Pitch, periodicity, and auditory organization

William M. Hartmann
Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824

(Received 12 June 1996; accepted for publication 2 October 1996)

The perception of pitch forms the basis of musical melody and harmony. It is also among the most precise of all our human senses, and with imagination, this precision can be used experimentally to investigate the functioning of the auditory system. This tutorial presents auditory demonstrations from the zoo of pitch effects: pitch shifts, noise pitch, virtual pitch, dichotic pitch, and the pitches of things that are not there at all. It introduces models of auditory processing, derived from contemporary psychoacoustics and auditory physiology, and tests these models against the experimental effects. It concludes by describing the critical role played by pitch in the important human ability to disentangle overlapping sources of sound. © *1996 Acoustical Society of America.*

PACS numbers: 43.10.Ln, 43.66.Ba, 43.66.Hg [DWM]

Some sounds are higher pitched, being composed of more frequent and more numerous motions.
Euclid (330–275 BC)

INTRODUCTION

This article is adapted from a tutorial entitled, ''Pitch, Periodicity, and the Brain,'' given at the 131st meeting of the Acoustical Society of America, at Indianapolis in the spring of 1996. Like others in this series, this tutorial presented a specialized topic in acoustics to an audience of acousticians from diverse fields. The Indianapolis tutorial was devoted to the psychoacoustical topic of pitch perception, and that is the topic of the present text, although many other aspects of human hearing, both physiological and psychological, are involved. Much of the tutorial and the article has been taken from a book, to be published by the AIP Press, called *Signals Sound, and Sensation*; this book goes into many of these topics more deeply than can be done in this brief review. The tutorial emphasized demonstrations, which, unfortunately, are not presented in this article. The demonstrations are described in words here. Many of them are not difficult to do and interested readers may want to try them.

It is evident that the pitch of a tone has a lot to do with its frequency. In fact, in informal discourse, pitch and frequency are often confused. The tutorial began by sneaking up on the subject of pitch by first dealing with the perception of frequency. Complex tones are composed of many frequencies, but in a sine tone there is only a single frequency. Therefore, it is logical to start with the perception of frequency for a sine tone.

I. FREQUENCY DISCRIMINATION

The most fundamental measurement that can be made on frequency perception is the just-noticeable difference or difference limen (DL). This is the smallest change in frequency that can be reliably detected by listeners. There are two parameters in a sine-tone DL experiment. The most important is the frequency range of the tone; the second is the level.

Wier *et al.* (1977) measured the frequency DL in a forced-choice task. They presented listeners with two 500-ms tones in succession. The frequency of one tone was slightly higher than the frequency of the other, and the listeners had to say whether the first or the second tone was higher.

The results of an experiment of this kind can be represented as a psychometric function with the frequency difference on the horizontal axis and the percentage of correct responses on the vertical axis. When the frequency difference is near zero listeners cannot tell the difference between the tones; they can only guess, and their scores are 50% correct, corresponding to random guessing. When the frequency difference is large listeners can always get the right answer. Therefore, the psychometric function extends from 50% correct to 100%. Halfway between guessing and perfect is 75%, and the frequency difference for which responses are 75% correct is taken to be the DL in frequency.

The difference limens obtained by Wier *et al.* (1977) are shown in Fig. 1, and there are several things to notice there. The first is that once the sensation level (the level above threshold for detecting the tone) gets greater than about 10 dB the difference limen hardly depends at all on the level of the tone. In fact, careful studies like this show that the difference limen always decreases as the level increases, but, except for the lowest levels, the improvement is small.

The next thing to notice is that the frequency DLs are quite small. In the best frequency range, from 1 to 2 kHz, DLs are normally 0.2%. That corresponds to distinguishing 1000 Hz from 1002 Hz, and some trained listeners can reduce the difference to only 1 Hz. A DL of 0.2 % corresponds to a musical interval of about three one-hundredths of a semitone, or three cents, as shown in Fig. 2. To give an impression of what a DL sounds like, the first demonstration illustrated a frequency difference of 0.2%.

Demonstration 1 presented two successive tones. The first had a frequency of 1000 Hz and the second had a frequency of 1002 Hz. These two tones are discriminable in a typical force-choice test.

The fact that the difference limen for frequency is so small is important from the point of view of information transmission. A communications channel where the receiver has good discrimination is capable of transferring informa-

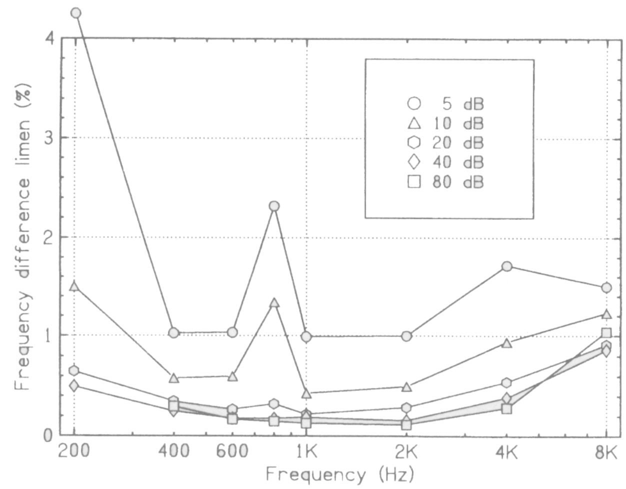

FIG. 1. Frequency difference limens (DLs) for sine tones with eight frequencies and five sensation levels given in dB. DLs are given as a percent of the reference frequency indicated on the horizontal axis. Keen acuity in discriminating frequencies is indicated by a small frequency DL. The peak at 800 Hz for low levels is thought to have no significance. It does not appear in other studies. By contrast, other studies agree with the rest of the data (adapted from Wier *et al.*, 1977, courtesy AIP Press).

tion efficiently. The idea is that if a receiver is capable of discriminating 1000 different signals, then when it receives one particular signal there are 999 things that the signal is not. That constitutes about 10 bits of information $2^{10}=1024\approx 1000$).

In order to determine the amount of information carried in the perceptual channel that encodes frequency, we only need to integrate the reciprocal of the DL. Letting N be the number of discriminable values within a frequency range, and letting Δf be the DL,

$$N=\int_{f_{\mathrm{bot}}}^{f_{\mathrm{top}}}\frac{df}{\Delta f}. \tag{1}$$

Weir *et al.* fitted their DL data to an analytic function of frequency,

$$\log_{10}(\Delta f)=a\sqrt{f}+b, \tag{2}$$

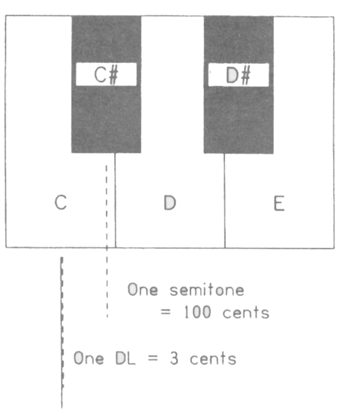

FIG. 2. Starting from the solid line on the left, the size of a difference limen (distance to the first dashed line) is compared with the size of a semitone on the familiar keyboard (distance to the second dashed line).

and, fortunately, the reciprocal of Δf is analytically integrable. For f_{bot} and f_{top} equal to 100 and 10 000 Hz, approximating most of the audible frequency range, N comes out to be 2034 for a and b parameters characteristic of an 80-dB sine. That is essentially 11 bits of information ($2^{11}=2048$).

The calculation of information capacity in this perceptual channel might be extended one more step by considering the time it takes to transmit the information. Frequency perception becomes more precise as the tone duration increases, up to a duration of about 50 to 100 ms. As tones are made longer than 100 ms, the DL for frequency decreases only slowly if at all. Therefore, the frequency channel can transmit 11 bits of information in about 0.1 s. There are temporal interaction effects that prevent us from taking the next step and concluding that this channel can transmit 110 bits per second. Still, we know that this channel has a high information capacity. Our ears are acutely tuned to fundamental-frequency inflections in the human voice and in music, as suggested by the second demonstration. *Demonstration 2* showed how the expressive variations in the fundamental frequency of human vocalization convey meaning. With the same level and duration, the recorded talker said "Oh?...Oh!...Oh."

It should be noticed that this entire discussion of the frequency difference limen has taken place without any mention of the word "pitch." Although it is believed that the way that listeners do the frequency discrimination task is on the basis of pitch, that assumption was in no way necessary in the measurement of the frequency DL.

We next make the common assumption that it is *pitch* perception that is at the basis of our exquisite sensitivity to frequency changes. The fact that trained listeners can distinguish 1000 Hz from 1002 Hz means that we can expect to get four significant figures of accuracy out of pitch experiments if we are careful. Before moving on to the science of pitch we ought to take some steps in the direction of defining the word.

II. THE DEFINITION OF PITCH

According to the ANSI standard of 1994,

> *"Pitch is that attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from low to high. Pitch depends mainly on the frequency content of the sound stimulus, but it also depends on the sound pressure and the waveform of the stimulus."*

This definition is all right so far as it goes. The emphasis on frequency content distinguishes the pitch sensation from the loudness sensation. But this definition is not exactly what psychoacousticians mean when they talk about pitch. Normally we further restrict the word to refer to a low-to-high ordering on the scale that is used for melody in music. That is not to say that pitch is restricted to musical sounds. It does say that the psychological dimension implied by the term "pitch" is the same as the psychological dimension of melody. The distinction made by this additional restriction is that it attempts to separate pitch from timbre. For example,

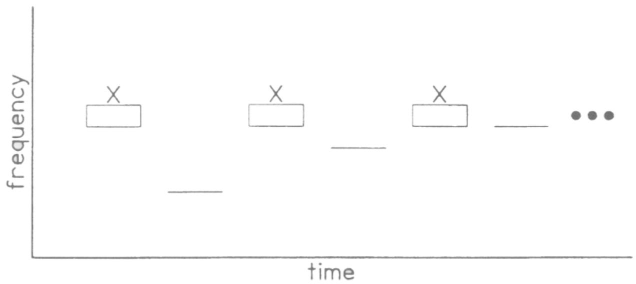

FIG. 3. Sequence of sounds to determine the pitch of X (courtesy AIP Press).

the sound /S/ is different from the sound /SH/. They differ because /S/ is higher than /SH/, and yet we would not normally say that /S/ has a higher pitch than /SH/; we say that it has a brighter timbre. The dimension of highness and lowness for these broadband-noise sounds is not the dimension of melody.

Better than a verbal definition of pitch is an operational definition. We say that a sound has a certain pitch if it can be reliably matched by adjusting the frequency of a sine wave of arbitrary amplitude. Figure 3 shows how this works. The sound, called X, whose pitch is in question alternates in a sequence with a sine tone. It is the task of the listener to adjust the frequency of the sine to match sound X in the melodic sense. This operational definition leads to a unit for measuring pitch; pitch is measured in units of Hertz.

There is a question about the level of the matching sine. For various tasks it might be different values. The listener might even adjust the level along with the frequency. But if we want to follow the convention, we use a sine tone with a level of 40 dB SPL (Fletcher, 1934; Zwicker and Fastl, 1990, p. 105).

This operational definition of pitch has several features: First, we are prepared to accept variability. If a listener matches a sound 10 times, we can make a histogram of the matches and say that the pitch is, for example, 400 Hz plus or minus 2 Hz. It might also happen that the distribution of pitch matches is bimodal, and in that case we are prepared to say that the pitch is ambiguous or that there are two pitches.

Naturally, this operational definition of pitch allows for individual differences. Then we conclude that a given *stimulus* has a certain pitch only if a great majority of listeners make similar matches.

The choice of measuring pitch in units of frequency is not an obvious one. It would be more sensible to measure pitch along a dimension that scales with the perceived magnitude of the pitch sensation, like the sone scale for loudness. Such pitch scales have been proposed in the past, and they are generally known as mel scales (Stevens *et al.*, 1937). However, the inability of researchers to agree on a single mel scale combined with strong competition from the logarithmic scale used in musical practice has resulted in a situation where the linear frequency scale is used by default.

III. PITCH OF SINES AND NOISE

There are two reasons for studying the pitch of sine tones. The first is that it is a powerful way to study the fundamental encoding process in the auditory system, particularly in the peripheral auditory system where a sine tone leads to the simplest possible excitation. The peripheral system, in turn, is interesting because more is known about its physiology than about other parts of the system, and correspondence between physiology and perception can be made. The second reason for studying the pitches of sine tones is that the pitches of complex tones, as in speech and music, appear to be determined by the pitches of their constituent harmonics, and the harmonics are sines.

A model of pitch perception is an attempt to understand pitch from elementary principals. It tries to explain why a signal with certain physical characteristics leads to a particular pitch sensation. There are two broad classes of pitch perception models, place models and timing models, and both of them are connected to auditory physiology. We introduce those models by applying them to sine tones.

A. The place theory of pitch

Place theories of pitch perception begin with the mechanics of the inner ear, or cochlea. The cochlea is a snail-shaped cavity in the temporal bone that is filled with fluids and is divided into several canals by membranes, as shown in Fig. 4.

At its large end, the cochlea is connected to the middle ear by means of the oval window. At the other end is the apex where the membranes terminate. The most important part of the cochlea is the basilar membrane, which stretches for about 35 mm along the cochlear duct. Distributed along the length of the basilar membrane are hair cells that transform sound vibrations into neural electrical impulses. The hair cells are, as it were, the brain's microphones. When the middle ear causes the cochlear fluids to move, the hair cells generate electrical spikes. These spikes are then transmitted along the auditory nerve as neural impulses to higher auditory centers in the brain. Therefore, the hair cells are not only microphones, they are like microphones coupled to analog-to-digital converters.

The basilar membrane is organized tonotopically. Low-frequency tones lead to membrane motion near the apex. High-frequency tones lead to motion near the oval window. Therefore, motion at different places along the basilar membrane leads to the excitation of different neurons in the auditory nerve. In this way, the frequency of a tone is represented in a code based on which neurons are active and which are silent.

All physiological studies show that the tonotopic organization of neurons is maintained throughout the ascending auditory system—all the way up to the auditory cortex. The frequency dependence established initially by the hydromechanics of the cochlea is retained at the highest levels. Nothing could be more logical than to assume that frequency discrimination, and pitch itself, is the result of this tonotopic encoding. This is the idea of place models of pitch.

To be more specific about place models, we can define the excitation pattern caused by a sine tone of given frequency and intensity. An excitation pattern represents the firing rate of neurons as a function of a tonotopic variable such as the place of origin on the basilar membrane. A hy-

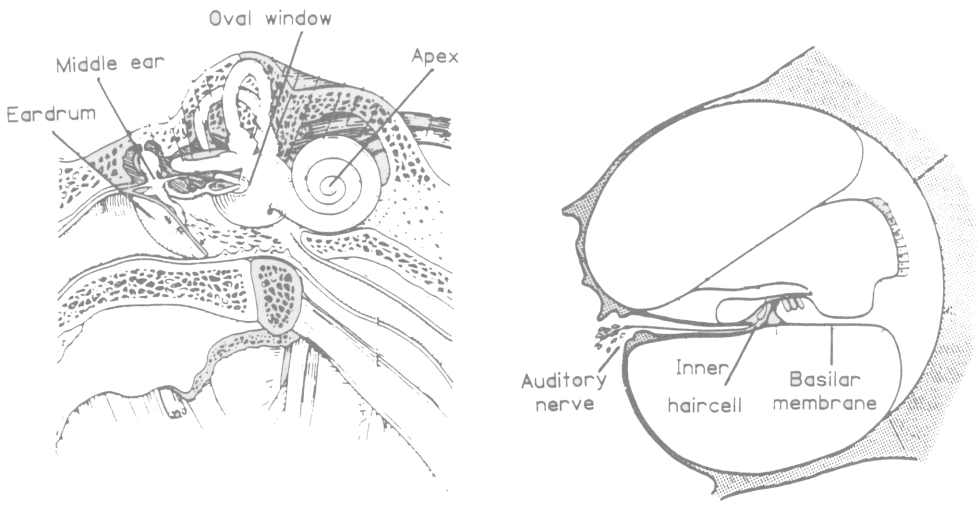

FIG. 4. (left) The snail-shaped cochlea gets an acoustical signal from the middle ear and converts it into neural spikes on the auditory nerve. (right) A cross section of the cochlea shows the three canals and the basilar membrane. Hair cells are distributed along the length of the basilar membrane.

pothetical excitation pattern is shown in Fig. 5.

The peak of the excitation pattern moves as the frequency changes. A combination of physiological measurements and psychoacoustical experiments has led to cochlear mappings giving the location of the peak along the basilar membrane as a function of the frequency of the tone. Greenwood (1990) derived a simple formula for the position of the peak measured in millimeters from the apex,

$$z = 7.24 \log_e(1 + f/165). \tag{3}$$

This function is shown in Fig. 6. Recent measurements of the resolving power of the excitation pattern at the auditory periphery suggest that it is adequate to account for our precise sense of pitch.

Demonstration 3 illustrated Eq. (3) by exciting listeners' basilar membranes at seven places that, according to the equation, are equally spaced. The first tone was 85 Hz, leading to a displacement peak 3 mm from the apex. Successive tones were separated by half a centimeter, as shown in Table I.

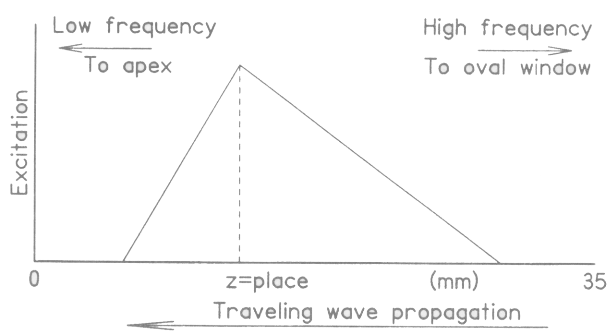

FIG. 5. An excitation pattern is the firing rate as a function of tonotopic coordinate. Here the tonotopic coordinate gives the place along the basilar membrane of the hair cell that originates the excitation. Tonotopic coordinates can be drawn for every site in the auditory system.

B. Evidence in favor of the place theory

The evidence for place models of pitch primarily comes from the malleability of the pitch sensation. The pitches of sine tones change in ways that might be explained by changes in peripheral excitation patterns. For example, the pitch of a sine tone is slightly different in each of the two ears. The effect is called "diplacusis," and it occurs in all normal ears. Diplacusis can be observed by using headphones that send a standard tone to one ear and then a matching tone to the other. When the subject matches the pitch of the standard by tuning the frequency of the matching tone, the tones in the two ears are found to have different frequencies. Different ears perceive pitch slightly differently because of small irregularities in individual cochleas (Brink, 1975a,b).

The pitch of a sine tone can be changed by changing the intensity of the tone. The effect can be seen in a matching or forced-choice experiment leading to a plot of pitch shift versus intensity called the pitch-intensity function. The pitch-intensity functions for particular sine tone frequencies and particular listeners show individual differences due to small cochlear differences. But averaged over a frequency region, there is an orderly behavior that has come to be known as Stevens rule (Stevens, 1935). Stevens rule is illustrated in Fig. 7. It says that when the intensity increases, the pitch of a low tone goes down, but the pitch of a high tone goes up. This effect was shown in the next demonstration.

TABLE I. List of successive tones for Demonstration 3.

Place (mm)	Frequency (Hz)
3.00	85
8.00	333
13.00	829
18.00	1818
23.00	3790
28.00	7725
33.00	15 575

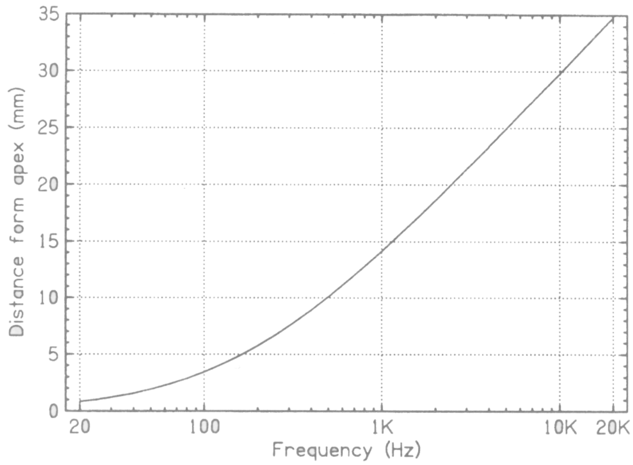

FIG. 6. Place of maximum excitation, referenced to the apex of a typical human cochlea, as a function of the frequency of the sine tone that generates the excitation.

Demonstration 4 presented two high-frequency tones in succession. The second tone was 30 dB more intense than the first, but the frequencies were equal. On the average, listeners should hear an upward change in pitch. The demonstration was presented three times with three different high frequencies (3, 4, and 5 kHz) because the effect is different for different frequencies, depending on the listener. Then two low-frequency tones with the same frequency were presented, again with the second tone 30 dB more intense. The demonstration was repeated with three different low frequencies (300, 250, and 200 Hz). Most listeners should hear the louder tone at a lower pitch, especially for 200 Hz.

Because the basilar membrane is a nonlinear neuromechanical system, the excitation pattern moves along the basilar membrane with increasing intensity. Physiological observations show that the pattern moves in the direction that agrees with Stevens rule for high frequencies, which actually tends to be the dominant effect perceptually. However, there is a problem because the peak of the excitation pattern shifts

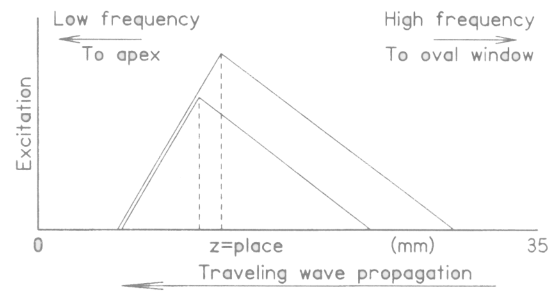

FIG. 8. Excitation patterns for two tones with the same frequency and a level difference of 30 dB. The low-frequency tail does not move much (courtesy AIP Press).

too much. It may shift by a place difference corresponding to almost an octave, whereas the pitch normally shifts by only a few percent. Figure 8 shows a cartoon of the excitation patterns that are observed physiologically for two different levels. Although the peak moves considerably, the lower-frequency tail of the excitation pattern does not (Zwislocki, 1991). This suggests a place theory model wherein pitch depends on the pattern in this low-frequency tail. A second reason to prefer a model in which pitch depends on the tail region instead of the peak is that the peak region of the excitation pattern tends to broaden at high levels as neural excitation saturates, but frequency difference limens are stable at high level.

The regularity represented by Stevens rule likely results from something fundamental in the hearing process, though we are not sure what it is. The individual differences from this regular behavior represent something significant too because each individual ear has a reproducible characteristic signature. Figure 9 shows a pitch-intensity relationship for the author's right ear. On the vertical axis is the change in the frequency of a sine wave that must be made to keep the pitch unchanged when the intensity increases from 10 dB to 40 dB SPL. The three curves were taken over a 4-month

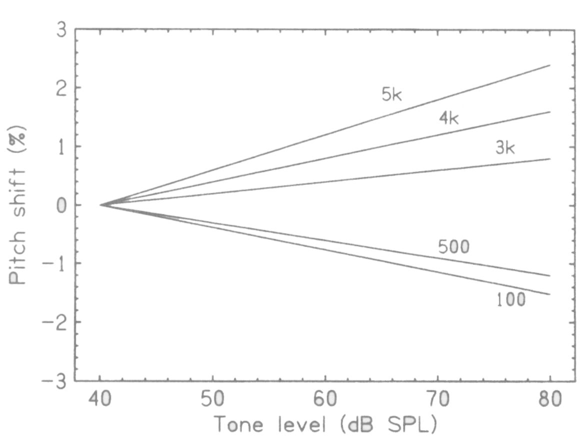

FIG. 7. The change in pitch, relative to a 40-dB standard, induced by increasing the level of a sine tone. The change is shown as a percent of the tone frequency, which appears as a parameter. The plot is calculated from a simple mathematical formula (Terhardt *et al.*, 1982b) that fits the average of many listeners.

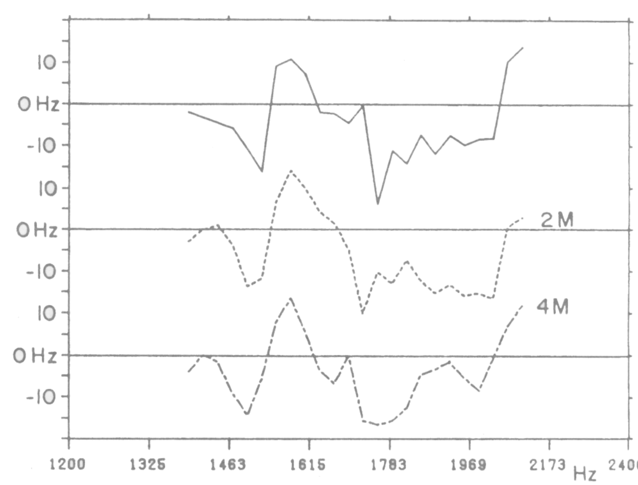

FIG. 9. Pitch shift as a function of sine tone frequency in one ear for a 30-dB increase in level. The baseline experiment is shown by the top line; the following lines show the result of repeating the experiment after 2 months and after 4 months (from Klein, 1981).

interval. The similarity between the functions shows that idiosyncratic variations are not caused by random error but result from systematic inhomogeneities in the particular ear.

It has recently become quite clear that the cochlea is not homogeneous and does not treat all frequencies the same. Instead, it behaves like a transducer with multiple fixed resonances, distributed inhomogeneously along its length and corresponding to different frequencies. These resonances originate in the action of the outer rows of hair cells that parametrically change the mechanical properties of the basilar membrane (Dallos, 1992). This action results in increased sensitivity, especially for sounds within 40 dB of the threshold of hearing. Evidence for these resonances can be found in emissions from the cochlea, spontaneous or stimulated tones that can be measured with a microphone in the ear canal (e.g., Stover and Norton, 1993).

The cochlear resonances are known to cause microstructure when the threshold of hearing is measured as a function of frequency. Hearing thresholds show minima, indicative of maxima in sensitivity at the frequencies of emissions. There is also evidence of correlation between the microstructure in the threshold and the fine structure in the pitch-intensity relationship. The most likely scenario is that nonlinear processing of the cochlea attracts the pitch of a tone down into a resonance when the tone intensity is low, but when the intensity is higher, the nonlinear cochlear action has less effect on the excitation pattern. Therefore, there is a differential pitch shift for the low level tone. Unfortunately, psychoacousticians have found that failures in relating pitch-intensity microstructure to threshold microstructure are about as frequent as successes. One possibility is that the successes are flukes; another is that we are not doing the experiments optimally because our hypotheses about the form of the relationship are not correct in detail.

One apparent connection between cochlear physiology and pitch is widely available. It is possible to change the frequency of a cochlear emission by pressurizing the middle ear, which then puts static pressure on the oval window. It is possible to change the pitch of a tone in this way too, and one way to generate a slight pressure is simply to clench the jaw (Corey, 1950).

Demonstration 5 presented three sine tones with different frequencies for listeners to try out the jaw effect, clenching and then relaxing two or three times while a tone was sounding. As expected for an inhomogeneous effect, it does not work for all people for all frequencies, and for some it may not work at all. Many listeners hear a change of about a semitone.

C. The timing theory of pitch

In addition to the place principle, whereby the frequency of a tone is encoded tonotopically, the timing pattern of neural pulses can also encode frequency. Timing can be observed in the auditory nerve as the intensity of a tone is gradually increased from zero.

Before the tone is on, the neurons of the auditory nerve fire spontaneously. High-spontaneous-rate neurons may produce 100 spikes per second. They come at random times, subject only to the refractory character of neurons. As the

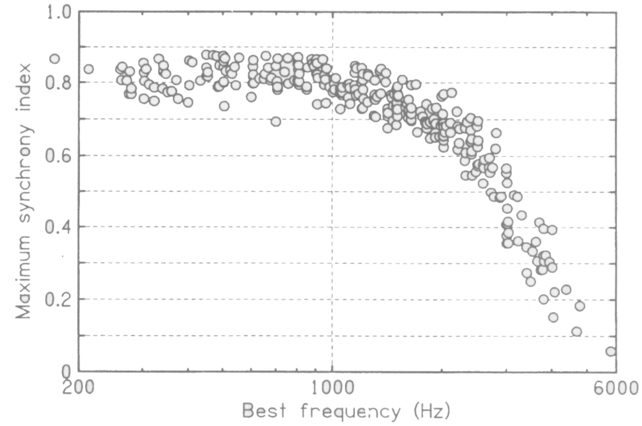

FIG. 10. Synchrony coefficient for 315 neurons as measured by D. H. Johnson who provided data for this plot (courtesy AIP Press).

tone intensity increases from zero, nothing happens at first. Then as the level grows, the neural spikes begin to order themselves in time so as to synchronize with the period of the signal. The number of spikes per second stays about the same. As the level of the tone increases further, the spikes become more numerous and the synchrony coefficient, which measures the extent to which the spikes follow the stimulus, also increases, finally peaking near 90%. The synchrony depends strongly on the frequency of the tone. Because of timing jitter intrinsic to neural firing, synchrony vanishes rapidly as the frequency increases from 2 to 5 kHz, as shown by Johnson's (1980) data in Fig. 10.

Timing models of pitch perception assume that the sensation of pitch is determined by neural synchrony. It is known that there is more than enough information in the timing pattern to account for the acuity of our sense of pitch, but because of pitch-shift effects, such as diplacusis and the pitch-intensity effect, we know that pitch cannot be determined by the period of the neural spike train. This period is too rigidly tied to the period of the stimulus to accommodate the pitch-shift effects. However, pitch could be determined by a derived quantity such as the neutral autocorrelation function, as will be discussed later.

D. Evidence in favor of the timing theory

Some of the strongest evidence for the timing theory comes from difficulties encountered by the place theory. Because the place theory is based on the excitation pattern established by tuned neural elements, it shares some of the limitations of tuned filters. Particularly, one would expect that the excitation pattern would grow wider as the duration of the tone grows shorter. Wider excitation patterns should, in turn, lead to larger frequency difference limens. For example, if a rectangularly gated tone has duration Δt, its power spectrum has a bandwidth Δf, where

$$\Delta f\ \Delta t = 1. \tag{4}$$

Based on the place theory, one would expect an uncertainty principle like this to hold for pitch perception. Experimentally, it is found that frequency difference limens do grow with decreasing tone duration, but in detail Eq. (4) is

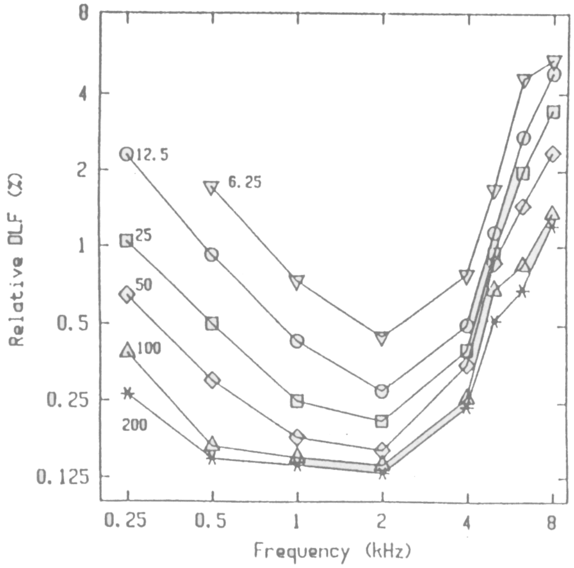

FIG. 11. Frequency difference limen for sine tones as a function of frequency as measured by Moore in 1973 and published in the 1989 edition of his book. The parameter gives the tone duration in ms (courtesy Academic Press).

all wrong. First, the pitch perception process beats the uncertainty principle by a factor of 5. Second, the uncertainty Δf, as measured by the difference limen, does not depend on the inverse first power of Δt. Instead it exhibits a complicated dependence that cannot be fitted with a single power. Then too, there is an unusual frequency dependence of the difference limen for short tones, as shown in Fig. 11. The DL rises abruptly as the frequency goes from 2 to 5 kHz.

The significance for timing models of the short-tone experiments is, first of all, that timing models do not need to obey the uncertainty principle; a reciprocal reading frequency counter can determine a frequency with arbitrary precision given only a single cycle of the signal. Second, the range of 2 to 5 kHz, where the difference limen increases, is precisely the range where neural synchrony disappears.

Timing theory also seems to apply to the phenomenon known as repetition pitch. When a broadband signal is added to a delayed version of itself, the result is a pitched sensation. Figure 12 shows how this is done.

Demonstration 6 illustrated repetition pitch by passing white noise through a delay-and-add system. As the demonstration progressed the delay changed in steps of 1 ms from

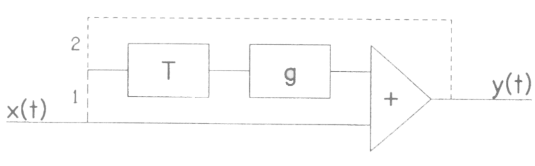

FIG. 12. Block diagram of a delay-and-add system. When connection "1" is made and connection "2" is broken, the input $x(t)$ is delayed by T and attenuated by gain g, then added back to itself to produce an output $y(t)$ having a pitch of $1/T$. When connection "1" is broken and connection "2" is made, the system becomes an infinite-impulse-response filter with the feedback configuration called "staircase." The pitch is still $1/T$, but it is stronger.

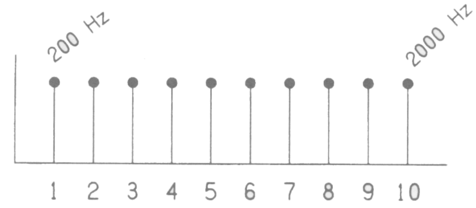

FIG. 13. A periodic complex tone with a fundamental frequency of 200 Hz and the first ten harmonics.

1 to 20 ms. The reciprocals varied from 1000 to 50 Hz, and these frequencies correspond to the pitches that are normally heard. Then feedback was added to the delay, as shown by connection "2" in Fig. 12. The sequence of increasing delays was repeated. The pitch was perceived to be stronger with the feedback.

In fact, one could explain the phenomenon of repetition pitch beginning with either a spectral model or a timing model. But the way in which the pitch strength changes as feedback is added agrees better with a timing model than with a model based on the spectrum (Yost *et al.*, 1996).

Further evidence for the role of timing in pitch perception is the case of sine-wave-amplitude-modulated white noise or SAM noise. The long-term spectrum of SAM noise is featureless and flat. There is no more long term spectral evidence for pitch than there is in any white noise, and yet, SAM noise has a pitch (Burns and Viemeister, 1981). Although the pitch can be heard with headphones, it is hard to hear using loudspeakers in a room. To improve the changes of success the next demonstration used square-wave modulation. That does not change the principle of the effect.

Demonstration 7 presented broadband noise amplitude modulated by a square wave. The modulation frequency alternated between 150 and 300 Hz, about once per second. Listeners close to the loudspeakers and mainly in the direct sound field could detect the pitch change.

Thus far, psychoacousticians have been unable to decide whether pitch is the result of a place process or a timing process. Often it appears that both mechanisms are involved, though timing mechanisms may dominate for low frequency and place mechanisms may dominate for high.

It is now time to move on to study the pitches of complex tones, which are more common in everyday life than sine tones. However, the matter of sine tone pitch will come back again, and we shall continue to ask whether place models or timing models are preferable.

IV. PITCH OF COMPLEX TONES

Complex tones come in two varieties, periodic and not periodic. The periodic tone is characteristic of the sounds of sustained-tone musical instruments such as the bowed strings, the brass, and the woodwind instruments. The vowel sounds of human speech are also periodic tones, and vowel sounds are the basis of singing. It should be evident that periodic complex tones are very important to music. Periodic complex tones have harmonics. Their spectra consist of regularly spaced line components, as shown in Fig. 13 for a 200-Hz tone.

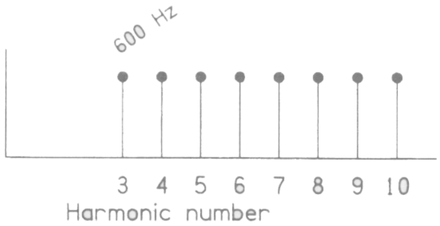

FIG. 14. A periodic complex tone with a fundamental frequency of 200 Hz and only harmonics 3 through 10. Its pitch is still 200 Hz.

A periodic complex tone has a pitch corresponding to the fundamental frequency. There is some evidence that the pitch may be slightly less than the fundamental (Walliser, 1969), but the difference is not a large one. Right away there is a problem for the place theory of pitch because different components in the tone will excite different places, and it is not clear which place should code for pitch. On the other hand, there would seem to be no problem for the timing model of pitch because the period of the tone is the reciprocal of the fundamental, in agreement with the pitch.

The place theory might escape from its problem with an arbitrary rule that says that so far as pitch is concerned, the lowest-frequency place wins. That rule would be found to work for the great majority of musical and vocal tones. However, with special experimental stimuli we can create more trouble for the place model by proving that the 200-Hz pitch of the tone in Fig. 13 does not depend on the existence of the fundamental component at 200 Hz. This is shown in Fig. 14 and in the next demonstration.

Demonstration 8 illustrated the effect of the missing fundamental with two tones. The first was a 200-Hz complex tone with ten harmonics of equal level. The second was the same except that the fundamental component was missing from the spectrum. The second harmonic was missing also. The usual response to these two tones is that their pitches are equal.[1]

The pitch of a tone with a missing fundamental is consistent with a timing model because the period is still 1/200 s. However, we have not said what, precisely, is timed. Perhaps it is the details of the peaks and valleys of the waveform, perhaps it is the overall structure as represented by the envelope. In fact, both choices lead to problems.

The problem with the peaks and valleys model is that these are waveform features that depend sensitively on the phases of the components. These phases change dramatically as one moves from place to place in a room, and yet we find that the pitch of a tone does not change as we walk around a room. A listener wearing headphones receives stable phases, but again we find that when the phases of the stimulus signal are changed the pitch does not change. Therefore, the envelope periodicity model is more attractive. However, we can create a problem for the envelope periodicity model by generating a complex tone with all its harmonics shifted by a common frequency difference, as shown in Fig. 15. The envelope of the shifted tone still has a period of 1/200 s because the envelope depends only on the spacing of the harmonics. Therefore, the envelope periodicity model predicts that the pitch should be unchanged by the frequency shift. The problem is that experiments show clearly that the pitch does change as the harmonic frequencies are all incremented in this way. Demonstration 9 showed a pitch shift that is perceived clearly by almost everyone.

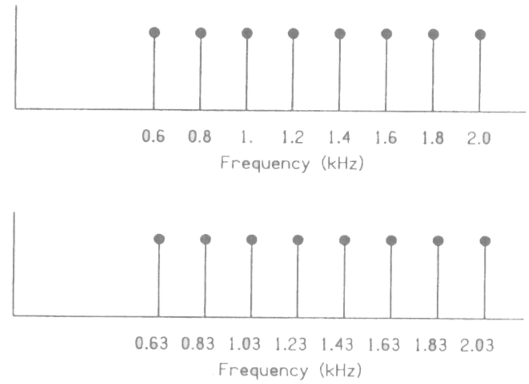

FIG. 15. The top graph shows the spectrum of Fig. 14, plotted as a function of frequency. The bottom graph shows the "shifted-spectrum" made by shifting each harmonic by a constant frequency increment, here 30 Hz.

Demonstration 9 consisted of two tones. The first was a 200-Hz complex tone composed of harmonics 3 through 10 at equal level. The second was the same except that all components were shifted upward by 30 Hz. Such a shift causes the pitch to increase.

The failure of the envelope model to account for the pitch of the shifted spectrum means that we need to look elsewhere for a valid timing model. The most popular choice is based on autocorrelation, and autocorrelator models have enjoyed some success, especially in dealing with wide-band stimuli like the repetition pitch (e.g., Meddis and Hewitt, 1991a,b). Autocorrelation models run into difficulty with low-frequency tones. An autocorrelator that can register a 100-Hz tone needs to have a neural delay line at least 20 ms long, and delays this long have not yet been encountered in the auditory physiology. Then too, autocorrelator models must cope in some way with the reality of auditory filtering. Current models deal with the problem by applying autocorrelation within filter channels and then summing the outputs of the array of autocorrelators.

An alternative approach to the pitch of complex tones stands this situation on its head and *emphasizes* the role of auditory filtering. It generates pitch from a spectral analysis of the signal. The analysis could actually be either in the place domain or in the time domain, but the important role played by spectrally resolved harmonics finds a natural representation in the place domain. This approach is particularly associated with the topic of auditory organization.

V. PITCH AND AUDITORY ORGANIZATION

The concept of auditory organization begins with the incontrovertible fact that the peripheral auditory system analyzes complex signals into different frequency bands. It then becomes the responsibility of processors at more central sites to reassemble the channels in some meaningful way. The process of reassembly is auditory organization. The remainder of this article attempts to show that pitch is fundamental to this organization.

The connection between pitch and auditory organization begins with the observation that the perception of pitch for complex tones resembles a pattern matching process that is done by the central auditory system. (Goldstein, 1973; Terhardt, 1974; Terhardt *et al.*, 1982a,b). The idea is that the central processor attempts to fit a harmonic template to the pattern of components in the tone. The template is restrictive because its harmonics are consecutive. For every input signal there will be some harmonic template, defined by a fundamental frequency, that will lead to the best fit. According to the model, it is this best-fitting fundamental frequency that corresponds to the perceived pitch of the tone. In the case that two templates with different fundamental frequencies lead to equally good fits, one would expect the stimulus to have an ambiguous pitch, or two pitches might be heard.

There is a great deal of evidence to support this template fitting model. One of the most dramatic bits of evidence for the central-processing aspect of the model is that pitches can be created by putting the necessary components into the different ears. Houtsma and Goldstein (1972) were able to generate a pitch of 200 Hz by putting 600 Hz into one ear and 800 Hz into the other. To create the 200-Hz pitch sensation the central auditory system must combine the signals in the two ears. It is said that the auditory system "synthesizes" the pitch. The reader should be warned that binaural synthesis is not the usual result of putting different tones into the two ears. The usual result is that one hears the two tones separately, a mode known as "analytic listening." However, by using low-level signals and by cueing the listener appropriately, Houtsma and Goldstein were able to get listeners to hear the binaural stimulus synthetically. Their experimental protocol found unambiguous evidence for synthetic listening once the listeners arrived in that state. Houtgast (1976) took the next logical step and showed that listeners can be cued in such a way as to hear a low pitch given only a *single* upper harmonic. The method depended on creating uncertainty by embedding the signal in noise and presenting the signals in a stimulus protocol that encouraged listeners to form a low-frequency template.

Further dramatic evidence for the synthesized character of complex tone pitch is that the harmonics do not have to be simultaneous. Hall and Peters (1981) created a low pitch using a stimulus in which the third, fourth, and fifth harmonics were sounded sequentially. Each harmonic lasted 40 ms, and there was a 10-ms gap between successive harmonics. This effect was shown in Demonstration 10.

Demonstration 10 began by encouraging the perception of a low pitch. A familiar melody, the Westminster chimes, was first played by adding up some high-frequency harmonics to make a low-pitched tone. Then the melody was played again with the same high-frequency harmonics, but the harmonics were not simultaneous. Listeners heard the same melody.

The template fitting procedure that leads to complex-tone pitch is held to be responsible for the pitch of chime tones. *Demonstration 11* made that point with a digitally generated chime tone in which seven components were added together, as shown in Fig. 16. These components represent the free vibrations of the chime bar, and they are not harmonic. What is interesting is that the pitch of the chime tone is not to be found among the modes of the chime. Instead, the pitch is synthesized by the fourth and fifth modes, which appear to act like second and third harmonics.

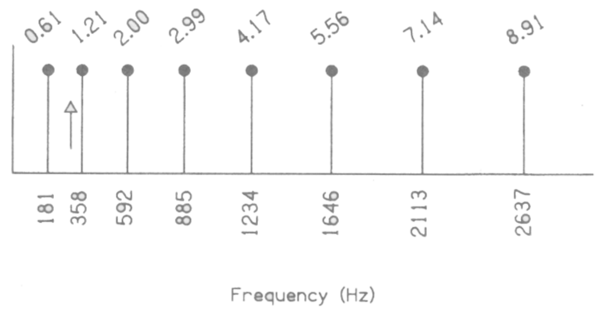

FIG. 16. The spectrum of a chime tone. Frequencies are given in Hertz and as a multiple of the "strike tone" frequency, which is what listeners say is the pitch of the tone. The strike tone is here 296 Hz, approximately the note D4, shown by the arrow.

According to the template model, the harmonics of a complex tone go together to synthesize a pitch. Just how the harmonics are combined is a question of great interest. A promising experimental approach is to mistune one or more harmonics and study the effect on the synthesized low pitch. Ritsma (1967) mistuned harmonics in blocks, Moore *et al.* (1985) mistuned individual harmonics. Both experiments led to the conclusion that the harmonics that are most important in determining the low pitch are those that are resolved into different tuned channels by the auditory periphery. The observed importance of the resolved harmonics supports the idea that pitch perception takes place at high levels where excitations from different peripheral channels are recombined. Pattern matching, or template fitting models are like that (Goldstein, 1973; Terhardt, 1974). The model by Goldstein derives the low pitch from a pattern match to the frequencies of the harmonics, and the levels of the individual harmonics do not play a role. There is some experimental evidence (Houtsma, 1981; Moore *et al.*, 1985) to support the alternative model by Terhardt which says that the relevant input to the pattern-matching model from the harmonics is not their frequencies but their pitches. As a result, any effect that shifts the pitch of a harmonic, such as the pitch intensity effect, will potentially shift the low pitch of the complex tone (Brink, 1975a, b). Terhardt's model also depends on the harmonic levels. In general, the template fitting process seems to give particular weight to harmonics with harmonic numbers 3 through 6. Such harmonics are said to be dominant, and it is significant that they are among the resolved harmonics. It is particularly the resolved harmonics that need to be reassembled in the process of auditory organization.

The concept of auditory organization emphasizes that a given stimulus may be heard in different ways as its component parts are put back together differently. For example, the individual harmonics of a complex tone are not necessarily combined to make a low pitch; in special circumstances they may be heard individually. This is shown by a demonstration that begins with a complex tone and then omits and reinserts one of the harmonics.

Demonstration 12 consisted of a 200-Hz complex tone with ten harmonics. While the tone was played, the fifth harmonic was removed and reinserted several times. In the end the tone was just the same as at the beginning, but at the end listeners heard the fifth harmonic standing out separately from the rest of the complex tone because the process had called attention to it. If the complex tone is turned off briefly, it is usual to find that listeners lose their hold on the fifth harmonic. A silent interval seems to reset the auditory system to a normal state in which it integrates all the components of a complex tone.

As a demonstration of tuning in the auditory periphery, Demonstration 12 is something of an embarrassment because it works all too well. It shows that individual harmonics can be heard out of a complex tone at least up to harmonic number 12, and probably higher. However, only the first five or six harmonics are thought to be resolved in the sense that the output of a peripheral auditory filter can be entirely dominated by one harmonic. The ability to reorganize the perception of a tone to hear out a high harmonic is not limited to the resolved harmonics.

What all this says about the pitch of a complex tone is that pitch is a grand consolidator. Starting with a large collection of harmonics, the pitch processor reduces the complexity by integrating them all into a single entity characterized by a pitch. But confronted by a component that asserts its independence by an anomalous temporal behavior, the processor will gladly assign it a pitch of its own. The auditory organization involved in pitch perception is thus an essential part of making sense out of the sounds of the world.

There are limits to the pitch processor's willingness to integrate. If a harmonic is not well in tune so that it does not fit the template then the pitch processor segregates it. Then the individual mistuned harmonic stands out, as shown in Demonstration 13.

Demonstration 13 consisted of a single tone, a 200-Hz complex tone with ten harmonics. The fifth harmonic was mistuned upward by 7%. Such a mistuned harmonic is normally perceived as a separate tone.

Typically, a tone with harmonic components is heard as a single entity whereas an inharmonic tone may break up into its individual components. The inharmonic tone in Demonstration 13 is representative of what happens when there are two or more sources that are not phase locked. The problem of disentangling several sources is currently an important element in artificial intelligence research. An example of the way in which pitch can disambiguate overlapping speech sounds was given in a demonstration by Assmann and Summerfield (1990).

Demonstration 14 presented two simultaneous synthetic vowels. On successive presentations, the fundamental frequency of one vowel was increased; the partials increased correspondingly because the vowel remained harmonic. An increase of only a few percent is enough to cause the individual vowels to be individually recognized.

Although spectral template matching seems to account for many important cases in which signals are organized according to pitch, there are some pitch effects that are not readily explicable with this standard model. One of them is the pitch of the stretched Shepard tone. The Shepard tone (without stretch) is a complex tone with octave components only. The frequencies of the components are given by $f_n=f_0 2^n$, where f_0 is a constant. In principle a Shepard tone has no upper or lower frequency; the components extend indefinitely to all positive and negative n. In a stretched Shepard tone, with an "octave" ratio of 2.1, the frequencies are given by $f_n=f_0 2.1^n$. If we imagine that a stretched Shepard tone is recorded on tape, and played back twice as fast, all the frequencies are, of course, doubled. The result is shown in Fig. 17. The curious perceptual effect is that doubling all the frequencies causes the pitch of the tone to go down. *Demonstration 15* presented a Shepard tone made with an "octave" ratio of 2.1 as shown in the top part of Fig. 17. When it was played back twice as fast the pitch decreased.

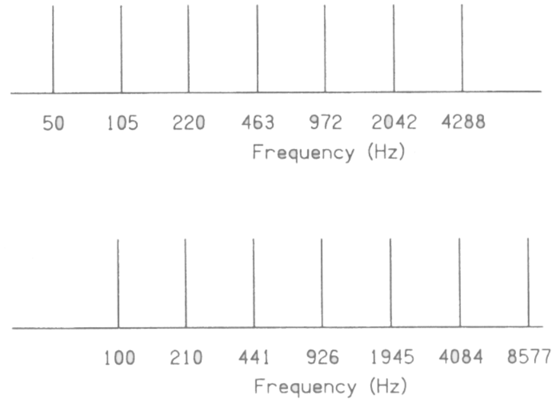

FIG. 17. The top line shows a stretched Shepard tone. The bottom line shows what happens when each frequency is doubled.

It is not hard to guess what is going on. In the neighborhood of each component of the first tone there is a component of the second tone that has a lower frequency. Every *local* impression indicates that the frequency of the complex is going down, and it is probably important that descending components occur in the spectrally dominant region near 800 Hz. It is also true that in every frequency region the difference between successive components becomes smaller when the speed is doubled. This demonstration also exploits the fact that we are not very sensitive to the spectral changes that occur at low frequencies, partly because of the rising equal-loudness curve and partly because of the low-frequency rise in frequency difference limen, as shown in Fig. 1. The standard model cannot explain this effect because the consecutive harmonic template does not apply to this unusual octave spectral structure. Correspondingly, the experimental results are quite different if the tone is composed of stretched consecutive harmonics; then doubling the playback rate increases the pitch in the usual way.

A final example shows a second case in which an unusual stimulus leads to an unusual auditory organization. It is the Duifhuis pitch, the pitch of the tone that is not there (Duifhuis, 1970, 1971). The stimulus spectrum is shown in

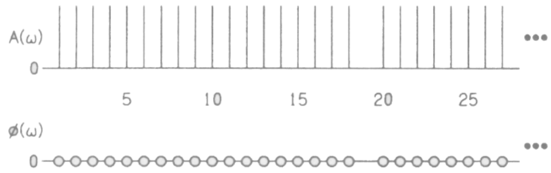

FIG. 18. Amplitude and phase spectra that produce a Duifhuis pitch. The components are spaced by 50 Hz and the 19th harmonic at 950 Hz is missing. The highest component number may be in the hundreds. All phase angles are the same, for example, all zero.

Fig. 18. It consists of a large number of harmonics of a low-frequency fundamental with all the harmonics in phase...they are all cosine waves. All harmonics are present except for one; in this example it is the 19th. The listener hears two tones: one is the low-frequency pulse, with a pitch corresponding to the lowest harmonic frequency; the other is a pure tone sound with a pitch corresponding to the missing harmonic.

At first, the Duifhuis pitch seems surprising. What is heard is precisely what is not present in the spectrum. On second thought, however, the effect is not so surprising. One simply has to think of the stimulus as a pulse train to which is added a sine tone that is exactly out of phase with the corresponding component of the pulse so as to cancel that component. The fact that one hears the sine tone when it is added to the pulse might have been expected. This explanation, however, is rather too simple. We know that the peripheral auditory system analyzes input signals into channels according to frequency. Therefore, the pulse *per se* never appears anywhere in the system. This makes the effect surprising again. The explanation that Duifhuis advanced for the effect is that if the missing harmonic is high enough and the components are close enough together then the output of an auditory filter centered on the missing harmonic frequency approximates a pulse plus oscillations at the missing harmonic. The missing harmonic can be heard during the gaps in the filtered pulse. This explanation predicts that the effect should only occur for high harmonic frequencies, which agrees with experiment. Therefore, the effect is not so surprising after all. The latest round in this cyclical tale comes from Lin (1996), who found that when the bandwidth of the stimulus is narrowed so that only components within a single auditory filter remain, the Duifhuis pitch disappears. Apparently the auditory organization required to synthesize the pitch of the missing harmonic requires the coordination of information in different auditory filters. It must be the kind of coordination that preserves phases across auditory filter channels. In sum, the effect has become surprising again.

The examples of this section show that the auditory system organizes sound input. It is capable of interpreting certain components as belonging to source A while regarding others as belonging to source B, and the basis for this segregation is pitch. The concept of a harmonic template has proved to be very useful in modeling this process. A template model can correctly predict the pitch of a wide variety of complex tone stimuli, including inharmonic signals, and it can recognize ``oddballs'' that do not fit the template and thus segregate them. However, there is a strong place theory bias, or at least a spectral bias, in the template model that may not always fit the facts. For example, experiments on the perceptual segregation of mistuned harmonics by Hartmann *et al.* (1990) exhibited a dramatic frequency dependence. The ability to identify the pitch of a mistuned harmonic decreased with increasing frequency in a way that exactly paralleled the neural synchrony index shown in Fig. 10. This result is not readily understood from a place model, but it agrees with the idea that auditory organization can also be mediated by consistencies and anomalies in neural timing.

VI. CONCLUSION

This tutorial has divided the topic of pitch perception into two main parts. First there are the basic processes that can be studied with sine wave excitation. These fundamental processes appear at the peripheral level of the auditory system, where frequency is encoded both in the tonotopic pattern of excitation and in the timing of neural pulses. Sine tone *psychoacoustical* studies show that the frequency encoding mechanisms are remarkably reliable. Frequency difference limens are small enough that one can get four-significant-figure precision from a pitch experiment. There are, of course, many pitch effects that prevent the human ear from being a perfect frequency meter. The pitch of a sine tone changes with changing intensity, with changing cochlear pressure, and with added masking. Pitch is different in the two ears, and there are other effects as well. Yet in normal ears these effects lead to variations of only a few percent. It is only because human listeners are so good at detecting small pitch changes that these effects can be observed at all, and observing them is one means of learning about processing in the auditory periphery. In addition, the basic mechanisms are capable of extracting a reliable pitch given only a brief exposure to a sound.

The second aspect of pitch perception deals with complex tones and with the real world of speech and music. This aspect cannot be studied meaningfully at the level of the peripheral auditory system. It involves processes of interpretation and synthesis that take place in the central nervous system, although it depends on the peripheral processes for its input, so that pitch shift effects at the peripheral level have consequences on the complex signal pitches derived centrally.

Psychoacoustical experiments discover that the central auditory system does some rather remarkable things in assigning a pitch to a complex tone. It assigns the fundamental pitch to a collection of harmonics, even when the fundamental is missing from the spectrum. The missing fundamental pitch can be synthesized from only two harmonics, even if the harmonics are in different ears. A missing fundamental pitch can be synthesized from harmonics that are not simultaneous, and given appropriate conditions of bias and uncertainty, the missing fundamental can be synthesized from a single upper harmonic. If the components of a tone are inharmonic, the auditory system can synthesize a low pitch anyway, apparently based on an internal harmonic template that best fits the components of the tone. These effects, together with repetition pitch, and many other pitch effects (e.g., the pitches of noise band edges) that could not be presented in the tutorial for want of time, show that pitch is ubiquitous. One would be justified in concluding that the central nervous system seems to be trying awfully hard to assign a pitch to anything that comes to its attention.

The importance of pitch to the auditory system is not accidental, and it is probably not the result of the universal tendency for human beings to make music either. Instead, the percept of pitch plays a central role in defining individual objects in the acoustical world and separating them from other objects. Given an environment of concurrent sounds, some interesting, some threatening, and some merely background noise, all mixed together and competing for attention, the auditory system bears the awesome responsibility of detecting and identifying the sources. Pitch is a primary identifier, and the remarkable processes of pitch perception make it so. So far as music is concerned, one might risk an aphorism: *Nature gave us limbs for fight or flight, and we invented athletics. Nature gave us pitch to sort out the world, and we invented music.*

ACKNOWLEDGMENTS

Jian-Yu Lin helped create the demonstrations for the tutorial. Peter Assmann kindly sent an interfering vowels DAT. The author is grateful to the tutorials committee of the Acoustical Society for the opportunity to present this topic. This work was supported by the NIDCD of the NIH.

[1]The missing fundamental experiment provides a good illustration of the difficulty caused by the overly inclusive pitch definition in ANSI 1994 Terminology. When the fundamental is removed the tone color becomes brighter. On a scale extending from low to high, the missing-fundamental tone is perceived to be higher because the spectral center is higher. However, the pitch, in a melodic sense, or as measured in a sine-tone matching test, is not changed.

ANSI (**1994**). ANSI S1.1-1994, ''American National Standard Acoustical Terminology'' (Acoustical Society of America, New York).

Assmann, P. F., and Summerfield, Q. (**1990**). ''Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies,'' J. Acoust. Soc. Am. **88**, 680–697.

Brink, G. van den (**1975a**). ''The relation between binaural diplacusis for pure tones and for complex sounds under normal conditions and with induced monaural pitch shift,'' Acustica **32**, 159–165.

Brink, G. van den (**1975b**). ''Monaural frequency-pitch relations as the origin of binaural diplacusis for pure tones and residue sounds,'' Acustica **32**, 166–173.

Burns, E. M., and Viemeister, N. F. (**1981**). ''Played again SAM,'' J. Acoust. Soc. Am. **70**, 1655–1660.

Corey, V. B. (**1950**). ''Change in pitch of sustained musical note by clenching jaws of observer,'' J. Acoust. Soc. Am. **22**, 640.

Dallos, P. (**1992**). ''The active cochlea,'' J. Neurosci. **12**, 4575–4585.

Duifhuis, H. (**1970**). ''Audibility of high harmonics in a periodic pulse,'' J. Acoust. Soc. Am. **48**, 888–893.

Duifhuis, H. (**1971**). ''Audibility of high harmonics in a periodic pulse. II. Time effect,'' J. Acoust. Soc. Am. **49**, 1155–1162.

Fletcher, H. (**1934**). ''Loudness, pitch, and timbre of musical tone and their relations to the intensity, the frequency and the overtone structure,'' J. Acoust. Soc. Am. **6**, 59–69.

Goldstein, J. L. (**1973**). ''An optimal processor theory for the central formation of the pitch of complex tones,'' J. Acoust. Soc. Am. **54**, 1496–1516.

Greenwood, D. (**1990**). ''A cochlear frequency-position function for several species—29 years later,'' J. Acoust. Soc. Am. **87**, 2592–2605.

Hall, J. W., and Peters, R. W. (**1981**). ''Pitch for nonsimultaneous successive harmonics in quiet and noise,'' J. Acoust. Soc. Am. **69**, 509–513.

Hartmann, W. M. (**1997**). *Signals, Sound, and Sensation* (AIP Press, New York).

Hartmann, W. M., McAdams, S., and Smith, B. K. (**1990**). ''Matching the pitch of a mistuned harmonic in an otherwise periodic complex tone,'' J. Acoust. Soc. Am. **88**, 1712–1724.

Houtgast, T. (**1976**). ''Subharmonic pitches of a pure tone at low S/N ratio,'' J. Acoust. Soc. Am. **60**, 405–409.

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

Houtsma, A. J. M., and Goldstein, J. L. (**1972**). ''The central origin of the pitch of complex tones: Evidence from musical interval recognition,'' J. Acoust. Soc. Am. **51**, 520–529.

Johnson, D. H. (**1980**). ''The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones,'' J. Acoust. Soc. Am. **68**, 1115–1122.

Klein, M. A. (**1981**). ''Two Studies of Pitch Perception,'' Ph.D. thesis, Michigan State University (unpublished).

Lin, J.-Y. (**1996**). ''Psychoacoustical theory and experiments on human auditory organization of complex sounds and the critical bandwidth,'' Ph.D. thesis, Michigan State University (unpublished).

Meddis, R., and Hewitt, M. J. (**1991a**). ''Virtual pitch and phase sensitivity of a computer model of the auditory periphery I,'' J. Acoust. Soc. Am. **89**, 2866–2882.

Meddis, R., and Hewitt, M. J. (**1991b**). ''Virtual pitch and phase sensitivity of a computer model of the auditory periphery II,'' J. Acoust. Soc. Am. **89**, 2883–2894.

Moore, B. C. J. (**1973**). ''Frequency difference limens for short duration tones,'' J. Acoust. Soc. Am. **54**, 610–619.

Moore, B. C. J. (**1989**). *An Introduction to the Psychology of Hearing* (Academic, London), 3rd ed.

Moore, B. C. J., Glasberg, B. R., and Peters, R. W. (**1985**). ''Relative dominance of individual partials in determining the pitch of complex tones,'' J. Acoust. Soc. Am. **77**, 1853–1859.

Ritsma, R. J. (**1967**). ''Frequencies dominant in the perception of pitch of complex sounds,'' J. Acoust. Soc. Am. **42**, 191–198.

Stevens, S. S. (**1935**). ''The relation of pitch to intensity,'' J. Acoust. Soc. Am. **6**, 150–154.

Stevens, S. S., Volkman, J., and Newman, E. B. (**1937**). ''A scale for the measurement of the psychological magnitude pitch,'' J. Acoust. Soc. Am. **8**, 185–190.

Stover, L., and Norton, S. J. (**1993**). ''The effects of aging on otoacoustic emissions,'' J. Acoust. Soc. Am. **94**, 2670–2681.

Terhardt, E. (**1974**). ''Pitch, consonance, and harmony,'' J. Acoust. Soc. Am. **55**, 1061–1069.

Terhardt, E., Stoll, G., and Seewann, M. (**1982a**). ''Pitch of complex tonal signals according to virtual pitch theory: Tests, examples and predictions,'' J. Acoust. Soc. Am. **71**, 671–678.

Terhardt, E., Stoll, G., and Seewann, M. (**1982b**). ''Algorithm for extraction of pitch and pitch salience from complex tonal signals,'' J. Acoust. Soc. Am. **71**, 679–688.

Walliser, K. (**1969**). ''Zusammenhange zwischen dem Schallreiz und der Periodentonhohe,'' Acustica **21**, 319–329.

Wier, C. C., Jesteadt, W., and Green, D. M. (**1977**). ''Frequency discrimination as a function of frequency and sensation level,'' J. Acoust. Soc. Am. **61**, 178–184.

Yost, W. A., Patterson, R., and Sheft, S. (**1996**). ''A time-domain description of the pitch strength for iterated rippled noise,'' J. Acoust. Soc. Am. **99**, 1066–1078.

Zwicker, E., and Fastl, H. (**1990**). *Psychoacoustics—Facts and Models* (Springer-Verlag, Berlin).

Zwislocki, J. J. (**1991**). ''What is the cochlear place code for pitch?,'' Acta Otolaryngol. (Stockholm) **111**, 256–262.