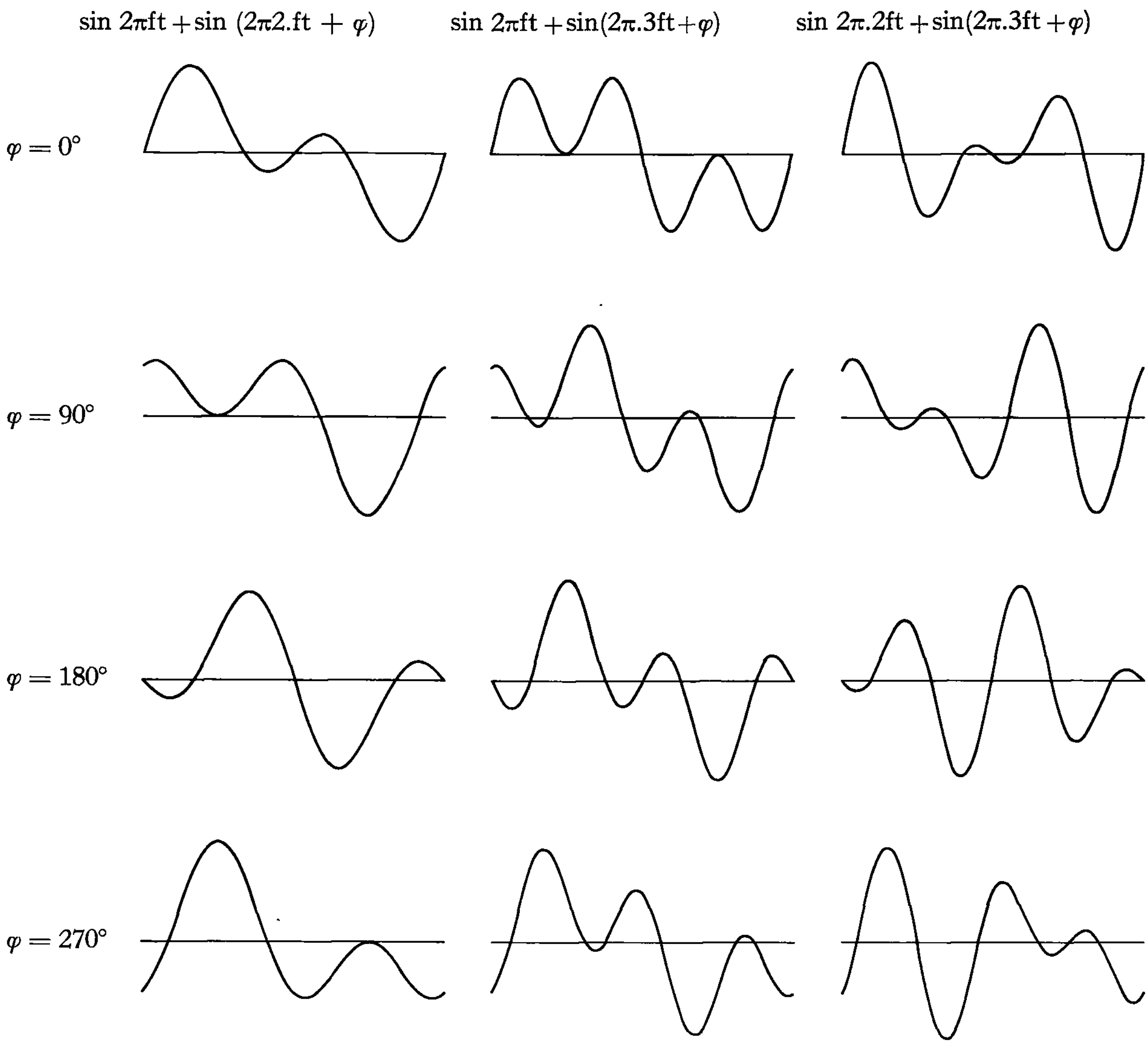


FIGURE 59. Waveforms of two superimposed sinusoids.

this explanation: (1) Schügerl's calculations show that the variation in envelope maximum of the superimposed sinusoids depends much more upon the amplitude ratio of the individual sinusoids for frequency ratios as 1:2, 1:3, and 1:4 than for ratios as 3:5, 4:5, and 5:6. This would imply that in the first case the sound-pressure level of the higher tone for best beats has much narrower limits than in the latter case. A check of this consequence by computing, on the basis of the data underlying Figure 47, the mean difference between the two determinations of the sound-pressure levels for best beats showed, on the contrary, a mean difference of 4.5 dB for 1:2, 1:3, and 1:4 together, and of 2.75 dB for 3:5, 4:5, and 5:6 together. (2) It is difficult to see why beats occur for many more frequency ratios at low than at high frequencies. (3) The sweep-tone effect remains unexplained. These objections suggest that, although variations in the maximum of the waveform's envelope may play a rôle in hearing beats, its main origin has to be sought elsewhere.

The reason why von Békésy did not consider this explanation as an adequate one had a different ground. He observed, using his cochlear model in contact with the surface of the arm, that the beats of 1:2, slightly mistuned, manifest themselves mostly as a change in the place of stimulation along the arm. This place shifted periodically over nearly the whole range between the places where the "tones" were localized when presented separately. Similar results were obtained for frequency ratio 1:3, with the difference that in that case the intermediate stages were more difficult to observe than for 1:2, giving the impression of a jump from one position to the other. Upon presenting tone intervals with the same frequency ratios to the ear, for instance the mistuned interval 200+601 cps, the same phenomena were observed as with the model. This correspondence was considered by von Békésy as an affirmation that in both cases the same mechanism is involved. In his opinion, it seems clear that the periodicity of the nerve impulses does not play an important part in the production of beats.

Although this reasoning appears rather attractive, I prefer an alternative explanation of the beats which is based on the assumption that pitch is related to the periodicity of nerve impulses. The arguments are: (1) The similarity in beat sensation between the ear and the skin seems to be smaller than was suggested by von Békésy in view of our finding that the pitch of the sweep tone for primary tones of 200 and 600 cps does not shift between the primary tones but over a rather small range beyond the higher tone (600-730 cps). (2) It is difficult to see how mistuned consonances give rise to periodic shifts of the maximum of the vibration pattern along the basilar membrane. (3) Also in the case that the lower tone is presented to one ear and the higher to the other, beats can be observed; in this condition, cochlear interactions are excluded.

Our explanation is as follows. Mistuned consonances give rise to interference of the vibration patterns along the basilar membrane and it is likely that the membrane's displacement as a function of time at a particular place can be compared with waveforms as represented in Figure 59. We saw in Chapter 7 that, for simple tones below 3000-4000 cps, nerve impulses are evoked at a particular moment of the sinusoidal wave, suggesting that these impulses are evoked when the displacement of the basilar membrane passes a critical value. The assumption that nerve impulses are triggered by the displacement of the basilar membrane implies that slow variations in waveform, corresponding to beats, result in slow variations in the time pattern of the impulses. The beat sensation may have its origin in these periodic variations.

This conception is supported by two experimental findings. In the first place, it was found that beats are audible for many more different frequency ratios at low than at high frequencies (Figures 47-51). This result is to be expected in view of the fact that synchronism of the nerve impulses decreases for increasing frequency. The reason why the beats of mistuned consonances are more distinct for simpler frequency ratios must be sought in the fact that the waveform variations of the superimposed sinusoids are more prominent for the simpler ratios.

The second point concerns the sweep-tone effect. If pitch is related to the time interval between successive nerve impulses, we have to take into account that the superposition of two sinusoidal vibrations affects the duration of these intervals. On the assumption that impulses are evoked when the displacement of the basilar membrane passes a critical value, we have made an electric model of this activity. This model showed that, for two superimposed sinusoids with equal amplitudes and with frequency ratio 1:3, slightly mistuned, the time interval between successive pulses varies periodically. Although the time pattern of the pulses is somewhat more complex, it can be considered mainly as consisting of two pulse trains with equal time intervals, and with a time delay between the trains shifting periodically in a saw-tooth like manner; the reciprocal of the period of these trains corresponds with the lower tone, whereas the reciprocal of the varying shorter interval between the impulses of one train with respect to the other corresponds roughly with the sweep tone.

It is of interest that phenomena similar to the sweep-tone effect have been investigated by some authors in experiments in which not two simple tones but two identical pulse trains with variable time delay were presented to the ear (Thurlow and Small, 1955; Thurlow, 1957, 1963; Jenkins, 1961; Small and McClellan, 1963; Nordmark, 1963; Kylstra, 1964). Some divergences between their results illustrate that in this case the extra tone, also corresponding to the reciprocal of the shorter time interval between the pulses of one train with respect to the other, is difficult to observe, just as in our experiments. The most satisfying explanation of this effect, proposed by Nordmark (1963), runs along the same lines as presented above for the sweep-tone effect.

The last part of our discussion of the experimental results in the light of the physiological knowledge of the hearing organ concerns the conclusion in Chapter 6 that pitch of both complex and simple tones seems to be based on periodicity rather than on frequency.

This conclusion involves a choice between Hypothesis II and Hypothesis IV of the preceding chapter. The first hypothesis stated that pitch is determined by the nerve fibres stimulated by the sound or, stated in more-modern terms, by the maximum of the vibration pattern along the basilar membrane. This conception accounts very easily for the pitch of simple tones. The pitch of complex tones, however, is much more difficult to explain by the place principle: we should expect that the lower harmonics are audible individually, so long as their frequency difference exceeds the critical band, whereas the interaction of the higher harmonics should give rise to additional sensations with an indefinite pitch. Instead of this, complex tones are always characterized in normal listening by one definite pitch, even in the case with the fundamental absent and despite the ear's ability to discriminate the lower harmonics.

According to the other hypothesis, however, pitch is determined by the rate of nerve impulses. Experimental evidence has demonstrated that this conception can be considered as physiologically possible up to 3000-4000 cps (Chapter 7). In fact, periodicity rather than frequency is used as the physical basis of pitch in this view; for simple tones this distinction looks trivial, but it is very essential in the case of complex tones as the following exposition shows.

If the hearing organ did not contain a frequency-analyzing mechanism, it would not be difficult to explain how the periodicity of complex tones could be used for deriving pitch: periodic nerve impulses would be initiated irrespective of the waveform of the signal. On the other hand, if the frequency selectivity was very great, it would be very difficult to understand how periodicity is preserved: the periodic wave would be split up in a large number of noninteracting sinusoidal components. We have seen, however, that both extremes do not occur but that the ear's frequency-analyzing power is rather limited. This means that, over a large distance, the displacement of the basilar membrane is not determined by one component of the complex tone but by a group of them.

The significance of such a limited-resolution filtering system for the preservation of periodicity was put forward for the first time by Schouten (1940a, 1940b). He illustrated this on the basis of the output signals of a set of band-pass filters which were stimulated by the same periodic signal. Since Schouten's filters were much narrower than the critical band, his illustration is replaced in the following one based upon 1/3-octave filters. The ratio between bandwidth and centre frequency of these filters is 0.23, whereas for the critical bands of the ear this ratio is 0.30-0.45 at 250 cps, 0.16-0.24 at

500 cps, 0.16 at 1000 cps, 0.15 at 2000 cps, 0.17 at 4000 cps, and 0.20 at 8000 cps. On the average, these ratios do not differ greatly from the value for 1/3-octave filters.

The left part of Figure 60 represents, on a vertical scale, the frequency-response curves of the filters, with centre frequencies of 160, 200, 250, . . . , 8000 cps, respectively. The filters were stimulated simultaneously by periodic pulses of 200 pps, pulse duration 0.1 msec. To the right, the waveforms of both the input signal and the output signals of the filters are reproduced. The Fourier spectrum of the periodic pulses consists of all harmonics of 200 cps, with equal amplitudes up to about 10000 cps. The 1/3-octave filters with centre frequencies of 200, 400, 630, 800, and 1000 cps, respectively, are sufficiently narrow to pass mainly one harmonic, but this is not so at higher frequencies. The higher-frequency filters pass more than one harmonic of the input signal and, as we see, these filter outputs have periodic envelope variations with the same period as the input signal. This periodicity becomes more prominent as the number of harmonics passed by the filter increases.

This model of the ear's frequency-analyzing mechanism must be considered as a simplified representation, since the ear does not contain a limited number of fixed filters but is a continuous system of overlapping filters. Nevertheless, Figure 60 elegantly accounts for both the discrimination of the lower harmonics and the preservation of periodicity. The way in which these waveforms give rise in the ear to periodic nerve impulses, how-

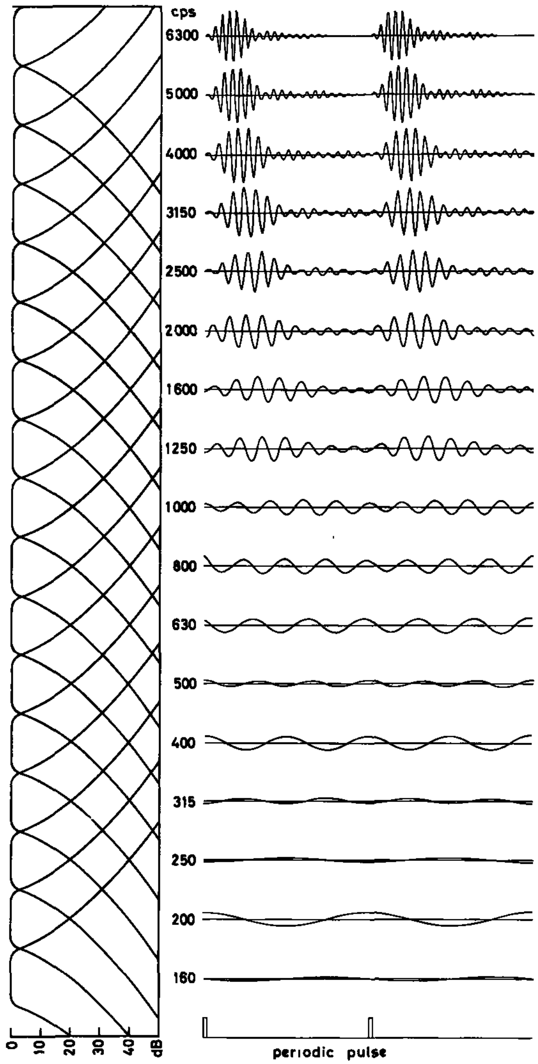


FIGURE 60. Illustration of the way in which periodicity is preserved when a periodic pulse (200 pps) stimulates a set of 1/3-octave filters.

ever, is still rather unknown. Psychophysical experiments by de Boer (1956), Schouten *et al.* (1962), and Ritsma and Engel (1964) strongly suggest that pitch is not determined by the envelope of the waveform but by its fine structure. Licklider (1956, 1959, 1962) developed a hearing theory in which autocorrelation processes are included to account for the detection of periodicity pitch (see also Ritsma, 1962).

The experiments described in Chapter 6 suggest that the frequency separation of the harmonics of complex tones below about 1400 cps is small enough to allow the periodicity of the sound wave to be retained in the output signals of many "cochlear filters". Apparently, all these filters together give rise to a stronger pitch sensation than the filters passed by only one of the lower harmonics. For fundamental frequencies above about 1400 cps, however, the situation is different. In this case, there are only a few filters with a bandwidth exceeding the frequency separation of the harmonics. Moreover, the synchronism of the nerve impulses is much poorer than at lower frequencies and may be better for sinusoids than for complex waveforms. As a consequence of both factors, the pitch of the lowest partial is the most distinct one of the complex.

In the writer's opinion, this explanation of pitch perception in terms of periodicity is more satisfactory than those based on the place principle. If the pitch of complex tones is derived from periodicity, it is difficult to see why this should not be the case for simple tones, too. Some arguments brought forward in favour of the view that pitch is derived from the place of stimulation along the basilar membrane are discussed below.

1. *Dependence of pitch upon loudness.* The pitch of low tones seems to decrease for increasing sound level, whereas at high frequencies a reverse shift has been noticed. This phenomenon was studied by Mach (1864), Burton (1895), Ewald and Jäderholm (1908), Zürmühl (1930), (Stevens 1935), and many others. Recent investigations have shown that the pitch shift is very small (Morgan *et al.*, 1951) and not significant in the majority of the cases (Cohen, 1961). Thurlow (1943) found that similar shifts can be obtained at high sound levels by introducing a tone in the other ear, indicating that we have to be rather cautious in explaining the effect on the basis of shift of the place of maximal stimulation along the basilar membrane. Furthermore, it might be that changes in brightness, which depends upon this place according to the conception expounded in the next section, play an important rôle in this type of experiment. It is of interest that the just-noticeable difference in frequency is based on change in brightness rather than in pitch, as is possible to conclude from recent experiments by Ritsma (1966).

2. *Binaural diplacusis.* In pathological cases, the same simple tone does not always give rise to equal pitch sensations for both ears. Diplacusis also occurs as a result of stimulation deafness caused by exposure to loud

sounds (Rüedi and Furrer, 1948; Davis *et al.*, 1950; Elliott *et al.*, 1964). These studies show that simple tones below about 4000 cps give rise to upward pitch shifts in the exposed ear; contradictory results were obtained at higher frequencies. Ward (1963) considered this phenomenon as a strong objection against periodicity pitch since the repetition rate of the nerve impulses must be the same in both ears. There is no reason, however, to exclude the possibility that diplacusis is introduced in a higher centre of the auditory pathway where the periodic impulses are decoded. This view is supported by a recent experiment by van den Brink (1965) in which the diplacusis of the pitch corresponding to the waveform period was measured for a stimulus consisting of three adjacent higher harmonics of an absent fundamental.

3. *Pitch shifts caused by masking.* On the basis of the place theory, von Békésy (1963b) predicted that the introduction of a masking band of noise should shift the pitch of a simple tone in the direction of the frequency spectrum of the noise band. Although he confirmed this effect experimentally, other investigators (see Webster and Muerdter, 1965) found that these shifts are generally always upwards, both for high- and low-passed noises. This may show that it is difficult to draw conclusions from this type of experiment in favour of one of the competing theories.

4. *Evidence from the cochlear model.* In some earlier papers, von Békésy presented an argument in favour of the place theory based on some experiments performed with his cochlear model (von Békésy, 1961a, 1961b, 1961c). He found that the artificial membrane of this model vibrates at the place corresponding to the absent fundamental of a periodic sound. Apart from the question of how a model spanning only two octaves can throw any light on the perception of a complex tone, we may ask how this observation can explain why the pitch of a complex tone is not altered when the lower harmonics are masked completely by noise (see Chapter 6). In that case, the contribution of the place corresponding to the fundamental is eliminated.

THE ORIGIN OF TIMBRE

The hypothesis that the pitch of a tone is based on the period of the sound waves may be criticized on the ground that this periodicity is preserved up to 3000-4000 cps, whereas we are able to distinguish tones up to about 16000 cps. This discrepancy is one of the most serious arguments against periodicity pitch. It is obviated by Wever's assumption (Chapter 7) that the ear is provided with two pitch-detecting mechanisms: one, based on periodicity, for low frequencies and one, based on place of maximal stimulation along the basilar membrane, for high frequencies. This conception is not very attractive, however.

In this connection, one of von Békésy's arguments against periodicity pitch is relevant. He compared the primary attributes of sensations on the skin and in the ear and objected that, from the point of view of the periodicity theory, we find no attribute of hearing analogous to the localization of a stimulus on the surface of the skin (von Békésy, 1963a). This would imply that, apart from its rôle in frequency analysis, the correlation between frequency and place of maximal stimulation along the basilar membrane is neglected as a source of information.

In the writer's opinion, the latter conclusion is not justified indeed. We have to modify Helmholtz's view on the rôle of the "specific energies" (Hypothesis II) instead of rejecting it. This will be clear when we consider the phenomenon of timbre. (It is remarkable that von Békésy did not include timbre in his list of auditory attributes.)

Traditionally, the timbre of complex tones is considered as depending upon the relative amplitudes of the harmonics (Brandt, 1861; Helmholtz, 1863). The question of how this relation should be understood has never been studied extensively. In practice, we don't hear the pitches of the harmonics individually but the presence of the harmonics is manifested implicitly by the way in which they affect timbre as an attribute of the tone as a whole. This would imply that timbre is a second attribute of the frequency spectrum of a sound in addition to pitch as has been suggested by some investigators (Stumpf, 1890; Schouten, 1940a; Davis *et al.*, 1951; Franssen, 1960; Schouten *et al.*, 1962).

In view of the fact that the frequency-analyzing power of the ear is limited by the critical bandwidth, we can make a further step by assuming that timbre is determined by the relative sound levels within successive critical bands. This means that the vibration pattern along the basilar membrane might play an important rôle, and that Hypothesis II must be modified in the following way:

Hypothesis V. The distribution of stimulation along the basilar membrane determines timbre.

There is much experimental evidence in favour of this hypothesis. Periodic sound waves with equal frequency spectra are very similar in timbre as can be demonstrated by passing periodic pulses of about 200 pps through a band-pass filter with a centre frequency above about 1000 cps and listening to the output signal when the period is altered gradually. In fact, we have the same effect in the perception of speech vowels. The different vowels are characterized by strong harmonics in some fixed frequency bands, known as the formants, with centre frequencies varying from vowel to vowel. The typical timbre of each vowel, by which it is recognized as an *a*, *e*, *o*, etc., is greatly independent of its pitch. The fact that, in whispering, the vowels are even recognized when the voice tract is stimulated by wide-band noise instead of periodic waves supports the view that the stimulation pattern along the basilar membrane as the physiological correlate of frequency spectrum must be considered as the basis of timbre.

A consequence of this view is that simple tones, too, are characterized by a typical frequency-dependent timbre. Apart from the unsuccessful attempt of Mach (1885) to explain tone sensation as a mixture of two components: dull and bright, Engel (1886) and Stumpf (1890) may be regarded as the first promotors of the idea that, in addition to pitch, simple tones have timbre (Stumpf's "Tonfarbe"). They realized that only on the basis of this assumption the timbre of complex tones can be understood.

From another point of view, Köhler (1909, 1911, 1915) came to the same conclusion. He gave attention to the fact, also observed by other investigators (Grassmann, 1877; von Wesendonk, 1909), that simple tones have some resemblance, depending upon frequency, with particular speech vowels. Köhler (1911) established that, from low to high frequencies, the character of a simple tone shifts from the German *u* over *o*, *a*, and *e* to *i*. He even went so far to accept that "ideal" vowels correspond with discrete frequencies in steps of an octave from about 260 cps for *u* up to about 4160 cps for *i*. Although the latter opinion could not stand the test of others (Weiss, 1920; Engelhardt and Gehrcke, 1930), in general the order of vowels attached to a tone moving from low to high frequency was confirmed. It is plain that this resemblance is related to the location of the formants in the frequency range. The fact that vowels are characterized by more than one formant shows the reason why the agreement is not very good. Recent research indicates that a better agreement can be achieved with two simultaneous simple tones (Morton and Carpenter, 1962). Since it is unlikely that the resemblance between simple tones and particular speech vowels is caused by pitch, it is, as Köhler rightly concluded, much more probable that simple tones are characterized by a distinct timbre.

As a name for this attribute of simple tones, the term *brightness*, already used by Mach as we saw above, appears to be the most appropriate one. The term is frequently used in literature as a description of the difference in sensation between low and high tones. Although it may seem rather peculiar that two psychological parameters, pitch and brightness, are related to the single frequency parameter of a simple tone, this is not so strange if we accept that pitch is correlated with periodicity and brightness with frequency. For simple tones, period is the reciprocal of frequency, but this is not the case for complex tones.

This conception presents a reasonable explanation of the discrepancy between the upper frequency limit for distinguishing tones and the highest frequency for which periodicity is preserved in the auditory nerve. The frequency of the highest tones produced by musical instruments is about 4000 cps and this frequency can also be considered as the limit of our ability to perceive pitch in its musical meaning. Still higher tones are discriminated on the basis of brightness rather than of pitch. This explanation is supported by the fact, already observed by Köhler (1915), that tones beyond about 4000 cps have a noise-like character.

The correlation between brightness and place of maximal stimulation along the basilar membrane can also account for another interesting experiment by von Békésy (1963a). He asked some subjects to localize the sensation of monaurally presented simple tones in their head and to reproduce the results graphically. The subjects succeeded to a certain degree, and it appeared that these loci were ordered after a spiral corresponding to the spiral of the organ of Corti. This phenomenon suggests that there is a correlation between frequency and place, but it does not imply that place is correlated to pitch. It fits easily in the conception that place is correlated to brightness.

For simple tones, a one-dimensional relation exists between frequency and timbre: low tones sound dull and high tones sound bright. This is not the case for complex tones, however. If timbre is based on the distribution of stimulation along the basilar membrane, which is closely correlated to the frequency spectrum of the sound, it has to be considered in this case as a multidimensional quantity. This means that the parameter brightness holds only in a strict sense for complex tones when most of their energy is concentrated in a limited frequency band; otherwise we need more than one category for distinguishing among different timbres. Actually, we don't use a set of continuous categories but we have learned to "label" verbally a number of characteristic timbres (which are adapted from speech, musical instruments, animal sounds, etc.) and use these terms for describing unfamiliar sounds.

The explanation of timbre presented in this section does not include all meanings sometimes attached to the word. If timbre has to be defined as the sum of all aspects left when loudness and pitch are "subtracted" from the total sensation of a sound, then this definition includes much more than has been considered above. It is justified, however, to consider frequency spectrum as the main component of timbre in its wide interpretation. The other ones are left out of consideration here.

RECAPITULATION AND CONCLUSIONS

As a result of the discussion of the various aspects of tone perception given in this chapter, the following description of the way in which complex tones are perceived is proposed.

The periodic sound wave stimulates the ear drum, which in turn transmits its displacements with very little nonlinearity through the ossicles of the middle ear to the stapes and the fluid in the cochlea. As a result of this, the basilar membrane in the cochlea is set into vibration, and the amplitude of the vibration pattern along the membrane is a rather good replica of the frequency spectrum of the sound. The frequency-analyzing power of this membrane is given by the critical bandwidth, implying that the first 5 to 8 harmonics give rise to separate peaks in the total vibration pattern. The higher harmonics produce no separate peaks, because their individual vibration patterns overlap extensively. Owing to this limited frequency resolution, the periodicity of the sound wave is preserved in the waveform of the basilar membrane's vibration. This periodicity is preserved also in the time pattern of the nerve impulses initiated in the hair cells (volley principle), up to 3000-4000 cps for sinusoidal waveforms and up to some lower limit greater than 1400 cps for complex periodic waveforms. The nerve impulses are transmitted to the brain and the psychological parameters of a complex tone: pitch, timbre and loudness are based upon the time-place pattern of nerve impulses. Pitch is derived from the repetition rate of the impulses, timbre from the relative number of nerve impulses evoked at the different places along the basilar membrane, and loudness from the total number of nerve impulses. In this conception, it is very valuable that the frequency-analyzing power of the ear is limited, because it enables us to perceive the complex tone as a unity with one definite pitch; apparently, this pitch "overshadows" the individual pitches of the first 5 to 8 harmonics.

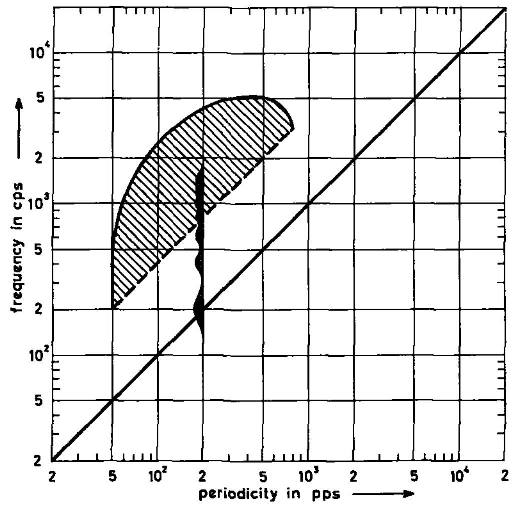


FIGURE 61. Illustration of the way in which pitch and timbre are related to the physical parameters of tones.

The way in which pitch and timbre are related to the physical parameters of tones is visualized in Figure 61. The horizontal axis represents periodicity in pps and the vertical axis frequency in cps. Since for simple tones both quantities are numerically equal, these tones are represented by a straight line with a positive slope of 45°. All complex tones are localized in the area above this line, because the period of a complex tone is equal to the reciprocal of the frequency of its fundamental. The hatched area represents the region over which Ritsma (1962) found that three adjacent harmonics (a simple tone of frequency nf that was 100% amplitude-modulated by another tone of frequency f) give rise to a pitch sensation equal to the pitch of the absent fundamental.

A complex tone can be represented in this graph by plotting the repetition rate of the sound waves along the horizontal axis and the frequency spectrum measured in pass bands equal to the critical bands along the vertical axis. In fact, we need a third axis, perpendicular to the other two, to represent the sound-pressure levels in the different pass bands. This is illustrated for a complex tone of 200 cps with harmonics up to 1600 cps; the width of the line represents sound-pressure level. The psychological parameters pitch and timbre are related in a simple way to the three axes: pitch is correlated with the periodicity and timbre with the frequency spectrum.

The graph elucidates why mistakes often

are made in comparing the pitch of a complex tone with the pitch of a simple tone. The sounds differ in two aspects and apparently it is quite difficult to isolate these aspects sufficiently in our perception. Matching for the correct pitch results in different timbres, so another pitch is often preferred as a compromise between the most similar pitch and the most similar timbre of the two tones. The fact that the pitches of tones at a distance of an octave resemble each other plays an important rôle in this comparison.

Timbre is not only an attribute of periodic but also of nonperiodic sounds. In the latter case, the vertical and the third axis have both their meaning but the horizontal has not. The whispering voice is a good example of this type of sounds. The phenomenon that appropriate successive variations of the centre frequency of a band filter passing noise or some higher harmonics of a low (absent) fundamental (Meyer-Eppler *et al.*, 1959) can give an auditory impression of a melody may be related to the fact that the mean time interval between successive zero crossings of the sound pressure is equal to the period of a simple tone coinciding with the centre frequency.

The experimental results concerning combination tones and beats can be understood in terms of this theory. Combination tones must be considered as resulting from nonlinearity for overlapping vibration patterns along the basilar membrane. Besides, the overlapping vibration patterns for tones with a frequency difference smaller than the critical band give rise to strong interference resulting in slow beats for very small frequency differences and a roughness sensation for larger ones, with a maximum for differences of about a quarter of critical bandwidth. The total roughness for a sound consisting of two complex tones has distinct minima when the frequency ratio of the fundamental frequencies of the complex tones can be expressed by small integral numbers (tonal consonance). At sound-pressure levels above about 60 dB, interference also occur over frequency distances larger than the critical bandwidth, as a result of the extensiveness of the individual vibration patterns. For mistuned consonances in the low-frequency range, this interaction, resulting in slow periodic variations in the waveform of the basilar membrane and corresponding variations in the time pattern of the nerve impulses, gives rise to slow beats and faint "sweep tones". Just as for nonlinear distortion, it is justified to state that this phenomenon plays only a minor rôle at usual listening levels of speech and music.

Finally, we have to pay attention to the fact that a complex of periodic sound waves, as produced by musical instruments and the human voice, is analyzed in such a way that both the pitches and timbres of the different complex tones are recognized. This fact was mentioned in Chapter 1 as a serious difficulty for any auditory theory in which the hearing organ is considered either as a frequency analyzer or as a periodicity analyzer. On the basis of our conclusions about the rôle of both analyzing processes and the way in which they are related, this problem has lost much of its enigmatic character. The following explanation is proposed.

The cochlea performs a frequency analysis of the complex sound with a frequency resolution that is limited by the critical bandwidth. If the sound-pressure level of each complex tone in a complex sound is high enough over one or more critical bands so as not to be masked by the other complex tones, then the corresponding region of the basilar membrane passes on the information required to derive both the pitch and the timbre of that complex tone. In particular, when the low harmonics of a tone are masked by the harmonics of a lower tone, which may be considered as the most frequent type of masking in practice, this is not an obstacle to the perception of the pitches of both tones. The timbre of a complex tone, however, should be influenced by the degree in which its harmonics are masked by other tones. The fact that, in listening to a concert, we are not aware of this effect should be understood as being similar to the ability to recognize familiar visual objects even when they are partly covered by other ones.

In fact, hearing is a dynamic process and the various laws of Gestalt psychology play an important rôle in it. All tones produced simultaneously by different sound sources are characterized by a typical time function in pitch, timbre, and loudness and this fact contributes much to their recognition. These aspects of hearing, based on central processes in the brain, are, however, beyond the scope of this study in which only some properties of isolated tones and multitone stimuli were investigated. Much more research, in particular on pattern recognition of auditory stimuli, must be done before all aspects of tone perception may become clear.

REFERENCES

Békésy, G. von: Sensations on the Skin Similar to Directional Hearing, Beats, and Harmonics of the Ear, *J. Acoust. Soc. Am. 29*, 489-501 (1957). [=G. von Békésy: *Experiments in Hearing*, McGraw-Hill Book Company, Inc., New York, pp. 568-590 (1960).]

Békésy, G. von: Pitch Sensation and Its Relation to the Periodicity of the Stimulus. Hearing and Skin Vibrations, *J. Acoust. Soc. Am. 33*, 341-348 (1961a).

Békésy, G. von: Concerning the Fundamental Component of Periodic Pulse Patterns and Modulated Vibrations Observed on the Cochlear Model with Nerve Supply, *J. Acoust. Soc. Am. 33*, 888-896 (1961b).

Békésy, G. von: Abweichungen vom Ohmschen Gesetz der Frequenzauflösung beim Hören, *Acustica 11*, *Akust. Beihefte*, 241-244 (1961c).

Békésy, G. von: Hearing Theories and Complex Sounds, *J. Acoust. Soc. Am. 35*, 588-601 (1963a).

Békésy, G. von: Three Experiments Concerned with Pitch Perception, *J. Acoust. Soc. Am. 35*, 602-606 (1963b).

Boer, E. de: *On the "Residue" in Hearing*, Doctoral Dissertation, University of Amsterdam (1956).

Brandt, S.: Ueber Verschiedenheit des Klanges (Klangfarbe), *Ann. Phys. Chem. 112*, 324-336 (1861).

Brink, G. van den: Pitch Shift of the Residue by Masking, *Intern. Audiol. 4*, 183-186 (1965).

Burton, C. V.: Some Acoustical Experiments. I. Subjective Lowering of Pitch. II. Objective Demonstration of Combination Tones, *Phil. Mag. 5th Ser. 39*, 447-453. (1895).

Cohen, A.: Further Investigation of the Effects of

Intensity upon the Pitch of Pure Tones, *J. Acoust. Soc. Am. 33,* 1363-1376 (1961).

Davis, H., C.T. Morgan, J.E. Hawkins, Jr., R. Galambos, and F.W. Smith: *Temporary Deafness Following Exposure to Loud Tones and Noise, Acta Oto-Laryngol. Suppl. 88,* pp. 44-51 (1950).

Davis, H., S.R. Silverman, and D.R. McAuliffe: Some Observations on Pitch and Frequency, *J. Acoust. Soc. Am. 23,* 40-42 (1951).

Elliott, D.N., J. Sheposh, and L. Frazier: Effect of Monaural Fatigue upon Pitch Matching and Discrimination, *J. Acoust. Soc. Am. 36,* 752-756 (1964).

Engel, G.: Ueber den Begriff der Klangfarbe, *Phil. Vorträge, Berlin, Neue Folge, II. Ser., Heft 12,* 311-355 (1886).

Engelhardt, V., and E. Gehrcke: Über die Vokalcharaktere einfacher Töne, *Z. Psychol. 115,* 16-33 (1930).

Ewald, J.R., and G.A. Jäderholm: Die Herabsetzung der subjektiven Tonhöhe durch Steigerung der objektiven Intensität, *Arch. ges. Physiol. 124,* 29-36 (1908).

Franssen, N.V.: *Some Considerations on the Mechanism of Directional Hearing,* Doctoral Dissertation, Technological University of Delft, Chapter 9 (1960).

Goldstein, J.L.: *An Investigation of Monaural Phase Perception,* Doctoral Dissertation, University of Rochester, Rochester, New York (1965).

Grassmann, H.: Ueber die physikalische Natur der Sprachlaute, *Ann. Phys. Chem. 1,* 606-629 (1877).

Greenwood, D.D.: Auditory Masking and the Critical Band, *J. Acoust. Soc. Am. 33,* 484-502 (1961a).

Greenwood, D.D.: Critical Bandwidth and the Frequency Coordinates of the Basilar Membrane, *J. Acoust. Soc. Am. 33,* 1344-1356 (1961b).

Helmholtz, H.L.F. von: *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik,* F. Vieweg & Sohn, Braunschweig, 1st Ed., Chapter 5 (1863).

Jenkins, R.A.: Perception of Pitch, Timbre, and Loudness, *J. Acoust. Soc. Am. 33,* 1550-1557 (1961).

Köhler, W.: Akustische Untersuchungen. I, *Beitr. Akust. Musikwiss. 4,* 134-182 (1909). [= *Z. Psychol. Physiol. Sinnesorgane 54,* 241-289 (1910).]

Köhler, W.: Akustische Untersuchungen. II, *Z. Psychol. 58,* 59-140 (1911). [= *Beitr. Akust. Musikwiss. 6,* 1-82 (1911).]

Köhler, W.: Akustische Untersuchungen. III, *Z. Psychol. 72,* 1-192 (1915).

Kylstra, P.H.: Can a Melody Be Heard in Monotonous Sounds?, *Proc. Kon. Nederl. Akad. Wetensch., Ser. C, 67,* 351-359 (1964).

Licklider, J.C.R.: Auditory Frequency Analysis, in C. Cherry (Ed.): *Information Theory,* Butterworths Scientific Publications, London, pp. 253-268 (1956).

Licklider, J.C.R.: Three Auditory Theories, in S. Koch (Ed.): *Psychology: A Study of a Science,* Vol. 1, McGraw-Hill Book Company, Inc., New York, pp. 41-144 (1959).

Licklider, J.C.R.: Periodicity Pitch and Related Auditory Process Models, *Intern. Audiol. 1,* 11-36 (1962).

Mach, E.: Über einige der physiologischen Akustik angehörige Erscheinungen, *Sitzungsber. Kaiserl. Akad. Wissensch., Mathemat. – Naturwissensch. Classe 50, II. Abtheilung,* 342-362 (1864).

Mach, E.: Zur Analyse der Tonempfindungen, *Sitzungsber. Kaiserl. Akad. Wissensch., Mathemat. – Naturwissensch. Classe 92, II. Abtheilung,* 1283-1289 (1885).

Meyer-Eppler, W., H. Sendhoff, and R. Rupprath: Residualton und Formantton, *Gravesaner Blätter 4, No. 14,* 70-91 (1959).

Morgan, C. T., W. R. Garner, and R. Galambos: Pitch and Intensity, *J. Acoust. Soc. Am. 23,* 658-663 (1951).

Morton, J., and A. Carpenter: Judgement of the Vowel Colour of Natural and Artificial Sounds, *Language and Speech 5,* 190-204 (1962).

Newman, E.B.: Review of E.G.Wever: Theory of Hearing, *J. Acoust. Soc. Am. 23*, 117-119 (1951).
Nordmark, J.: Some Analogies between Pitch and Lateralization Phenomena, *J. Acoust. Soc. Am. 35*, 1544-1547 (1963).
Ritsma, R.J.: Existence Region of the Tonal Residue. I, *J. Acoust. Soc. Am. 34*, 1224-1229 (1962).
Ritsma, R.J.: Personal Communication (1966).
Ritsma, R.J., and F.L.Engel: Pitch of Frequency-Modulated Signals, *J. Acoust. Soc. Am. 36*, 1637-1644 (1964).
Rüedi, L., and W.Furrer: *Das akustische Trauma*, Verlag S. Karger, Basel, pp. 148-155 (1948).
Schouten, J.F.: The Residue and the Mechanism of Hearing, *Proc. Kon. Nederl. Akad. Wetensch. 43*, 991-999 (1940a).
Schouten, J.F.: The Perception of Pitch, *Philips Techn. Review 5*, 286-294 (1940b).
Schouten, J.F., R.J.Ritsma, and B.Lopes Cardozo: Pitch of the Residue, *J. Acoust. Soc. Am. 34*, 1418-1424 (1962).
Schügerl, K.: Zur Theorie der Schwebungen verstimmter Konsonanzen, *Hochfrequenztechn. Elektroakust. 73*, 169-173 (1964).
Six, P.D.: *De cochleaire verschiltoon*, Doctoral Dissertation, University of Amsterdam (1956).
Small, A.M., Jr., and M.E.McClellan: Pitch Associated with Time Delay between Two Pulse Trains, *J. Acoust. Soc. Am. 35*, 1246-1255 (1963).
Stevens, S.S.: The Relation of Pitch to Intensity, *J. Acoust. Soc. Am. 6*, 150-154 (1935).
Stumpf, C.: *Tonpsychologie*, Vol. 2, Verlag S. Hirzel, Leipzig, pp. 514-543 (1890).
Thurlow, W.R.: Studies in Auditory Theory. I. Binaural Interaction and the Perception of Pitch, *J. Exptl. Psychol. 32*, 17-36 (1943).
Thurlow, W.R.: Further Observation on Pitch Associated with a Time Difference between Two Pulse Trains, *J. Acoust. Soc. Am. 29*, 1310-1311 (1957).
Thurlow, W.R.: Perception of Low Auditory Pitch: A Multicue, Mediation Theory, *Psychol. Rev. 70*, 461-470 (1963).
Thurlow, W.R., and A.M.Small, Jr.: Pitch Perception for Certain Periodic Auditory Stimuli, *J. Acoust. Soc. Am. 27*, 132-137 (1955).
Ward, W.D.: Diplacusis and Auditory Theory, *J. Acoust. Soc. Am. 35*, 1746-1747 (1963).
Webster, J.C., and D.R. Muerdter: Pitch Shifts Due to Low-Pass and High-Pass Noise Bands, *J. Acoust. Soc. Am. 37*, 382-383 (1965).
Weiss, A.P.: The Vowel Character of Fork Tones, *Am. J. Psychol. 31*, 166-193 (1920).
Wesendonk, K. von: Über die Synthese der Vokale aus einfachen Tönen und die Theorien von Helmholtz und Grassmann, *Physikal. Z. 10*, 313-316 (1909).
Wever, E.G., C.W.Bray, and M.Lawrence: The Origin of Combination Tones, *J. Exptl. Psychol. 27*, 217-226 (1940).
Wever, E.G., and M.Lawrence: *Physiological Acoustics*, Princeton University Press, Princeton, N.J., pp. 163-172 (1954).
Zürmühl, G.: Abhängigkeit der Tonhöhenempfindung von der Lautstärke und ihre Beziehung zur Helmholtzschen Resonanztheorie des Hörens, *Z Sinnesphysiol. 61*, 40-86 (1930).
Zwicker, E.: Der ungewöhnliche Amplitudengang der nichtlinearen Verzerrungen des Ohres, *Acustica 5*, 67-74 (1955).

SUMMARY

The way in which simultaneously sounding tones are perceived and discriminated by the ear suggests that both frequency analysis and periodicity analysis are involved. Frequency analysis can explain the fact that we are able to distinguish the tones produced by different musical instruments, but does not answer the question why each of these tones, usually consisting of a series of harmonics, manifests itself as a unity with one definite pitch. Periodicity analysis, on the contrary, accounts for the latter phenomenon very neatly, but does not indicate how simultaneous tones can be heard separately. In the writer's opinion, these two approaches need not be considered incompatible. The experiments described in this study were performed to determine the rôle of both types of analysis in audition and the way in which they are related.

Firstly, the limit of the ear's ability to discriminate the harmonics of a complex tone is treated (Chapter 2). The experiments showed that, even under most favourable conditions, not more than the first five to eight harmonics were perceived individually. This limit of the ear's frequency-analyzing power agrees with the so-called critical bandwidth, that is to say that harmonics are distinguished only if their frequency separation exceeds this bandwidth. The same limit was found for the discrimination of the partials of inharmonic series of tones. Experiments in which the masking pattern of a complex tone was investigated confirmed this significance of critical bandwidth.

Secondly, the appearance of combination tones is studied (Chapter 3). These tones are created in the ear when it is stimulated by two loud tones. When the frequencies of the higher and lower primary tones are represented by h and l, respectively, the most important combination tones appeared to be $h-l$, $2l-h$, and $3l-2h$. Experiments in which 18 subjects participated showed large individual differences in the detectability threshold for combination tones, defined as the minimum sensation level of primary tones at which combination tones are audible; all mean detectability thresholds exceeded 40 dB. Combination tones cannot be explained either by frequency analysis or by periodicity analysis, but must be attributed to nonlinear distortion. The finding that the detectability threshold for combination tones was significantly lower for small than for large frequency differences between the primary tones implies that the distortion cannot be represented by

a frequency-independent nonlinear characteristic. Comparison of the ear's distortion with the distortion that can be tolerated in sound-reproduction equipment before it is noticed suggests that the ear's distortion is so low that no combination tones are observed at usual listening levels of speech and music. Therefore, combination tones cannot be considered as a constitutive basis for tonal consonance nor as an explanation of the fact that the pitch of a complex tone without fundamental is equal to the pitch of the latter tone; on the average, sensation levels of more than 50 dB were required to reintroduce the missing fundamental.

Thirdly, in relation to the origin of tonal consonance, the interference of simple tones with a small frequency difference is investigated (Chapter 4). By tonal consonance here is meant the singular nature of tone intervals consisting of complex tones with simple frequency ratios as 1:2, 2:3, 3:4, etc. Experiments indicated that, over the frequency range between 125 and 2000 cps, simple-tone intervals were judged as consonant for frequency differences exceeding critical bandwidth and as most dissonant for frequency differences equal to about a quarter of this bandwidth. These results support the view that tones interfere when their frequency difference is smaller than critical bandwidth, manifesting itself in beats for very small differences and roughness sensation or dissonance for larger differences. In this respect, they link up with the results of Chapter 2. By taking into account that, normally, all tones in practice consist of a series of harmonics, which may also give rise to interference, the phenomenon of tonal consonance is explained in a way to be considered as a confirmation of Helmholtz's consonance theory. Statistical analyses of compositions of J.S.Bach and A.Dvořák, in which the interval distributions of simultaneous tones were computed as a function of frequency and number of harmonics, illustrate the important rôle of critical bandwidth in music.

Fourthly, the phenomenon that two simple tones may also interfere for frequency differences much larger than critical bandwidth is discussed (Chapter 5). This interference manifests itself in beats for mistuned consonances. For tone intervals with frequency ratio $1:n$ ($n=2, 3, \ldots$), slightly mistuned, these beats are usually explained as resulting from harmonics created in the ear by distortion. It appeared that combination tones or harmonics cannot be the main origin of the beats of mistuned tone intervals with frequency ratio $m:n$ ($m=2, 3, \ldots, n-1$), because the beats do not disappear when the secondary tones mentioned are masked by bands of noise. Experiments in which, for a constant lower tone, the sound-pressure level of the higher tone for most distinct beats was determined showed that this level decreases gradually as a function of frequency difference between the tones, irrespective of frequency ratio. Moreover, it was found that the beats for 1:3, slightly mistuned, manifest themselves in a weak tone sensation with a pitch shifting periodically in a sawtooth-like manner (sweep-tone effect). These experimental results strongly suggest that, also for the case $1:n$, the beats are not due to combination tones or aural harmonics but related to periodic variations of the waveform of the superimposed sinusoidal vibrations. This implies that, mainly for low frequencies, simple tones also interfere at frequency differences much larger than critical bandwidth.

Fifthly, the question is studied whether the pitch of complex tones is based on the frequency of the fundamental or on the periodicity of the sound as a whole (Chapter 6). Pitch-judgement experiments in which the fundamental of a complex tone was shifted to a 10% lower frequency and all other harmonics to a 10% higher frequency demonstrated that for fundamental frequencies up to about 1400 cps the pitch changed according to the harmonics, whereas beyond this frequency the pitch followed the fundamental. This holds both for tones with harmonics of equal amplitude and tones with harmonics of which the amplitudes decrease with 6 dB per octave. These results support the view that the pitch of complex tones is based on periodicity rather than on frequency; it is reasonable to suppose that this also holds for simple tones.

The experiments outlined above inform us about the over-all characteristics of the hearing organ; comparison with the physiological data can give us a better insight on how tones are actually perceived by the organ. As an introduction to this, a survey of the ear's physiology is given (Chapter 7). It appears that, as far as tone perception is concerned, the history of hearing theory is governed mainly by the following four hypotheses, formulated in modern terms: (1) the inner ear performs a frequency analysis of the sound in such a way that low tones give rise to a stimulation pattern along the basilar membrane with a maximum near the helicotrema and high tones with a maximum near the oval window; (2) pitch is based on the location of the maximum of the stimulation pattern; (3) the sound transmission in the ear is characterized by nonlinear distortion; (4) pitch is based on the repetition rate of the nerve impulses evoked in the hair cells. The first three hypotheses were originally proposed by Helmholtz, the last one by Wundt. Physiological evidence supports hypotheses 1 and 3, whereas it does not exclude one of the alternatives 2 and 4.

Finally, the experimental results are discussed in the light of the physiological data (Chapter 8). They confirm hypotheses 1, 3, and 4 but do not support hypothesis 2. As an alternative for the latter one, the following hypothesis is proposed: (5) the distribution of stimulation along the basilar membrane determines timbre.

On the basis of the discussion in Chapter 8, the perception of complex tones is described as follows. For a complex tone, the individual vibration patterns along the basilar membrane of the harmonics overlap such that the first five to eight harmonics give rise to separate peaks, but the higher harmonics do not. As a result of this, the periodicity of the sound wave is reproduced in the time course of the membrane's vibration; up to a frequency between 1400 and 4000 cps the periodicity is preserved in the time pattern of the nerve impulses. The time-place pattern of the nerve impulses is the basis of the psychological parameters pitch, timbre and loudness: pitch is derived from the repetition rate of the impulses (volley principle), timbre from the impulse distribution along the basilar membrane, and loudness from the total number of nerve impulses. In this conception, it is very valuable that the frequency-analyzing power of the ear is limited, because it enables us to perceive the complex tone as a unity with one definite pitch; apparently, this pitch "overshadows" the individual pitches of the first five to eight harmonics.

The experimental results concerning combination tones and beats can be understood in terms of this theory. Combination tones are attributed to the appearance of nonlinear distortion for overlapping vibration patterns of the basilar membrane. Besides, this overlapping explains the interference phenomena, not only for tones with a small frequency difference, but also for tones with a frequency difference much larger than the critical bandwidth. The beats of mistuned consonances and the sweep-tone effect are considered as resulting from periodic variations in the time pattern of the nerve impulses.

When the ear is stimulated with two or more complex tones, each of them can be perceived when its sound-pressure level is high enough over one or more critical bands so as not to be masked by the other tones. In this way we can understand our ability to discriminate simultaneous tones as produced by musical instruments and the human voice.

RÉSUMÉ

En observant la façon dont des sons simultanés sont perçus et distingués les uns des autres par l'oreille on a l'impression que l'analyse de fréquence ainsi que l'analyse de périodicité y jouent un rôle. L'analyse de fréquence peut expliquer le fait que nous sommes à même de distinguer les sons produits par des instruments différents, mais ne répond pas à la question pourquoi chacun de ces sons, se composant en général d'un nombre d'harmoniques, se manifeste comme une unité d'une seule hauteur (*pitch*). L'analyse de périodicité par contre explique assez facilement ce dernier fait, mais n'indique pas comment deux sons simultanés peuvent être entendus séparément. A mon avis ces deux façons d'aborder le problème ne sont toutefois pas incompatibles. Les expériences décrites dans cette étude furent exécutées dans le but de définir le rôle que jouent ces deux méthodes d'analyse de l'audition ainsi que la façon dont elles se rapportent l'une à l'autre.

En premier lieu on traite la question de savoir en quelle mesure l'oreille est à même de distinguer les harmoniques d'un son complexe périodique (Chapitre 2). Les expériences démontrèrent que même dans les conditions les plus favorables, l'oreille ne peut percevoir séparément que les premières cinq à huit harmoniques. Cette limite de la capacité de l'ouïe d'analyser la fréquence est en rapport avec la largeur de la soi-disant bande critique (*critical bandwidth*); les harmoniques ne peuvent donc être distinguées que lorsque leur différence mutuelle de fréquence dépasse cette largeur de bande. Quant aux sons consistant en un grand nombre de composantes inharmoniques, on obtint les mêmes résultats. Les expériences où l'on mesura la répartition de masque (*masking pattern*) d'un son complexe périodique confirmèrent le rôle des bandes critiques.

En second lieu on étudie l'apparition de sons subjectifs (Chapitre 3). Ces sons (*combination tones*) sont produits dans l'oreille lorsque deux sons intenses sont fournis. h et l représentant respectivement le plus aigu et le plus grave de ces deux sons, les sons subjectifs les plus importants se trouvent être $h-l$, $2l-h$ et $3l-2h$. Par des expériences où prirent par 18 sujets, on démontra que le seuil de perception de sons subjectifs défini comme le niveau minimum de pression acoustique des sons primaires par rapport à leur seuil absolu d'audibilité, auquel les sons subjectifs sont percevables, diffère beaucoup

selon l'individu; la moyenne des seuils de perception se trouva être au-dessus de 40 dB. On ne peut expliquer les sons subjectifs en se basant sur *l'analyse* de fréquence ou l'analyse de périodicité, mais on doit les attribuer à une distorsion non-linéaire. Le fait que le seuil de perception de sons subjectifs se trouve être beaucoup plus bas en cas de petites différences de fréquence que lorsque celles-ci sont grandes, implique que la distorsion ne peut pas être représentée par une caractéristique non-linéaire nondépendant de la fréquence. Lorsqu'on compare la distorsion dans l'oreille à celle qui peut être tolérée dans les appareils d'emission avant qu'on s'en aperçoive, on a raison à croire que la distorsion dans l'oreille est si minime que dans les niveaux usuels de la parole et de la musique qu'on écoute on ne perçoit pas de sons subjectifs. Ainsi ni la consonance tonale, ni le fait que la hauteur d'un son complexe périodique sans fondamentale est la même que celle de cette dernière, ne peuvent être expliqués au moyen des sons subjectifs; en moyenne il fallait des niveaux de pression acoustique de plus de 50 dB au-dessus du seuil d'audibilité pour réintroduire à l'oreille la fondamentale manquante.

En troisième lieu on examina l'interférence de sons simples avec une petite différence de fréquence, ceci afin de trouver l'origine de la consonance tonale (Chapitre 4). Par consonance tonale on entend ici le caractère singulier d'intervalles musicaux formés par des sons composés avec des rapports de fréquence simples, comme 1:2, 2:3, 3:4, etc. Les expériences démontrèrent que, sur une gamme de 125 à 2000 Hz, les intervalles formés par deux sons simples avec une différence de fréquence dépassant la bande critique, furent jugés consonants; lors de différences de fréquence d'à peu près un quart de cette largeur de bande, ils furent jugés le plus dissonants. Ces résultats confirment le point de vue que les sons interfèrent lorsque la différence de fréquence est plus petite que la bande critique; ceci se manifeste par des battements lors de petites différences de fréquence et par une sensation de rugosité ou de dissonance lors de différences plus étendues. A cet égard les résultats sont en rapport avec ceux du Chapitre 2. En admettant que normalement tous les sons se composent d'une série d'harmoniques qui peuvent aussi provoquer une interférence, le phénomène de la consonance tonale est expliqué d'une façon qui doit être considérée comme une confirmation de la théorie de consonance de Helmholtz. Les analyses statistiques de compositions de J.S.Bach et de A.Dvořák, où les distributions des intervalles de sons simultanés furent calculées en fonction de la fréquence et du nombre d'harmoniques, forment une illustration du rôle important de la bande critique dans la musique.

En quatrième lieu on traite le phénomène que deux sons simples peuvent interférer lors de différences de fréquence dépassant de beaucoup la bande critique (Chapitre 5). Cette interférence se manifeste par des battements lorsqu'il s'agit de consonances désaccordées. Dans des intervalles avec un rapport de fréquence de $1:n$ ($n=2, 3,\ldots$), légèrement en désaccord, l'explication de ces battements se trouve en général dans le fait qu'ils proviennent d'harmoniques produites par des distorsions dans l'oreille. Les sons subjectifs ou les harmoniques parurent ne pas pouvoir former la principale origine des battements d'intervalles désaccordés, avec un rapport de fréquence de $m:n$ ($m=2, 3,\ldots, n-1$) puisque les battements ne disparaissent pas lorsque les sons secondaires sont masqués par des bandes de bruit. Les expériences où, dans des intervalles où le son grave était constant, on détermina le niveau de pression acoustique du son aigu auquel les battements sont les plus distincts, démontrèrent que ce niveau, nonobstant le rapport de fréquence, baisse graduellement en fonction de la différence de fréquence entre les sons. On trouva en outre qu'à un rapport de 1:3, légèrement désaccordé, les battements se manifestent par une faible sensation de son avec une hauteur qui se déplace périodiquement en dents de scie (*sweep-tone effect*). Ces résultats expérimentaux font supposer que les battements pour le cas $1:n$ aussi ne peuvent pas être attribués aux sons subjectifs ou aux harmoniques produits dans l'oreille, mais sont principalement en relation avec les modifications périodiques de la forme d'onde des vibrations superposées sinusoïdales. Ceci implique que les sons simples, principalement à fréquences basses, peuvent interférer aussi lors de différences de fréquence qui dépassent de beaucoup la bande critique.

En cinquième lieu on étudie la question de savoir si la hauteur de sons complexes périodiques est basée sur la fréquence de la fondamentale ou bien sur la périodicité du son comme une unité (Chapitre 6). Les expériences où l'on examina la modification de hauteur d'un son complexe périodique dont la fondamentale fut déplacée jusqu'à une fréquence de 10% plus basse et toutes les autres harmoniques jusqu'à une fréquence de 10% plus haute, démontrèrent que pour les fréquences de la fondamentale plus basses que 1400 Hz environ, la hauteur du son se modifiait selon les harmoniques, tandis que, pour les fréquences plus élevées, la hauteur du son s'adaptait à la fondamentale. Ceci est en rigueur aussi bien pour des sons ayant des harmoniques de la même amplitude que pour des sons avec des harmoniques dont l'amplitude diminue de 6 dB par octave. Ce résultat confirme l'idée que la hauteur de sons complexes périodiques n'est pas basée sur la fréquence mais sur la périodicité; il est raisonnable de supposer qu'il en est de même pour les sons simples.

Les expériences décrites ci-dessus nous renseignent sur les caractéristiques de l'organe de l'ouïe comme totalité; une comparaison avec les données physiologiques peut nous renseigner davantage sur la façon réelle dont les sons sont perçus par l'organe. Comme introduction on donne un aperçu de la physiologie de l'oreille (Chapitre 7). Il paraît que l'histoire de la théorie de l'audition, en tant qu'elle se rapporte à la perception des sons, se compose en principe de quatre hypothèses, mentionnées en forme moderne: 1. l'oreille interne exécute une analyse de fréquence de façon à ce que les sons graves amènent une répartition de vibration (*vibration pattern*) sur la membrane basilaire, avec un maximum vers l'hélicotréma, tandis que les sons aigus amènent une répartition avec un maximum vers la fenêtre ovale; 2. la hauteur du son est basée sur l'endroit du maximum de la répartition de vibration; 3. la transmission acoustique de l'oreille est caractérisée d'une distorsion non-linéaire; 4. la hauteur du son est basée sur la fréquence de répétition des influx nerveux éveillés dans les cellules ciliées. Les trois premières hypothèses furent formulées pour la première fois par Helmholtz, la dernière par Wundt. Les données physiologiques confirment les hypothèses 1 et 3 tout en n'excluant aucune des alternatives 2 et 4.

En dernier lieu on traite les résultats expérimentaux par rapport aux données physiologiques (Chapitre 8). Ils confirment les hypothèses 1, 3 et 4 mais ne soutiennent pas l'hypothèse 2. Comme alternative de cette dernière on propose l'hypothèse suivante: 5. la distribution de la vibration sur la membrane basilaire détermine le timbre.

En se basant sur la discussion du Chapitre 8 on décrit la perception des sons de la façon suivante. Lors d'un son complexe périodique les répartitions de vibration sur la membrane basilaire individuelles des harmoniques se couvrent en partie; les premières cinq à huit harmoniques produisent des maxima séparés, mais les harmoniques plus élevées n'en produisent pas. Ceci implique que la périodicité du son se retrouve dans la forme d'onde de la vibration de la membrane; jusqu'à une fréquence d'entre 1400 et 4000 Hz cette périodicité se retrouve aussi dans la distribution dans le temps des influx nerveux. La distribution temps-lieu (*time-place pattern*) des influx nerveux est à la base des paramètres psychologiques hauteur, timbre et niveau: la hauteur est déduite de la fréquence de répétition des influx nerveux (*volley principle*), le timbre de leur distribution sur la membrane basilaire et le niveau du nombre total des influx nerveux. L'auteur insiste sur le fait que selon cette conception la limitation de la capacité de l'oreille d'analyser la fréquence est une qualité précieuse parce qu'elle nous rend à même de percevoir le son complexe périodique comme une unité avec une seule hauteur. Apparemment cette hauteur "couvre" les hauteurs individuelles des premières cinq à huit harmoniques.

Les résultats expérimentaux concernant les sons subjectifs et les battements sont expliqués selon cette theorie. Les sons subjectifs proviennent de l'apparition de la distorsion non-linéaire lorsque les répartitions de vibration sur la membrane basilaire se couvrent. En outre le fait que ces répartitions se couvrent explique les phénomènes d'interférence non seulement pour les sons d'une petite différence de fréquence mais encore pour les sons d'une différence de fréquence dépassant la bande critique. Les battements de consonances désaccordées et l'effet *sweeptone* sont considérés comme résultant des modifications périodiques dans la distribution dans le temps des influx nerveux.

Lorsqu'on fournit à l'oreille deux ou plusieurs sons complexes périodiques chacun de ces sons peut être perçu lorsque son niveau de pression acoustique sur une ou plusieurs bandes critiques est assez élevé pour ne pas être masqué par d'autres sons. De cette façon s'explique notre faculté de distinguer les uns des autres des sons simultanés, produits par des intruments et par la voix humaine.

ZUSAMMENFASSUNG

Die Art und Weise wie gleichzeitig klingende Töne durch das Ohr wahrgenommen und unterschieden werden, festigt den Eindruck, daß hier sowohl Frequenzanalyse als auch Periodizitätsanalyse eine Rolle spielen. Frequenzanalyse kann die Tatsache erklären, daß wir fähig sind die von verschiedenen Musikinstrumenten erzeugten Töne zu unterscheiden; sie beantwortet aber nicht die Frage, weshalb jeder dieser Töne, gewöhnlich aus einer Anzahl Harmonischen zusammengesetzt, sich als eine Ganzheit mit einer bestimmten Tonhöhe manifestiert. Dagegen bietet Periodizitätsanalyse keine Schwierigkeiten letztgenanntes Phänomen zu deuten; sie zeigt aber nicht wie es möglich ist gleichzeitige Töne getrennt wahrzunehmen. Nach meiner Ansicht brauchen diese beiden Annäherungsmethoden nicht miteinander in Widerspruch zu stehen. Die in vorliegender Arbeit behandelten Experimente dienten dazu, die Funktion beider Analysemethoden beim Hören zu bestimmen, sowie auch die wechselseitigen Beziehungen festzustellen.

Erstens wird das Problem behandelt in welchem Maße das Ohr fähig ist die Harmonischen eines zusammengesetzten Tones von einander zu unterscheiden (Kapitel 2). Aus Versuchen ging hervor, daß sogar unter günstigsten Umständen nicht mehr als die ersten fünf bis acht Harmonischen getrennt wahrgenommen werden konnten. Diese Grenze des frequenzanalysierenden Vermögens des Gehörorgans stimmt mit der sogenannten Frequenzgruppe (*critical bandwidth*) überein; Harmonischen können also nur unterschieden werden wenn ihr gegenseitiger Frequenzabstand größer ist als die Frequenzgruppe. Für Laute, aufgebaut aus einer großen Anzahl nichtharmonischer Töne, gilt das Gleiche. Messungen der Mithörschwelle eines zusammengesetzten Tones befestigten diese Bedeutung der Frequenzgruppe.

Zweitens wird das Auftreten von Kombinationstönen studiert (Kapitel 3). Diese Töne entstehen im Ohr durch Darbietung von zwei lauten Tönen. Wenn h und l die Frequenzen der höheren b.z.w. tieferen dieser zwei Töne darstellen, ergibt sich, daß $h-l$, $2l-h$ und $3l-2h$ die wichtigsten Kombinationstöne sind. Experimente mit 18 Versuchspersonen zeigten, daß die Wahrnehmungsschwelle von Kombinationstönen, definiert als der minimale Schalldruckpegel von den primären Tönen über der Hörschwelle, bei denen Kombinationstöne hörbar sind, individuell stark unterschiedlich ist; die durchschnittlichen Wahrnehmungsschwellen lagen alle über

40 dB. Kombinationstöne sind zu betrachten als verursacht von nichtlineare Verzerrung und können nicht auf Grund der Frequenz- b.z.w. Periodizitätsanalyse erklärt werden. Da die Wahrnehmungsschwelle von Kombinationstönen mit geringen Frequenzdifferenzen zwischen den primären Tönen signifikant niedriger ist als bei großen Differenzen, ist eine frequenzunabhängige nichtlineare Verzerrungskurve ausgeschlossen. Gegenseitige Vergleichung der Verzerrung im Ohr und die erlaubte, nicht wahrnehmbare Verzerrung im Wiedergabegerät, lassen eine dermaßen geringfügige Verzerrung im Ohr annehmen, daß beim üblichen Hörpegel von Sprache und Musik keine Kombinationstöne wahrnehmbar sind. Kombinationstöne kommen deshalb zur Deutung der tonalen Konsonanz nicht in Frage und ebensowenig zur Erklärung der Tatsache, daß die Höhe eines zusammengesetzten Tones ohne Grundton gleich die des letztgenannten Tones ist. Um den fehlenden Grundton im Ohr von neuem zu erzeugen, waren durchschnittlich Schalldruckpegeln von mehr als 50 dB über der Hörschwelle notwendig.

Drittens wird, im Zusammenhang mit dem Ursprung tonaler Konsonanz, die Interferenz von Sinustönen mit geringem Frequenzunterschied erforscht (Kapitel 4). Unter tonaler Konsonanz wird der singuläre Charakter von Tonintervalle, bestehend aus zusammengesetzten Tönen mit einfachen Frequenzverhältnissen wie 1:2, 2:3, 3:4, usw., verstanden. Aus Versuchen, durchgeführt in dem Frequenzbereich von 125 bis 2000 Hz, ging folgendes hervor: Intervalle bestehend aus Sinustönen mit einem Frequenzunterschied größer als die Frequenzgruppe, wurden als konsonant, mit einem Frequenzunterschied von etwa einem Viertel dieser Bandbreite dagegen als maximal dissonant beurteilt. Diese Ergebnisse stützen die Auffassung, daß Töne interferieren wenn der Frequenzunterschied geringer ist als die Frequenzgruppe. Dieses offenbart sich in Schwebungen bei sehr geringen, b.z.w. einem Rauhigkeitseindruck oder Dissonanz bei grösseren Frequenzunterschieden. In dieser Hinsicht schliessen diese Ergebnisse sich bei den Ergebnissen von Kapitel 2 an. Wenn man berücksichtigt, daß im allgemeinen alle Töne in der Praxis aus einer Reihe Harmonischen bestehen, die gleichfalls Interferenz verursachen können, wird das Phänomen der tonalen Konsonanz auf eine Weise gedeutet die als eine Bestätigung der Helmholtzschen Konsonanztheorie aufzufassen ist. Die wichtige Rolle der Frequenzgruppe in der Musik wird an Hand statistischer Analysen von Kompositionen von J. S. Bach und A. Dvořák gezeigt. Hier wurden Intervallverteilungen gleichzeitiger Töne als Funktion von Frequenz und Anzahl Harmonischen berechnet.

Viertens wird das Phänomen besprochen, daß zwei Sinustöne auch interferieren können wenn der Frequenzunterschied viel größer ist als die Frequenzgruppe (Kapitel 5). Dieses kommt in Schwebungen bei verstimmten Konsonanzen zum Ausdruck. Für Tonintervalle mit Frequenzverhältnis $1:n$ ($n=2, 3, \ldots$), leicht verstimmt, betrachtet man meistens durch Verzerrungen im Ohr erzeugten Harmonischen als Ursache dieser

Schwebungen. Es stellte sich heraus, daß Kombinationstöne oder Harmonischen nicht die Hauptursache der Schwebungen von verstimmten Tonintervallen mit Frequenzverhältnis $m:n$ ($m=2, 3, \ldots, n-1$) sein können, da diese Schwebungen nicht verschwinden wenn genannte Sekundärtöne durch Bandpaß-Rauschen verdeckt werden. In Versuchen mit konstantem tieferen Ton wurde der Schalldruckpegel des höheren Tones bestimmt, wobei Schwebungen am deutlichsten hörbar sind. Es zeigte sich, daß dieser Pegel bei steigendem Frequenzunterschied zwischen den Tönen allmählig abnimmt, ungeachtet des Frequenzverhältnisses. Außerdem wurde ermittelt, daß Schwebungen von Tonintervallen mit einem Frequenzverhältnis 1:3, leicht verstimmt, sich als eine Tonempfindung mit einer Tonhöhe, die in einer zackigen Weise periodisch verschiebt, bemerkbar machen (*sweep-tone effect*). Diese Versuchsergebnisse erwecken den Eindruck, daß die Schwebungen, auch im Falle $1:n$, nicht durch im Ohr erzeugte Kombinationstöne oder Harmonischen verursacht werden, sondern sich auf die periodischen Wellenformänderungen der überlagerten sinusförmigen Schwingungen beziehen. Daraus geht hervor, daß Sinustöne, hauptsächlich bei tiefen Frequenzen, auch bei Frequenzunterschieden interferieren können, die viel größer sind als die Frequenzgruppe.

Fünftens wird die Frage studiert, ob die Tonhöhe zusammengesetzter Töne auf die Grundfrequenz oder auf die Periodizität des Schalles als Ganzes gegründet ist (Kapitel 6). Experimente bei denen die Veränderung der Tonhöhe eines zusammengesetzten Tones beurteilt wurde, dessen Grundton nach einer 10% tieferen und alle übrigen Harmonischen nach einer 10% höheren Frequenz verschoben wurden, zeigten, daß bei Grundtonfrequenzen unter etwa 1400 Hz die Tonhöhe sich entsprechend den Harmonischen änderte, während die Tonhöhe bei höheren Frequenzen dem Grundton folgte. Dieses gilt sowohl für Töne mit Harmonischen von gleicher Amplitude, als auch für Töne mit Harmonischen derer Amplituden um 6 dB pro Oktave abnehmen. Dieses Ergebnis bestätigt die Auffassung, daß die Tonhöhe eines zusammengesetzten Tones *nicht* auf Frequenz, sondern auf Periodizität beruht; es ist anzunehmen, daß dieses auch für Sinustöne gilt.

Oben erwähnte Versuche geben uns Aufschluß über die Eigenschaften des Gehörorgans als Ganzes; eine Vergleichung mit den physiologischen Vorgängen kann unsere Einsicht vertiefen wie Töne vom Gehörorgan faktisch perzipiert werden. Hierzu ist als Einleitung eine Übersicht der Physiologie des Gehörorgans gegeben (Kapitel 7). Es zeigt sich, daß die Geschichte der Gehörtheorie, hinsichtlich die Tonempfindung, hauptsächlich von folgenden vier Hypothesen beherrscht wird, die in moderner Form folgendermaßen lauten: 1. das Innenohr führt in solcher Weise eine Frequenzanalyse vom Schalle durch, daß für tiefe Töne eine Vibrationsverteilung (*vibration pattern*) über der Basilarmembran mit einem Maximum bei dem Helicotrema und für hohe Frequenze mit einem Maximum beim ovalen

Fenster die Folge ist; 2. die Tonhöhe beruht auf der Stelle vom Maximum der Vibrationsverteilung; 3. die Schallübertragung im Ohr wird durch nichtlinearer Verzerrung gekennzeichnet; 4. die Tonhöhe beruht auf der Wiederholungsfrequenz von in den Haarzellen erregten Nervenimpulsen. Die ersten drei Hypothesen wurden zum ersten Mal von Helmholtz, die letzte von Wundt formuliert. Die physiologische Daten stützen die Hypothesen 1 und 3, während sie keine der Alternativen 2 und 3 ausschließen.

Schließlich werden die Versuchsergebnisse im Lichte der physiologischen Daten besprochen (Kapitel 8). Sie bestätigen die Hypothesen 1, 3 und 4 aber stützen Hypothese 3 nicht. Als Alternative für letztgenannte wird folgende Hypothese vorgeschlagen: 5. die Verteilung der Erregung über der Basilarmembran bestimmt die Klangfarbe (*timbre*).

Auf Grund der Diskussion in Kapitel 8 wird die Perzeption von Tönen wie folgt beschrieben. Bei einem zusammengesetzten Ton überdecken die individuellen Vibrationsverteilungen der Harmonischen über der Basilarmembran sich zum Teil, wobei die ersten fünf bis acht Harmonischen getrennte Maxima zeigen, die höheren Harmonischen aber nicht. Dieses hat zur Folge, daß die Schallperiodizität in der Wellenform der Membranschwingung zurückgefunden wird; diese Periodizität wird in dem Zeitbild (*time pattern*) der Nervenimpulse bis zu einer Frequenz zwischen 1400 und 4000 Hz in Stand gehalten (*volley principle*). Die psychologischen Parameter, Tonhöhe, Klangfarbe und Lautheit beruhen auf dem Zeit-Ortbild (*time-place pattern*) der Nervenimpulse: die Tonhöhe wird von der Wiederholungsfrequenz der Impulse abgeleitet, die Klangfarbe von ihrer Verteilung über der Basilarmembran und die Lautheit von der Gesammtzahl der Nervenimpulse. Nach dieser Auffassung ist es sehr wertvoll, daß das frequenzanalysierende Vermögen des Gehörorgans begrenzt ist, weil es uns die Möglichkeit bietet den zusammengesetzten Ton als eine Ganzheit mit einer bestimmten Tonhöhe wahrzunehmen; offenbar werden die individuellen Tonhöhen der ersten fünf bis acht Harmonischen von dieser Tonhöhe "überschattet".

Die Versuchsergebnisse bezüglich die Kombinationstöne und Schwebungen werden im Rahmen dieser Theorie erklärt. Kombinationstöne treten infolge nichtlinearer Verzerrung von einander überdeckenden Vibrationsverteilungen der Basilarmembran auf. Nebenbei erklärt dieses Überdecken die Interferenzphänomene, nicht nur für Töne mit einem geringen Frequenzunterschied, sondern auch für Töne mit einem Frequenzunterschied größer als die Frequenzgruppe. Die Schwebungen verstimmter Konsonanzen und der *sweep-tone* Effekt werden als Folge periodischer Änderungen im Zeitbild der Nervenimpulse aufgefaßt.

Wenn dem Gehörorgan zwei oder mehr zusammengesetzte Töne dargeboten werden, ist jeder dieser Töne wahrnehmbar wenn dessen Schalldruckpegel über eine oder mehrere Frequenzgruppen ausreicht um nicht von anderen Tönen verdeckt zu werden. Auf diese Weise kann erklärt werden, daß wir fähig sind gleichzeitige Töne, z.B. von Musikinstrumenten oder der menschlichen Stimme erzeugt, von einander zu unterscheiden.

SAMENVATTING

De wijze waarop gelijktijdig klinkende tonen door het oor worden waargenomen en onderscheiden, geeft de indruk dat hierbij zowel frequentieanalyse als periodiciteitsanalyse een rol spelen. Frequentieanalyse kan het feit verklaren dat we in staat zijn de door verschillende muziekinstrumenten geproduceerde tonen van elkaar te onderscheiden maar geeft geen antwoord op de vraag waarom ieder van deze tonen, gewoonlijk bestaande uit een aantal harmonischen, zich manifesteert als een eenheid met één bepaalde toonhoogte. Daarentegen heeft periodiciteitsanalyse geen moeite het laatstgenoemde verschijnsel te verklaren, maar geeft zij niet aan hoe gelijktijdige tonen afzonderlijk kunnen worden gehoord. Deze twee benaderingswijzen behoeven echter mijns inziens niet als onverenigbaar te worden beschouwd. De in deze studie beschreven experimenten werden uitgevoerd met het doel de rol van beide analysemethoden bij het horen vast te stellen alsmede de wijze waarop zij met elkaar zijn verbonden.

In de eerste plaats wordt de vraag behandeld in welke mate het oor in staat is de harmonischen van een samengestelde toon van elkaar te onderscheiden (Hoofdstuk 2). De experimenten toonden aan dat, zelfs onder de gunstigste omstandigheden, niet meer dan de eerste vijf tot acht harmonischen afzonderlijk konden worden waargenomen. Deze grens van het frequentieanalyserend vermogen van het gehoororgaan komt overeen met de zgn. kritieke bandbreedte; harmonischen kunnen dus alleen onderscheiden worden wanneer hun onderling frequentieverschil groter is dan deze bandbreedte. Dezelfde grens werd gevonden voor de onderscheidbaarheid van de deeltonen van nietharmonische series van tonen. Experimenten waarin het maskeringspatroon van een samengestelde toon werd gemeten, bevestigden deze betekenis van de kritieke bandbreedte.

Ten tweede wordt het optreden van combinatietonen bestudeerd (Hoofdstuk 3). Deze tonen worden in het oor gevormd bij aanbieding van twee luide tonen. Wanneer h en l de frequenties van de hoogste resp. de laagste van deze twee tonen voorstellen, dan blijken de belangrijkste combinatietonen $h-l$, $2l-h$ en $3l-2h$ te zijn. Experimenten waaraan 18 proefpersonen deelnamen, toonden aan dat de waarnemingsdrempel van combinatietonen, gedefinieerd als het minimale geluiddrukniveau van de primaire tonen t.o.v. de gehoordrempel, waarbij combinatietonen hoorbaar zijn, individueel sterk verschilt;

de gemiddelde waarnemingsdrempels lagen alle boven 40 dB. Combinatietonen kunnen niet worden verklaard op basis van frequentieanalyse of van periodiciteitsanalyse maar moeten worden toegeschreven aan nietlineaire vervorming. Het feit dat de waarnemingsdrempel van combinatietonen bij kleine frequentieverschillen tussen de primaire tonen significant lager is dan bij grote verschillen, impliceert dat de *vervorming* niet kan worden voorgesteld door een frequentie-onafhankelijke niet-lineaire karakteristiek. Vergelijking van de vervorming in het oor met de vervorming die in geluidsweergaveapparatuur kan worden getolereerd voordat zij wordt waargenomen, geeft reden te veronderstellen dat de vervorming in het oor zo gering is dat bij gebruikelijke luisterniveaus van spraak en muziek geen combinatietonen worden waargenomen. Combinatietonen komen daarom niet in aanmerking ter verklaring van tonale consonantie en evenmin van het feit dat de toonhoogte van een samengestelde toon zonder grondtoon gelijk is aan die van de laatstgenoemde toon; gemiddeld waren geluiddrukniveaus van meer dan 50 dB boven de gehoordrempel vereist om de ontbrekende grondtoon in het oor opnieuw te introduceren.

Ten derde wordt, met het oog op de oorsprong van tonale consonantie, de interferentie van enkelvoudige tonen met een klein frequentieverschil onderzocht (Hoofdstuk 4). Onder tonale consonantie wordt hierbij verstaan het singuliere karakter van toonintervallen bestaande uit samengestelde tonen met eenvoudige frequentieverhoudingen als 1:2, 2:3, 3:4, enz. Experimenten toonden aan dat, over het frequentiebereik van 125 tot 2000 Hz, intervallen bestaande uit enkelvoudige tonen met een frequentieverschil groter dan de kritieke bandbreedte als consonant werden beoordeeld en als het meest dissonant bij frequentieverschillen van ongeveer een kwart van deze bandbreedte. Deze resultaten steunen de opvatting dat tonen interfereren wanneer het frequentieverschil kleiner is dan de kritieke bandbreedte, tot uiting komend in zwevingen bij zeer kleine frequentieverschillen en een ruwheidssensatie of dissonantie bij grotere ververschillen. In dit opzicht sluiten zij aan bij de resultaten van Hoofdstuk 2. Door in aanmerking te nemen dat normaliter alle tonen in de praktijk uit een serie harmonischen bestaan, die eveneens aanleiding kunnen geven tot interferentie, wordt het verschijnsel van tonale consonantie verklaard op een wijze die als een bevestiging van de consonantietheorie van Helmholtz dient te worden beschouwd. Statistische analyses van composities van J.S. Bach en A. Dvořák, waarbij de intervaldistributies van gelijktijdige tonen werden berekend als functie van frequentie en aantal harmonischen, illustreren de belangrijke rol van de kritieke bandbreedte in de muziek.

Ten vierde komt het verschijnsel aan de orde dat twee enkelvoudige tonen ook bij veel grotere frequentieverschillen dan de kritieke bandbreedte kunnen interfereren (Hoofdstuk 5). Deze interferentie uit zich in zwevingen bij ontstemde consonanten. Voor toonintervallen met frequentieverhouding $1:n$ ($n=2, 3,\ldots$), iets ontstemd, worden deze zwevingen gewoonlijk verklaard als gevolg van door vervorming in het oor geproduceerde harmonischen. Het bleek dat combinatietonen of harmonischen niet de belangrijkste oorzaak van de zwevingen van ontstemde toonintervallen met frequentieverhouding $m:n$ ($m=2, 3,\ldots, n-1$) kunnen zijn, daar de zwevingen niet verdwijnen wanneer de genoemde secundaire tonen worden gemaskeerd door ruisbanden. Experimenten waarin bij constante laagste toon het geluiddrukniveau van de hoogste toon werd bepaald, waarbij de zwevingen het duidelijkst te horen zijn, toonden aan dat dit niveau, ongeacht de frequentieverhouding, bij toenemend frequentieverschil tussen de tonen geleidelijk afneemt. Bovendien werd gevonden dat de zwevingen bij 1:3, iets ontstemd, zich manifesteren in een zwakke toonsensatie met een toonhoogte die periodiek op een zaagtandachtige wijze verschuift (*sweep-tone effect*). Deze experimentele resultaten doen ten sterkste vermoeden dat de zwevingen, ook voor het geval $1:n$, niet een gevolg zijn van in het oor geproduceerde combinatietonen of harmonischen, maar te maken hebben met de periodieke veranderingen van de golfvorm van de gesuperponeerde sinusvormige trillingen. Dit houdt in dat enkelvoudige tonen, voornamelijk bij lage frequenties, ook kunnen interfereren bij frequentieverschillen die veel groter zijn dan de kritieke bandbreedte.

Ten vijfde wordt de vraag bestudeerd of de toonhoogte van samengestelde tonen gebaseerd is op de frequentie van de grondtoon of op de periodiciteit van het geluid als een geheel (Hoofdstuk 6). Experimenten waarin werd beoordeeld de verandering in toonhoogte van een samengestelde toon waarvan de grondtoon naar een 10% lagere frequentie werd verschoven en alle andere harmonischen naar een 10% hogere frequentie, wezen uit dat voor frequenties van de grondtoon lager dan ongeveer 1400 Hz de toonhoogte veranderde overeenkomstig de harmonischen terwijl de toonhoogte bij hogere frequenties de grondtoon volgde. Dit geldt zowel voor tonen met harmonischen van gelijke amplitude als voor tonen met harmonischen waarvan de amplituden met 6 dB per octaaf afnemen. Dit resultaat steunt de opvatting dat de toonhoogte van samengestelde tonen niet op de frequentie maar op de periodiciteit is gebaseerd; het is redelijk te veronderstellen dat dit eveneens voor enkelvoudige tonen geldt.

De hierboven geschetste experimenten lichten ons in over de karakteristieken van het gehoororgaan als geheel; een vergelijking met de fysiologische gegevens kan ons een beter inzicht geven hoe tonen in feite door het orgaan worden gepercipieerd. Als inleiding hiertoe is een overzicht van de fysiologie van het oor gegeven (Hoofdstuk 7). Het blijkt dat de geschiedenis van de gehoortheorie, voorzover deze de waarneming van tonen betreft, hoofdzakelijk beheerst wordt door vier hypothesen die in moderne vorm luiden: 1. het binnenoor voert een frequentieanalyse van het geluid uit, op zodanige wijze dat lage tonen een trillingspatroon over de basilaire membraan met een maximum bij het helicotrema tot gevolg hebben en hoge tonen met een maximum bij het ovale venster; 2. de toonhoogte is gebaseerd op de plaats van het maximum van het stimulatiepatroon; 3. de geluidoverdracht in het oor wordt gekenmerkt door niet-lineaire vervorming; 4. de toonhoogte is gebaseerd op de herhaalfrequentie van de in de haarcellen opgewekte zenuwimpulsen. De eerste drie hypothesen werden voor het eerst geformuleerd door Helmholtz, de laatste door Wundt. De fysiologische gegevens steunen hypothesen 1 en 3, terwijl zij geen van de alternatieven 2 en 4 uitsluiten.

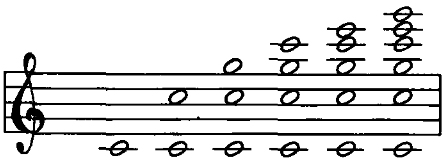
Tenslotte worden de experimentele resultaten besproken in het licht van de fysiologische gegevens (Hoofdstuk 8). Zij bevestigen de hypothesen 1, 3 en 4 maar steunen hypothese 2 niet. Als alternatief voor de laatstgenoemde wordt de volgende hypothese voorgesteld: 5. de distributie van de stimulatie over de basilaire membraan bepaalt het timbre (klankkleur).

Op grond van de discussie in Hoofdstuk 8 wordt de perceptie van tonen als volgt beschreven. Bij een samengestelde toon vallen de individuele trillingspatronen van de harmonischen langs de basilaire membraan gedeeltelijk over elkaar heen, waarbij de eerste vijf tot acht harmonischen nog gescheiden maxima geven, maar de hogere niet. Dit heeft tot gevolg dat de periodiciteit van de geluidsgolf teruggevonden wordt in de golfvorm van de trilling van de basilaire membraan; deze periodiciteit wordt, tot een frequentie tussen 1400 en 4000 Hz, eveneens teruggevonden in het tijdspatroon van de zenuwimpulsen. Het tijd-plaatspatroon van de zenuwimpulsen is de basis van de psychologische parameters toonhoogte, timbre en luidheid: toonhoogte wordt afgeleid van de herhaalfrequentie van de impulsen, timbre van hun distributie over de basilaire membraan en luidheid van het totale aantal zenuwimpulsen. Volgens deze opvatting is het zeer waardevol dat het frequentieanalyserend vermogen van het oor beperkt is, omdat het ons in staat stelt de samengestelde toon als een eenheid met één bepaalde toonhoogte waar te nemen; blijkbaar "overschaduwt" deze toonhoogte de individuele toonhoogten van de eerste vijf tot acht harmonischen.

De experimentele resultaten betreffende combinatietonen en zwevingen worden in het kader van deze theorie verklaard. Combinatietonen zijn het gevolg van het optreden van niet-lineaire vervorming bij over elkaar heen vallende trillingspatronen van de basilaire membraan. Daarnaast verklaart dit over elkaar heen vallen de interferentieverschijnselen, niet alleen voor tonen met een klein frequentieverschil, maar ook voor tonen met een frequentieverschil groter dan de kritieke bandbreedte. De zwevingen van ontstemde consonanten en het *sweep-tone* effect worden beschouwd het gevolg te zijn van periodieke veranderingen in het tijdspatroon van de zenuwimpulsen.

Wanneer het gehoororgaan twee of meer samengestelde tonen worden aangeboden, kan ieder van deze tonen worden waargenomen wanneer zijn geluiddrukniveau over één of meer kritieke banden hoog genoeg is om niet door andere tonen te worden gemaskeerd. Op deze wijze kan ons vermogen worden verklaard om gelijktijdige tonen, zoals deze worden voortgebracht door muziekinstrumenten en de menselijke stem, van elkaar te onderscheiden.

Tones produced by musical instruments and the human voice consist nearly always of a number of harmonics. The frequencies of these harmonics are multiples of the frequency of the first harmonic or fundamental. In the first demonstration, such a complex tone is built up by adding to a fundamental (262 cps) the second, third, fourth, fifth, and sixth harmonic, successively.



(demonstration)

Our ability to hear these harmonics separately is limited to the first five to eight harmonics, as the experiments described in Chapter 2 have shown. This limit agrees with the so-called critical bandwidth of the hearing organ.

When the ear is stimulated by two loud tones, secondary tones may appear, the so-called combination tones. In the next demonstration, at first a sequence of four simple tones is presented, and after that a sequence of four tone intervals, each con-

Les sons produits par des instruments de musique et la voix humaine se composent presque toujours d'un nombre d'harmoniques. Les fréquences de ces harmoniques sont des multiples de la fréquence de la première harmonique ou fondamentale. Dans la première démonstration on édifie un tel son complexe en ajoutant successivement à une fondamentale (262 Hz) la seconde, troisième, quatrième, cinquième et sixième harmonique.

Notre capacité d'entendre séparément ces composantes se restreint aux premières cinq à huit harmoniques, comme il a été démontré par les expériences décrites au Chapitre 2. Cette limite correspond à la soi-disant bande critique de l'organe de l'ouïe.

Lorsqu'on offre à l'oreille deux sons intenses des sons secondaires peuvent se produire, les soi-disant sons subjectifs. Dans la démonstration suivante on offre d'abord une série de quatre sons simples, et après cela quatre intervalles de sons se composant chacun de

Die Töne, von Musikinstrumenten und der menschlichen Stimme erzeugt, sind fast immer aus einer Anzahl Harmonischen zusammengesetzt. Die Frequenzen dieser Harmonischen sind Vielfachen der Frequenz der ersten Harmonischen oder Grundton. Im ersten Beispiel wird solch ein zusammengesetzter Ton aufgebaut durch an einen Grundton (262 Hz) nacheinander die zweite, dritte, vierte, fünfte und sechste Harmonische hinzu zu fügen.

Unsere Fähigkeit diese Teiltöne getrennt zu hören, beschränkt sich auf die ersten fünf bis acht Harmonischen; dieses wurde durch die in Kapitel 2 behandelten Experimente gezeigt. Diese Grenze stimmt mit der sog. Frequenzgruppe des Gehörorgans überein.

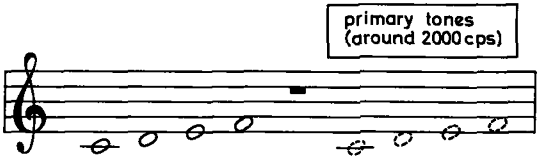
Wenn dem Ohr zwei laute Töne dargeboten werden, können Sekundartöne auftreten, die sog. Kombinationstöne. Im nächsten Beispiel wird zuerst eine Reihe von vier Sinustönen dargeboten und danach eine Reihe von vier Tonintervalle aus je zwei lauten Tönen.

De tonen voortgebracht door muziekinstrumenten en de menselijke stem bestaan bijna altijd uit een aantal harmonischen. De frequenties van deze harmonischen zijn veelvouden van de frequentie van de eerste harmonische of grondtoon. In de eerste demonstratie wordt zo'n samengestelde toon opgebouwd door aan een grondtoon (262 Hz) achtereenvolgens de tweede, derde, vierde, vijfde en zesde harmonische toe te voegen.

Ons vermogen om deze componenten afzonderlijk te horen, is beperkt tot de eerste vijf tot acht harmonischen, zoals de in Hoofdstuk 2 beschreven experimenten hebben aangetoond. Deze grens komt overeen met de zgn. kritieke bandbreedte van het gehoororgaan.

Wanneer aan het oor twee luide tonen worden aangeboden, kunnen secundaire tonen optreden, de zgn. combinatietonen. In de volgende demonstratie wordt eerst een serie van vier enkelvoudige tonen aangeboden en sisting of two loud tones. These intervals are composed in such a manner that in the ear a sequence of four combination tones is created with frequencies equal to the frequencies of the simple tones first presented. The demonstration is repeated once.

deux sons intenses. Ces intervalles ont été composés d'une telle façon qu'il se forme dans l'oreille une série de quatre sons subjectifs dont les fréquences sont les mêmes que celles des quatre premiers sons. La démonstration est répétée une fois.



(demonstration)

The audibility of these combination tones depends upon the extent to which the primary tones are distorted in your ears, and may be influenced by the distortion produced in the sound-reproduction equipment. The experiments treated in Chapter 3 suggest that combination tones do not play a rôle at normal sound levels of speech and music.

Two simple tones with a very small frequency difference give rise to slow beats changing into a roughness sensation at larger frequency differences. In the next demonstration, a constant tone of 250 cps and a variable tone increasing gradually from 250 up to 550 cps are presented simultaneously. The tones are interrupted briefly at a frequency difference of a quarter of the critical bandwidth and at a frequency difference equal to the critical bandwidth (25 and 100 cps, respectively). These two frequency differences were judged by the subjects in the experiments of Chapter 4 as corresponding with the most dissonant interval and the

L'audibilité de ces sons subjectifs dépend de la mesure en laquelle les sons primaires sont distordus dans les oreilles et peut avoir subi l'influence de la distorsion produite par l'appareil d'émission. Les expériences traitées au Chapitre 3 donnent lieu à supposer que les sons subjectifs ne jouent aucun rôle lorsqu'il s'agit de niveaux usuels de la parole et de la musique.

Deux sons simples avec une différence de fréquence minime produisent des battements lents qui se transforment en une sensation de rugosité lorsque la différence de fréquence est plus grande. Dans la démonstration suivante on offre simultanément un son constant de 250 Hz et un son variable qui monte graduellement de 250 Hz à 550 Hz. On soumet les sons à une brève interruption lorsque la différence de fréquence correspond à un quart de la bande critique (25 Hz) et lorsque celle-ci correspond à la bande critique (100 Hz). Ces deux différences de fréquence furent jugées par les sujets mentionnés au Chapitre 4

Diese Intervalle sind derartig zusammengesetzt, daß im Ohr eine Reihe von vier Kombinationstönen erzeugt wird, deren Frequenzen den zuerst dargebotene vier Tonfrequenzen gleich sind. Das Beispiel wird einmal wiederholt.

daarna een serie van vier toonintervallen, ieder bestaande uit twee luide tonen. Deze intervallen zijn zodanig samengesteld dat in het oor een serie van vier combinatietonen wordt gevormd waarvan de frequenties gelijk zijn aan de frequenties van de eerst aangeboden vier tonen. De demonstratie wordt één keer herhaald.

Die Wahrnehmbarkeit dieser Kombinationstöne ist von dem Maße abhängig in dem die primären Töne in Ihren Ohren verzerrt werden und kann durch die im Wiedergabegerät erzeugte Verzerrung beeinträchtigt sein. Die in Kapitel 3 behandelten Experimente lassen annehmen, daß Kombinationstöne keine Rolle spielen bei normalen Schallpegeln von Sprache und Musik.

Zwei Sinustöne mit einem sehr geringen Frequenzunterschied veranlassen langsame Schwebungen die bei größeren Frequenzunterschieden in einen Rauhigkeitseindruck übergehen. Im nächsten Beispiel werden gleichzeitig ein konstanter Ton von 250 Hz und ein von 250 bis 550 Hz allmählich zunehmender Ton dargeboten. Die Töne sind kurz unterbrochen bei einem Frequenzunterschied von einem Viertel der Frequenzgruppe und bei einem Frequenzunterschied gleich der Frequenzgruppe (25 b.z.w. 100 Hz). Diese zwei Frequenzunterschiede wurden von den Versuchspersonen in den Experimenten

De waarneembaarheid van deze combinatietonen hangt af van de mate waarin de primaire tonen in uw oren vervormd worden en kan beïnvloed zijn door de vervorming die de geluidsweergaveapparatuur produceert. De experimenten behandeld in Hoofdstuk 3 geven grond aan de veronderstelling dat combinatietonen geen rol spelen bij normale geluidniveaus van spraak en muziek.

Twee enkelvoudige tonen met een zeer klein frequentieverschil geven aanleiding tot langzame zwevingen die in een ruwheidssensatie overgaan bij grotere frequentieverschillen. In de volgende demonstratie worden tegelijkertijd een constante toon van 250 Hz en een geleidelijk van 250 tot 550 Hz toenemende toon aangeboden. De tonen zijn kort onderbroken bij een frequentieverschil van een kwart van de kritieke bandbreedte en bij een frequentieverschil gelijk aan de kritieke bandbreedte (25 resp. 100 Hz). Deze twee frequentieverschillen werden door de proefpersonen in de experimenten van smallest consonant interval, respectively.

(demonstration)

For complex tones, however, also the harmonics give rise to beats and roughness. This is shown in the next example. Now, both the constant and the varying tones consist of six harmonics with equal amplitudes. The demonstration is given twice; in the first one, the tones are interrupted for frequency ratios corresponding with the musical fourth (3:4), fifth (2:3), and octave (1:2), respectively.

(demonstration)

Simple tones may also give rise to faint beat sensations for frequency differences much larger than the critical bandwidth, namely for slightly mistuned consonances. The most distinct beats appear for a mistuned octave with a lower frequency of, for example, 125 cps. They sound as follows.

(demonstration)

In general, these beats are not due to distortion in the ear but to interference of the tones themselves, as is argued in Chapter 5.

We saw in the beginning that the hearing organ is able to distinguish the first five to eight harmonics of a complex tone. In listening to music, however, we are not aware of the existence of these harmonics. On the contrary, complex tones appear to have one definite pitch equal to the pitch of the fundamental. This is shown in the following illus-

comme correspondant avec l'intervalle le plus dissonant, respectivement avec l'intervalle consonant le plus petit.

Dans les sons complexes périodiques les harmoniques aussi donnent lieu à des battements et à une sensation de rugosité. L'exemple suivant le montre. La constante ainsi que la variable se composent de six harmoniques avec les mêmes amplitudes. La démonstration est faite deux fois; la première fois les sons sont interrompus lorsque le rapport de fréquence correspond avec respectivement la quarte (3:4), la quinte (2:3) et l'octave musicale (1:2).

Les sons simples peuvent aussi donner de faibles sensations de battement lorsque la différence de fréquence dépasse de beaucoup la bande critique, c. à d. lorsque les consonances sont légèrement désaccordées. Les battements les plus distincts se présentent dans une octave désaccordée dont la fréquence la plus basse est par exemple de 125 Hz. Elle sonne ainsi:

En général ces battements ne proviennent pas de la distorsion dans l'oreille, mais d'une interaction des sons eux-mêmes, ce qui fut démontré au Chapitre 5.

Nous avons vu que l'organe de l'ouïe est à même de distinguer les unes des autres les premières cinq à huit harmoniques d'un son complexe périodique. En écoutant la musique nous ne nous rendons pas compte de l'existence de ces harmoniques. Les sons complexes périodiques paraissent par contre avoir une in Kapitel 4 beurteilt als dem dissonantesten Intervall b.z.w. dem kleinsten konsonanten Intervall entsprechend.

Hoofdstuk 4 beoordeeld als corresponderend met het meest dissonante resp. het kleinste consonante interval.

Bei zusammengesetzten Tönen veranlassen jedoch auch die Harmonischen Schwebungen und Rauhigkeit. Dieses wird im folgenden Beispiel gezeigt. In diesem Fall sind sowohl der konstante als auch der veränderliche Ton aus sechs Harmonischen mit gleichen Amplituden zusammengesetzt. Das Beispiel wird zweimal gegeben; das erste Mal sind die Töne unterbrochen bei Frequenzunterschieden der musikalischen Quart (3:4), Quint (2:3) und Oktave (1:2) entsprechend.

Bij samengestelde tonen hebben echter ook de harmonischen zwevingen en ruwheid ten gevolge. Dit toont het volgende voorbeeld. Nu bestaan zowel de constante als de variabele toon uit zes harmonischen met gelijke amplituden. De demonstratie wordt twee keer gegeven; de eerste keer zijn de tonen onderbroken bij frequentieverhoudingen corresponderend met resp. de muzikale kwart (3:4), de kwint (2:3) en het octaaf (1:2).

Sinustöne können auch bei Frequenzunterschieden viel größer als die Frequenzgruppe, schwache Schwebungsempfindungen geben, nämlich bei leicht verstimmten Konsonanzen. Die deutlichsten Schwebungen treten bei einer verstimmten Oktave mit einer tieferen Frequenz von z.B. 125 Hz auf. Sie klingen wie folgt.

Enkelvoudige tonen kunnen ook bij frequentieverschillen die veel groter zijn dan de kritieke bandbreedte, zwakke zwevingssensaties geven, namelijk bij iets ontstemde consonanten. De duidelijkste zwevingen treden op bij een ontstemd octaaf met een laagste frequentie van bijv. 125 Hz. Zij klinken als volgt.

Im allgemeinen sind diese Schwebungen nicht die Folge von Verzerrungen im Ohr sondern einer Wechselwirkung der Töne selbst, wie in Kapitel 5 gezeigt wurde.

Am Anfang haben wir gesehen, daß das Gehörorgan fähig ist die ersten fünf bis acht Harmonischen eines zusammengesetzten Tones von einander zu unterscheiden. Beim Hören von Musik sind wir uns der Existenz dieser Harmonischen jedoch nicht bewußt. Es stellt sich im Gegenteil heraus, daß zu-

In het algemeen zijn deze zwevingen niet het gevolg van vervorming in het oor maar van wisselwerking van de tonen zelf, zoals in Hoofdstuk 5 is aangetoond.

We zagen in het begin dat het gehoororgaan in staat is de eerste vijf tot acht harmonischen van een samengestelde toon van elkaar te onderscheiden. Bij het luisteren naar muziek zijn we ons echter niet bewust van het bestaan van deze harmonischen. Samengestelde tonen blijken integendeel één bepaalde tration. Each tone of this melody consists of a large number of harmonics.

(demonstration)

In the next demonstration, a complex tone of 250 cps is presented first, and after that a second one of which the fundamental is shifted to a 10% lower frequency and the other harmonics to a 10% higher frequency. The twelve harmonics have equal amplitudes. The short tone pulses are repeated once.

(demonstration)

In the experiments described in Chapter 6, the pitch of the second tone pulse of each pair was nearly always judged as higher than the pitch of the first tone pulse. The same result was obtained for the following tone pulses in which the amplitudes of the harmonics decrease gradually with 6 dB per octave.

(demonstration)

For fundamental frequencies above 1400 cps, however, the direction of the pitch shift corresponded to the shift in the fundamental. These results strongly suggest that pitch perception is based on the temporal periodicity of the complex tone as a whole rather than on the frequency of the fundamental.

Another illustration of the significance of this waveform periodicity for pitch per-

hauteur définie égale à la hauteur de la fondamentale. L'illustration suivante le démontre. Chaque son de cette mélodie se compose d'un grand nombre d'harmoniques.

Dans la démonstration suivante on offre d'abord un son complexe périodique de 250 Hz et ensuite un second son dont la fondamentale a été déplacée à une fréquence plus basse de 10% et les autres harmoniques à une fréquence plus élevée de 10%. Les douze harmoniques ont les mêmes amplitudes. Les brèves impulsions de son sont répétées une fois.

Dans les expériences décrites au Chapitre 6 la seconde impulsion de chaque paire de sons fut presque toujours jugée comme plus haute que la première. Le même résultat fut obtenu dans des impulsions de son suivantes, dans lesquelles les amplitudes des harmoniques diminuent graduellement de 6 dB par octave.

Lorsque la fréquences de la fondamentale fut au-dessus de 1400 Hz, la direction de la modification de la hauteur correspondirent avec la modification de la fondamentale. Ces résultats donnent la forte impression que la perception de la hauteur du son n'est pas basée sur la fréquence de la fondamentale mais sur la périodicité du son complexe comme unité.

sammengesetzte Töne eine bestimmte Tonhöhe haben, gleich der Tonhöhe des Grundtones. Dieses wird im folgenden Beispiel gezeigt. Jeder Ton dieser Melodie besteht aus einer großen Anzahl Harmonischen.

toonhoogte te hebben, gelijk aan de toonhoogte van de grondtoon. Dit toont de volgende illustratie. De tonen van deze melodie bestaan ieder uit een groot aantal harmonischen.

Im nächsten Beispiel wird zuerst ein zusammengesetzter Ton von 250 Hz dargeboten und danach ein zweiter dessen Grundton nach einer 10% tieferen Frequenz und die anderen Harmonischen nach einer 10% höheren Frequenz verschoben sind. Die zwölf Harmonischen haben gleiche Amplituden. Die kurzen Tonstöße werden einmal wiederholt.

In den in Kapitel 6 beschriebenen Versuchen wurde die Tonhöhe des zweiten Tonstoßes jedes Paares fast immer höher beurteilt als die Tonhöhe des ersten Tonstoßes. Dasselbe Ergebnis wurde ermittelt bei den folgenden Tonstößen in denen die Amplituden der Harmonischen allmählich mit 6 dB pro Oktave abnehmen.

Bei Grundtonfrequenzen über etwa 1400 Hz stimmte die Richtung der Tonhöhenänderung jedoch mit der Grundtonverschiebung überein. Diese Ergebnisse deuten stark darauf hin, daß die Tonhöhenempfindung nicht auf der Grundtonfrequenz beruht sondern auf der Periodizität des zusammengesetzten Tones als Ganzes.

Ein anderes Beispiel der Bedeutung der

In de volgende demonstratie wordt eerst een samengestelde toon van 250 Hz aangeboden en daarna een tweede waarvan de grondtoon naar een 10% lagere frequentie is verschoven en de andere harmonischen naar een 10% hogere frequentie. De twaalf harmonischen hebben gelijke amplituden. De korte toonstoten worden één keer herhaald.

In de in Hoofdstuk 6 beschreven experimenten werd de toonhoogte van de tweede toonstoot van ieder paar bijna altijd als hoger beoordeeld dan de toonhoogte van de eerste toonstoot. Hetzelfde resultaat werd verkregen voor de volgende toonstoten waarin de amplituden van de harmonischen geleidelijk met 6 dB per octaaf afnemen.

Voor grondtoonfrequenties boven ongeveer 1400 Hz correspondeerde de richting van de toonhoogteverandering echter met de verschuiving van de grondtoon. Deze resultaten geven sterk de indruk dat de waarneming van toonhoogte niet gebaseerd is op de frequentie van de grondtoon maar op de periodiciteit van de samengestelde toon als geheel.

Een andere illustratie van de betekenis van ception can be given by presenting a sequence of complex tones and simultaneously bands of noise which mask completely all harmonics except those in an octave band with a centre frequency of 1500 cps. Still the melody can be heard faintly.

(demonstration; same melody as before)

The fact that timbre is determined by the frequency spectrum of the sound, discussed in Chapter 8, is demonstrated in the following examples. At first a sequence of complex tones is presented. After that the same sequence is repeated with all harmonics below 500, 1000, 2000, 4000, and 8000 cps removed, successively.

(demonstration; same melody as before)

As is expounded in the last chapter of the study, it is likely that the critical band governs the frequency analysis performed in the inner ear. Owing to this limited frequency resolution, the temporal periodicity of complex tones is preserved in the time pattern of the nerve impulses, and this gives rise to pitch sensation. The phenomena demonstrated on this record are in accordance with this representation of the ear's analyzing mechanism.

Afin d'illustrer autrement le rôle de la périodicité dans la perception de la hauteur du son, on peut offrir une série de sons complexes périodiques en même temps que des bandes de bruit qui rendent absolument inaudibles toutes les harmoniques à l'exception d'une bande d'octave avec une fréquence moyenne de 1500 Hz. La mélodie est toutefois faiblement audible.

Le fait que le timbre est déterminé par le spectre de fréquence du son, ce qui a été traité au Chapitre 8, est démontré par les exemples suivants. On offre d'abord une série de sons complexes périodiques. Ensuite on répète la même série, en supprimant successivement toutes les harmoniques au-dessous de 500, 1000, 2000, 4000 et 8000 Hz.

Comme il a été exposé au dernier chapitre de cette étude, il est probable que l'analyse de fréquence dans l'oreille interne est déterminée par la bande critique. Grâce à cette capacité limitée de résoudre la frequence, la périodicité de la forme d'onde de sons complexes périodiques est conservée dans la distribution dans le temps des influx nerveux, ce qui amène la perception de la hauteur du son. Les phénomènes démontrés par ce disque correspondent avec nos idées sur le mécanisme analytique de l'organe de l'ouïe.

Wellenformperiodizität für die Wahrnehmung von Tonhöhe kann gegeben werden durch eine Reihe zusammengestetzter Töne darzubieten und gleichzeitig Rauschbände die alle Harmonischen unhörbar machen mit Ausnahme einer Oktave mit einer Mittelfrequenz von 1500 Hz. Die Melodie ist trotzdem schwach zu hören.

Die Tatsache, daß Klangfarbe vom Frequenzspektrum des Schalles bestimmt wird, wie in Kapitel 8 besprochen, wird in folgenden Beispielen gezeigt. Zuerst wird eine Reihe zusammengesetzter Töne dargeboten. Danach wird dieselbe Reihe wiederholt wobei nacheinander alle Harmonischen unter 500, 1000, 2000, 4000 und 8000 Hz entfernt sind.

Wie im letzten Kapitel dieser Arbeit erörtert wurde, ist es wahrscheinlich, daß die Frequenzanalyse im Innenohr durch die Frequenzgruppe bestimmt wird. Es ist diesem beschränkten frequenzauflösenden Vermögen zu verdanken, daß die Periodizität der Wellenform zusammengesetzter Töne im Zeitbild der Nervenimpulse erhalten bleibt; dieses führt zur Tonhöhenempfindung. Die auf dieser Schallplatte dargebotenen Phänomene sind in Übereinstimmung mit dieser Darstellung des analysierenden Mechanismus des Gehörorgans.

de periodiciteit van de golfvorm voor de waarneming van toonhoogte kan worden gegeven door het aanbieden van een serie samengestelde tonen gelijktijdig met ruisbanden die alle harmonischen volkomen onhoorbaar maken, met uitzondering van een octaafband met een middenfrequentie van 1500 Hz. Toch is de melodie nog zwak te horen.

Het feit dat timbre (klankkleur) bepaald wordt door het frequentiespectrum van het geluid, hetgeen in Hoofdstuk 8 is besproken, wordt in de volgende voorbeelden gedemonstreerd. Eerst wordt een serie samengestelde tonen aangeboden. Daarna wordt dezelfde serie herhaald waarbij achtereenvolgens alle harmonischen beneden 500, 1000, 2000, 4000 en 8000 Hz verwijderd zijn.

Zoals in het laatste hoofdstuk van deze studie is uiteengezet, is het waarschijnlijk dat de frequentieanalyse in het binnenoor door de kritieke bandbreedte bepaald wordt. Dank zij dit beperkte frequentieoplossende vermogen blijft de periodiciteit van de golfvorm van samengestelde tonen behouden in het tijdspatroon van de zenuwimpulsen, hetgeen leidt tot gewaarwording van de toonhoogte. De op deze grammofoonplaat gedemonstreerde verschijnselen zijn in overeenstemming met deze voorstelling van het analyserend mechanisme van het gehoororgaan.