4.9, 4.11

Received 22 September 1971; revised 8 November 1971

The Central Origin of the Pitch of Complex Tones: Evidence from Musical Interval Recognition *

A. J. M. Houtsma and J. L. Goldstein†

Center for Communications Sciences, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

The human auditory system's ability to recognize simple melodies that correspond to fundamental periods in sequences of periodic sounds devoid of fundamental energy was studied through musical interval identification experiments. Stimuli comprising two randomly chosen successive upper harmonics were presented both monotically (two harmonics to one ear) and dichotically (one harmonic to each ear). Subjects could recognize melodies equally well with both modes of stimulus presentation. The results imply that the pitch of these complex tones is mediated by a central processor operating on neural signals derived from those effective stimulus harmonics that are tonotopically resolved.

INTRODUCTION

An essential basis for practicing the art of music is the human auditory system's ability to perceive melody in a sequence of periodic or near-periodic sounds. In Western culture, a musical note scale has been developed which characterizes musical sounds solely by their fundamental frequencies, and a melody comprises a sequence of notes on such a scale. Musical instruments are designed in such a way that when a note is played a periodic or quasiperiodic sound is generated whose fundamental frequency is designated by the note. On listening to a sequence of such musical sounds one can easily retrieve the original series of notes from the auditory sensation, regardless of the particular Fourier spectrum generated by the instrument. Some instruments (e.g., strings) produce sounds with a rich spectrum comprising the fundamental and many partials; others (e.g., flute) have little more than the bare fundamental; and still others (French horn) generate lower notes with relatively little energy at the lower partials, including the fundamental (Saunders, 1937, 1946). Also melodies played on any given instrument can easily be recognized after the sound has been passed through a bandpass filter with a bandwidth much narrower than the spectra of the original sounds; one can readily observe this when listening to music from an inexpensive transistor radio. This phenomenon of melody invariance over a large class of spectral transformations has had profound impact on auditory science ever since the concept of Fourier spectrum was introduced to audition by Ohm in 1843 (Plomp and Smoorenburg, 1970). This earlier work is discussed more fully in Secs. V and VI. We report in this paper a number of new psychophysical experiments that were designed within a musical framework to explore how the auditory system retrieves and encodes the information of the fundamental frequency from a sequence of periodic sounds.

I. METHODS

A. General Paradigm

In order to simulate musical behavior as closely as possible, the basic procedure was to present a subject with a simple musical message consisting of a number of musical sounds, and to have him respond by identifying it. In cases where the set of musical messages or melodies was known to the subject beforehand, a response box was used; otherwise the subject was asked to write down the perceived melody using a relative note scale.

The musical sounds were complex tones comprising two successive upper harmonics of equal intensity; no energy at the fundamental frequency itself was present. Despite the fact that more complex stimuli might have provided better simulation of musical-instrument sounds, these relatively simple synthetic stimuli were preferred because they minimize the number of free stimulus parameters, and offer better control over the location of spectral energy in the effective stimulus. In addition, they were selected because a large body of

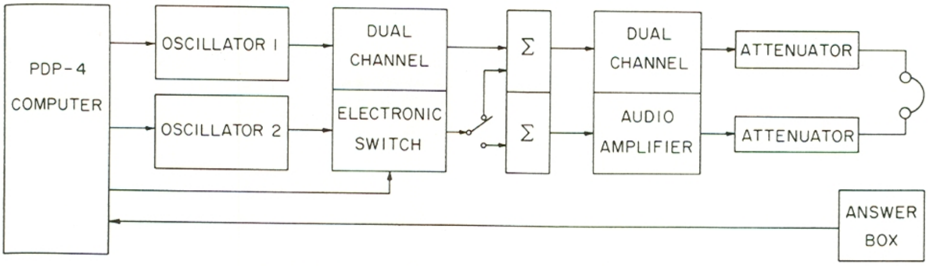

FIG. 1. Equipment diagram for all two-tone experiments.

possibly relevant psychophysical and physiological data already existed for these stimuli (Goldstein, 1967, 1970; Goldstein and Kiang, 1968; Kiang, 1965; Sachs and Kiang, 1968; Sachs, 1969).

If the same harmonic numbers were chosen for each note, one could easily recover the melody described by the (missing) fundamentals merely by tracking one of the two partials, since the frequencies of harmonic partials are always integral multiples of the fundamental frequency. In this situation, both partials would describe the same melody as the fundamental, merely transposed upward in key. In order to prevent the subject from using this trivial cue, the harmonic numbers were randomized from note to note, and hence the subject was required to retrieve the notes from hearing a sequence of pairs of randomly chosen successive upper harmonics.

B. Apparatus

The apparatus consisted of two programmable oscillators (General Radio 1161-A and KrohnHite 4031 R), an electronic switch (Grason–Stadler), a two-channel audio amplifier, and a set of TDH-39 headphones. Harmonic and intermodulation distortion measured at the headphones was better than 50 dB below stimulus level. (In one experiment, where four successive harmonics were needed, we used a third programmable oscillator, KrohnHite 4031 R, and a dual multiplier, CBL 47.) Subjects were seated in an IAC model 1200 sound-insulated chamber. Switches and oscillators were controlled by a DEC PDP-4 computer which generated all random events, performed stimulus computations, stored responses, and controlled feedback. An equipment diagram is shown in Fig. 1.

C. Subjects

Three subjects participated in the basic experiments. All were male, between 18 and 31 years of age, and had quite extensive musical training and experience (majors in organ, viola, and singing). All subjects could perform the required tests with little special training; test sessions were limited to a maximum of 2 h daily. An additional seven subjects, four male and three female, participated in some more descriptive and qualitative experiments. All had at least some degree of musical training and were familiar with musical notation and dictation. A few could not perform the required identification or recognition tests initially and were given some training (less than 1 h) starting with tone complexes of up to six successive harmonics, which were then gradually reduced to the two-tone stimuli used in the test experiments.

II. BASIC EXPERIMENTS: IDENTIFICATION OF EIGHT KNOWN MUSICAL INTERVALS

A. General Procedures

Five basic experiments were carried out having many features in common. The task of the subject was to identify on each trial which out of eight known two-note melodies (or musical intervals) was presented [Fig. 2(a)]. These simple melodies had identical envelope time structure [Fig. 2(b)], and all began with the same note. The notes were approximately tuned to the natural scale, that is, frequency ratios of 16/15, 9/8, 6/5, and 5/4 for the minor and major second, and minor and major third, respectively. The stimuli representing the notes comprised two successive harmonics, the number of the lower harmonic being chosen randomly for each note over a range of three successive integers [Fig. 2(c)]. The middle of the lower harmonic number range, $\bar{n}$, and the fundamental frequency of the first note, f_0, were chosen as independent parameters in measuring identification performance expressed as the percentage of correct responses. We chose this particular set of musical intervals, not because we assume they have any inherent significance, but merely because they provide, at least in the Western culture, a convenient language which musically trained people can understand.

The subjects were first tested in preliminary experiments for their ability to identify musical intervals by using periodic sounds with fundamental energy; a control run of 50 trials was carried out, the stimuli being square waves with fundamental frequencies equal to the notes in Fig. 2(a). Subjects were given a "key" ascribing a number to each interval and were instructed

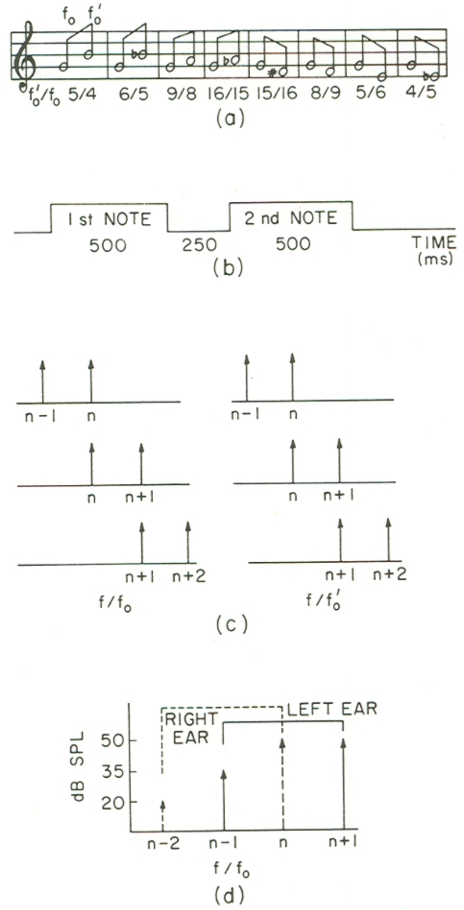

FIG. 2. Experimental paradigm for musical interval identification experiments: (a) musical intervals to be identified, (b) time-envelope of the total stimulus, (c) the three possible two-tone stimuli for each of the two notes; for each note a random choice was made among the three possible stimuli, (d) stimulus configuration for one note of a dichotically presented melody, with the addition of simulated combination tones.

to push a corresponding button on an answer box within 4 sec after each stimulus presentation, after which feedback was provided. All three subjects scored perfectly, which indicates that they did possess the required musical skills, and that performance limitations so often encountered in identification experiments of this kind and which are commonly attributed to imperfect memory (Miller, 1956) were apparently of little significance here.

Next, in the basic experiments, subjects were tested individually. AH and SW were given feedback after each trial, while NH preferred no feedback. A typical run consisted of 50 trials for subject AH and 25 trials for the other subjects. Fundamental frequencies were incremented in steps of 100 Hz (or sometimes 200 Hz), and for each fundamental frequency a number of runs was taken, each time incrementing the average harmonic number $\bar{n}$ by one. As many runs were taken as were necessary to make performance drop from perfect (100% correct) to chance (12.5% correct, corresponding to one out of the eight intervals guessed correctly); typically about six runs at each fundamental frequency were required. The resulting psychometric functions were then remapped into "equal performance contours," relating performance level to both fundamental frequency, f_0, and average harmonic number, $\bar{n}$.

B. Monotic Experiments

Stimuli were presented monotically at a sensation level of 20 dB. Equal performance contours are shown in Fig. 3(a). On one subject (AH) a control experiment was performed with the partials of each note presented in time sequence, rather than simultaneously. We chose values for f_0 and $\bar{n}$ which had previously yielded perfect performance. Despite the fact that f_0 can easily be computed from this stimulus, the subject was not able to perform better than chance, even after many concentrated efforts. We concluded that randomization of harmonic number forces the subject to use both partials simultaneously.

The experimental results show clearly that the best performance is achieved with the lowest harmonic numbers, that is, the larger the spacing between harmonics on a log–frequency scale, the better the performance. This finding, together with what is known about the physiology of cochlear frequency analysis (Kiang, 1965; Goldstein, Baer, and Kiang, 1971) and the psychophysics of resolution of partials in complex tones (Plomp, 1964) suggested the hypothesis that both harmonics of the sounds employed in this experiment are processed through separate channels of the cochlear output to obtain successful identification.

C. Dichotic Experiments and Relation to Monotic Results

The separate-channel hypothesis was tested employing probably the most extreme channel separation one can think of, namely separate ears. We thought that, while negative results would not necessarily invalidate the hypothesis, positive results would prove it to be correct. The experimental paradigm was the same as in the monotic experiment, except that the stimulus partials were presented dichotically, one to each ear. The results are shown in Fig. 3(b). Comparison with the monotic results shows that each subject's performance is essentially the same under both stimulus conditions, suggesting that a central mechanism integrates and processes information from both cochleas,

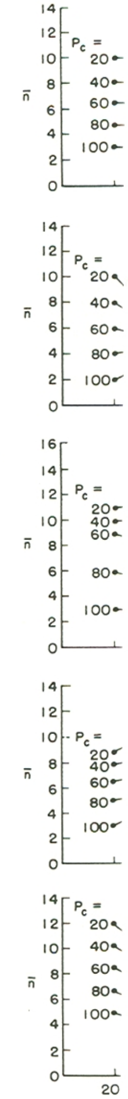

FIG. 3. Perforn
Stimulus conditi
dichotic, 50 (40)

and that the i
both monotic
The effect
and dichotic

. A typical
ıd 25 trials
:ncies were
:s 200 Hz),
ber of runs
e harmonic
:n as were
m perfect
·esponding
:orrectly);
frequency
functions
:ontours,"
ıdamental
, n̄.

sensation
ıre shown
:periment
presented
ısly. We
y yielded
an easily
was not
er many
mization
ıse both

the best
armonic
between
tter the
what is
:quency
Kiang,
rtials in
oothesis
in this
nnels of
ication.

nploy-
on one
t that,
alidate
to be
me as
mulus
h ear.
with
form-
mulus
anism
hleas,

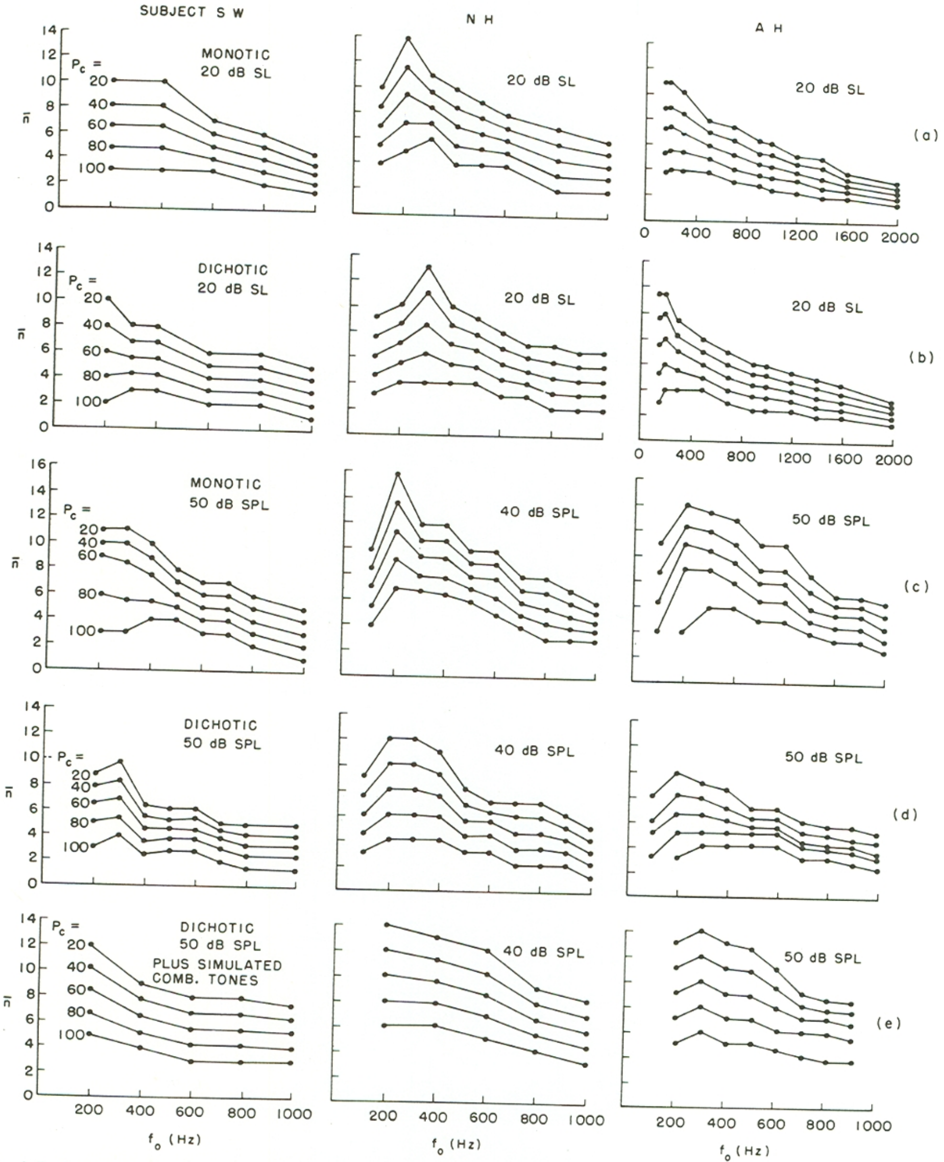

FIG. 3. Performance contours for musical interval identification experiments with pairs of randomly chosen successive upper harmonics. Stimulus conditions: (a) monotic, 20 dB SL, (b) dichotic, 20 dB SL, (c) monotic, 50 (40) dB SPL, (d) dichotic, 50 (40) dB SPL, (e) dichotic, 50 (40) dB SPL, with two simulated combination tones.

and that the inputs to this mechanism are similar under both monotic and dichotic stimulation.

The effect of stimulus intensity upon both monotic and dichotic performance was investigated in a third and fourth experiment undertaken at higher stimulus intensity levels, 50 dB SPL for subjects AH and SW, and 40 dB SPL for subject NH. These sound-pressure levels (SPLs) were chosen to avoid the aural difference tone, f_2-f_1, in the monotic case (Plomp, 1965; Goldstein, 1967), and interaural crosstalk in the dichotic case (Zwislocki, 1953). Equal performance contours are shown in Figs. 3(c) and 3(d).

D. Influence of Aural Combination Tones

Comparing results from all four experiments, it appears that each subject's performance is essentially the same under all four stimulus conditions, except that for higher-intensity monotic stimulation [Fig. 3(c)] performance contours are shifted upwards by approximately 2 or 3 harmonic numbers. Such an upward shift might be expected because of the presence of aural combination tones of the type $f_1-k(f_2-f_1)$ generated in the peripheral ear for a monotic stimulus comprising the frequencies f_1 and f_2 (Zwicker, 1955; Plomp, 1965; Goldstein, 1967, 1970; Goldstein and Kiang, 1968). These combination tones provide the ear with two or three harmonics below those contained in the stimulus, which are probably very useful since all the results thus far indicate that performance improves with decreasing harmonic number. These combination tones could make the *effective* average harmonic number about 2 or 3 lower than the actual value of $\bar{n}$ in Fig. 3. In the dichotic experiments, combination tones are not present; in the monotic experiment at 20 dB SL they would be near or below threshold (Goldstein, 1967).

There are two obvious ways to test this combination-tone hypothesis. One approach would be to repeat the third experiment (medium intensity, monotic) and add a number of pure tones to the stimulus having the exact frequency, amplitude, and phase to cancel all aurally generated combination tones (Goldstein, 1967); identification performance should then approach that for dichotic experiments. The technical and experimental difficulties of providing the correct cancellation amplitude and phase for each note caused us to eschew this approach. Instead, an experiment was carried out using dichotic stimuli at 50 dB SPL (40 dB for subject NH) with the addition of two more tones that approximately simulated the aural combination tones which the ear generates under monotic conditions. The stimulus paradigm is shown in Fig. 2(d) and the experimental results in Fig. 3(e). Performance contours showed the same upward shift as is seen in the third experiment, and the similarities of Figs. 3(c) and 3(e) are considered strong evidence that such performance differences as occur between monotic and dichotic stimulus conditions can be attributed to combination tones generated in the peripheral auditory system.

III. EXPERIMENTAL TESTS FOR MUSICAL BEHAVIOR

The purpose of the experiments reported in the previous section was to investigate melody perception, which was taken operationally as the ability of the human listener to retrieve the notes represented by a sequence of periodic sounds. While these experiments provide information about the conditions under which subjects can perform a certain interval identification task, they do not show that the subject "hears" melody. The subject could have learned to discriminate and identify eight stimuli merely on the basis of some unspecified, subjective criterion, rather than by recognizing each stimulus as a particular musical interval, although each subject asserted that the latter was true.

To obtain more objective evidence that the measured performance truly reflects the musical behavior of melody perception, a control experiment was carried out in which subjects were required to name previously unknown four-note melodies using standard musical notation. On each trial four periodic sounds at 50 dB SPL were presented in regular time sequence, each sound being a two-tone complex with frequencies at successive harmonics of some (missing) fundamental frequency in the range of 200–400 Hz. The lower harmonic number was chosen randomly from 3, 4, and 5 for each sound. Since each note could be represented by three different sounds, a four-note melody could be presented in 3^4 (i.e., 81) different ways. Four different four-note melodies were used, which were unknown to the subjects prior to the experiment. In the first part of the experiment, subjects could determine which of the four melodies was presented, and their task was to recognize all four by writing them down using a relative note scale. Since each melody could be represented in 81 different ways, subjects were allowed to listen to several presentations of each melody before answering. In the second part, all four of the four-note melodies used in the first part were presented at random, and subjects were asked to identify each melody presented by pushing the appropriate button on a response box. This experiment was performed both with monotically and dichotically presented tone complexes, using 10 subjects all of whom were familar with musical dictation. Since our principal interest was to investigate whether the subject's response reflected a natural musical behavior rather than some ingeniously acquired skill for performing a complicated task, very little time was spent on training.

The results of the monotic experiment were that nine out of ten subjects characterized each sequence of four sounds by the missing fundamentals in the first part and scored perfectly or near perfectly in the second part. The same was true for six subjects in the dichotic experiment; three had some difficulties in both parts, and the one subject who could not perceive consistent melodies in the first part under monotic conditions was also unable to do so for dichotic stimuli. The three subjects who participated in our previous basic experiments had no difficulties with either the monotic or

---

dichotic tests
that the phe
normal musi
learned skill
perception.

IV. EVIDEN
PERC

Since rand
one note to
musical pract
in melody re
in Sec. III if t
ically rather
choice of the
would lead t
frequency of
upper harmo
melody. A mo
the fundame
melodies. Su
sequences use
were not all
sequences us
melody, one
successive no
same harmo
similar to th
was performe
used, which
experiment.
four-sound s
but different
subject retrie
A number o
stimuli and t
associate wit
sequence w
fundamental
to one of th
melodies des

Helmholtz
sounds can
complex is p
or "analytic
each one ha
showed that
be controlled
indeed suppo
that listener
when this pr
a particular

V. RE
F

For well o
mental retri

dichotic tests. These results provide direct evidence that the phenomenon under study does indeed reflect normal musical behavior rather than a sophisticated learned skill that is not directly relevant to melody perception.

IV. EVIDENCE FOR TWO MODES OF MUSICAL PERCEPTION OF COMPLEX TONES

Since random variations in harmonic number from one note to the next are highly unusual in normal musical practice, we might wonder what would happen in melody recognition experiments like those discussed in Sec. III if the harmonics of each note varied systematically rather than randomly. As was noted earlier, a choice of the same harmonic number for every note would lead to a trivial situation where retrieving the frequency of either the missing fundamental or a single upper harmonic would lead to a report of the same melody. A more interesting stimulus would be one where the fundamental and harmonics describe different melodies. Such stimuli are given by any of the sound sequences used earlier for which the harmonic numbers were not all equal. Of the 81 possible four-sound sequences used in Sec. III to represent each four-note melody, one sequence was chosen for which no two successive notes were represented by sounds with the same harmonic numbers. A recognition experiment similar to the first part of the experiment in Sec. III was performed. Four different four-note melodies were used, which were unknown to the subjects prior to the experiment. Each melody was represented by one four-sound sequence. These stimuli provide consistent but different information depending on whether the subject retrieves the missing fundamental or a partial. A number of subjects were asked to listen to such stimuli and to report which note sequence they would associate with each of them. Some reported a note sequence which was consistent with the missing fundamentals; others reported a melody corresponding to one of the partials; one subject was able to report melodies described by both the upper and lower partial.

Helmholtz (1863) mentioned that complex periodic sounds can be perceived "synthetically," i.e., the complex is perceived as one sound, having one pitch; or "analytically," i.e., partials are heard individually, each one having its own pitch. Cross and Lane (1963) showed that these two modes of musical behavior can be controlled by previous training. Our experiments do indeed support these observations and show, moreover, that listeners can switch from one mode to the other when this provides them with relevant information for a particular task.

V. RELATION TO EARLIER WORK ON FUNDAMENTAL RETRIEVAL

For well over a century, the phenomenon of fundamental retrieval in the perception of periodic musical sounds has been studied actively and reported in terms of musical pitch perception evoked by single sounds with fixed harmonics (Seebeck, 1841; Helmholtz, 1863; Hermann, 1912; Fletcher, 1924; Schouten, 1940b; Thurlow and Small, 1955; deBoer, 1956; Flanagan and Guttman, 1960; Schouten, Ritsma, and Cardozo, 1962; Ritsma, 1962; Smoorenburg, 1970). Reviews of this literature have recently been given by Plomp, 1967; Small, 1970; Schouten, 1970; and Ritsma, 1970. We have chosen to perform experiments using sequences of sounds with random harmonic number because we found that successive notes evoked in subjects a sense of musical interval and provided a context for that feature of each sound that was being contrasted. In addition, this choice minimized the opportunities for behavioral responses that are directly correlated with tracking of partials. Finally, it enabled us to demonstrate directly the role of fundamental retrieval in musical behavior.

Despite our departures from earlier procedures, we believe that our investigations involve the same phenomenon as has been described in earlier work. De Boer (1956) reported pitch matches in which inharmonic complex tones comprising five or seven partials with uniform frequency spacing were aurally matched to periodic complex tones with a fundamental that differed systematically from the spectral spacing of the inharmonic sound. Schouten, Ritsma, and Cardozo (1962) produced similar data for AM complexes (three partials) which are reproduced in Fig. 4, and Smoorenburg (1970) found essentially the same results using complexes consisting of only two tones. All these tone complexes were presented monotically or diotically (i.e., one ear stimulated, or both ears identically stimulated). Using a musical-interval matching paradigm, we obtained very similar results with two-tone stimuli presented both monotically and dichotically (Houtsma and Goldstein, 1971). Two additional experiments were performed to examine the relation between the findings of pitch-matching experiments and our experiments on interval identification.

The first was an interval identification experiment with the same paradigm as was used in the basic experiments, except that for each note the two frequencies (with successive random harmonic numbers) were uniformly shifted up or down in frequency by some random increment less than one-fourth of the fundamental frequency. This limit was imposed to avoid pronounced pitch ambiguities that are to be expected for larger inharmonic frequency shifts from the data shown in Fig. 4. Stimuli were presented monotically at 50 dB SPL to one subject (AH) and at 40 dB to another (NH). We chose values for $\bar{n}$ and f_0 that had yielded perfect identification performance without the random frequency shift. The results are shown in Fig. 5. It is obvious that performance is considerably less than perfect, although the relevant

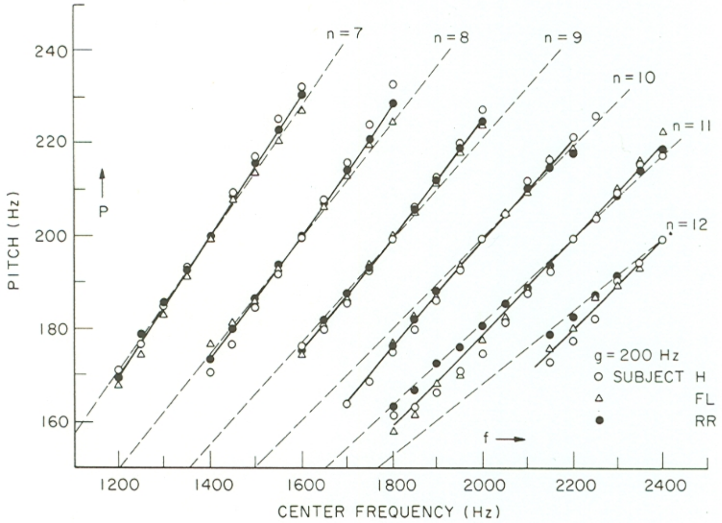

FIG. 4. Pitch as a function of the center frequency for a three-component complex tone. Stimulus intensity is 35 dB SL. The fundamental frequency, p, of an amplitude-modulated sound comprising the harmonic frequencies $(n-1)p$, np, and $(n+1)p$ is adjusted to match aurally the musical pitch of another amplitude-modulated sound comprising the inharmonic frequencies $f-200$ Hz, f, and $f+200$ Hz. The first effect is defined by the dashed lines of slope $1/n$. (From Schouten, Ritsma, and Cardozo, 1962, with permission from *The Journal of the Acoustical Society of America*).

message was preserved in the difference frequencies of the complex tones, indicating that inharmonic frequency shift of a complex tone does indeed alter its pitch. Performance was better than chance (12.5% correct) and improved with increasing harmonic number $\bar{n}$. This was predictable from the particular choice of musical intervals to be identified and the pitch uncertainty introduced by the uniformly distributed random frequency shift by using the simple decision model outlined in the legend of Fig. 5.

In the second experiment we examined the musical relevance of the empirical relation between pitch of a complex tone and inharmonic frequency shift of its partials by applying this relation to our musical-interval identification paradigm. The experiment differed from the previous one in that the inharmonic sounds were chosen to preserve the original musical intervals. For the first note a random frequency shift, positive or negative, was chosen with a maximum of one-fourth of the fundamental frequency; then the empirical "first effect" relation (Schouten, Ritsma, and Cardozo, 1962) shown in Fig. 4 was used to compute the inharmonic frequency shift for the second note required to give the original musical interval. Figure 5 shows the results; identification performance is almost perfect. Occasional mistakes are easily accounted for by the additional difficulty for the subject of having to identify eight intervals in a roving key, since intervals no longer started with the same note because of the random inharmonic frequency shift introduced in the first note.

Having established the relation between earlier work on musical pitch and our basic experiments, we make the following conclusions.

(1) Despite differences in experimental procedure, there is clear evidence that our studies on fundamental retrieval behavior using interval identification reflect the same basic phenomenon as the work of many other investigators on "periodicity pitch" or "residue pitch."

(2) Inharmonic tone complexes do have a definite musical value, provided that the deviation from the harmonic situation is not so large that it will cause pitch ambiguities. This value is, at least for the pitch resolution required in our experiments, well described by what is known in the literature as the "first effect" of inharmonic frequency shift.

(3) The large departures from the first effect which have been reported to occur under certain monotic and diotic stimulus conditions (de Boer, 1956; Schouten, Ritsma, and Cardozo, 1962; Ritsma, 1970; Smoorenburg, 1970; van den Brink, 1970) reflect only a stimulus modification attributable to aural combination tones (Goldstein and Kiang, 1968; Ritsma, 1970; Smoorenburg, 1970), rather than essential features of the central processor of musical pitch.

VI. DISCUSSION

The two currently popular theories of musical pitch perception are the frequency detection theory (Helmholtz, 1863; Fletcher, 1924) and the periodicity-detection theory (Schouten, 1940a; Licklider, 1962). According to the former theory, the cochlear spectrum analyzer maps the energy of the fundamental frequency tonotopically, and pitch is associated with spatial position of the fundamental; if the stimulus has no energy at the fundamental frequency, nonlinearities in the peripheral ear will introduce it. According to the

---

FIG. 5. Effect
quency shift on
tification perform
The solid line in
the random fre
predicted by a si
which assumes th
for each note is
with a rectangul
(2) the means of
2(a) are uniform
and (3) the first
ignored, while the
for maximum ex

latter theory, s
temporal perio
spectrum anal
temporal interv

All the dicho
prove directly
frequency nor
output are nece
fundamental f
sounds. This
detection theor
reconcile the es
mental (Seebec
tection theory
tones (Fletche
quate. Musical
with stimuli for
be ruled out be
ten, 1938; Plo
masking at the
Furthermore,
the musical
(Hermann, 191
Schouten, Rits

Considering
point, the resu
paper show, m
fundamental t
the fundament
strated, but pr
the cochlear
fundamental p
following reaso

(1) Monotic
fundamental p
dichotic stimu
monotic and d
are well accoun

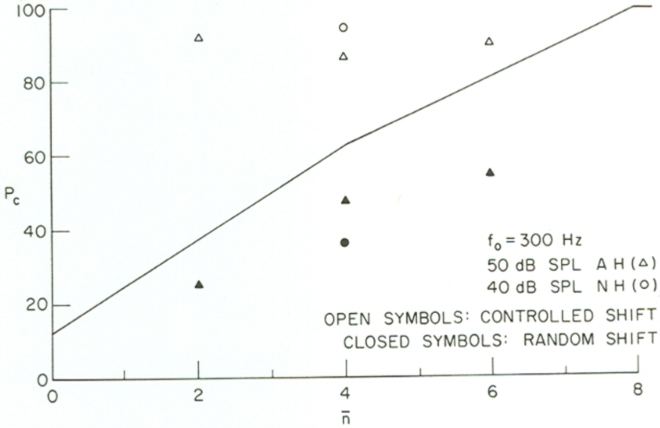

FIG. 5. Effects of inharmonic frequency shift on musical interval identification performance of two subjects. The solid line indicates performance for the random frequency shift paradigm predicted by a simplified decision model which assumes that (1) the musical value for each note is randomly distributed with a rectangular pdf of width $f_0/2\bar{n}$, (2) the means of the nine notes in Fig. 2(a) are uniformly separated by $f_0/16$, and (3) the first note of each interval is ignored, while the second note is detected for maximum expected correct answers.

latter theory, some mechanism is presumed to measure temporal periodicities at the output of the cochlear spectrum analyzer and then associate pitch with temporal interval.

All the dichotic experiments discussed in this paper prove directly that neither energy at the fundamental frequency nor fundamental periods in the cochlear output are necessary conditions for a subject to retrieve fundamental frequencies in a sequence of periodic sounds. This definite inadequacy of the periodicity-detection theory is a new finding. Earlier attempts to reconcile the established sensation of a missing fundamental (Seebeck, 1841) with Helmholtz's frequency-detection theory by means of aurally generated distortion tones (Fletcher, 1924) had already proved inadequate. Musical pitch perception could be established with stimuli for which difference tone distortion had to be ruled out because of low stimulus intensity (Schouten, 1938; Plomp, 1967; Goldstein, 1967) or external masking at the difference frequency (Licklider, 1954). Furthermore, difference frequency did not designate the musical pitch of inharmonic tone complexes (Hermann, 1912; Schouten, 1940b; de Boer, 1956; Schouten, Ritsma, and Cardozo, 1962).

Considering these theories from a converse viewpoint, the results of the experiments reported in this paper show, moreover, that a sufficient condition for fundamental tracking is indeed given by energy at the fundamental, which can very easily be demonstrated, but probably not by fundamental periods in the cochlear output. The data suggest that such fundamental periods are probably irrelevant for the following reasons:

(1) Monotic stimuli, which can provide cochlear fundamental periods, give no better performance than dichotic stimuli: performance differences between monotic and dichotic conditions at higher intensities are well accounted for by aural combination tones.

(2) With monotic stimuli the possibilities for fundamental periods in the cochlear output are enhanced as the harmonic number n is increased. Nevertheless, interval identification performance always deteriorates with increasing harmonic number.

(3) All monotic stimuli for which identification performance was better than chance either consisted of behaviorally resolvable tones, or generated such tones as combination tones (Plomp, 1964; Goldstein, 1967).

The similarity of the monotic and dichotic behavior and the close correlation between the limits on fundamental retrieval and behavioral frequency resolution strongly suggest that fundamentals of complex-tone stimuli are retrieved by means of a central mechanism which operates on those stimulus tones or combination tones that can be resolved in the cochlea. This conclusion is a radical departure from the theory of the "residue," which is defined as: "the joint perception of those higher Fourier components which the ear fails to resolve" (Schouten, Ritsma, and Cardozo, 1962). The experiments reported here imply that the unresolved *remainder* of the harmonic *series*, or *residue*, is not responsible for fundamental retrieval. The finding that performance for dichotic stimuli is bounded by harmonic number similarly as under monotic stimulus conditions cannot be accounted for by cochlear frequency resolution alone; the cause must be more central.

The finding that the only effective conveyors of musical pitch are those partials which can be resolved behaviorally is actually not new. Ritsma (1962, 1963, 1967a, 1967b) found that the existence region of the "tonal residue" is bounded by an upper harmonic number whose magnitude depends on the complexity of the stimulus and the experimental paradigm; in addition, he found that the most effective or dominant pitch conveyors are the harmonics 3, 4, and 5. Our finding that there is no noticeable drop in performance for harmonic numbers below 4 is not inconsistent with

Ritsma's findings, since the effect is probably too small to make performance in our experiments deteriorate below perfect (Bilsen, 1971). Smoorenburg (1970) concluded from his pitch matching data for two-tone stimuli that the effective harmonic numbers for fundamentals of 200 Hz have an upper bound of about 9; this conclusion is supported and extended to other fundamental frequencies by our results.

The effectiveness of the lower partials alone in conveying musical pitch is also consistent with some readily observed musical phenomena. First, there are some musical instruments, e.g., piano strings (Young, 1952), which produce sounds whose higher partials are inharmonic. From our results we conclude that this is not necessarily disturbing, since only the lower partials are effective in communicating the note. Second, Mach's (1881) observation—that dissonance is equally pronounced when the sounds from two mistuned simple tone sources are presented to one ear or to separate ears, and can therefore not be attributed to beating effects in the cochlea—is also quite consistent with our experimental findings. Simple tones whose frequencies are related by ratios of small integers are effective conveyers of musical pitch independently of whether they are presented monotically or dichotically. Presumably, stimulus conditions that evoke no definitive musical pitch contribute to a musician's judgment of dissonance (cf. Plomp and Levelt, 1965).

Little specific can be said at this point about the neural mechanism that mediates fundamental retrieval. Its close connection with behavioral frequency resolution suggests that neural signals that allow one to recognize simple tones and retrieve the missing fundamental of harmonic complex tones are common in early stages of processing. Whatever kind of neural mechanism is postulated, for complex-tone stimuli it will operate on neural signals derived from peripherally resolved tones. It is known that information about the frequency of such tones is preserved after the neural transformation in the form of the place of active nerve fibers and, at least for lower frequencies, in the temporal firing patterns of individual fibers (Kiang, 1965; Whitfield, 1970). From our work it appears that all the constraints could conceivably be met by mechanisms based on either time or place information. An important simplification for further work on musical pitch is offered by the empirical finding that fundamental retrievals for dichotically and monotically presented two-tone stimuli are essentially identical. Thus, the basic properties of the central processor of musical pitch can be investigated with the two-tone dichotic stimulus that avoids confounding and irrelevant phenomena caused by peripheral interactions among stimulus partials. Earlier efforts to find cochlea-generated place (von Békésy, 1961) and time (Schouten, 1940a) cues in response to the periodicity or missing fundamental of complex input stimuli now appear to be irrelevant; mechanisms more central than the cochlea must be investigated.

ACKNOWLEDGMENTS

The authors are deeply indebted to W. F. Kelley, D. J. Callahan, and L. D. Braida for their valuable assistance in developing the instrumentation and computer software that made these experiments possible. Informative discussions were contributed by H. S. Colburn, N. I. Durlach, D. M. Epstein, and W. M. Siebert. This work was supported by the National Institutes of Health (Grant 5 P01 GM14940-05 and Grant 5 T01 GM01555-05).

---

[*] This report is based on a PhD thesis submitted to the Dept. of Electrical Engineering, Mass. Inst. of Technol., June 1971, by A. J. M. Houtsma, and supervised by J. L. Goldstein. The basic results were presented at the 79th Meeting of the Acoustical Society of America in Atlantic City, April 1970, [A. J. M. Houtsma and T. L. Goldstein, J. Acoust. Soc. Amer. 48, 88(A) (1970)].

[†] Also at the Eaton Peabody Laboratory of Auditory Physiology, Massachusetts Eye and Ear Infirmary, Boston, Massachusetts 02114.

REFERENCES

von Békésy, G. (**1961**). "Concerning the Fundamental Component of Periodic Pulse Patterns and Modulated Vibrations Observed on the Cochlear Model with Nerve Supply," J. Acoust. Soc. Amer. *33*, 888–896.

Bilsen, F. A. (**1971**). "Pitch Perceptibility for Multi-Component Signals," J. Acoust. Soc. Amer. *51*, 113(A).

de Boer, E. (**1956**). "On the Residue in Hearing," PhD thesis, Univ. of Amsterdam.

de Boer, E. (**1956**). "Pitch of Inharmonic Signals," Nature *178*, 535–536.

van den Brink, G. (**1970**). "Two Experiments on Pitch Perception: Diplacusis of Harmonic AM Signals and Pitch of Inharmonic AM Signals," J. Acoust. Soc. Amer. *48*, 1355–1365.

Cross, D., and Lane, H. (**1963**). "Attention to Single Stimulus Properties in the Identification of Complex Tones," in *Experimental Analysis of the Control of Speech Production and Perception*," Univ. of Michigan ORA Rep. No. 05613-1-P.

Flanagan, J. L., and Guttman, N. (**1960**). "Pitch of Periodic Pulses without Fundamental Components," J. Acoust. Soc. Amer. *32*, 1319–1328.

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

Goldstein, J. L. (**1967**). "Auditory Nonlinearity," J. Acoust. Soc. Amer. *41*, 676–689.

Goldstein, J. L. (**1970**). "Aural Combination Tones," in *Frequency Analysis and Periodicity Detection in Hearing*, R. Plomp and G. F. Smoorenburg, Eds. (Sythoff, Leiden, The Netherlands).

Goldstein, J. L., Baer, T., and Kiang, N. Y. S. (**1971**). "A Theoretical Treatment of Latency, Group-delay and Tuning Characteristics for Auditory Nerve Responses to Clicks and Tones," in *Physiology of the Auditory System*, M. B. Sachs, Ed. (National Educational Consultants, Baltimore, Md.)

Goldstein, J. L., and Kiang, N. Y. S. (**1968**). "Neural Correlates of the Aural Combination Tone $2f_1-f_2$," Proc. IEEE *56*, 981–992.

---

von Helmholtz, H
dungen als physiol
(F. Vieweg & Sohn.

Hermann, L. (1912
ungstone," Arch. ge

Houtsma, A. J. M.
Musical Intervals:
of Complex Tones,"

Kiang, N. Y. S. (1
Cat's Auditory Ner

Licklider, J. C.
Pitch," J. Acoust.

Licklider, J. C.
Auditory Process

Mach, E. (**1881**).
zungsber. Kaisers
sensch., Klasse 92,

Miller, G. A. (1
Minus Two," Psy

Ohm, G. S. (**1843**)
geknüpfter Theor
Vorrichtungen," A

Plomp, R. (**1964**).
Soc. Amer. *36*, 162

Plomp, R. (**1965**
Tones," J. Acoust

Plomp, R. (**1967**)
Amer. *41*, 1526–1

Plomp, R., and
and Critical Band

Plomp, R., and
Analysis and Pe
Leiden, The Neth

Ritsma, R. J. (19
I," J. Acoust. So

Ritsma, R. J. (19
II," J. Acoust. S

Ritsma, R. J. (19
of the Pitch of
191–198.

Ritsma, R. J. (1
tion of Periodic
Rep. 2, Eindhove

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

Hermann, L. **(1912)**. "Neue Versuche zur Frage der Unterbrechungstone," Arch. ges. Physiol. *146*, 249–294.

Houtsma, A. J. M., and Goldstein, J. L. **(1971)**. "Perception of Musical Intervals: Evidence for the Central Origin of the Pitch of Complex Tones," Res. Lab. of Electron., MIT Tech. Rep. 484.

Kiang, N. Y. S. **(1965)**. *Discharge Patterns of Single Fibers in the Cat's Auditory Nerve* (MIT Press, Cambridge, Mass.).

Licklider, J. C. R. **(1954)**. "Periodicity Pitch and 'Place' Pitch," J. Acoust. Soc. Amer. *26*, 137(A).

Licklider, J. C. R. **(1962)**. "Periodicity Pitch and Related Auditory Process Models," Int. Audiol. *1*, 11–35.

Mach, E. **(1881)**. "Zur Analyse der Tonempfindungen," Sitzungsber. Kaisersl. Akad. Wissensch., Mathemat.-Naturwissensch., Klasse 92, II Abt.

Miller, G. A. **(1956)**. "The Magical Number Seven Plus or Minus Two," Psych. Rev. *63*, 81–97.

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

Plomp, R. **(1964)**. "The Ear as a Frequency Analyzer," J. Acoust. Soc. Amer. *36*, 1628–1636.

Plomp, R. **(1965)**. "Detectability Threshold for Combination Tones," J. Acoust. Soc. Amer. *37*, 1110–1123.

Plomp, R. **(1967)**. "Pitch of Complex Tones," J. Acoust. Soc. Amer. *41*, 1526–1533.

Plomp, R., and Levelt, W. J. M. **(1965)**. "Tonal Consonance and Critical Bandwidth," J. Acoust. Soc. Amer. *38*, 548–560.

Plomp, R., and Smoorenburg, G. F., Eds. **(1970)**. *Frequency Analysis and Periodicity Detection in Hearing* (A. W. Sijthoff, Leiden, The Netherlands).

Ritsma, R. J. **(1962)**. "Existence Region of the Tonal Residue. I," J. Acoust. Soc. Amer. *34*, 1224–1229.

Ritsma, R. J. **(1963)**. "Existence Region of the Tonal Residue. II," J. Acoust. Soc. Amer. *35*, 1241–1245.

Ritsma, R. J. **(1967a)**. "Frequencies Dominant in the Perception of the Pitch of Complex Sounds," J. Acoust. Soc. Amer. *42*, 191–198.

Ritsma, R. J. **(1967b)**. "Frequencies Dominant in Pitch Perception of Periodic Pulses of Alternating Polarity," IPO Ann. Progr. Rep. 2, Eindhoven, pp. 14–24.

Ritsma, R. J. **(1970)**. "Periodicity Detection," in *Frequency Analysis and Periodicity Detection in Hearing*, R. Plomp and G. F. Smoorenburg, Eds. (A. W. Sijthoff, Leiden, The Netherlands).

Sachs, M. B. **(1969)**. "Stimulus Response Relation for Auditory Nerve Fibres: Two-Tone Stimuli," J. Acoust. Soc. Amer. *45*, 1025–1036.

Sachs, M. B., and Kiang, N. Y. S. **(1968)**. "Two-Tone Inhibition in Auditory-Nerve Fibres," J. Acoust. Soc. Amer. *43*, 1120–1128.

Saunders, F. A. **(1937)**. "The Mechanical Action of Violins," J. Acoust. Soc. Amer. *9*, 81–98.

Saunders, F. A. **(1946)**. "Analysis of the Tones of a Few Wind Instruments," J. Acoust. Soc. Amer. *18*, 395–401.

Schouten, J. F. **(1938)**. "The Perception of Subjective Tones," Proc. Kon. Nederl. Akad. Wetenschap *41*, 1086–1093.

Schouten, J. F. **(1940a)**. "The Residue and the Mechanism of Hearing," Proc. Kon. Nederl. Akad. Wetenschap *43*, 991–999.

Schouten, J. F. **(1940b)**. "The Perception of Pitch," Philips Tech. Rev. *5*, 286–294.

Schouten, J. F., Ritsma, R. J., and Cardozo, B. L. **(1962)**. "Pitch of the Residue," J. Acoust. Soc. Amer. *34*, 1418–1424.

Schouten, J. F. **(1970)**. "The Residue Revisited," in *Frequency Analysis and Periodicity Detection in Hearing*, R. Plomp, and G. F. Smoorenburg, Eds., (A. W. Sijthoff, Leiden, The Netherlands).

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

Small, A. M., Jr. **(1970)**. "Periodicity Pitch" in *Foundations of Modern Auditory Theory*, J. R. Tobias, Ed. (Academic, New York), Vol. I.

Smoorenburg, G. F. **(1970)**. "Pitch Perception of Two-Frequency Stimuli," J. Acoust. Soc. Amer. *48*, 924–942.

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

Whitfield, I. C. **(1970)**. "Central Nervous Processing in Relation to Spatio-Temporal Discrimination of Auditory Patterns," in *Frequency Analysis and Periodicity Detection in Hearing*, R. Plomp, and G. F. Smoorenburg, Eds. (A. W. Sijthoff, Leiden, The Netherlands).

Young, R. W. **(1955)**. "Inharmonicity of Plain Wire Strings," J. Acoust. Soc. Amer. *24*, 267–273.

Zwicker, E. **(1965)**. "Der Ungewöhnliche Amplitudengang der nichtlinearen Verzerrungen des Ohres," Acustica *5*, 67–74.

Zwislocki, J. **(1953)**. "Acoustic Attenuation between the Ears," J. Acoust. Soc. Amer. *25*, 752–759.