Circularity in Judgments of Relative Pitch

Roger N. Shepard

Bell Telephone Laboratories, Inc., Murray Hill, New Jersey
(Received 23 July 1964)

A special set of computer-generated complex tones is shown to lead to a complete breakdown of transitivity in judgments of relative pitch. Indeed, the tones can be represented as equally spaced points around a circle in such a way that the clockwise neighbor of each tone is judged higher in pitch while the counterclockwise neighbor is judged lower in pitch. Diametrically opposed tones—though clearly different in pitch—are quite ambiguous as to the direction of the difference. The results demonstrate the operation of a "proximity principle" for the continuum of frequency and suggest that perceived pitch cannot be adequately represented by a purely rectilinear scale.

INTRODUCTION

The construction of one-dimensional psychological scales of pitch corresponding to the one-dimensional physical scale of frequency for tones has been accomplished by Stevens, Volkman, and Newman[1] and others.[2] Still others, however, have suggested that the perception of pitch may be too complex to be adequately represented by a single, linear scale. They have pointed out, for example, that the resemblance in pitch between two notes may actually be greater if they are exactly an octave apart than if they are somewhat less than an octave apart.[3] Indeed, there is some evidence that even the rat hears an increase in similarity at the interval of an octave.[4] Moreover, I have observed that, at least for harmonically rich tones, a sequence that increases in intervals of a major seventh (i.e., intervals that are just one half-tone short of an octave) has a peculiarly ambiguous character. Although it tends, naturally, to be heard as increasing in pitch, it can in some sense be heard, also, as a chromatically *decreasing* sequence. Clearly, such effects are difficult to reconcile with a rectilinear, unidirectional scale of pitch.

As early as 1846, Drobisch attempted to accommodate some of these effects by distorting the continuum of pitch into a helical curve in such a way that tones just an octave apart would be represented by corresponding points on successive turns of the helix.[5] Such a representation has the advantage of bringing tones an octave apart into closer spatial proximity. At the same time, it permits an explanation of other seemingly anomalous phenomena by providing for the analysis of pitch into two distinct dimensions: namely, "height" (or over-all pitch level), represented by the vertical axis of the helix and "tonality" ("tonal quality" or "tone chroma"), represented by the circular scale at the base of the helix. The ambiguous character of the sequence of notes increasing in major sevenths, for example, can then be explained by noting that, while these notes are indeed increasing in height, they are simultaneously moving in a contrary direction with respect to tonality (as reflected in the inverse alphabetical order of the letters by which these notes are designated).

The fact that one of the two components of pitch is circular in the helical model raises the remarkable possibility that, by appropriately exaggerating that component (viz., tonality), one might be able to bring about

[1] S. S. Stevens, J. Volkman, and E. B. Newman, "A Scale for the Measurement of the Psychological Magnitude of Pitch," J. Acoust. Soc. Am. 8, 185–190 (1937); see also, S. S. Stevens and J. Volkman, "The Relation of Pitch to Frequency; A Revised Scale," Am. J. Psychol. 53, 329–353 (1940).

[2] E.g., J. Beck and W. A. Shaw, "The Scaling of Pitch by the Method of Magnitude Estimation," Am. J. Psychol. 74, 242–251 (1961).

[3] See E. G. Boring, *Sensation and Perception in the History of Experimental Psychology* (Appleton-Century, New York, 1942), particularly pp. 376, 380; or J. C. R. Licklider, "Basic Correlates of the Auditory Stimulus," in *Handbook of Experimental Psychology*, S. S. Stevens, Ed. (John Wiley & Sons, Inc., New York, 1951), pp. 985–1039, particularly pp. 1003–1004.

[4] H. R. Blackwell and H. Schlosberg, "Octave Generalization, Pitch Discrimination, and Loudness Thresholds in the White Rat," J. Exptl. Psychol. 33, 407–419 (1943).

[5] See C. A. Ruckmick, "A New Classification of Tonal Qualities," Psychol. Rev. 36, 172–180 (1929).

a breakdown of transitivity in judgments of relative pitch. In the extreme case, if the dimension of height could somehow be suppressed altogether, all tones an octave apart would be mapped into the same tone; that is, the tonal helix would be collapsed into a tonal circle. Judgments of relative pitch should then become completely circular in the sense that there would be no highest or lowest tone in the set but only an isotropic ring in which every tone has both a clockwise neighbor that is judged higher in pitch and a counterclockwise neighbor that is judged lower in pitch.

In the experiments described here, this curious situation was realized by means of a specially contrived set of complex tones. The generation of these tones was made possible by a computer program developed by M. V. Mathews for the synthesis of musical sounds in an extremely flexible and precisely controlled manner.

I. GENERATION OF THE TONES

Each tone consisted of many sinusoidal components locked at successive intervals of an octave and sounded simultaneously. Thus the frequency of each component above the lowest was exactly twice the frequency of the one just below. The amplitudes were large for the components of intermediate frequency only, however, and tapered off gradually to subthreshold levels for the components at the highest and lowest extremes of frequency. The sound-pressure level (in decibels) contributed by each component of one of these multicomponent tones is represented, graphically, by the height of the heavy vertical line corresponding to that component in Fig. 1. At the beginning of each tone, all components were started in the same phase relation (upward from a zero crossing).

The most important aspect of the scheme adopted for the generation of these tones was that the spectral "envelope" of the sound levels, as represented by the light curve in the Figure, was identical for all tones in the set. Consider, for example, a second tone in which all components are shifted up (in log frequency from the corresponding components of the first tone) the same fraction of the way toward the next higher octave. The spectral composition of this second tone would differ from that for the original tone in the manner in which the dashed vertical lines in Fig. 1 differ from the original solid vertical lines. The essential point to notice here is that the upward shift in frequency has been offset, in some measure at least, by increasing the contributions of the lower components while decreasing the contributions of the higher components. Indeed, if the second tone is shifted up one whole octave, it becomes identical to the original tone. For, at this point, the highest component (which has already faded below threshold) is dropped out and a new component (which will also be below threshold) is introduced one octave below the previously lowest component.

The tones generated according to this cyclic scheme

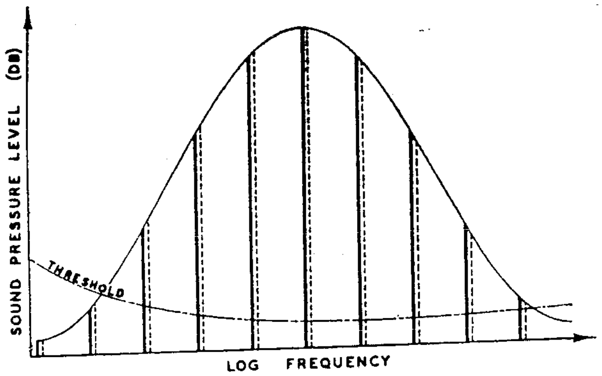

FIG. 1. Sound-pressure levels (in dB) of 10 simultaneously sounded sinusoidal components spaced at octave intervals. (The dotted lines correspond to an upward shift in the frequencies of all components.)

were confined to a discrete set of positions equally spaced (in log frequency) between the original tone and the return (one octave "above") to that same tone. In a set consisting of t_{max} different tones, the frequency of the cth component of the tth tone is

$$f(t,c)=f_{min}\cdot 2^{[(c-1)\cdot t_{max}+t-1]/t_{max}}, \qquad (1)$$

where f_{min} is the frequency of the lowest component of the first tone. The sound-pressure level of any one component (in decibels) is

$$L(t,c)=L_{min}+[L_{max}-L_{min}]\cdot[1-\cos\theta(t,c)]/2, \qquad (2)$$

where L_{min} and L_{max} fix the range of sound levels for the individual components and where the function θ is defined by

$$\theta(t,c)=2\pi[(c-1)\cdot t_{max}+t-1]/[t_{max}\cdot c_{max}]. \qquad (3)$$

Here, c_{max} denotes the total number of components in each tone.

The trigonometric "envelope" defined by Eq. 2 (and illustrated in Fig. 1) was chosen as the most convenient function of the general shape desired that is "smooth" and, at the same time, satisfies the relation

$$\sum_{c=1}^{c_{max}} L(t,c)=\text{const} \quad (\text{over all } t). \qquad (4)$$

This relation ensured that the different tones in the set did not vary appreciably in over-all loudness. For, if the components of a complex tone are widely separated (as here), the over-all loudness of the tone is approximated by the sum of the loudnesses of its separate components.[6–7]

[6] H. Fletcher and W. A. Munson, "Loudness, Its Definition, Measurement, and Calculation," J. Acoust. Soc. Am. 5, 82–108 (1933); D. H. Howes, "The Loudness of Multicomponent Tones," Am. J. Psychol. 63, 1–30 (1950).

[7] The invariance of loudness was of course only approximate since the threshold curve is not flat (Fig. 1) and since subjective loudness is not strictly linear with sound-pressure level. In fact, however, there was no noticeable variation in loudness among the different tones.

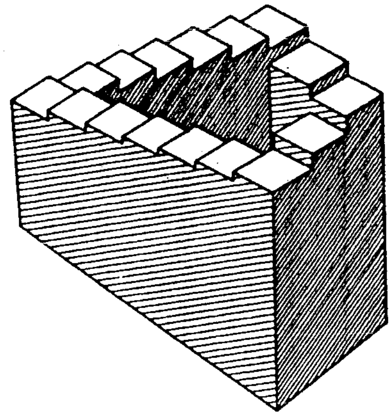

FIG. 2. "Circular" staircase illusion.

The waveforms of these specially tailored multicomponent tones were readily constructed on an IBM-7094 computer by means of Mathews' program for synthesizing quite arbitrarily specified musical sounds.[8] The only addition that had to be made to Mathews' program was a short FORTRAN subroutine for the computation of the parameters of frequency and sound level as prescribed, above, in Eqs. (1)–(3). For each successive tone, the output consisted of a long sequence of digitally coded numbers on magnetic tape. These numbers specified displacement as a function of time at a rate of 10 000 samples/sec of subsequent playing time. (This sampling rate corresponds to a bandwidth of 5000 cps.) An analog tape suitable for playback (at 7½ ips) on a standard audio tape player was then obtained from the digital tape by off-line equipment for digital-to-analog conversion.

Although each of the complex tones generated in this way was really made up of many sinusoidal tones, most listeners did not hear these sinusoids as subjectively separate components—i.e., as in a chord. Rather, they tended to describe the total complex as a single tone with a sonorous, rather organlike timbre. This is perhaps not surprising, since all higher components occurred at harmonic intervals of the lowest audible component. Moreover, the spectral-amplitude envelope modified these components much as a broad-band resonance would give rise to a formant in the spectrum of a musical instrument.

In addition to these complex tones, pure sinusoidal tones and bursts of wide-band noise were also generated by means of Mathews' computer program. The pure tones were used in some initial tests of each subject's ability to discriminate differences in pitch, and the noise bursts were used for certain signaling purposes in some of the experiments described here. The sequences of tones, noise bursts, delays, as well as other parameters of the sounds themselves, were communicated to the computer program by a deck of punched cards (called the "score"). In order to reduce transients, [illegible] attack and decay envelopes (lasting about [illegible] each) were imposed at the beginning and end of each tone.

II. PRELIMINARY DEMONSTRATION OF CIRCULARITY

Owing to the cyclic nature of the complex tones considered here, they can conveniently be represented as regularly spaced points around a circle. The convention followed is that a clockwise displacement between neighboring tones represents an upward shift in the frequencies of all corresponding components. For the purposes of preliminary tests, the tones were generated in the order in which they are encountered in passing around this circle in one direction. Since the steps between adjacent tones were produced by half-tone shifts in all corresponding components, the sequence resembled a chromatic scale (of equal temperament) and 12 steps were required to complete exactly one revolution about the circle.

Subjective circularity was clearly demonstrated in a sequence of this kind that made many of these revolutions. Each tone in this sequence was sounded for 0.1[illegible] sec and successive tones were always separated by a 0.84-sec period of silence. The components of each tone spanned 10 octaves and the specific values of the parameters were as follows: $t_{max}=12$, $c_{max}=10$, $f_{min}=4.863$ cps, $L_{min}=22$ dB, and $L_{max}=56$ dB. (The overall sound-pressure level of each of the complex tones was about 66 dB, but this measure is meaningful only in a relative sense; the audio gain of the tape player was adjusted only to achieve a comfortable listening level.)[9]

The resulting audio tape, or minor modifications of it have now been played to over 60 listeners and, so far all have described the sequence as progressing monotonically in pitch (upward or downward, depending upon whether the tape is played forward or backward). With the forward order, that is, each tone evidently wa[s]

[8] M. V. Mathews, "The Digital Computer as a Musical Instrument," Science 142, 553–557 (1963).

[9] The main interest here is in the (psychological) problem of demonstrating circularity of judgments rather than in the (psychophysical) problem of systematically exploring the effect of physical parameters on auditory perception. Since the desire[d] demonstration of circularity can be accomplished by an analys[is] of the pattern of judgments alone, it does not really depend upo[n] any physical measurements of the stimuli. For this reason, determination of the absolute energies of the various component[s] of a tone (as well as of the tone as a whole) at the ear was no[t] required. Also, owing to the relative insensitivity of the ear t[o] tones of low frequency, the lowest three or more components wer[e] probably entirely inaudible with the particular values chosen fo[r] the parameters (especially since the lowest components range[d] down to only 4.86 cps). Thus, the symmetry of the distribution o[f] sound-pressure levels of the components with respect to lo[g] frequency (Fig. 1) did not entail a strict symmetry of the distribution of psychological loudnesses. However, a strict symmetry turne[d] out to be unnecessary for the phenomenon and, probably, almo[st] any smooth distribution that tapers off to subthreshold levels a[t] low and high frequencies would have done as well as the cosin[e] curve actually employed.

always heard as higher in pitch than the preceding. Toward the end of the sequence, some of the listeners became puzzled by the fact that the tones (which clearly had been going up for so long) did not really seem to be getting much "higher." Other listeners, however, did not notice this stationarity of height. Indeed, these latter subjects were astonished to learn that the sequence was cyclic rather than monotonic and that it in fact repeatedly returned to precisely the tone with which it had begun. Several listeners likened the auditory effect to the "stairway to heaven" visual illusion reproduced in Fig. 2. *Note added in proof.* At the time of writing, I was unaware of the origin of the visual "staircase" illusion shown in Fig. 2. J. F. Schouten subsequently called my attention to the fact that it was presented (in slightly different form) by L. S. Penrose and R. Penrose, "Impossible Objects: A Special Type of Visual Illusion," Brit. J. Psychol. **49**, 31–33 (1958).

The 0.84-sec delays in the sequence just described (or perhaps merely the retardation of the progression that resulted from these delays) seem to be essential to the phenomenon. In any case, the illusion of a monotonic progression failed to emerge in an earlier attempt, in which the sequence proceeded from tone to tone in rapid succession without any delays. That sequence was heard to drop abruptly down an octave at one point during each revolution about the "circle" of 12 tones. Curiously, though, the point at which this downward jump was heard was not completely determined by the physical stimulus but depended, in part, upon the point at which the cyclic sequence was begun. Some subjects with musical training indicated that the perceived drop occurred when the sequence was heard to return to the tonic. Thus it may be an attentional phenomenon governed by the key to which the subject has been set (perhaps by the first note) to relate the ensuing tones.

III. EXPERIMENTAL PROCEDURE

In order to examine circularity of pitch judgments in a more extensive and objective manner, it seemed desirable to randomize the order of isolated pairs of tones so that a subject would have to make a separate judgment (of "up" or "down") with respect to each pair separately. Moreover, such pairs would not have to be confined to adjacent tones in the circular representation. Pairs of diametrically opposed tones, for example, should provide an interesting case of ambiguity, for the second tone could then be regarded equally well as clockwise or counterclockwise from the first.

Accordingly, four sequences of tone pairs were generated for the purposes of a more formal experiment of this kind. The duration of each tone was $\frac{1}{8}$ sec and the two tones within the same pair were also separated by a $\frac{1}{8}$-sec gap. A relatively soft burst of broad-band noise (54 dB) preceded and followed each pair of tones. This served to alert the subject that a pair was about to be presented, whether the tape was played forward or backward. These noise bursts lasted $\frac{1}{8}$ sec each and were both separated from the intervening tones by gaps of that same duration. A single presentation thus took about 1 sec for each pair. Between such presentations there was a 3-sec silence during which the subject was to write down his response to the immediately preceding pair—a "U" (if the second tone in the pair seemed to go *up* from the first) or a "D" (if the second tone seemed to go *down* from the first).

In order to reduce the number of pairs to be judged, the tonal circle was divided into 10 rather than 12 intervals. Hence, adjacent tones were slightly more than a tempered half-tone apart, and only the diametrically opposed tones were separated by a conventional musical interval (the diminished fifth or "tritone"). A total of 90 distinct ordered pairs could therefore be formed from this set of 10 tones.

The first two sequences of pairs that were presented to each subject were included only as a test of the subject's ability at discriminating differences in pitch. For this purpose, the pairs were made up from 10 pure sinusoidal tones rather than from the 10 complex tones described above. The frequency spacing between these pure tones was the same as that between the complex tones, however. That is, the 10 tones divided one octave into 10 equal intervals on a scale of log frequency. In each of these two sequences, only adjacent tones were presented for comparison and each of the 18 ordered pairs of these appeared just once (in a randomly determined position) within each sequence. In the first sequence, which was intended to be relatively easy, the tones ranged in frequency from 311.2 to 580.8 cps at 62 dB. The second sequence was made more difficult by lowering all tones two octaves (so that the lowest frequency was 77.8 cps) and by reducing the sound level to 56 dB.

The third sequence was designed to provide conclusive evidence of circularity in pitch judgments with the 10 complex tones. For this purpose, only adjacent tones (in the circular representation) were presented in each pair and each of the 20 such ordered pairs was presented twice within the total resulting sequence of 40 pairs. The positions in which the ordered pairs were presented within this sequence were chosen at random with the restrictions that (a) the same tone never appeared in both of two consecutive pairs and (b) each combination to two tones appeared together as a pair just once within each successive block of 10 pairs. The parameters of these tones were as follows: $t_{max}=10$, $c_{max}=10$, $f_{min}=4.863$ cps, $L_{min}=22$ dB, and $L_{max}=56$ dB.

The fourth sequence that was presented to each subject included all 90 ordered pairs of the 10 complex tones. The purpose of this final sequence was to investigate the possible ambiguity of pairs in which the second tone is on the opposite side of the circle from the first. The parameters of the individual tones were the same as given above, for the third sequence. (In both of these last two sequences, the over-all sound-pressure level for

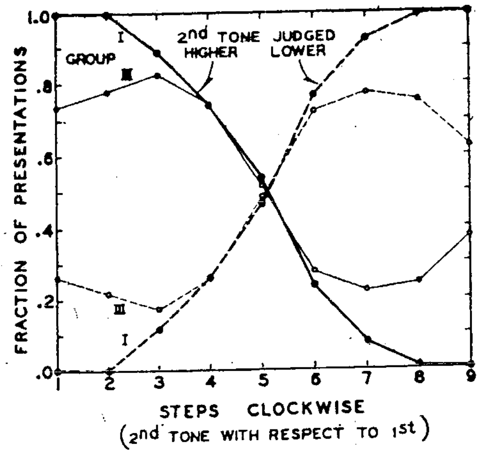

FIG. 3. Fraction of presentations for which the second tone of a pair was judged higher (or lower) than the first as a function of the number of steps by which the second tone was displaced from the first in a clockwise direction in the circular representation. (The heavy and light lines are for the subjects who were previously found to be most or least accurate, respectively, in judging relative pitches of pure tones.)

each of the complex tones was 66 dB.) Each of the 90 ordered pairs was presented once, and the order of the pairs was essentially random except for appropriate counterbalancing restrictions (similar to those already described). The tape for the fourth sequence was played backward for half of the subjects in order to achieve a further counterbalancing of possible sequential effects.

The subjects were 50 employees of the Bell Telephone Laboratories. They listened to the four taped sequences of tones through earphones and marked their responses ("U" and "D") on multilithed answer sheets. The subjects were run individually and the entire session, including the tape-recorded instructions, required 20 min for each subject.

IV. RESULTS

The first thing to be noted about the results is that the ability to hear differences in pitch varies markedly from one subject to another. This, of course, has been known for some time. In 1919, for example, Seashore reported that the minimum difference in frequency that can be detected in two successive tones can vary as much as 200-fold between different listeners.[10] Even so, it was a little surprising to find that only 62% of the present 50 subjects responded without error to the 18 "easy" pairs in the first sequence of sinusoidal tones. In the second, more difficult sinusoidal sequence, this figure dropped still further to only 30%. The mean numbers of errors on the first and second series were 1.6 and 3.2, respectively, while the maximum numbers of errors made by any one subject on these same series were 9 and 11, respectively. These last two figures [illegible] level expected on the basis of random respo[illegible] 9 errors per sequence) and, so, provide supp[illegible] notion that the term "tone deaf" may well be [illegible] to an appreciable fraction of the population. [illegible] prisingly, subjects who expressed an interest in music or who played a musical instrument generally made few, if any, errors.

For the purposes of analyzing the judgments of the complex tones, it seemed unwise to combine, indiscriminately, the data from such diverse subjects. Instead, the 50 subjects were first divided into three groups on the basis solely of their performance on the sequences of pure tones. Group I (the most accurate) consisted of the 26 subjects who made no errors on the first sequence of pure tones and no more than 2 errors on the first two sequences taken together. (The weaker criterion, of two errors, was adopted for the second, harder sequence in order to keep Group I reasonably large.) Group II then consisted of those 10 of the remaining subjects who made no more than 2 errors on the first sequence and no more than 5 on both sequences together. Group III, finally, consisted of the remaining 14 subjects (who made the largest number of errors on the pure tones).

Consider, now, the first sequence of complex tones (in which the two tones in each pair were always adjacent in the circular representation). In the case of the subjects who were most accurate on the pure tones (Group I), each pair of adjacent complex tones was presented 4×26 or a total of 104 times. Yet altogether there were only two presentations for which the clockwise tone was not judged higher, and these two presentations were of different pairs. Thus, for each of the 10 pairs of adjacent tones, over 99% of the judgments were in the direction required by the hypothesis of perfect circularity. There was, in short, a complete breakdown of transitivity for the most accurate subjects.

The responses of the subjects in the other two groups conformed to this same circular pattern—though less perfectly, as might be expected on the basis of their lesser accuracy on the pure tones In Group II, the percent choices that conformed with the circular scheme ranged between 87% and 100% for different adjacent pairs. In Group III, the range was between 63% and 82%. Moreover, the reversals from the expected circularity that did occur were roughly uniformly distributed around the circle. This suggests that these reversals resulted from essentially random variability rather than from any consistent departure from circularity.

Consider, next, the final sequence, which contained all possible pairs of complex tones. The data from this sequence are presented in Fig. 3 for the two most extreme groups (I and III). The heavy lines (with filled circles) are for Group I while the light lines (with open circles) are for the less accurate group, III. The solid lines indicate the fraction of presentations for which the

[10] C. E. Seashore, *Psychology of Musical Talent* (Silver, Burdett, New York, 1919), p. 50.

second tone was judged higher in pitch, while the dashed lines indicate the fraction of presentations for which the second tone was judged lower. (These two fractions are complimentary, of course, and do not represent independent data.) The heavy lines show that, for Group I, the second tone was always judged higher when it was only 1 or 2 steps clockwise from the first tone, but was almost always judged lower when it was 8 or 9 steps clockwise (i.e., one or two steps counterclockwise). In the vicinity of 5 steps, the solid and dashed curves cross. As anticipated, then, diametrically opposed tones are ambiguous and the second tone is judged to be higher about as often as it is judged to be lower.

The curves for Group III depart from those for Group I in an instructive manner. Apparently, when the two tones are only a tenth of an octave apart, these subjects have difficulty in telling whether the second tone went up or down. In fact, they give the "correct" response on only about 68% of the presentations. However, when the two tones are further apart (say three-tenths of an octave) their performance improves to about 80%. Then, as the two tones become still further apart, ambiguity sets in (as it did for the more accurate subjects) and the solid and dashed curves again cross (at 50%) near the 5-step point.

In order to avoid unnecessary complexity, the curves for Group II were omitted from Fig. 3. As might be expected, these curves fell between the curves plotted for Groups I and III. They were, however, closer to those for Group I.

An interesting aspect of the diametrically opposed tones is that their ambiguity is usually not *subjective*. That is, the second tone is nearly always heard as clearly higher or clearly lower than the first tone (by a diminished fifth in either case). The ambiguity shows up, rather, in the discrepancy between the judgments made on different occasions. This, of course, is in keeping with other types of ambiguous stimuli—like the well-known Necker cube in vision. Such figures are called reversible only because they are seen in different ways by different subjects or by the same subject at different times; on any one occasion, however, they are seen in only one way. Occasionally, the present subjects would report a case in which the ambiguity did seem to register subjectively. That is, they would say that although the second tone was clearly different in pitch they were unsure of the direction of the difference or, even, that the first tone seemed to "split up" and go in both directions at once.

V. DISCUSSION

The demonstration of a consistent circularity in the judgment of relative pitch raises a serious question, I think, as to the adequacy of a single, linear scale for this subjective attribute. Indeed, the particular manner in which this circularity was achieved seems to provide some reinforcement for the notion that pitch can be analyzed into two distinguishable attributes: viz., "height" and "tonality." Appropriately, band-limited noises with different center frequencies might provide an example of sounds that differ in height but have no clearly defined tonality. The complex tones studied here, however, represent the opposite extreme; they preserve tonality but evidently cannot be ordered with respect to height.

To some extent, the possibility of this latter extreme was already foreshadowed in investigations of subjects who possess the rare faculty known as "absolute pitch." For, while such subjects are capable of remarkable accuracy in giving the conventional letter name for an arbitrarily sounded musical note (e.g., "C", "F#", etc.), they occasionally place it in the wrong octave.[11] One individual who has been demonstrated to have absolute pitch (S. D. Speeth) was included in the present sample of subjects. However, as might be expected from the nature of the particular task, his performance closely resembled that of the other subjects with good relative pitch only.

Except in the case of individuals with absolute pitch, tonality seems to be a purely relative property. That is, one tone can be heard to be the same as or different from another in tonality, but it does not seem to have a subjectively unique quality that can be identified absolutely (i.e., in the absence of any comparison tone). In this respect, too, tonality seems quite analogous to the attribute of being clockwise or counterclockwise. One of two nearby points on a circle can be said to be clockwise from the other; but it makes no sense to say how clockwise a single point is absolutely. For this reason, such terms as "tone color" or "tone chroma," which have been proposed for the circular component of pitch, do not seem to have quite the right connotation for subjects who do not have absolute pitch. Surely, true spectral colors present a rather different situation. They can be identified absolutely (e.g., as "blue," "red," "yellow," etc.) without reference to any other color.

Even among those subjects who do not possess absolute pitch, there appear to be wide variations in the sense of relative pitch. When two successive sinusoidal tones were separated by a tenth of an octave, most of the subjects easily and accurately reported which was higher in pitch. Some of these subjects even remarked that such a difference is so obvious that they find it difficult to conceive that any but a deaf person could be in any doubt. Yet among the present subjects, who all appeared to have perfectly normal hearing in all other respects, there were many who experienced considerable difficulty in making these judgments and, indeed, some who gave responses that had no correlation with the true directions of change in frequency. It would probably be an oversimplification, though, to classify or order subjects on the basis, solely of the *number* of

[11] A. Bachem, "Time Factors in Relative and Absolute Pitch Determination," J. Acoust. Soc. Am. 26, 751–753 (1954).

errors that they made. Possibly quite different processes were responsible for the errors made by different subjects. (In the domain of vision, certainly, there are known to be several distinct types of color weaknesses.)

A simple hypothesis that might account for some of the present errors in the auditory realm would be simply that the difference limen for pitch is much larger for some subjects than for others. This hypothesis is suggested by the fact that the solid curve for Group III in Fig. 3—instead of decreasing monotonically, like the solid curve for Group I—first increases, then decreases, and finally increases again. It also receives some support from those subjects who complained that they had to guess on many trials because the two tones in the pair sounded exactly alike. Some of these subjects, who did indeed make large numbers of errors, also remarked that they were already aware of their inability to distinguish among notes, to reproduce a tune, etc.

This hypothesis seems less plausible for other subjects who said that, although the experimental tones sounded *different*, they needed more time to reach a decision about the *direction* of the difference for each pair. And, indeed, when some of the subjects subsequently were given more time, they did prove able to match a sequence of two tones by singing it. The difficulty here, then, apparently was not at the sensory level so much as at a higher level of interpretation and response. Even when one tone is clearly heard as higher than the other, the subject could become confused as to which tone came first—particularly since the entire two-tone sequence lasted less than half a second. A similar phenomenon arises in the response of unpracticed subjects to the dot-and-dash signals of the Morse Code, for they are relatively accurate in distinguishing the number of dots and number of dashes in a given signal, but they often mix up the order.[12] Then, too, subjects are sometimes apt to confuse binary attributes such as left versus right or, in the present case, up versus down. Some of the errors, then, may represent confusions between the responses and, so, may not stem from any failure to distinguish among tones.

The fact that each of the complex tones described here sounded like a single tone rather than like a chord indicates that there is something special about the interval of an octave. Another tape was prepared in which each tone consisted of 40 sinusoidal components conforming to the same bellshaped weighting function [given in Eqs. (2) and (3) and in Fig. 1] but in which these components were separated by minor thirds rather than by octaves. (There were thus four components per octave and only three tempered half-tone steps around the circle.) Each of these tones definitely sounded like a chord rather than like the single, organlike tone that resulted when the components were separated by octaves.

Even when the tones were built up from octaves, however, one could to some extent attend to the separate sinusoidal components. Indeed, some unconscious fluctuations of attention of this kind may account for the fact that the same pair of diametrically opposed tones led to opposite but highly confident judgments by different subjects. (When several trained musical observers listened together to the same ambiguous pair, such disagreements in interpretation could be quite startling to the observers themselves.)

In order to explore some of these aspects a little further, my colleague, P. D. Bricker, arranged a pitch-matching task in which subjects could alternately listen to the same complex tone (from a continuous tape loop) and a variable sinusoidal oscillator. As might be expected from the results already reported, the three subjects who were tried on this task generally set the oscillator, with considerable accuracy, at the frequency of any of the two or three strongest sinusoidal components of the complex tone. Again, the "tonality" of the complex tone was rather precisely determined, but the octave in which it fell (or its "height") remained ambiguous.

Perhaps the clearest implication that the present findings have for auditory perception is that a powerful "proximity principle" operates in the auditory dimension of frequency just as in the visual dimension of spatial location. Indeed, the perceived changes in pitch between the complex tones described here are quite analogous to a well-known visual phenomenon. Consider, in particular, a long row of identical light bulbs in which, say, every tenth bulb is illuminated. Imagine, now, that all of these lights are simultaneously turned off and the adjacent lights to the right of these are simultaneously turned on. The invariable result, of course, is that the pattern of illuminated spots is seen to shift, as a whole, one position to the right. Note, however, that the physical change would be equally consistent with a shift of nine positions to the *left*. But this is never perceived.

The direction of the perceived shift is entirely determined by how the illuminated spots in the second display are identified with those in the first, and this identification tends to be established on the basis of proximity. That is, a light in the second display is interpreted to be identical to that light in the preceding display that appeared in the closest position. Of course, if a jump of five positions is made, this proximity rule gives rise to ambiguity. The pattern of lights can then be seen as shifting by equal amounts to the left or right.

The situation with the complex tones seems quite parallel. In order to determine whether the second complex tone is higher or lower, the subject must make an identification between the sinusoidal components. If all components shift up by a tenth of an octave, the over-all tone could be regarded as having jumped *down*

[12] R. N. Shepard, "Analysis of Proximities as a Technique for the Study of Information Processing in Man," Human Factors 5, 33–48 (1963).

nine-tenths of an octave. But, again, this would violate the proximity principle, and it seldom if ever occurs. Instead, each component tends to be identified with the closest one to it in the preceding tone and, hence, the tone is usually heard to shift the short way around the circle. When the components of the second tone occur exactly between the components of the first tone, the distance around the circle is the same in both directions and, as Fig. 2 shows, roughly half the subjects make each of the two possible interpretations.

ACKNOWLEDGMENTS

I am indebted to M. V. Mathews for instructing me in the use of his program for the computer synthesis of musical sounds; the availability of that program both suggested the present experiment and made it possible. I am also indebted to P. D. Bricker, who generously donated his time and equipment in exploring alternative methods of presenting the audio tapes, and to Maureen Sheenan, who carried out the bulk of the tabulation and plotting of the data.