Pitch identification and discrimination for complex tones with many harmonics

A. J. M. Houtsma and J. Smurzynski
Institute for Perception Research (IPO), P. O. Box 513, 5600 MB Eindhoven, The Netherlands

(Received 27 June 1989; accepted for publication 18 September 1989)

Four experiments are reported that deal with pitch perception of harmonic complex tones containing up to 11 successive harmonics. In particular, the question is raised whether the pitch percept of the missing fundamental is mediated only by low-order resolvable harmonics, or whether it can also be conveyed by high-order harmonics that the cochlea fails to resolve. Melodic interval identification performance was found to remain significantly above chance level even if the range of harmonics extended from the 20th to the 30th. Just-noticeable differences (jnd) in the pitch of the missing fundamental were found to increase with increasing harmonic order, but to level off when all harmonics are above the 12th. These results are consistent with the notion of the existence of two distinct neural pitch mechanisms in the auditory system, but are, in principle, also compatible with a single central-spectrum mechanism that uses the interspike interval histograms of auditory-nerve fibers as inputs.

PACS numbers: 43.66.Hg, 43.66.Ba, 43.75.Cd [WAY]

INTRODUCTION

The perceptual phenomenon "pitch of the missing fundamental" has been known for a century and a half since Seebeck (1841) first reported it. It concerns the ability of our auditory system to perceive pitches that correspond to the fundamental frequencies of complex tones while those fundamentals are physically absent. This curious phenomenon can be noticed when one listens to music through severely bandlimited channels like a small transistor radio or a telephone. One can still follow melodies of low-pitched passages without ambiguity, despite the total absence of acoustic energy at the fundamental frequencies of the notes. The phenomenon has received much attention in the psychoacoustical literature, theoretical as well as experimental. Reviews of this literature have been presented by de Boer (1976) and Scharf and Houtsma (1986).

The oldest explanation of the phenomenon, except perhaps Ohm's (1843), who dismissed it as an acoustic illusion, was based on the distortion properties of the auditory periphery producing a difference tone (Pipping, 1895; Fletcher, 1924; Hoogland, 1953). It became apparent, however, that the observed pitch did not always follow the difference-tone frequency (Schouten et al., 1962). The aural-distortion theory was gradually abandoned in favor of Schouten's residue theory (Schouten, 1940). This theory accounted for the percept of the missing fundamental as follows: (a) The cochlea performs a spectral analysis with limited frequency resolution; (b) unresolved partials yield a signal with a periodic or quasiperiodic envelope at the cochlear output; (c) auditory-nerve fibers trigger on peaks of the cochlear output signal, preserving the (quasi) periodicity of the signal's fundamental frequency in the neural firing pattern. The name residue theory underscored the idea that the fundamental pitch percept is based on those frequency components that the cochlea fails to resolve. Physiological support for the theory was the evidence of finite cochlear resolution (Békésy, 1943/1949; Galambos and Davis, 1943) and neural phase-locking to stimulus period in 8th-nerve fibers (Kiang et al., 1965; Brugge et al., 1969) and in cochlear-nucleus units (Möller, 1970).

There was also mounting evidence against the residue theory, however. As early as 1956, de Boer suggested, based on his experimental results, that two complementary pitch mechanisms might exist, one for closely spaced unresolved harmonics, and one for widely spaced resolved harmonics. Low-order harmonics, particularly the third through the fifth, were found to be the most effective conveyors of missing fundamental pitch (Plomp, 1967; Ritsma, 1967). The pitch percept turned out to be rather insensitive to the phases of partials, contrary to what the residue theory predicted (Patterson, 1973; Wightman, 1973a). Dichotic distribution of harmonics was found to yield the same pitch recognition scores as monotic or diotic distribution (Houtsma and Goldstein, 1972). Subsequently, formulated theories of pitch perception (Terhardt, 1972; Wightman, 1973b; Goldstein, 1973) were, accordingly, based on some form of central pattern recognition of aurally resolved tone components, rather than on central detection of a specific signal feature caused by tone interference in the auditory periphery.

This study is concerned with the question whether the abandonment of Schouten's residue principle in favor of a central pattern recognition principle is entirely justified. In particular, we will investigate the possibility of two neural mechanisms or, at least, two different modes of operation by which our auditory system extracts fundamental pitch from resolved low-order harmonics on the one hand, and from unresolved high-order components on the other. There are several reasons for such a dual-mechanism hypothesis. De-
spite the aforementioned evidence against the residue principle, it can be shown that a periodic pulse train retains some pitch quality even if all low-order harmonics that could be thought of as resolvable in the cochlea have been removed (Moore and Rosen, 1979). Hoekstra (1979) found that the just-noticeable difference (jnd) in the missing fundamental of an octave-band wide complex tone remains finite (about 5 Hz) when the missing fundamental becomes very low and the octave band contains many closely spaced harmonics. Stimuli with a flat long-term average spectrum but with a periodic temporal envelope, such as periodically gated or sine-amplitude-modulated white noise, are found to evoke weak pitch sensations (Miller and Taylor, 1948; Burns and Viemeister, 1976), although there is doubt about the effectiveness of such stimuli to convey frequency ratio or musical interval information (Houtsma et al., 1980; Houtsma, 1984).

There is also ample physiological evidence that, with a complex tone with only upper harmonics, correlates of the missing fundamental are clearly present in the firing patterns of 8th-nerve fibers (Brugge et al., 1969; Evans, 1983; Horst et al., 1986) and units of the cochlear nucleus (Evans, 1977; Kim et al., 1986). One therefore wonders why the central auditory nervous system would never use such information to retrieve the missing fundamental in the form of a pitch percept.

This study consists of two pitch identification and two pitch discrimination experiments, all with tone complexes that have up to 11 successive harmonics. In order to keep identification and discrimination results as comparable as possible, the same subjects are used and experimental conditions are kept closely matched. Pitch identification and discrimination behavior is studied as a function of (a) the harmonic order of the stimulus (i.e., the number of the lowest harmonic present in the stimulus), (b) the total number of the harmonics present in the stimulus, and (c) different phase conditions of the stimulus harmonics. The four experiments are subsequently reported and their results discussed. In a general discussion, their mutual implications are explored and conclusions are drawn.

I. EXPERIMENT I

A. Procedure

Four musically experienced subjects listened to pairs of sequential complex tones comprising only upper harmonics. Two of these subjects were the authors of this paper. All four had also considerable experience with pitch experiments on complex tones.

The missing fundamental of the first tone was always 200 Hz; the fundamental of the second was one out of seven possible frequencies: 211.9, 224.5, 237.9, 252.0, 267.0, 282.9, and 299.7 Hz, presented in random order. The stimuli thus contained all seven melodic intervals between a minor second and a perfect fifth, all in the upward direction. Subjects had to identify the melodic interval or, equivalently, the last note they heard by pressing the appropriate key on a keyboard.

All complex tones had 11 successive harmonics of equal amplitude added in sine phase. For each sound (or note), the lowest harmonic number \( N \) was chosen at random from a range of three successive integers. The middle of this range for the second note of each pair, referred to as \( \bar{N} \), was the independent variable. Sounds were 512 ms in duration, including a 40-ms linear rise and fall. Between sounds of a pair, there was a 500-ms silent interval and after each pair an indefinite response interval. The subject's response triggered feedback of the correct answer and presentation of the next melodic interval.

The lower harmonic numbers \( N \) were randomized in order to remove pitch cues formed by the upper and lower spectral edges. If \( N \) were kept constant, these edges would trace exactly the same melodic interval as the (missing) fundamental, which would render interpretation of melodic interval identification results rather ambiguous. Randomization of harmonic order does cause some "smear" in the value of the independent variable but, as long as the randomization range is kept small, the average \( \bar{N} \) of this range remains a meaningful parameter (Houtsma and Goldstein, 1972).

Stimuli were synthesized digitally on a Microvax II computer using a 16-bit D/A converter operating at a sampling rate of 20 kHz. They were presented diotically through Etymotic ER-2 insert earphones. A 30-dB SL pink-noise signal formed a constant background to the tone pairs that were raised 20 dB above masked threshold. The noise was added to mask possibly confounding aural combination tones.

B. Results

Identification runs were made for the conditions \( \bar{N} = 7 \) through \( \bar{N} = 19 \) in increments of three. For each value of \( \bar{N} \) five runs of 63 trials were taken for each subject. The results are presented in Fig. 1 that shows the percentage of correct melodic interval identifications plotted against the average lowest harmonic number \( \bar{N} \). Each data point shown represents the average score of five runs of a subject. The dashed function represents the subject-averaged score. One can see that scores drop sharply from perfect at \( \bar{N} = 7 \) to a more or less asymptotic value of about 60% correct for values of \( \bar{N} > 13 \). This score level is still far above chance level, how-

![Fig. 1. Melodic interval identification scores of four subjects plotted against the average lowest harmonic number \( \bar{N} \) of 11-tone harmonic complexes. The drawn function represents subject-averaged scores.](image-url)
ever, which is one-seventh or 14.3% correct (if one ignores the possible additional spectral edge clue that is present one-third of the time, when lower harmonic numbers of successive complex tones are identical).

Seven-by-seven confusion matrices for each value of $\bar{N}$, not shown in this paper, revealed that, even for large numbers of $\bar{N}$, responses were tightly clustered around the main diagonal. This implies that fundamental pitch, evoked by groups of 11 high-order harmonics, is still sufficiently clear to allow recognition with an accuracy of about a semitone. More specific measurements of this accuracy are carried out later in Exp. III.

C. Discussion

The most interesting feature of the results shown in Fig. 1 is the sharp and monotonic score drop in the region $7 < \bar{N} < 13$. This implies that, for complexes containing low-order harmonics, the salience of the fundamental pitch percept degrades when harmonic order increases (i.e., when less and less aurally resolved harmonics are present). This result is entirely consistent with earlier findings (Houtsma and Goldstein, 1972). The new element in the data of Fig. 1 is that performance does not degrade all the way down to chance level as harmonic order increases, but reaches an asymptotic level well above chance and independent of harmonic order.

The existence of two clearly different regions of harmonic order may very well reflect the presence of two distinct and separate pitch mechanisms in the auditory system. On the other hand, the two different kinds of behavior may also be mediated by a single neural mechanism that simply behaves differently for aurally resolved and for unresolved frequencies.

The apparent fact that melodic interval identification with 11-tone complexes does not degrade to chance level for large $\bar{N}$ is somewhat surprising since a very similar experiment with two-tone complexes did show such a degradation (Houtsma and Goldstein, 1971, 1972). One obvious difference between the two experiments is, of course, the number of successive harmonics used to represent notes. It was therefore decided to examine the effect of the number of harmonics in a complex tone on the salience of its (fundamental) pitch. This is done in the next experiment.

II. EXPERIMENT II

A. Procedure

The experimental procedure was the same as the one used in Exp. I, except that the total number of successive harmonics in each sound $M$ was made variable and that the average lowest harmonic number $\bar{N}$ was kept fixed at the values 10 and 16. The actual values of $M$ investigated were 2, 3, 5, and 7, to which were added the results for $M = 11$ from the previous experiment. Two subjects, who also participated in Exp. I performed five runs of 63 trials for each of the eight conditions on $M$.

B. Results and discussion

The results are shown in Fig. 2. Three features of the data are worth noting. The first is that identification performance for $\bar{N} = 10$ is uniformly better than for $\bar{N} = 16$, the second that in both cases performance degrades with decreasing $M$, and the third that this degradation of performance is more abrupt for $\bar{N} = 10$ (somewhere around $M = 4$) than for $\bar{N} = 16$. This behavior may be explained as follows.

When the lowest harmonic number of a complex tone equals 10, some combination tones of the series $f_n = f_1 - n(f_2 - f_1)$, where $f_1$ and $f_2$ are the 10th and 11th harmonics (Goldstein, 1967), may very well be audible despite the noise background that was present. For harmonic tone complexes, these combination tones simply act as other low-order harmonics (9th, 8th, etc.), so that the lower end of the tonal spectrum may be partially resolved in the auditory system. If the number and the strength of these aural combination tones decreases when the number of primary frequencies becomes less than about 5, the number of aurally resolvable frequencies will decrease as well, resulting in a poorer pitch identification score for $M < 5$. If the average lowest harmonic number $\bar{N}$ equals 16, however, all tone components including possible aural combination tones will be beyond the limit of aural frequency resolution. Pitch perception and identification, relatively poor as it may be, must then be mediated by a neural mechanism that operates on temporal properties of the signal as a whole. The rather smooth and monotonic score function is consistent with the fact that periodicity of a complex tone signal becomes better defined, particularly in a noise background, when more harmonics are present. Finally, the fact that both score functions degrade when $M$ decreases provides a general consistency with our earlier 2 and 3-tone results (Houtsma and Goldstein, 1971), although the present scores for the case $M = 2$ are somewhat higher than those found in the earlier study.

III. EXPERIMENT III

The purpose of this experiment was to investigate the basis of limitations to pitch identification performance ob-
observed in the data of Exps. I and II. If limitations are purely perceptual (i.e., due to vagueness or ambiguity of the pitch percept), the same trends observed in the data of Fig. 1 should be present in just-noticeable pitch difference functions measured with comparable sounds.

### A. Procedure

Differential thresholds for the (missing) fundamental were measured with an adaptive two-interval forced-choice two-down, one-up procedure (Levitt, 1970). Two complex tones, each having 11 successive harmonics and a duration of 512 ms with a 500-ms silent interval in between, were presented diotically in each trial. The value of $N$, defined the same way as in Exps. I and II, was the independent variable. In order to make sure that in none of the trials subjects could use the spectral edges as discrimination cues, the actual values of $N$ were not allowed to be the same for the two sounds of one trial. One complex tone had a fundamental frequency of 200 Hz, the other a frequency of $200 + \Delta f$ Hz. The temporal order of the tones was random with equal probabilities. Four subjects, the same as in Exp. I, were instructed to indicate whether the missing fundamental (or perceived pitch) went up or down by pressing the appropriate key on a keyboard. Visual feedback of the correct response was provided immediately following the subject's response, after which the next trial was presented.

The fundamental frequency difference $\Delta f$ of the initial pair of each adaptive run was chosen at about twice the expected difference limen (DL) estimated from pilot experiments. The initial step size was 0.2 Hz but was halved after the first five reversals. The DL was estimated from the mid-run average of ten reversals with the 0.1-Hz step size.

### B. Results

Five DL estimates were collected from each subject for $N$ values of 7, 10, 13, 16, 19, and 25. The results, averaged for the four subjects, are shown in Fig. 3 by the solid function. Each data point along this function represents an average of 20 adaptive runs. Vertical bars indicate ranges of ± one standard deviation of the five-run average scores of individual subjects. (The other function shown in Fig. 3 is discussed in the next experiment.)

### C. Discussion

The first thing to notice in the solid function of Fig. 3 is its rather sharp rise between $N = 7$ and $N = 13$, and its leveling off to the rather constant value of about 5 Hz for $N > 13$. This behavior is exactly the inverse of what is seen in Fig. 1, and therefore very consistent.

The data of Fig. 3 are also quite consistent with results obtained by Hoekstra (1979) during forced-choice fixed-difference pitch discrimination experiments with 1/3-octave-filtered pulse trains. For tones of low harmonic order ($2 < N < 10$), he found thresholds just under 0.1%, which translates to 0.2 Hz for our case, for tones of high harmonic order ($20 < N < 100$) a threshold of 2% (4.0 Hz for our case), and in the region $10 < N < 20$ a steep transition. The slight differences in absolute threshold values could be attributable to the fact that Hoekstra did not randomize harmonic numbers, making the discrimination task perhaps somewhat easier than ours. The difference in harmonic number of the transition region simply reflects the fact that our $N$ represents the low edge and Hoekstra's $n$ represents the center of the spectrum.

Our data are also consistent with results of Cullen and Long (1986), who measured rate jnd's for high-pass-filtered pulse trains as a function of pulse rate and cutoff frequency. They found rate jnd's of about 0.6 Hz at a pulse rate of 200 Hz and without high-pass filtering, which is in good agreement with our present results at $N = 7$. At a cutoff frequency of 2.5 kHz, which is comparable to our case of $N = 13$, they found a rate jnd of about 4.5 Hz, which is also very close to our present finding. Finally, with increasing cutoff frequency, and therefore with a decreasing number of stimulus harmonics, they found that the rate jnd climbed from 4.5 to 15 Hz (at a cutoff frequency of 10 kHz). This is similar to what we found in Exp. II, where identification scores dropped monotonically with decreasing number of harmonics $M$.

Our results obtained for $N = 7$ also agree well with recent data by Moore and Glasberg (1988), who performed DL measurements for the missing fundamentals of complex tones in normal and impaired ears. If one considers their DLs for the case of a complex tone with a fundamental of 200 Hz and with harmonics 6–12 added in cosine phase, and if one averages them for the four normal ears they tested, one obtains a DL of 0.4% or 0.8 Hz. The fairly precise agreement with out data provides additional reassurance that also, in Moore and Glasberg's experiment, DLs for the missing fundamental were measured, and not frequency DLs for spectral edges. The latter could have been the case since, like Hoekstra, they did not randomize harmonic order.

In the same study Moore and Glasberg found that in normal ears, there was no consistent influence of phase on the size of the fundamental pitch DL. The largest value of "lowest harmonic" they tested was six, however, which is according to our findings so far well within the region of aurally resolved frequencies. It was thought interesting to find out whether this apparent phase insensitivity also holds
for complexes of nonresolvable frequencies (i.e., tones with \( N \) larger than 13). This is attempted in the next experiment.

**IV. EXPERIMENT IV**

**A. Procedure**

The experiment was in every respect identical to Exp. III, except for the phase relations between stimulus harmonics. Instead of sine-phase relations, used in the previous experiments, phases were calculated according to Schröder's (1970) formula:

\[
\phi_n = -\pi n(n + 1)/M,
\]

where \( \phi_n \) represents the phase of the \( n \)-th order harmonic and \( M \) represents the total number of harmonics of the stimulus, which is 11 in our case. This condition will be referred to as “Schröder-phase.” From masking experiments done with complex tones having this kind of phase relation (Smith et al., 1986), it appears that such stimuli cause a very different excitation pattern on the basilar membrane compared with complexes of comparable harmonics added in sine or in cosine phase. The same four subjects from Exp. III participated.

**B. Results**

Subject-averaged results are shown in Fig. 3 by the dashed function, with vertical bars indicating the ± one standard deviation range of individual mean scores. As can be seen, DLs for sine phase and Schröder phase are virtually identical up to \( N = 10 \). For \( N > 10 \), the curves begin to diverge until at \( N = 25 \) the DL for the Schröder phase is 70% larger than the DL obtained with the sine-phase condition.

**C. Discussion**

The results of Exps. III and IV, as shown in Fig. 3, support the idea suggested in the discussion of Exp. I that our auditory system distinguishes, at least from a behavioral point of view, between two regions of stimulus partials. There is the region of low-order, partially resolved harmonics that yields clear fundamental pitch percepts and small pitch DLs that are independent of the phase relations between partials. There is also the region of unresolvable partials that yields weak or ambiguous pitch percepts and relatively large pitch DLs that do depend on phase. Pitch behavior's independence of phase relations between aurally resolvable harmonics was reported by Houtsma and Goldstein (1971), Terhardt (1972), Patterson (1973), and Wightman (1973a). Influence of phase on the pitch percept of harmonic 3-tone complexes of order \( 9 < N < 14 \) was reported by Ritsma and Engel (1964), but could not be reproduced by Wightman (1973a).

It should perhaps be pointed out that the two phase conditions chosen in Exps. III and IV do not necessarily represent extreme cases. They were merely chosen as two rather convenient phase configurations for which some elementary masking data exist and on the basis of which one might expect different pitch behavior. It is quite possible that, with the proper phase manipulation, even large pitch DL differences can be found than those shown in Fig. 3.

**V. GENERAL DISCUSSION**

**A. Identification versus discrimination**

Performance functions of absolute identification and discrimination experiments on similar stimuli should, in general, have an inverse relationship as long as memory limitations play an insignificant role. The better identification performance is, the smaller the expected jnd’s and vice versa. Such an inverse relationship is clearly present in the results of Exps. I and III. With the aid of signal detection theory, it is even possible to predict one from the other and to see to what extent memory limitations in Exp. I are really negligible. A Thurstonian decision model similar to the one developed by Braida and Durlach (1970) was applied to the identification data of Exp. I, which were pooled for all four subjects. The average sensitivity \( d' \) between all intervals differing by a semitone (i.e., all adjacent intervals) was calculated. This \( d' \) represents the typical sensitivity for the pitches of two complex tones whose (missing) fundamentals are a semitone apart. From this \( d' \), one can easily compute the DL as that frequency difference for which \( d' \) equals unity by linear interpolation. Resulting DLs are shown in Fig. 4 as a function of \( N \) and marked as “EXP. I.” The DLs found directly in Exp. III, which were measured with a 2-12AFC adaptive procedure, were also transformed to equivalent DLs for a fixed one interval 2AFC paradigm by compensating for the 70.7% vs 75% correct threshold levels of adaptive versus nonadaptive tests and for the \( \frac{1}{2} \) ratio of \( d' \)'s between 1- and 2-interval paradigms. These transformed DLs are also shown on the same set of coordinates in Fig. 4, as the function marked “EXP. III.” As one can see, there is hardly any discrepancy between the two functions, except perhaps at \( N = 16 \), which strengthens our belief that exactly the same limitations of our auditory system are encountered in the identification paradigm of Exp. I and the discrimination paradigm of Exp. III.

One could question whether it is fair to use the term “pitch perception” for the case of the high-order harmonics, as was done throughout this paper. Plomp (1976) rightly pointed out that subject's ability to discriminate a higher from a lower tone or to match a tonal comparison sound to
some complex test tone does not necessarily imply that such
a test sound evokes pitch in the musical sense. That argu-
ment holds, in principle, also for our discrimination exper-
iments (Exps. III and IV). The musical task and context of
the interval identification experiments I and II, however, do
support the notion that the identification clue was musical
pitch. Although subjects were never formally asked what
cue they used, both author-subjects have little doubt about
the answer. Moreover, Houtsma (1984) has shown with a
formal experiment that, for some complex sounds, simple
up-down “pitch” motion detection can be very good, while
pitch interval identification is very poor. The well above
chance level scores obtained with the musical interval identi-
fication paradigm of Exps. I and II provide a reasonable
assurance that also for high-order harmonic complex tones
we may speak of pitch sensations in the musical sense. The
similarity of transformed thresholds shown in Fig. 4 then
implies that the same may be said for the discrimination
experiments III and IV.

B. One or two neural pitch mechanisms?

The data represented in Figs. 1 and 3 strongly suggest a
behaviorally based separation of a complex tone’s harmonics:
those of low order and those of high order, with the
separation somewhere between the 10th and 13th harmonic.
One could argue that this observed behavior implies that
there are two distinct and separate neural pitch mechanisms
in the auditory system. One mechanism would operate on
neural signals, stochastically derived from frequencies that
are resolved in the cochlea, by finding a best-matching har-
monic template (Houtsma and Goldstein, 1972; Goldstein,
1973), or, equivalently, by finding a best-fitting common
subharmonic (Terhardt, 1972). Such a mechanism would
look only at frequency, and would be rather insensitive to
amplitude and phase differences among harmonics. It would
be the primary pitch mechanism because it apparently leads
to the smallest DLs and to perfect identification of complex-
tone pitches that are separated by a semitone or more. The
other mechanism would operate on neural transformations
of harmonic clusters that are not resolved in the cochlea,
similar to the “residue” mechanism proposed by Schouten
(1940). This mechanism is a secondary or backup mecha-
nism: Because of the vagueness and ambiguity of the pitch
percept and the large pitch DLs it yields, it cannot account
by itself for normally observed musical behavior. Because
neural transformations of unresolved harmonic clusters are
expected to be sensitive to phase relations between the har-
onics, this secondary pitch mechanism should be phase
sensitive, which was found to be the case. Such a dual pitch
mechanism was, in more general terms, already proposed by
de Boer (1956).

On the other hand, one can argue that there is only one
neural pitch mechanism in the central auditory system that
yields different performance dependent on its neural input.
As an example, let us take the central spectrum pitch proces-
sor proposed by Srulovicz and Goldstein (1983) as our cen-
tral mechanism. In this processor, a central spectrum magni-
tude at each frequency $f$ is determined by the response of the
auditory-nerve fiber with characteristic frequency $f_c$ that
matches the frequency $f$. The interspike interval histogram
(ISIH) from each fiber, in response to a tone burst, is passed
through a filter matched to the characteristic frequency $f_c$
of the fiber (i.e., is multiplied by a temporal weighting function
$\cos^2\pi f_c t$ and then integrated to yield a single-valued output
for the central spectrum at the frequency $f$).

If the ear’s input is a complex tone of low-order resolvable
harmonics, the central spectrum will be approximately
represented by the sum of responses to individual harmonics,
except for some synchrony-suppression interactions. A
case of a five-tone harmonic complex, comprising harmonics
2–6, is, in fact, worked out in the cited paper by Srulovicz
and Goldstein, showing a central spectrum with five distinct
and rather sharp peaks at the five harmonic frequencies.
From this central spectrum, the missing fundamental or
pitch is computed by matching it with a harmonic template
(Goldstein, 1973).

If the stimulus is a tone complex with only high-order
unresolvable harmonics, added in sine or cosine phase, 8th-
nerve fibers responding to such a complex typically exhibit
ISIHs in which the interval $\tau = 1/f_0$ is clearly dominant,
where $f_0$ is the complex’s missing fundamental frequency
(Horst et al., 1986). The matched filters of those fibers that
are tuned to multiples of $f_0$ will catch this interval compo-
nent of the ISIH and fully count it in their respective contrib-
utions to the central spectrum. The matched filters of those
fibers not tuned to multiples of $f_0$, however, will not count
these $\tau = 1/f_0$ intervals. The central spectrum, therefore,
will again show periodic peaks at frequencies that are multi-
plies of $f_0$, from which $f_0$ itself can be found with a harmonic
template as before. If instead of sine and cosine phases, other
phase relations are chosen that diminish the crest factor of
the acoustic signal, the ISIHs of 8th-nerve fibers exhibit a
more even distribution of intervals at the expense, of course,
of the 1/f_0 interval (Horst, 1989). This will result in less
pronounced periodic peaks in the central spectrum, leading
to a less salient pitch percept, as was found to happen in Exp.
IV. It must be pointed out, however, that the Srulovicz–
Goldstein model does not seem very efficient for complex-
tone stimuli with high-order harmonics. Whereas measured
ISIHs show that the 1/f_0 interval is very dominant in just
about every fiber that responds to the tone, the central mecha-
nism uses all this information only very indirectly.

Arguments used so far, either in favor of a single or two
separate neural pitch mechanisms, have been very qualita-
tive. From the work of Horst et al., it has become evident
that models of 8th-nerve timing behavior for one or two-tone
stimuli (Siebert, 1970; Colburn, 1973; Johnson, 1974), used
by Srulovicz and Goldstein to predict pitch behavior for
some complex tones, should not be used for complexes that
have high-order partials. A more thorough quantitative
treatment, in which expected performance limits are com-
puted from the stochastic properties of neural data, will have
to wait until 8th-nerve behavior for complex-tone stimula-
tion has been more fully explored and modeled.

ACKNOWLEDGMENTS

This study was made possible by a postdoctoral fellow-
ship grant from Eindhoven University of Technology, which
allowed Dr. Smurzynski to spend a year at IPO. The dedicated and patient assistance of subjects Michelle Heinen and Niek Versfeld is gratefully acknowledged.


