

Psychological Review

VOLUME 89 NUMBER 4 JULY 1982

Geometrical Approximations to the Structure of Musical Pitch

Roger N. Shepard
Stanford University

Rectilinear scales of pitch can account for the similarity of tones close together in frequency but not for the heightened relations at special intervals, such as the octave or perfect fifth, that arise when the tones are interpreted musically. Increasingly adequate accounts of musical pitch are provided by increasingly generalized, geometrically regular helical structures: a simple helix, a double helix, and a double helix wound around a torus in four dimensions or around a higher order helical cylinder in five dimensions. A two-dimensional "melodic map" of these double-helical structures provides for optimally compact representations of musical scales and melodies. A two-dimensional "harmonic map," obtained by an affine transformation of the melodic map, provides for optimally compact representations of chords and harmonic relations; moreover, it is isomorphic to the toroidal structure that Krumhansl and Kessler (1982) show to represent the psychological relations among musical keys.

A piece of music, just as any other acoustic stimulus, can be physically described in terms of two time-varying pressure waves, one incident at each ear. This level of analysis has, however, little correspondence to

I first described the double helical representation of pitch and its toroidal extensions in 1978 (Shepard, Note 1; also see Shepard, 1978b, p. 183, 1981a, 1981c, p. 320). The present, more complete report, originally drafted before I went on sabbatical leave in 1979, has been slightly revised to take account of a related, elegant development by my colleagues Krumhansl and Kessler (1982).

The work reported here was supported by National Science Foundation Grant BNS-75-02806. It owes much to the innovative researches of Gerald Balzano and Carol Krumhansl, with whom I have been fortunate enough to share the excitement of various attempts to bridge the gap between cognitive psychology and the perception of music. Less directly, the work has been influenced by the writings of Fred Attneave and Jay Dowling. Finally, I am indebted to Shelley Hurwitz for assistance in the collection and analysis of the data and to Michael Kubovy for his many helpful suggestions on the manuscript.

Requests for reprints should be sent to Roger N. Shepard, Department of Psychology, Jordan Hall, Building 420, Stanford University, Stanford, California 94305.

the musical experience. Because the ear is responsive to frequencies up to 20 kHz or more, at a sampling rate of two pressure values per cycle per ear, the physical specification of a half-hour symphony requires well in excess of a hundred million numbers. Clearly, our response to the music is based on a much smaller set of psychological attributes abstracted from this physical stimulus. In this respect the perception of music is like the perception of other stimuli such as colors or speech sounds, where the vast number of physical values needed to specify the complete power spectrum of a stimulus is reduced to a small number of psychological values, corresponding, say, to locations on red-green, blue-yellow, and black-white dimensions for homogeneous colors (Hurvich & Jameson, 1957) or on high-low and front-back dimensions for steady-state vowels (Peterson & Barney, 1952; Shepard, 1972). But what, exactly, are the basic perceptual attributes of music?

Just as continuous signals of speech are perceptually mapped into discrete internal representations of phonemes or syllables,

continuous signals of music are perceptually mapped into discrete internal representations of tones and chords; just as each speech sound can be characterized by a small number of distinctive features, each musical tone can be characterized by a small number of perceptual dimensions of pitch, loudness, timbre, vibrato, tremolo, attack, decay, duration, and spatial location. In addition, much as the internal representations of speech sounds are organized into higher level internal representations of meaningful words, phrases, and sentences, the internal representations of musical tones and chords are organized into higher level internal representations of meaningful melodies, progressions, and cadences.

The Fundamental Roles of Pitch and Time in Music

The perceptually salient attributes of the tones making up the musically significant chords and melodies are not equally important for the determination of those higher order units. In fact, for the music of all human cultures, it is the relations specifically of pitch and time that appear to be crucial for the recognition of a familiar piece of music. Other attributes, for example, loudness, timbre, vibrato, attack, decay, and apparent spatial location, although contributing to audibility, comfort, and aesthetic quality, can be varied widely without disrupting recognition or even musical appreciation.

The reasons for the primacy of pitch and time in music are both musical and extramusical. From the extramusical standpoint there are compelling arguments, recently advanced by Kubovy (1981), that pitch and time alone are the attributes that are "inispensable" for the perceptual segregation of the auditory ensemble into discrete tones. From the musical standpoint a case can be made that the richness and power of music depends on the listener's interpretation of the tones in terms of a discrete structure that is endowed with particular group-theoretic properties (Balzano, 1980, in press, Note 2). In the case of the human auditory system, moreover, the requisite properties appear to be fully available only in the dimensions of

pitch (Balzano, 1980; Krumhansl & Shepard, 1979; Shepard, 1982) and time (Jones, 1976; Pressing, in press), the dimensions within which higher order musical units such as melodies and chords are capable of structure-preserving transformations.

In this paper I confine myself to the case of pitch and to the question of how a single psychological attribute corresponding, in the case of pure sinusoidal tones, to a simple one-dimensional physical continuum of frequency affords the structural richness requisite for tonal music. One can perhaps readily conceive how structural complexity is achievable in the dimension of time, through overlapping patterns of rhythm and stress, but the structural properties of pitch seem to be manifested even in purely melodic sequences of purely sinusoidal tones (Krumhansl & Shepard, 1979). In the absence of physical overlap of upper harmonics of the sort considered by Helmholtz (1862/1954) and by Plomp & Levelt (1965), wherein does this structure reside?

Cognitive Versus Psychoacoustic Approaches to Pitch

Until recently, attempts to bring scientific methods to bear on the perception of musical stimuli have mostly adopted a psychoacoustic approach. The goal has been to determine the dependence of psychological attributes, such as pitch, loudness, and perceived duration, on physical variables of frequency, amplitude, and physical duration (Stevens, 1955; Stevens & Volkmann, 1940) or on more complex combinations of physical variables (de Boer, 1976; J. Goldstein, 1973; Plomp, 1976; Terhardt, 1974; Wightman, 1973).

By contrast, the cognitive psychological approach looks for structural relations within a set of perceived pitches independently of the correspondence that these structural relations may bear to physical variables. This approach is particularly appropriate when such structural relations reside not in the stimulus but in the perceiver—a circumstance that is well known to students of music theory, who recognize that an interval defined by a given physical difference in log frequency may be heard very differently in different musical contexts. To elaborate on

an example mentioned by Risset (1978, p. 526), in a C-major context the interval B-F is heard as having a strong tendency to resolve by contraction into the smaller interval C-E, whereas in an F[#]-major context the physically identical interval (now called B-E[#], however) is heard as having a similarly strong tendency to resolve by expansion into the larger interval A[#]-F[#]. Moreover, Krumhansl (1979) has now provided systematic, quantitative evidence that the perceived relations between the various tones within an octave do indeed depend on the context-induced musical key with respect to which those tones are interpreted. How, then, are we to represent the relations of pitch between tones as those relations are perceived by a listener who is interpreting the tones musically?

Previous Representations of Pitch

Rectilinear Representations

The simplest representations that have been proposed for pitch have been unidimensional scales based on judgments made in nonmusical contexts. Examples are the "mel" scale that Stevens, Volkmann, and Newman (1937) and Stevens and Volkmann (1940) based on a method of fractionation or the similar scale that Beck and Shaw (1961) later based on a method of magnitude estimation. Pitches are represented in such scales by locations along a one-dimensional line. In these scales, moreover, the location of each pure tone is related to the logarithm of its frequency in a nonlinear manner dictated by the psychoacoustic fact that pairs of low-frequency tones are less discriminable than are pairs of high-frequency tones separated equally in log frequency. Such scales thus deviate from musical scales (or from the spacing of keys on a piano keyboard) in which position is essentially logarithmic with frequency. From a musical standpoint such scales are in fact anomalous in two respects.

First, because of the resulting nonlinear relation between pitch and log frequency, the distance between tones separated by the same musical interval, such as an octave or a fifth, is not invariant under transposition up or down the scale. The failure of musi-

cally significant relations of pitch to emerge as invariant in these scales seems to be a direct consequence of the assiduous avoidance, by the psychoacoustic investigators, of any musical context or tonal framework within which the listeners might interpret the stimuli musically. Attneave and Olson (1971) showed that when familiar melodies, as opposed to arbitrarily selected nonmusical tones, were to be transposed in pitch, listeners required that the separations between the tones be preserved on the musically relevant scale of log frequency—not on a nonlinearly related scale such as the mel scale. That the scale underlying judgments of musical pitch must be logarithmic with frequency is in fact now supported by several kinds of empirical evidence (Dowling, 1978b; Null, 1974; Ward, 1954, 1970).

Second, because of the unidimensionality of scales of pitch such as the mel scale, perceived similarity must decrease monotonically with increasing separation between tones on the scale. There is, therefore, no provision for the possibility that tones separated by a particularly significant musical interval, such as the octave, may be perceived as having more in common than tones separated by a somewhat smaller but musically less significant interval, such as the major seventh. Indeed, even the musically more relevant log-frequency scale, also being unidimensional, is subject to this same limitation. Yet, in the case of the octave, which corresponds to an approximately two-to-one ratio of frequencies, the phenomenon of augmented perceptual similarity at that particular interval (a) has long been anticipated (Boring, 1942, pp. 376, 380; Licklider, 1951, pp. 1003-1004; Ruckmick, 1929), (b) was in fact empirically observed at about the time that the unidimensional mel scale was being perfected (Blackwell & Schlosberg, 1943; Humphreys, 1939), (c) has since been much more securely established (Allen, 1967; Bachem, 1954; Balzano, 1977; Dowling, 1978a; Dowling & Hollombe, 1977; Idson & Massaro, 1978; Kallman & Massaro, 1979; Krumhansl & Shepard, 1979; Thurlow & Erchul, 1977), and (d) probably underlies the remarkable precision and cross-cultural consistency with which listeners are able to adjust a variable tone so that it stands

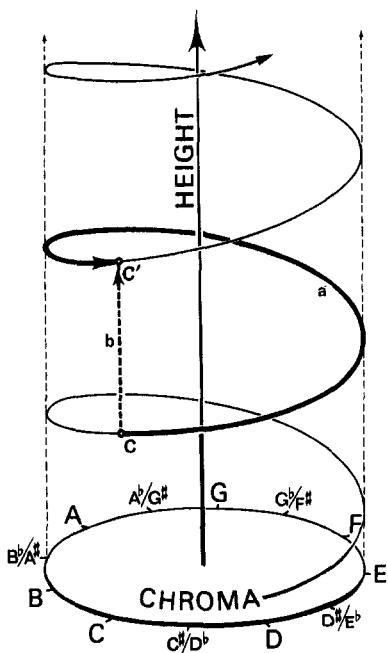


Figure 1. A simple regular helix to account for the increased similarity between tones separated by an octave. (From "Approximation to Uniform Gradients of Generalization by Monotone Transformations of Scale" by Roger N. Shepard. In D. I. Mostofsky (Ed.), *Stimulus Generalization*. Stanford, Calif.: Stanford University Press, 1965, p. 105. Copyright 1965 by Stanford University Press. Reprinted by permission.)

in an octave relation to a given fixed tone (Burns, 1974; Dowling, 1978b; Sundberg & Lindqvist, 1973; and, originally, Ward, 1954).¹

Simple Helical Representations

We can obtain geometrical representations that are consistent with an increased similarity at the octave by deforming a rectilinear representation of pitch into a higher dimensional embedding space to form a helix, as proposed for this purpose by Drobisch as early as 1846, or a spiral, as proposed by Donkin in 1874 (see Pikler, 1966; Ruckmick, 1929; and for a recently proposed planar spiral, Hahn & Jones, 1981). For, unlike a straight line, a helix or spiral that completes one turn per octave achieves the desired increase in spatial proximity between points an octave apart—at least if the slope of the curve is not too steep. (Compare Paths a and

b between the tones C and C', an octave apart, in Figure 1.) Moreover, this is true whether the curve is embedded in a cylinder, as proposed by Drobisch, a flat plane, as suggested by Donkin, or a bell-shaped surface of revolution, as advocated by Ruckmick (1929). Despite their differences, the representations proposed by these authors were alike in having adjacent turns more closely spaced toward the low-frequency end, where given differences in log frequency are less discriminable. (See the figures reproduced in Pikler, 1966; Ruckmick, 1929.) These representations were analogous, in this respect, to the unidimensional psychophysical representations of Stevens and his colleagues (Stevens et al., 1937; Stevens & Volkman, 1940).

In 1954, before learning of these early proposals, I had attempted to accommodate a heightened similarity at the octave by means of a tonal helix (Note 3) that differed, however, from the ones proposed by Drobisch, Donkin, and Ruckmick in being geometrically uniform or regular. (See Figure 1, which, except for relabeling, is reproduced from Shepard, Note 3—as it later appeared in Shepard, 1965). Because it is geometrically regular, this is the helical analog of the unidimensional scale having the musically more relevant logarithmic structure advocated by Attneave and Olson (1971); in it the distance corresponding to any particular musical interval is invariant throughout the representation. The uniform helix also possesses an additional advantage over curves embedded in variously shaped surfaces of revolution, such as Ruckmick's (1929) "tonal bell." Only when the helix is regular, and hence embedable in a cylindrical surface, will tones standing in the octave relation, in addition to coming into closer proximity with each other, fall on a common straight line parallel to the axis of the helix. Such lines

¹ Incidentally, these studies agree in showing that the ratio of physical frequencies of pure tones that subjects set in the octave relationship is about 2.02:1 and not exactly 2:1. This small "stretched-octave" effect (Bazzano, 1977; Burns, 1974; Dowling, 1973a; Sundberg & Lindqvist, 1973; Ward, 1954) still leads to an approximately logarithmic relation between pitch and frequency and does not vitiate any of the conclusions to be drawn here.

can be thought of as projecting all tones with the same musical name but differing by octaves (e.g., the tones C, C, C", etc.) down into a single corresponding point in a "chroma circle" on a plane orthogonal to the axis of the helix (see Figure 1). Moreover, this projective property, unlike the property of augmented proximity, holds regardless of the slope of the helix.

Physical Realization of the Chroma Circle

It is, in fact, the projective property of the regular helix that subsequently led me to a method of physically separating the two underlying components of pitch implicit in the helical representation, namely, the rectilinear component called *pitch height*, corresponding to the axis of the helix (or of the cylinder in which it lies), and the circular component variously called *tone quality* (Revesz, 1954) or *tone chroma* (Bachem, 1950, 1954), corresponding to the circumference of that cylinder (see Shepard, 1964b). What was required was the specification of two physical operations corresponding to the geometrical projection of the entire helix onto the central axis, in the one case, and onto an orthogonal plane, in the other. The auditory realization of the required operations was greatly facilitated by the development, at just this time, of computer techniques for the additive synthesis of arbitrarily specified tones (Mathews, 1963).

For the first operation I proposed a broadening of the band of energy around the center frequency of each tone until the resulting narrow-band noise encompassed about an octave. Because the different sounds generated in this way have different center frequencies, they still differ over the whole range of pitch height. But, because they have all been spread alike around the chroma circle, they can no longer be discriminated with respect to chroma. This operation thus corresponds to collapsing the helix onto its central axis.

For the second operation I proposed, instead, a harmonic elaboration of each tone until it included, alike, all multiples and submultiples of the original frequency (i.e., all tones standing in octave and multiple-octave relations to that original tone), with the am-

plitudes of the component frequencies determined by a fixed bell-shaped spectral envelope that was at its maximum near the middle of the standard musical range and that gradually fell away in both directions to below-threshold levels for very low and very high frequencies. The different sounds generated in this way remain fully distinct in chroma but are all equivalent in height. Thus, in shifting through chroma, from C to C[#] to D to D[#] and so on to B, the next step (though still heard as a step up in pitch), instead of leaving one an octave higher at C', leaves one back at the original starting tone C (Shepard, 1964b).² Indeed, application of multidimensional scaling to measures of similarity derived from judgments of relative pitch between tones varied in this manner (Shepard, 1964b) yielded the almost perfectly circular solution displayed in Figure 2 (a) (see Shepard, 1978a). (A similarly circular representation for tones generated in this way has also been reported by Charbonneau and Risset, 1973.)

Although this circular component, chroma, emerges most compellingly when the rectilinear component, height, is artificially suppressed, as with these special, computer-generated tones, the claim is that this circular component is necessarily present in all musical tones for which tones separated by an octave are perceived as more closely related than tones separated by a somewhat smaller interval. Circular multidimensional scaling solutions have in fact been obtained for ordinary musical tones. Figure 2 (b), for example, reproduces a similarly circular pattern subsequently obtained by Balzano (1977, 1982) from a multidimensional scaling analysis of his own discriminative reaction time data for melodic intervals.³

² The illusion of circular or "endlessly ascending" tones can be heard on a commercial record ("Shepard's Tones," 1970) or on a short 16-mm film (Shepard & Zajac, 1965), which we believe to be the first film in which both the sound and the animation were generated by computer. I demonstrated the independent variation of linear pitch height and circular tone chroma at the 1978 meeting of the Western Psychological Association (Shepard, Note 1).

³ One should, however, exercise caution in basing the inference of circularity solely on a multidimensional scaling solution. The curvature evident in some obtained solutions (e.g., the one reported by Levelt, Van de Geer,

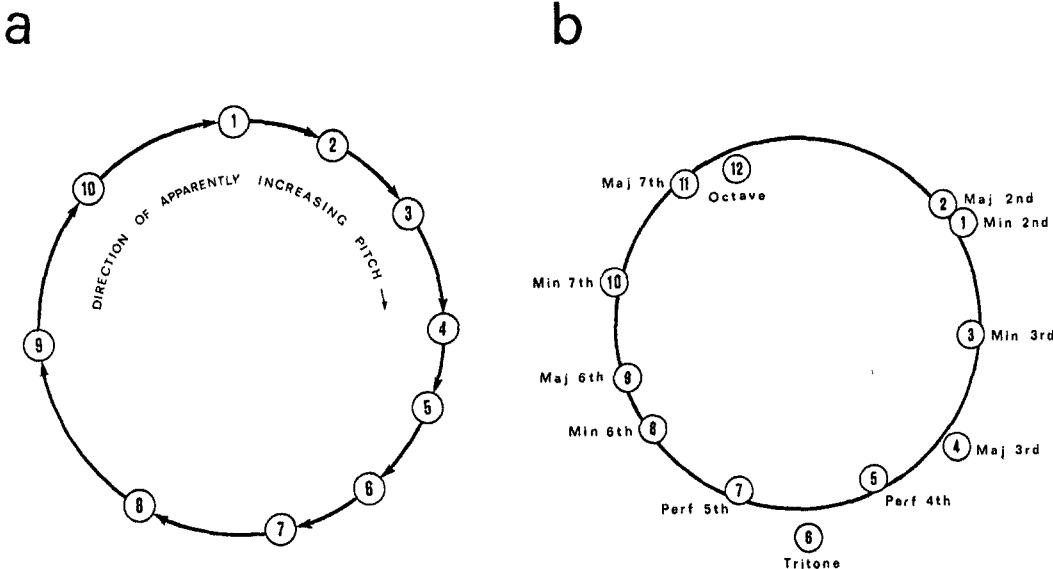


Figure 2. Chroma circles recovered by multidimensional scaling (a) for 10 computer-generated tones especially designed to eliminate differences in pitch height and (b) for twelve ordinary musical tones. (Panel a is from "The Circumplex and Related Topological Manifolds in the Study of Perception" by R. N. Shepard. In S. Shye (Ed.), *Theory Construction and Data Analysis in the Behavioral Sciences*. San Francisco, Calif.: Jossey-Bass, 1978. Copyright 1978 by Jossey-Bass, Inc. Reprinted by permission. Panel b is from *Chronometric Studies of the Musical Interval Sense* by G. J. Balzano. Unpublished doctoral dissertation, Stanford University, 1977. Reprinted by permission.)

By now the possibility of analyzing perceived pitch into the rectilinear and circular components of height and chroma has been accepted by a number of researchers (e.g., Bachem, 1950, 1954; Balzano, 1977; Deutsch, 1972, 1973; Jones, 1976; Kallman & Massaro, 1979; Pikler, 1955; Revesz, 1954; Risset, 1978). Even among advocates of a helical representation, though, opinions may still differ concerning the relative merits of a geometrically regular structure such as I proposed versus a more or less distorted variant such as Ruckmick (1929) advocated. Here, again, one's predilection may depend on whether one takes a more psychoacoustic or a more cognitive and musical point of view.

Argument for a Geometrically Regular Structure

From the psychoacoustic standpoint it seems natural to suggest that the spacing

between points in the representational structure, quite apart from whether that structure is basically rectilinear or helical in overall shape, should be adjusted to reflect how the operating characteristics of the sensory transducers shift as we move from low to high input frequencies. Someone preoccupied with such sensory considerations might even see some significance in the resemblance of a distorted spiral or helix to the anatomical conformation of the cochlea.

By contrast, a more cognitively and musically oriented approach to pitch is likely to regard such considerations of automatic peripheral transduction (and anatomy) as largely irrelevant. Adopting something like Chomsky's (1965) competence-performance distinction, I suggest that if it is *musical* pitch that interests us, the representation should reflect the deeper structure that underlies a listener's competence to impose a musical interpretation on a stream of acoustic inputs under favorable conditions. Such an interpretive structure continues to exist whether or not the acoustic stimuli in a particular stream fall within the range of

& Plomp, 1966) may simply reflect an artifact that almost always arises when basically one-dimensional data are fit in a higher dimensional space (Shepard, 1974, pp. 386-388).

frequencies, amplitudes, or durations that can be adequately transduced by a particular ear or whether or not a preceding context has been provided that is sufficient to activate and to orient or tune the internal structure required for a musical interpretation.

From this cognitive standpoint the already cited evidence for invariance under transposition (Attneave & Olson, 1971; Dowling, 1978b) would require not only that the structure be helical (rather than spiral, say) but also that the helix be geometrically regular and, so, be carried into itself by a rigid screwlike motion under the musical transposition that maps any tone into any other. Only then will the geometrical structure preserve the property, essential to music, that all pairs of tones separated by a given interval, such as an octave or a fifth, have the same musical relation regardless of their overall pitch height. Indeed, from this standpoint the reason that this structure must take a helical form (just as much as the reason that the DNA molecule must take a helical form) is a consequence of the fundamental geometrical fact that the most general rigid motion of space into itself is a combination of a rotation and a translation along the axis of the rotation, that is, a screw displacement (Coxeter, 1961, pp. 101, 321; H. Goldstein, 1950, p. 124; Greenwood, 1965, p. 318; Hilbert & Cohn-Vossen, 1932/1952, pp. 82, 285).

One should not suppose, here, that the octave-related tones that project onto the same point on the chroma circle share some absolutely identifiable quality of chroma any more than the projection of any tone onto the central axis of the helix corresponds to an absolutely identifiable quality of pitch height. Only those rare individuals possessing "absolute pitch" can recognize a chroma absolutely (Bachem, 1954; Siegel, 1972; Ward, 1963a, 1963b). For most individuals (including most musicians), in the case of pitch just as in the case of many other perceptual dimensions (loudness, brightness, size, distance, duration, etc.), it is the relations between presented values that have well-defined internal representations, not the values themselves (Shepard, 1978c, 1981b).

In fact, as Attneave and Olson (1971) have so persuasively argued, pitch is perhaps

best thought of not as a stimulus but as a "medium" in which auditory patterns (chords or tunes) can move about while retaining their structural identity, just as visual space is a medium in which luminous patterns can move about while retaining their structural identity. We are not very good at recognizing the position of an isolated point of light in otherwise empty space, but we can say with great precision whether one such point is above or below another or whether three points form a straight line or a right triangle. Likewise, many of us are poor at identifying the pitch of an isolated tone but quite good at saying whether one tone is higher or lower than another or whether three tones stand in octave relations or form a major triad. In either the visual or the auditory case, any such pattern moved to a different location or pitch continues to be recognized as the same pattern.

I go beyond the formulations of Attneave and Olson (1971) and of Dowling (1978b), however, in saying that in the case of pitch, such a motion must be helical. For example, suppose we alternate one well-defined pattern, say a major triad, with exactly the same pattern but each time with the second pattern displaced farther away from the first in pitch height. As the two patterns are moved apart from their initially complete coincidence, each will retain its own structural identity. But the perceived degree of relation between the two patterns will first decrease and then increase again as all three components of one come into the octave relation with corresponding components of the other. (The perceived relation may also increase somewhat at certain intermediate positions as the two triads pass through other special tonal relations, just as the visual similarity to the original of a rotating copy of a triangle will increase at certain intermediate orientations before coming again into complete coincidence at 360°. See Shepard, 1982.) Such a phenomenon is not explicable solely in terms of increasing separation within a purely rectilinear medium; it implies a medium with a circular component. Motions in such a medium are thus auditory analogs of the operations of mental rotation and rotational apparent motion in the visual domain (see Shepard & Cooper, 1982).

Limitations of the Simple Helix

The structure of the simple regular helix pictured in Figure 1 was dictated by two considerations: invariance under transposition and increased similarity at the octave. In such a regular helix, moreover, the special octave relation is represented by unique collinearity or projectability as well as by augmented spatial proximity. As it stands, however, the helical structure does not provide either augmented proximity or collinearity for tones separated by any other special musical interval. Yet, beginning with Ebbinghaus, Drobisch's proposal of a helical representation has been criticized for its failure to account for the special status of the interval of a perfect fifth (see Ruckmick, 1929).

There are, indeed, a number of converging reasons for supposing that the fifth should, like the octave, have a unique status. (a) As has been known at least since Pythagoras, after the 2-to-1 ratio in the lengths of a vibrating string that corresponds to the octave (which as we now know determines a 1-to-2 ratio in the resulting frequency of vibration), the fifth corresponds to the next simplest, 3-to-2, ratio (Helmholtz, 1862/1954). (b) In the case of musical and, therefore, harmonically rich tones, those separated by a fifth also have, within the octave, the fewest upper harmonics that deviate from coincidence by an amount expected to produce noticeable beats (Helmholtz, 1862/1954; Plomp & Levelt, 1965). (c) Moreover, such beats can be subjectively experienced even when the harmonics that most contribute to them are not physically present (Mathews, Note 4; see also Mathews & Sims, 1981). (d) Correspondingly, simultaneously sounded tones differing by a fifth—even pure, sinusoidal tones—tend to be heard as particularly smooth, harmonious, or consonant, and the fifth, together with the similarly harmonious major third, completes the uniquely stable and tonally centered chord, the major triad (Meyer, 1956; Piston, 1941; Ratner, 1962; Schenker, 1906/1954, p. 252). (e) In addition, the interval of the fifth plays a pivotal role in tonal music, being the interval that separates musical keys that share the greatest number of tones and between which

modulations of key most often occur (Forte, 1979; Helmholtz, 1862/1954; Schenker, 1906/1954). (f) Finally, according to Balzano's (1980, Note 2) group-theoretic analysis, the preeminence of the fifth in tonal music has a basis in abstract structural constraints independent of the psychoacoustic facts noted under (a) and (b).

Despite these diverse indications of the importance of the perfect fifth, the fifth has largely failed to reveal its unique status in psychoacoustic investigations for the same reason, I believe, that the octave often revealed its unique status only weakly, if at all. In the absence of a musical context, tones—particularly the pure sinusoidal tones favored by psychoacousticians—tend to be interpreted primarily with respect to the single, rectilinear dimension of pitch height. Without a musical context there is insufficient support for the internal representation of more complex components of pitch—components that may underlie the recognition of special musical intervals and that (like the chroma circle) are necessarily circular because, again, the musical requirement of invariance under transposition entails that each such component repeat cyclically through successive octaves.

Recent Evidence for a Hierarchy of Tonal Relations

Motivated by the considerations just given, Carol Krumhansl and I initiated a new series of experiments on the perception of musical intervals within an explicitly presented musical context. In this way we were in fact able to obtain clear and consistent evidence that the perfect fifth is at the top of a whole hierarchy of special relations within each octave. In these experiments we established the necessary musical context, just prior to presenting the to-be-judged test tone or tones, simply by playing, for example, the sequence of tones of a major diatonic scale (the tones named *do, re, mi, fa, sol, la, and ti* and corresponding, in the key of C major, to the white keys of the piano keyboard). Ratings of the ensuing test tones yielded highly consistent orderings of the musical intervals across listeners having equivalent musical backgrounds. This was true both in

the initial experiments in which listeners rated, in effect, the extent to which each individual test tone out of the 13 within one complete octave was substitutable for the tonic tone (*do*) that would normally have completed the major scale presented as context (Krumhansl & Shepard, 1979; also see Krumhansl & Kessler, 1982) and in further experiments in which listeners rated the similarities between the two test tones in all possible pairs selected from one complete octave (Krumhansl, 1979).

Only for listeners with little musical background did the obtained orderings of the musical intervals agree with previous psychoacoustic results in which similarity was determined primarily by proximity in pitch height between the two tones making up the interval, that is, in which the ranking of the intervals with respect to similarity or mutual substitutability of their two component tones was, from greatest to least, unison, minor second, major second, minor third, major third, and so on. For the more musically oriented listeners, the results tended, instead, toward the entirely different ranking: unison and octave (nearly equivalent to each other), followed by the fifth and sometimes the major third, followed by the other tones of the diatonic scale, followed by the remaining, nondiatonic tones (those corresponding in the key of C major to the sharps and flats or black keys of the piano).

In short, data collected from listeners who invest the test tones with a musical interpretation consistently reveal a whole hierarchy of tonal relations that cannot be accommodated within previously proposed geometrical representations of pitch whether rectilinear, helical, or spiral. Accordingly, it now appears justified to present some alternative, generalized helical structures together with the steps that led to their construction and some evidence that such generalized structures are indeed capable of accommodating the musically primary tonal relations. The following is intended, therefore, as the first full account of these new representations of pitch—first briefly described in 1978 (Shepard, Note 1; also see Shepard, 1981a, or, for a description concurrent with that presented here but following a different derivation, Shepard, 1982).

New Representations for Musical Pitch

The Diatonic Scale as an Interpretive Schema

In their characteristic eschewal of musical context, psychoacoustically oriented investigators missed the essential musical aspect of pitch. By failing to elicit, within the listeners, the discrete tonal schema or "hierarchy of tonal functions" (Meyer, 1956, pp. 214–215; Piston, 1941; Ratner, 1962) associated with a particular musical key, these investigators left the listeners with no unique cognitive framework within which to interpret the test tones. Even musically sophisticated listeners therefore had little choice but to make their judgments on the basis of the simplest attribute of tones differing in frequency—pitch height.

Cognitively oriented researchers are now recognizing that interpretive schemata play an essential role in the perception of musical pitch, just as they do in perception generally. In the case of pitch, the primary schema seems to be the musical scale—usually, in the case of Western listeners, the familiar major diatonic scale (*do, re, mi*, etc.). As noted by Dowling (1978b, *in press*), even though the most commonly used musical scales differ somewhat from culture to culture, they all share certain basic properties. Regardless of the total number of tones permitted by each scale, most are organized around five to seven "focal pitches" per octave. Moreover, the steps between such pitches rather than being constant in log frequency are almost always arranged according to a particular asymmetric pattern that repeats exactly within every octave.

The cyclic repetition of the pattern from octave to octave can be explained in terms of the perceived equivalence of tones differing by an approximately 2-to-1 ratio of frequencies, which led to the proposed simple helix for pitch. The other structural universals of musical scales have been attributed to pervasive cognitive constraints on the number of absolutely identifiable categories per perceptual dimension (7 ± 2 , as enunciated by Miller, 1956; cf. Dowling, 1978b) and to the requirement that the scale have a structure that affords reference points or tonal centers to which a melody can move

or come to rest (Balzano, 1980; Zuckerkandl, 1956, 1972). In this view, evenly spaced scales, such as the whole-tone scale or the chromatic (half-tone or twelve-tone) scale, have not been widely used because, in the absence of a reference point, music lacks the tension, motion, and resolution that engages the suitably tuned mind with such dynamic force.

Balzano's (1980, in press, Note 2) group-theoretic analysis has carried this line of argument to a new level of formal elegance and detail. He showed that the requirements that the organization of the scale be invariant under transposition and that each tone have a unique functional role within the scale jointly constrain both the number of scale tones within each octave and the approximate spacing between those tones. In fact, he showed that among possible scales presupposing a division of the octave into less than 20 steps (a restriction I take to be desirable, if not necessary, to avoid overloading the human cognitive system), the diatonic scale is the only one satisfying these requirements. Thus, it may not be accidental that this scale has become the basis of Western music, in which the structural complexities of harmony and counterpoint (though not of other aspects such as, particularly, rhythm; see Pressing, in press) have reached their greatest development. Nor is it a coincidence that the diatonic scale has arisen, with but slight variations, "independently in different ages and geographical locations" (Pikler, 1955, p. 442) and, according to recent archeological evidence, can be traced back over 3,000 years to the earliest decipherable records (Kilmer, Crocker, & Brown, 1976).

In any case, as Dowling (1978b, in press) notes, there are many reasons to believe that every culture has some such discrete tonal schema, which through early (and possibly irreversible) tuning provides a framework with respect to which listeners interpret all music they hear. Music of another culture may therefore be misinterpreted by assimilation to the listener's own, somewhat different tonal schema (Francès, 1958, p. 49). Within a culture, children evidently internalize the prevailing schema by 8 years of age (Imberty, 1969; Zenatti, 1969). The in-

terpretation of continuously variable tones with respect to this discrete internalized schema appears to be a kind of "categorical perception" (Burns & Ward, 1978; Krumhansl & Shepard, 1979; Locke & Kellar, 1973; Siegel & Siegel, 1977a, 1977b; Zatorre & Halpern, 1979; Blechner, Note 5) much like that originally reported for the perception of speech sounds (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967).⁴ Indeed, during a child's development the process whereby the abstract and seemingly universal structure of the underlying tonal schema becomes tuned to the particular musical scale entrenched in that child's culture may be like the process whereby, according to Chomsky (1968), the innate schematism that underlies all human languages becomes tuned to the particular language of that culture.

Laboratory studies substantiate the role of the diatonic scale for Western listeners. Cohen (1975, Note 6) showed that after hearing a short excerpt from a piece of music, listeners can generate the associated diatonic scale with considerable accuracy. Furthermore, listeners recognize or reproduce a series of tones more accurately when the tones conform to a diatonic scale than when they do not (e.g., Attneave & Olson, 1971; Cohen, 1975; Dewar, 1974; Dewar, Cuddy, & Mewhort, 1977; Francès, 1958, Experiments 3 and 4; Krumhansl, 1979). Moreover, as I have already noted, Krumhansl (1979) established that the perception of musical intervals depends on the key or

⁴ From the cognitive musical standpoint taken here, the important issues concern the structural relations between tones as they are represented in a discrete internal structure, such as the diatonic schema. To the extent that the perception of musical tones is categorical, the question of whether the physical stimuli that are thus categorically associated with nodes in this structure came from a physical scale of equal temperament (in which all chromatic steps are uniform with respect to log frequency) or a scale of just intonation (in which the frequency ratios of the musical intervals take the simplest integer forms), though affecting the timbral quality of simultaneously sounded harmonically rich tones (Helmholtz, 1962/1954) to a small extent (Mathews & Sims, 1981; Mathews, Note 4), is largely irrelevant. The same consideration leads me to disregard the distinction between a sharp of one note and the flat of the note just above (e.g., C[#] versus D^b).

tonality of the diatonic scale with respect to which the listeners interpret the test tones. Finally, in the informal experiment mentioned earlier in which a major triad was alternated with the same major triad displaced in pitch height, I noticed that between the unison and the octave displacement, the greatest perceived relation was at the displacement of a perfect fourth and a perfect fifth. But these intervals, which are adjacent to the unison around the circle of fifths, do not have the greatest number of component tones in the octave or unison relation; rather, for these displacements alone, all tones in either triad are in the diatonic scale determined by the alternately presented triad.

The rectilinear and the simple helical and spiral representations of pitch bear little relationship to the diatonic or related musical scales found in human societies. It is not surprising, therefore, that these previously proposed geometrical representations fail to provide an account of the various phenomena of culture-specific, context-dependent, and apparently categorical perception.

My own approach to the representation of pitch grew out of an informal observation concerning the diatonic scale: In listening to the eight successive tones of the major scale (*do, re, mi, . . . , do*), I tended to hear the successive steps as equivalent, even though I knew that with respect to log frequency, some of the intervals (viz., the interval *mi* to *fa* and the interval *ti* back to *do* an octave higher) are only half as large as the others. This apparent equivalence of successive steps of the diatonic scale could not be dismissed as an inability to discriminate between major and minor seconds.

When I then used a computer to generate a series of eight tones that divided the octave into seven equal steps in log frequency (a series that does not correspond to any standard musical scale), the successive steps sounded oddly nonequivalent. Apparently, just as we cannot voluntarily override the visual system's tendency to interpret parallelograms projected on the two-dimensional retina as rectangles in three-dimensional space (Shepard, 1981c, p. 298), I could not wholly override my auditory system's tendency to interpret tones in terms of the diatonic scale. Thus, after they had been made

physically equal, the steps that should have been half as large if the scale had been diatonic (viz., the steps between Tones 3 and 4 and between Tones 7 and 8) sounded too large. Moreover, in just completed, more formal experiments, a student and I have now obtained strong quantitative confirmation of this phenomenon (Jordan & Shepard, Note 7).

Our tendency to hear the successive intervals of the diatonic scale as uniform, which I take to underlie this auditory illusion, probably depends on perceptual set. That tendency may be weakened in musicians such as singers, trombonists, and string players who, unlike passive listeners or those who primarily play keyboard or other wind instruments, must learn to make vocal or motor adjustments essentially proportional to actual differences in log frequency. (Such differences in set may in part account for the departure from equality of successive scale steps implied by the results of Francès, 1958, Experiment 2, or Krumhansl, 1979). Nevertheless, our own results (Jordan & Shepard, Note 7) indicate that there is a tendency to hear scale steps as more nearly equivalent than they physically are. The work of Balzano (1980, 1982; Balzano & Liesch, in press; Balzano, Note 2) has also provided support for the notion that in addition to pitch height and tone chroma, the discrete steps or "degrees" of the musical scale are psychologically real.

Suppose, then, that listeners interpret successive tones of the major scale (e.g., C, D, E, F, G, A, B, C', in the key of C major) by assimilating each tone to a node in an internalized representation of the diatonic scale. If the steps in this internal representation are functionally equivalent (as steps), then the perception of uniformity would follow, despite the fact that the physical differences are only half as great for the steps E to F and B to C' as for the other steps.

Derivation of a Double Helix

I propose to represent musical tones by points in space and, as in the case of the simple regular helix, to represent the underlying relations between their pitches by geometrical relations of distance and collinear-

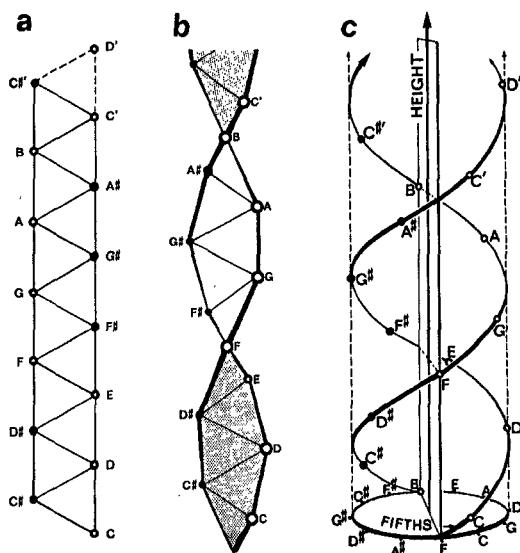


Figure 3. Stages in the construction of a double helix of musical pitch: (a) the flat strip of equilateral triangles, (b) the strip of triangles given one complete twist per octave, and (c) the resulting double helix shown as wound around a cylinder. (Panel c is from "Structural Representations of Musical Pitch" by R. N. Shepard. In D. Deutsch (Ed.), *Psychology of Music*. New York: Academic Press, 1982. Copyright 1982 by Academic Press, Inc. Reprinted by permission.)

ity between those points. Thus, I propose that the points corresponding to tones of the same chroma name (e.g., C, C', C'', etc.) but located in different octaves should fall on the same line within the geometrical structure and, hence, should project down into the same point on an orthogonal plane.

Now, however, I propose that the geometrical structure must also reflect certain subjective properties of the diatonic scale. Because steps between adjacent tones of the diatonic scale are of only two physical sizes—half-tone steps (minor seconds) and whole-tone steps (major seconds)—I start by considering the geometrical constraints that are entailed by perceived distances between tones separated by half- and whole-tone steps. My initial attempt to erect a new geometrical structure is then based on the following three requirements, the first two of which are the same as for the earlier simple helix but the third of which is new, having its origin in the just-mentioned informal observation about subjective uniformity of successive steps of the diatonic scale. (Later, I allow departures from the strict uniformity

imposed by this third, perhaps arguable, requirement.)

1. Invariance under transposition: Tones separated by a given musical interval must be separated by the same distance in the underlying structure regardless of the absolute location (height) of those two tones.

2. Octave equivalence: All the tones standing in octave relationships to any given tone, and only those tones, must fall on a unique (chroma) line passing through the given tone.

3. Uniformity of scale steps: The steps between tones that are adjacent within the major diatonic scale in any particular key should be represented as equal distances in the underlying geometrical structure.

According to Requirement 1 the equivalence between half- and whole-tone steps imposed within any one key by Requirement 3 must hold for all keys. Consequently, all major and minor seconds must be represented by equal distances in the underlying invariant structure. Thus, the geometrical implication of Requirements 1 and 3 together is that the points corresponding to any three successive tones of the chromatic scale must form an equilateral triangle, and these triangles must be connected together in an endless series as is shown for one octave (and in flattened form) in Figure 3 (a).

For illustration, the tones included in one particular key, namely, the key of C, are indicated by open circles in the figure, whereas the tones not included in that key (i.e., the tones corresponding to the black keys on the piano) are indicated by filled circles. The pattern for any other key would be the same except for a translation and, in the case of half the keys, a reflection about the central axis of the strip. Two things should be noted about this pattern. First, steps between adjacent open circles are of uniform size, in accordance with Requirement 3. Second, the pattern of open circles as a whole possesses the asymmetry necessary to confer on each tone within any one octave of the scale the unique structural role or tonal function required by music theory. For example, the most stable tone within the scale, the tonic, is always the lowest tone in the set of three adjacent scale tones along one edge of the strip, whereas the second

most stable tone, the dominant, a fifth above, is always the next-to-lowest tone in the set of four adjacent scale tones along the other edge of the strip.

Thus, by going to a more complex representation than a simple unidimensional scale of pitch height, we avoid an objection to the subjective equality of the intervals of the diatonic scale, namely, that such uniformity would "deny a major source of melodic variety" (Dowling, 1978b, p. 350). The structural uniqueness of the diatonic scale that underlies the desired variation of melody and modulation of key can be embodied in a qualitative asymmetry rather than in a purely quantitative one. No such structural uniqueness is possible within scales that are both quantitatively and qualitatively symmetric, such as the whole-tone scale, represented by just one edge of the strip of triangles, or the chromatic scale, represented by the symmetrical zigzag path that alternates between the two edges of the strip.

Still, although the strip of triangles shown in Figure 3 (a) is consistent with Requirements 1 and 3, it is not consistent with Requirement 2, according to which any two tones standing in an octave relation (such as C-C', C'-C'', etc.) must fall on their own unique chroma line. For, in this flattened form, the line passing through C and C' also passes through D, E, F#, G#, and A#, which are not octave or chroma equivalents to C and C'. Likewise, the line passing through C# and C'' also passes through D#, F, G, A, and B. However, there is no requirement that this structure remain flat. In fact, any folding of the strip of triangles along the sides of the triangles will preserve the equilaterality of the triangles imposed by Requirements 1 and 3. But in order to ensure the full satisfaction of Requirement 1, the folding must be done in a uniform manner throughout the strip. Only then will the transformation of the strip into itself induced by transposition into a different key consist only of rigid translations, rotations, and reflections of the structure as a whole and, thus, preserve all distances within it. Specifically, all folds must be made at the same angle.

The case in which all folds are also made in the same direction can be ruled out be-

cause the resulting structure then curves back into itself, merging points corresponding to distinct tones and eliminating the rectilinear component of pitch height. The case in which alternate folds are made in opposite directions does not lead to these undesirable consequences, however. As might be expected from the fact that the most general rigid motion of three-dimensional space into itself is a combination of a rotation and a translation along the axis of rotation, folding in alternating directions produces an endless helical structure with the amount of twist per octave determined by the uniform angle of folding. Because there are two edges to the originally flat strip, each corresponding to one of the two distinct whole-tone scales, such folding leads to a double helix.

In order to achieve the collinearity of tones that are equivalent except for height, in accordance with Requirement 2, there must be an integer number of full twists of the structure per octave. The flat version displayed in Figure 3 (a) corresponds to the trivial 0° twist and, as I noted, must be excluded because it does not segregate the chroma lines: They collapse into the two whole-tone scales. At the other extreme, two full twists per octave lead to a different kind of degeneracy in which all the triangles collapse into a single triangle. This different kind of flat configuration must also be excluded because in it not only octaves but also minor thirds map onto each other and, again, we lose the component of pitch height. The single remaining case of just one full twist per octave is the unique solution we seek. It is the nondegenerate double-helical structure of which one octave is portrayed in Figure 3 (b). In three-dimensional space it alone satisfies Requirements 1-3.

Emergent Properties of the Double Helix

Musically significant properties emerge from the double-helical representation that were not explicitly used in its derivation. First, successive tones of the chromatic scale project onto the axis of the helix in equal steps of pitch height. More remarkably, as is illustrated in Figure 3 (c), the same tones project down onto the plane orthogonal to that axis to form a circle, the "cycle of

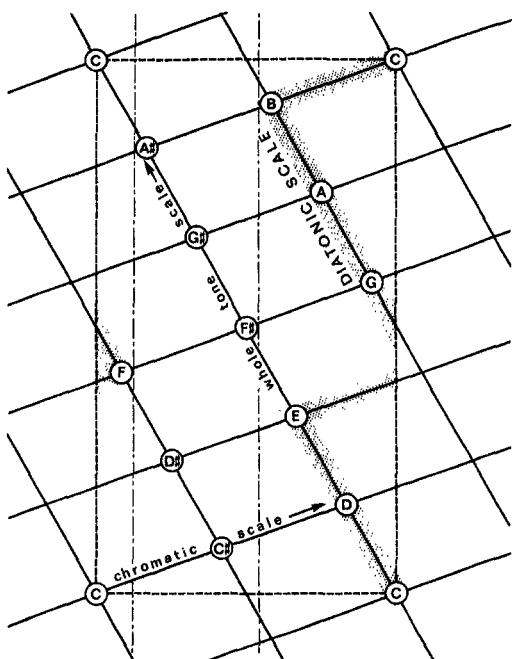


Figure 4. The two-dimensional melodic map of the double helix obtained by cutting and unwrapping the cylinder in Figure 3 (c). (From Shepard, 1981a. Copyright 1981 by Music Educators National Conference. Reprinted by permission.)

fifths" fundamental in music theory. In this projection tones separated by an octave map onto each other, and tones separated by a perfect fifth are closest neighbors around the circle.

As I noted, it was the possibility of projecting the geometrically regular version of the single helix (Figure 1) down into the chroma circle (Figure 2) that originally led me to devise a method of generating tones that in going endlessly around the chroma circle, give the illusion of ascending endlessly in pitch (Shepard, 1964b). Similarly, here, the emergence of the cycle of fifths as a projection of the double helix led me, more recently, to devise a method of generating tones that glide continuously around the circle of fifths, that is, that pass continuously from C to G to D, and so forth, without passing through intermediate chromas along the way.⁵

The emergence of the cycle of fifths also has two important consequences (cf. Balzano, 1980). The first is that a diametral plane passing through the central axis of the

helix partitions all tones into two disjoint sets: a set containing all the tones included in a particular diatonic key (two of which, in each octave, fall on the plane) and the complement of that set, containing all the tones not included in that key. The second consequence is that partitionings corresponding to the major diatonic keys can be obtained by rotating the plane about the central axis, with modulations between more closely related keys obtained by smaller angles of rotation.

In Figure 3 the tones of the diatonic scale in C major are indicated by open circles, whereas the corresponding nondiatonic tones are represented by filled circles. In the view portrayed in Figure 3 (b) and, slightly tilted, in Figure 3 (c), the three-dimensional structure has been positioned so that the plane partitioning the tones into those included in the key of C major and those not included is seen almost edge on, with the tones belonging to C major falling to the right of the diametral plane. As can be seen from the 3-2-3-2-... grouping of the black circles on the left, the arrangement on the piano keyboard of the black keys, which correspond in the key of C to the sharps and flats or "accidentals," is not itself an accident.

The Two-Dimensional Melodic Map of the Double Helix

Because the tones in the double helix fall on the surface of a right circular cylinder (Figure 3 [c]), we can make a vertical cut in this surface and spread it on a plane along with the embedded double helix. The resulting two-dimensional "map" of the cylinder, illustrated in Figure 4, facilitates visualization of some of the musically significant properties of the double helix. The rectangle bounded by the heavy dashed line is from the region of the surface between one C and the C an octave above. The horizontal and vertical axes of this rectangular map correspond to the circle of fifths and to pitch height, respectively. (The order in which the tones project onto the axis corresponding to the circle of fifths is reversed

⁵ These tones, too, were demonstrated at the 1978 meeting of the Western Psychological Association (Shepard, Note 1).

because the unwrapped surface of the cylinder is viewed in Figure 4 as if from the inside of the cylinder.) In order to represent the unbounded character of the cylindrical surface, the rectangle can be endlessly repeated in the plane as indicated in the figure.

The chromatic scale is represented in this two-dimensional map by the sequence of notes on any of the straight lines directed upward and to the right. The two whole-tone scales are represented by the two distinct sequences of notes on the somewhat steeper straight lines directed upward and to the left. The diatonic scale, designated by the stippled band, exhibits a 3-4-3-4-. . . zigzag pattern in which strings of whole-tone steps are asymmetrically broken by single half-tone steps.

Corresponding to the division of the tones into those that are and those that are not in a particular key by a plane pivoted about the axis of the helix, the tones belonging to a particular key fall, in the flattened representation, within a particular vertical band demarcated in the figure for the key of C by the lighter dashed lines. Modulation to another key corresponds, here, to a horizontal shift of the vertical band with, again, more closely related keys obtainable by smaller shifts. Alternatively, transpositions from any major key to any other can be thought of as translations of the stippled zigzag pattern of the diatonic scale from one location to another within this two-dimensional plane. That the two keys most closely related to a given key (e.g., C) are the two obtained by a shift of a fifth up (to G) or down (to F) is reflected in the geometrical fact that this zigzag pattern overlaps most with itself when the straight group of three adjacent points is superimposed on the straight group of four, slipped into either of the two alternative positions within that group.

Other commonly used scales take similar zigzag patterns within this space. In one of its versions, the relative minor scale is identical to its associated major scale except that the sequence is started and ended on a different tone in the sequence (e.g., on A rather than on C in the example illustrated). Indeed, the seven so-called authentic church modes (which derive rather directly from the earlier Greek modes) all correspond exactly

to the diatonic scale, differing only in which tone is taken as the principal (beginning, final, or tonic) tone in the scale. Moreover, the most common pentatonic scale is given by the complement of the diatonic scale (e.g., by C[#], D[#], F[#], G[#], and A[#] in the figure or by the black keys on the piano). Notice that apart from the always permissible key-changing translation, such a pentatonic scale is equivalent to a diatonic scale in which the two most outlying tones within the vertical band have simply been deleted (e.g., B and F in the diatonic key of C). The resulting 2-3-2-3-. . . pattern preserves many of the desired structural properties of the 3-4-3-4-. . . diatonic scale and, again, changes of key correspond to horizontal shifts of the (now narrower) vertical band.

The adjacency of tones differing by half- and whole-tone steps in the embedded two-dimensional lattice preserves proximity in pitch height. Thus, this two-dimensional structure is particularly suited for the representation of melodies as well as scales. For, as might be expected on the basis of Gestalt principles of good continuation and grouping by proximity (e.g., see Deutsch & Feroe, 1981), transitions in pitch between successive tones of a melody are most commonly a single step in the diatonic scale (Dowling, 1978b; Fucks, 1962; Merriam, 1964; Philippot, 1970, p. 86; Piston, 1941, p. 23). For example, in an analysis of nearly 3,000 melodic intervals in 80 English folk songs, Dowling (1978b, p. 352) found, even after omitting the unison, that 68% of the transitions were no larger than one step on the diatonic scale, and 91% no larger than two steps.⁶ On the basis of these considerations,

⁶ There may be more than one reason for this striking predilection for small melodic steps. As Dowling (1978b) notes, it probably stems, in part, from basic limitations of human memory capacity. It seems to be related to Deutsch's (1978) finding that accuracy of recognition of a repeated tone falls off inversely with the average size of the intervals in an interpolated sequence of tones. As I have suggested, however, its close connection to Gestalt principles of visual perception (Koffka, 1935; Köhler, 1947), to phenomena of "melodic fission" or "auditory stream segregation" (Bregman, 1978; Bregman & Campbell, 1971; Dowling, 1973b; McAdams & Bregman, 1979; van Noorden, 1975), and to the closely allied phenomena of the "trill threshold" (Miller & Heise, 1950) and apparent movement in pitch (Shep-

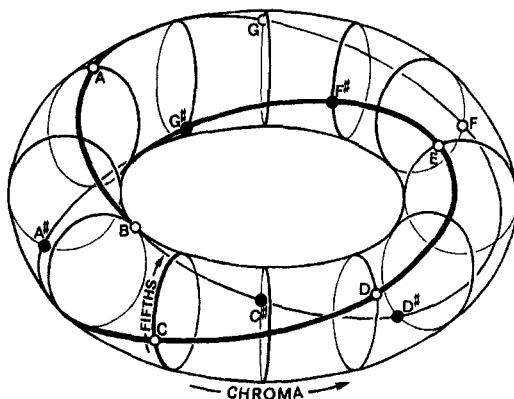


Figure 5. The double helix wound around a torus. (From "Structural Representations of Musical Pitch" by R. N. Shepard. In D. Deutsch (Ed.), *Psychology of Music*. New York: Academic Press, 1982. Copyright 1982 by Academic Press, Inc. Reprinted by permission.)

I propose that the unwrapped version of the double helix presented in Figure 4 be called the *melodic map*.

The Double Helix Wound Around a Torus

The double helix has of course retained the rectilinear component of the earlier simple helix, namely, the component called pitch height. Moreover, the circle of fifths, which has replaced the earlier chroma circle, is still closely related to the chroma circle, being obtainable from it simply by interchanging every other point with its diametrically opposite point around the circle. This is why the single helix becomes a double helix following such a substitution. Hence (a) chroma-equivalent tones still project into the same point on any plane orthogonal to the central axis and (b) chroma-adjacent tones still project into adjacent points on the central axis. Even so, these vestiges of the chroma circle in the double helix do not seem to reflect adequately the robustness of that circle as it has emerged from studies using ordinary musical tones (see Figure 2 [b]), or, certainly, the computer-synthesized circular tones (Shepard, 1964b; also see Figure 2 [a]).

ard, 1981c, p. 319) suggest a somewhat broader, partially perceptual basis. Possibly, too, melodies consisting of small steps provide a quicker and more effective indication of the underlying tonality or key of the piece.

Thus, we need a still more complex structure that somehow incorporates the essential features of both the double and the single helix. Such structures can be constructed, but only at the cost of increased dimensionality. We have to resort to an embedding space of four dimensions—or even five, if we wish to retain the linear component of pitch height. However, such higher dimensional structures are to some extent accessible to three-dimensional visualization.

Instead of regarding the chroma circle as obtained by projection of the simple helix, we could regard it as obtained by first cutting a one-octave segment out of either a rectilinear or a simple helical representation of pitch and then bending that segment until its two ends coincide to form a complete circle. Likewise, we could cut a one-octave segment out of the double helix portrayed in Figure 3 (c) and, by bending the cylindrical tube in which it is embedded until its two ends coincide, obtain a double helix wound around a torus as illustrated in Figure 5.

The two strands of the double helix (originally the two edges of the flattened strip of triangles) have become two interlocking rings embedded in the surface of the torus, and the surface of the original strip of triangles bounded by these two edges has become a twisted band. Unlike the only half-twisted band of Möbius, however, the two ends differed by a full 360° twist before they were joined, and hence, the band retains its two-sided or "orientable" topology (cf. Shepard, 1981c, p. 316).

Only in a four-dimensional embedding space does the torus possess the full degree of symmetry implicit in the fact that it is the "direct," Cartesian, or Euclidean product of two circles (Blackett, 1967)—in this case the chroma circle and the circle of fifths. (In a sense, to be illustrated in a later section, the embedded double helix can be thought of as a kind of Euclidean sum of those same two circles.) As a consequence there exists a rigid rotation of the whole structure such that the points corresponding to the 12 tones of the octave project onto the plane defined by one pair of orthogonal axes as a chroma circle and onto the plane defined by the remaining pair of orthogonal axes as a circle of fifths. In other words, the two circular components

enter into the structure in completely analogous ways, and in four-dimensional space, where rotations are about planes rather than about lines, we can rigidly rotate the whole structure about either of the two mentioned planes in such a way that the circle projected onto the other plane is carried into itself through any desired angle.

The complete symmetry between the two circular components of the toroidal version of the double helix is revealed more clearly in the unwrapped version already displayed in Figure 4. Such a planar map represents the toroidal surface just as it did the straight cylindrical surface (Blackett, 1967). The horizontal axis of the map still corresponds to the circle of fifths, but the vertical axis now corresponds to the chroma circle rather than to the rectilinear dimension of pitch height. Also, all four corners of the rectangular map now correspond to the same tone and, hence, the same point (C) in the torus. The lattice pattern of the melodic map is now strictly repeating vertically as well as horizontally, and the two types of rotations just described correspond, respectively, to horizontal and vertical translations in the two-dimensional plane.

The Double Helix Wound Around a Helical Cylinder

Actually, because tones can be synthesized such that those differing by an octave realize any specified degree of perceptual similarity or "octave equivalence," we need some way of continuously varying the structure representing those tones between the double helix with the rectilinear axis (Figure 3 [c]) and the variant with the completely circular axis (Figure 5). What modification of the latter, toroidal structure will bring back a separation of chroma-equivalent tones on a dimension of pitch height? Again the analogy with the original simple helix is helpful. Just as a cut and vertical displacement of the earlier chroma circle can yield one loop of the simple helix, a cut and vertical displacement of the torus portrayed in Figure 5 yields one loop of a higher order tubular helix. By attaching identical copies of such a loop, end to end, we can extend this higher order helical structure indefi-

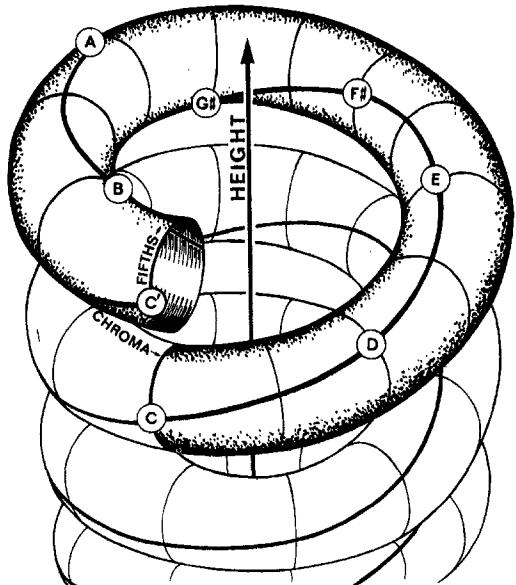


Figure 6. The double helix wound around a helical cylinder. (From "Structural Representations of Musical Pitch" by R. N. Shepard. In D. Deutsch (Ed.), *Psychology of Music*. New York: Academic Press, 1982. Copyright 1982 by Academic Press, Inc. Reprinted by permission.)

nitely in pitch height, as shown in Figure 6.

As before, we are distorting the true metric structure of this geometrical object by visualizing it in only three dimensions. It achieves its full inherent symmetry only in a space of five dimensions, where it exists as the Euclidean "sum" of the two-dimensional circle of fifths, the two-dimensional chroma circle, and the one-dimensional continuum of pitch height. In this way the structure continues to obey the already-stated principle that the most general rigid motion of space into itself is the product of rigid rotations and rectilinear translations. It is in this sense, too, that the structure crudely depicted as if three-dimensional in Figure 6 can rightly be regarded as a higher dimensional generalization of the helix.

A final point to be made about this generalized helical structure will be relevant when, in a following section, I compare it against empirical data. Whereas the double helix as it was first derived (Figure 3 [b]) was regarded as rigid (in order to preserve the equilateral character of its constituent triangles), the more general structure illustrated in Figure 6 can be regarded as ad-

justable through variation of three parameters: a weight for the circular component for fifths, a weight for the circular component for chroma, and a weight for the rectilinear component for height. Thus, we can accommodate the relations between musical pitches as they are perceived by different listeners who may vary, for example, in the extent to which they represent the cognitive structural component of the circle of fifths versus the purely psychoacoustic component of pitch height.

From this standpoint the original double helix (Figure 3) was perhaps too rigid. In present terms it can be seen to be the Euclidean sum of just the circle of fifths and the dimension of pitch height. But we now know that listeners vary widely in their responsiveness to these two attributes (Krumhansl & Shepard, 1979; Shepard, 1981a). So, in fitting the helix to data, we should perhaps make what might be regarded as a concession to performance, as opposed to competence, and allow a differential stretching or shrinking of the vertical extent of an octave of the helix relative to its diameter. This implies, of course, a departure from the constraint imposed by my third requirement (uniformity of scale steps), but I noted at the time that such a requirement may be appropriate only under a certain "perceptual set." If listeners differ in their judgments of the relations between tones, they are not all operating under identical perceptual sets.

In terms of the two-dimensional map of the double helix, differences in relative salience of pitch height and the circle of fifths would be accommodated by a certain class of linear transformations of the rectangle, namely, those restricted to relative stretching or shrinking of the rectangle along its vertical and horizontal axes only. In the limiting cases in which there is a degenerate collapse of the rectangle in the horizontal or vertical direction, we obtain a rectilinear (and logarithmic) dimension of pitch height or a simple circle of fifths, respectively.

Recovery of the Geometrical Representation From Empirical Data

The probe technique introduced by Krumhansl and Shepard (1979) is continuing to

yield robust, orderly, and informative data concerning the effects of musical contexts on the interpretive schema induced within the listener (Krumhansl & Kessler, 1982; Jordan & Shepard, Note 7). In this technique the context (e.g., a musical scale, a melody, a chord, a sequence of chords, or some richer musical passage) is immediately followed by a probe tone (selected, for example, from the 13 chromatic tones inclusively spanning a one-octave range), and the listener is asked to rate (on a 7-point scale) how well the probe tone "fits in" with the preceding context.

For any context, the average ratings from trials using different probe tones form a profile over the octave that in the case of musical listeners, reveals the hierarchy of tonal functions induced by that context—with the ratings highest for a tone interpreted as the tonic, next highest for a tone interpreted as the dominant, and so on (see Krumhansl & Shepard, 1979; and, especially, Krumhansl & Kessler, 1982). Indeed, Krumhansl and Kessler demonstrate that the circularly shifted position (or phase) of the profile permits one to infer which of the 24 possible major or minor diatonic keys is instantiated as the listener's momentary interpretive framework.

The rating of how well a given probe tone fits in with the preceding context can be interpreted as a measure of the spatial proximity of that probe to the ideal tonal center or tonality implied by that context. When applied to a suitably complete set of such proximity measures, techniques of multidimensional scaling (see Shepard, 1980) should therefore enable one to reconstruct the underlying spatial structure.

In the original experiment by Krumhansl and Shepard (1979), however, only the diatonic scale of a single key (C major) was presented as context. As a result, the obtained rating profiles directly provided information about the spatial proximities of the 13 probe tones to a single tonal center, C. However, there is every reason to believe that under transposition into any other key, the profile would be essentially invariant except for random fluctuations in the data—and this assumption already has some empirical support (Krumhansl & Kessler, 1982;

Jordan & Shepard, Note 7; also see Krumhansl, Bharucha, & Kessler, 1982). Accordingly, it seemed reasonable to approximate the complete matrix of proximity measures needed for the application of multidimensional scaling by simply duplicating the profile of ratings in each row of a square matrix, after circularly shifting each succeeding row by one cell so that the highest number (corresponding to the functional identity of the tonic tone and the ideal tonal center) fell on the principal diagonal of the matrix. Then, as is customary in multidimensional scaling, each entry in the matrix was averaged with its diagonally opposite counterpart to yield a symmetric matrix of proximity measures.

Because large individual differences, which are related to extents of musical background, characteristically emerge in these experiments (Krumhansl & Shepard, 1979; Shepard, 1981a), a symmetric matrix of the re-

quired sort was obtained separately from the average rating profile obtained from each of the 23 subjects in the experiment by Krumhansl and Shepard (1979). Application of individual-difference multidimensional scaling (INDSCAL; Carroll & Chang, 1970) to the entire resulting set of 23 individual matrices then yielded the four-dimensional solution presented in Figure 7.

Panel a shows the projection of the solution onto the plane of Dimensions 1 and 2, whereas Panel b shows its projection onto the plane of Dimensions 3 and 4. As suggested by the circular dashed line, the first projection (a) is essentially the chroma circle, going clockwise from C (through C[#], D, D[#], etc.) around to C' an octave above. The configuration departs from the chroma circle, however, in that the spacing is wider near C and C' and, particularly, in that C and C' do not coincide as they should if oc-

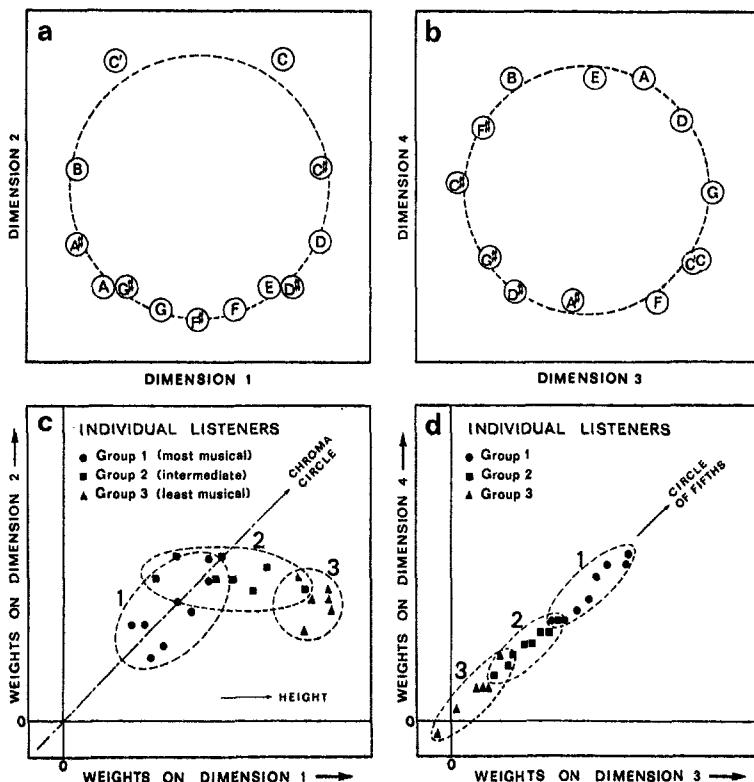


Figure 7. A four-dimensional solution obtained by application of INDSCAL to the data collected by Krumhansl & Shepard (1979). Panels a and b show the projections of the obtained configuration on the plane of Dimensions 1 and 2 and the plane of Dimensions 3 and 4, respectively; Panels c and d show the weights that these same dimensions have for listeners differing in musical background. (From Shepard, 1981a. Copyright 1981 by Music Educators National Conference. Reprinted by permission.)

tave equivalence had been complete for all listeners. Dimension 1 thus seems to combine one dimension of the chroma circle with the dimension of pitch height. The second projection (b), however, is an almost perfect circle of fifths, with the points representing C and C' nearly superimposed, indicating complete octave equivalence. Apart from the stretching and separation between the points representing C and C' on Dimension 1, the four-dimensional configuration is, in fact, the double helix on the torus depicted in Figure 5, which corresponds to the Euclidean product of the chroma circle and the circle of fifths.

Panels c and d display the INDSCAL weights for each of the listeners on each of the four dimensions. Panel c shows that the listeners with the least extensive musical backgrounds (represented by the triangles) had the heaviest weights on Dimension 1, which separated the tones with respect to pitch height, whereas Panel d shows that the listeners with the most extensive musical backgrounds (represented by the circles) had the heaviest weights on Dimensions 3 and 4, which contained the circle of fifths and implied complete equivalence between octaves. Moreover, the fact that the points for all listeners fall on a 45° line in Panel d means that the circle of fifths emerges, to whatever extent that it does for any one listener, as an integrated whole and never one dimension at a time. That the points for Group 1 listeners also fall close to the (broken) 45° line in Panel c indicates that under the complete octave-equivalence characteristic of the most musical listeners, the chroma circle, too, comes and goes as an integrated unit.

I interpret the obtained four-dimensional structure in Figure 7 as a one-octave piece of the endless five-dimensional theoretical structure portrayed in Figure 6. But because it includes only one octave, the gap between the two ends, which should have been represented by a displacement in a separate fifth dimension, has (with a small distortion) been accommodated in the four dimensions of the embedding space of the torus. In other words, the separation in pitch height between C and C' an octave above has been achieved by cutting through the torus in

Figure 5 at C and springing it slightly apart (with respect to chroma) in that same four-dimensional space rather than, as in Figure 6, in an orthogonal fifth dimension. Presumably, if similar data were collected and analyzed for tones spanning two or three octaves, the data could no longer be fit by a small distortion of this sort in the four-dimensional space and, thus, the truer, five-dimensional structure would emerge.

Further support for these conclusions comes from a linear regression used to assess the importance of the various proposed geometrical components of pitch, including—in addition to the one-dimensional component for height and the two circular components for chroma and for perfect fifths already discussed here—a two-dimensional component for major thirds. The use of linear regression for this purpose is made possible by the principle of “Euclidean composition,” according to which the squared distance between any two points in the final, higher dimensional configuration is a weighted sum of the squared distances between the corresponding points in each of the component configurations. What the regression analysis thus yields is the set of weights, one for each component configuration, that provides the best fit to the data. (See Shepard, 1982, for a fuller explanation of the analysis and the results.) The results indicated that the circles of chroma and of fifths did indeed account for a significant portion of the variance but that the factors of height and of major thirds were also significant for the least and the most musical listeners, respectively. More specifically, the principal factors (with fractions of variance accounted for in the symmetrized data) were for the most musical listeners, fifths (.43), chroma (.21), and thirds (.19); for the intermediate listeners, chroma (.36), fifths (.21), and thirds (.16); and for the least musical listeners, chroma (.39) and height (.31) only (Shepard, 1982, Table 1).

Clearly, then, pitch is multidimensional, and the different dimensions differ in salience for different listeners (as well as in different musical contexts). Moreover, because the final configuration (whether fitted by multidimensional scaling or Euclidean composition) contained circular compo-

nents, the obtained structures provide further support for the claim that some of the dimensions of pitch are circular. The structure as a whole thus appears to be consistent with the theoretical expectation of a helical or, under complete octave equivalence, toroidal character.

The Problem of Harmonic Relations

The generalizations of the double helix for pitch presented in the two preceding sections depended on an implicit weakening of Requirement 3 underlying the original derivation of that double helix, namely, the requirement that steps within a diatonic scale correspond to equal distances within the model. Only by accepting a weakening of this strong constraint can we provide for the quantitative variations in the relative weights of underlying components (height, chroma, fifths, etc.) necessary to fit the data of different listeners. In terms of the two-dimensional melodic map of the manifold in which the double helix is wound, changes in the relative weights correspond to linear stretchings or shrinkings of the rectangular melodic map along either or both of its two principal axes. One of these two axes corresponds to the circle of fifths, whereas the other corresponds to pitch height, chroma, or a mixture of these in the case of the straight cylindrical model (Figure 3 [c]), the toroidal model (Figure 5), or the cylindrical helical model (Figure 6), respectively.

The coefficients of this linear transformation determine the relative importance of certain musical intervals, namely, the perfect fifth, the octave, and (through emphasis of either pitch height or chroma) the minor second or chromatic step. Within this restricted class of linear transformations of the melodic map there is, however, no way to emphasize the intervals of the major or minor third. Yet these two intervals, though of limited importance for melodic structure, are fundamental to harmonic structure (Piston, 1941, p. 10). Thus, whereas the most common intervals between successively sounded tones (melodic intervals) are, as we noted, the minor or major seconds, which are adjacent in the melodic map (Figure 4), such intervals are dissonant when sounded si-

multaneously (that is, as harmonic intervals). Harmony, which governs the selection of tones to be sounded simultaneously in chords, is therefore based on the consonant intervals of the major and minor third, along with the consonant perfect fifth, which together make up the particularly harmonious, stable, and tonality-defining major triad (e.g., C-E-G, in the key of C).

There are two directions in which the geometrical models proposed here might be generalized in order to accommodate harmonic relations. One possibility, already suggested at the end of the preceding section, is to add further components to the structural representation. In addition to the one-dimensional component of pitch height, the two-dimensional component of chroma, and the two-dimensional component of the circle of fifths, we could include another two-dimensional component of major thirds. As I noted, such an extended model will indeed permit a somewhat better fit to the data (for the most musical listeners, an increase from 63% to 83% of the variance explained; see Shepard, 1982, Table 1). However, the prospect of increasing the number of dimensions of the embedding space from five to seven is not terribly attractive.

A second possibility is to remove the restriction that the linear expansions or compressions of the melodic map are to be permitted only along the two orthogonal axes of that map. If we allow elongations and contractions along arbitrarily oriented directions in the plane of that map, we can bring tones separated by other musical intervals into closer proximity without increasing dimensionality. Linear transformations of this more general type are called *affine* transformations (Coxeter, 1961); they preserve straight lines and parallelism, which is desirable if we are to continue to use collinearity and parallel projection to represent musical relations. Even so, we are not free to choose any affine transformation of the plane but must confine our choice to one that is compatible with the toroidal interpretation of the plane. Expansions or contractions along the orthogonal axes of the rectangular map can be of any magnitudes, corresponding to changes in the relative sizes (or weights) of the two circular components

(chroma and fifths) that generate the torus. But expansions or contractions along other directions can take on only certain discrete values, corresponding to operations of cutting through the torus in a plane parallel to one of the two generating circles, giving one free end of the resulting cylindrical tube an integer number of 360° twists relative to the other end, and then reattaching the two ends. Twists that are not multiples of 360° would fail to rejoin the two ends of each line on the plane (or helical path on the torus) and, hence, would disrupt continuity required for an affine transformation and collinearity desired for our musical interpretation.

The Harmonic Map and Its Affinity to the Melodic Map

In the present case we seek an affine transformation (of the admissible sort) that will bring tones separated by the harmonically important major and minor thirds and perfect fifth into close mutual proximity. In Figure 4 we see that for any given tone, say C at the lower left corner of the rectangle, the major third (E), the minor third (D $^{\#}$), and the perfect fifth (G), upward and to the right, form the vertexes of an elongated parallelogram in which the lower triangular half corresponds to the major triad built on C (viz., C-E-G) and the upper, complementary triangular half corresponds to the minor triad built on C (viz., C-D $^{\#}$ -G). Because no other tones fall within such parallelograms, an appropriate affine transformation should bring all such sets of four tones into the desired, more compact form. Fortunately, this can be achieved in a way that is compatible with the toroidal interpretation.

First, we cut through the original torus (Figure 5) at some point on the circle of fifths. Then we give the chroma circle at one end of the resulting cylindrical tube two full 360° twists relative to the chroma circle at the other end. Finally, we reattach the two ends, forming a new torus. These operations are most adequately visualized in terms of the two-dimensional map of the torus, as shown in Figure 8. The first twist takes the rectangle of the original melodic map (a, which is a simplified version of the earlier Figure 4) into the sheared parallelogram

(b). However, because the transformation is on the torus, the lower triangular half of the resulting parallelogram wraps around to fill in the vacated upper triangular half of the rectangular map (as shown in b). The second 360° twist operates in the same way (to take b into c). Finally, a relative expansion on the horizontal axis and a circular shift of the entire pattern around the torus to bring the tone C into the center of the map yields the final, transformed map (d) shown at the bottom of Figure 8.

Because the shearing transformation induced by the double twist was confined to the vertical axis corresponding to the chroma circle, the horizontal axis was unaffected and still corresponds to the circle of fifths. As a consequence the tones falling within any particular key continue to form a vertical band (illustrated for the key of C major in the figure), and modulations between keys still correspond to horizontal shifts of this band and, hence, to rotations around the torus.

However, owing to the twist in the torus entailed by the affine transformation, the harmonic map is not best described as a double helix wrapped around a torus. In it the series of perfect fifths now forms a single helix wrapped three times around the torus; the three series of major thirds form a triple helix wound once around the torus crosswise to the series of fifths; and the four series of minor thirds form four separate circular rings around the torus crosswise to both of the first two types of series. (See the horizontal and diagonal rows of letter names for the tones in Figure 8 [d].)

The important result is that tones related to any tone by major and minor thirds, as well as by perfect fifths, have now been brought into spatial proximity to that tone, whereas the formerly proximal tones related by major and minor seconds have been displaced to greater distances. This is illustrated for the tone C in Figure 8 (d), where, as can be seen, all the tones forming major or minor triads with C constitute a compact hexagonal cluster around C. Moreover, the tones traditionally considered to be consonant when sounded simultaneously with C now fall in a compact circular region around C. These same relations also hold for any other tone chosen as a reference point. Ac-

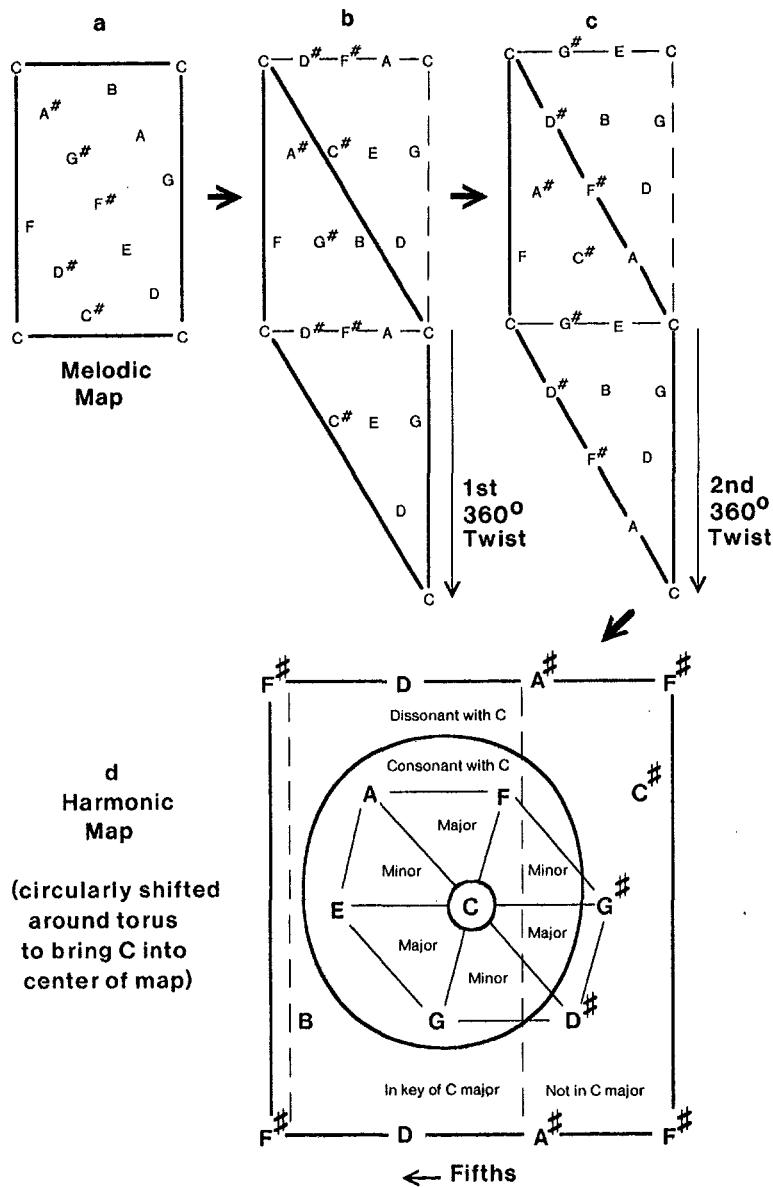


Figure 8. The harmonic map (d) as obtained from the melodic map (a) by an affine transformation consisting of two 360° twists of the torus (b and c).

cordingly, I propose that this transformed map be called the *harmonic map*. Whereas the melodic map provided for compact representations of musical scales and, hence, the most common melodies, the harmonic map provides for compact representation of consonant intervals and, hence, the most common chords. Although not explicitly illustrated, chords consisting of more than three

tones, including augmented and diminished chords, though more complex, also have compact representations in this map.⁷

⁷ As an amusing instance of "converging evidence," in the version of this space that Attnave presented at the symposium (Note 8), which was reflected with respect to the version shown here, the pattern for the seventh chord took, as he observed, the unmistakable shape of the numeral 7.

Representations that more or less resemble the harmonic map portrayed in Figure 8 have been independently proposed by a number of music theorists (e.g., Balzano, 1980; Hall, 1974; Lakner, 1960; Longuet-Higgins, 1962, 1979; O'Connell, 1962; Attneave, Note 8; Balzano, Note 2). What I have added here is the demonstration that this two-dimensional space, which is so well suited to the representation of harmonic relations, is related to the map of the double helix, which was not in any way based on harmonic considerations, by a simple mathematical transformation—specifically, an affinity.

Two recent developments provide additional converging evidence for the importance of this harmonic structure. One is Balzano's erection of essentially this structure from group-theoretic considerations, which make no use of acoustic notions of consonance (Balzano, 1980, Note 2). The other is Krumhansl and Kessler's (1982) striking demonstration that this same toroidal structure can be recovered by the multidimensional scaling of correlations between profiles obtained by applying the probe technique in musically rich contexts.

By presenting contexts in both major and minor keys, Krumhansl and Kessler have shown that this toroidal manifold can represent—in addition to the harmonic relations between individual tones considered here—the relations between musical keys, with the 12 minor tonalities falling in regular diagonal rows between the diagonal rows shown in Figure 8, which under their interpretation would represent the 12 major tonalities. Specifically, each major tonality (e.g., C major) is flanked, in their representation, by its corresponding relative and parallel minors (e.g., A minor and C minor).

Discussion and Conclusions

I have argued that musical pitch, if not psychophysical pitch height, should be represented by a geometrically regular structure that is capable of rigidly mapping into itself under transposition from one musical key into another. I have further argued that such a structure is a generalized helix—probably a double helix—in which the rectilinear axis of the helix, called pitch height, corresponds

to a logarithmic scale of frequency and in which each of two or more circular components completes one turn within a certain musical interval: the octave, the perfect fifth, or, possibly, the major third.

I have illustrated how techniques of multidimensional scaling can reconstruct such a structure from data collected from musical listeners within a musical context. I have also noted that unless certain constraints are imposed, the results of multidimensional scaling may not be strictly invariant under transposition. For example, if tones spanning only one octave are scaled, pitch height may be approximately accommodated by a distortion of the chroma circle. This distortion could be eliminated by scaling tones spanning several octaves or by means of a linear regression analysis (Shepard, 1982). It could also be eliminated by using the circular tones that I devised to achieve complete octave equivalence (Shepard, 1964b). In this latter case, however, we forgo the component of pitch height.

These considerations suggest two comments concerning the multidimensional scaling representations for musical pitch obtained by Krumhansl (1979) and by Krumhansl and Kessler (1982).

Krumhansl's (1979) analysis of judged similarities between tones presented in the context of a diatonic key yielded a conically shaped representation of the tones within an octave. The disposition of the tones on the conical surface corresponded to the hierarchy of tonal functions in this way: First, the tonic of the contextually established key and the other two tones making up the major triad based on that tonic (i.e., the major third and the perfect fifth above it) were clustered together near the apex of the cone. Second, this central cluster was surrounded by an intermediate band of the remaining four tones that belong to the contextually established key. Third, the remaining five tones not belonging to that key were widely spaced around the perimeter of the cone.

I believe that Krumhansl's multidimensional solution, like my own illustrated in Figure 7 (a), here, is somewhat distorted because tones from only one octave were included. The conical representation cannot be extended into higher and lower octaves in

such a way that musical transposition corresponds to rigid motions of the structure into itself. However, if we cut off the apex of the cone and then cut through the resulting closed band, we obtain a strip that with a slight twist, can be continued upward and downward as a helical structure (much as the torus of Figure 5 was cut through, twisted, and continued to form the higher order helix of Figure 6). A further consideration is raised by Tversky's (1977) demonstration that perceptually more salient stimuli are judged to be both more similar to each other when the judgments are of similarity and more different from each other when the judgments are of dissimilarity. It is possible that the tones that are perceived as closely related to the tonal center are rated as more similar when similarity is judged because they are more salient and not because they are closer in the underlying representational space. If so, the underlying structure might be more similar, though evidently not isomorphic, to one of the generalized helical structures proposed here. Alternatively, it may be that the departures from helical regularity in Krumhansl's conical solution reflect real differences in the relative stabilities of the tones within the diatonic context. (These possibilities are further considered in Shepard, 1982, pp. 383-384.)

The toroidal manifold of possible tonalities obtained more recently by Krumhansl and Kessler (1982) is, apart from its inclusion of the minor tonalities, isomorphic to the toroidal manifold whose two-dimensional harmonic map is shown in Figure 8. However, whereas the map in Figure 8 allows the inclusion of pitch height and, hence, the extension of the structure into higher and lower octaves in the manner illustrated in Figure 6, Krumhansl and Kessler's solution is a closed torus with complete octave equivalence and, hence, no component of pitch height.

This closure of the torus was ensured in Krumhansl and Kessler's experiment by generating the tones presented to the listeners according to a scheme that achieves complete equivalence of octaves and suppression of pitch height (after Shepard, 1964b). Moreover, such a closure is appropriate in

their case because pitch height is not relevant to tonality. To say that a piece is in F major is not to say whether it is to be played by a piccolo or a tuba. Krumhansl and Kessler's work elegantly uses the tone-probe technique to show how a listener's internally represented tonal center shifts in this closed toroidal surface under the influence of an evolving musical context.

As I mentioned earlier, there may be fundamental parallels between the cognitive structural constraints governing visual-spatial and auditory-musical representation. Pitch is the "morphophoric" medium that is most analogous to physical space in the case of vision (Attneave, 1972; Attneave & Olson, 1971; Kubovy, 1981; Shepard, 1981c, 1982). Musically meaningful objects, such as melodies and chords, are the closest auditory analogs of visual shapes because such objects preserve their structure under rigid transformations within their respective morphophoric media. Moreover, because the medium for the rigid transformation of melodies and chords has circular components of chroma and fifths, these rigid transformations include rotations as well as translations. Just as in the case of visual cognition, the time needed to carry out such transformations mentally is expected to depend on the angular extent of these spatial transformations (Shepard, 1978a, 1981c; Shepard & Cooper, 1982). Likewise, modulations from one key to another correspond to rotations (Balzano, 1980; Shepard, 1982; Balzano, Note 2; Shepard, Note 1). The toroidal space in which these latter rotations occur has now been fully elaborated by Krumhansl and Kessler (1982), who have given all major and minor keys angular coordinates within this same two-dimensional manifold.

Finally, although I have in this paper focused exclusively on just one of the two indispensable musical attributes, namely, pitch, there are reasons to believe that the structure described here may carry over to the other indispensable musical attribute, namely, time. Relations of beat and rhythm have the same structural properties of augmented correspondence at the simplest ratios of tempo, such as the 2-to-1 or 3-to-2 ratios, which in the case of frequency of tones cor-

respond to the octave and perfect fifth, respectively. Accordingly, illusions of continually accelerating tempo can be constructed that are analogous to the illusion of continually ascending pitch (Risset, Note 9) and, hence, that demonstrate the existence of a temporal analog of the circle of chroma. Further, recent ethnomusicological studies of the highly developed rhythms of Indian and West African cultures have even revealed temporal analogs of the diatonic scale (Pressing, in press). By the same group-theoretic arguments already given for pitch by Balzano (1980, in press), this finding implies (as Pressing notes) the existence of a temporal analog of the circle of fifths. Remembering that we must add the rectilinear time line to these two circular components, we arrive at the conclusion that a structure that is isomorphic to the five-dimensional generalized helix portrayed in Figure 6 will be needed for the representation of rhythmic relations as well.

Reference Notes

- Shepard, R. N. The double helix of musical pitch. In R. N. Shepard (Chair), *Cognitive structure of musical pitch*. Symposium presented at the meeting of the Western Psychological Association, San Francisco, April 1978.
- Balzano, G. J. The structural uniqueness of the diatonic order. In R. N. Shepard (Chair), *Cognitive structure of musical pitch*. Symposium presented at the meeting of the Western Psychological Association, San Francisco, April 1978.
- Shepard, R. N. Unpublished doctoral dissertation proposal, Yale University, 1954.
- Mathews, M. V. Personal communication, January 7, 1982.
- Blechner, M. J. *Musical skill and the categorical perception of harmonic mode* (Status Report on Speech Perception SR-51/52). New Haven, Conn.: Haskins Laboratories, 1977, pp. 139-174.
- Cohen, A. J. Inferred sets of pitches in melodic perception. In R. N. Shepard (Chair), *Cognitive structure of musical pitch*. Symposium presented at the meeting of the Western Psychological Association, San Francisco, April 1978.
- Jordan, D., & Shepard, R. N. *Effects on musical cognition of a distorted diatonic scale*. Manuscript in preparation, 1982.
- Atneave, F. Discussant's remarks. In R. N. Shepard (Chair), *Cognitive structure of musical pitch*. Symposium presented at the meeting of the Western Psychological Association, San Francisco, April 1978.
- Risset, J.-C. *Demonstration tape of auditory illusions*. (Available from J.-C. Risset, UER de Luminy, 13008 Marseille, France.)

References

- Allen, D. Octave discriminability of musical and non-musical subjects. *Psychonomic Science*, 1967, 7, 421-422.
- Atneave, F. The representation of physical space. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory*. Washington, D.C.: V. H. Winston, 1972, pp. 283-306.
- Atneave, F., & Olson, R. K. Pitch as a medium: A new approach to psychophysical scaling. *American Journal of Psychology*, 1971, 84, 147-166.
- Bachem, A. Tone height and tone chroma as two different pitch qualities. *Acta Psychologica*, 1950, 7, 80-88.
- Bachem, A. Time factors in relative and absolute pitch determination. *Journal of the Acoustical Society of America*, 1954, 26, 751-753.
- Balzano, G. J. Chronometric studies of the musical interval sense (Doctoral dissertation, Stanford University, 1977). *Dissertation Abstracts International*, 1977, 38, 2898B. (University Microfilms No. 77-25, 643).
- Balzano, G. J. The group-theoretic description of twelvefold and microtonal pitch systems. *Computer Music Journal*, 1980, 4, 66-84.
- Balzano, G. J. Musical versus psychoacoustical variables and their influence on the perception of musical intervals. *Bulletin of the Council for Research in Music Education*, 1982, 70, 1-11.
- Balzano, G. J. The pitch set as a level of description for studying musical pitch perception. In M. Clynes (Ed.), *Music, mind, and brain*. New York: Plenum Press, in press.
- Balzano, G. J., & Liesch, B. W. The role of chroma and scalestep in the recognition of musical intervals in and out of context. *Psychomusicology*, in press.
- Beck, J., & Shaw, W. A. The scaling of pitch by the method of magnitude estimation. *American Journal of Psychology*, 1961, 74, 242-251.
- Blackett, D. W. *Elementary topology*. New York: Academic Press, 1967.
- Blackwell, H. R., & Schlosberg, H. Octave generalization, pitch discrimination, and loudness thresholds in the white rat. *Journal of Experimental Psychology*, 1943, 33, 407-419.
- Boring, E. G. *Sensation and perception in the history of experimental psychology*. New York: Appleton-Century-Crofts, 1942.
- Bregman, A. S. The formation of auditory streams. In J. Requin (Ed.), *Attention and performance VII*. Hillsdale, N.J.: Erlbaum, 1978.
- Bregman, A. S., & Campbell, J. Primary auditory stream segregation and the perception of order in rapid sequences of tones. *Journal of Experimental Psychology*, 1971, 89, 244-249.
- Burns, E. M. Octave adjustment by non-Western musicians. *Journal of the Acoustical Society of America*, 1974, 56, 525. (Abstract)
- Burns, E. M., & Ward, W. I. Categorical perception—phenomenon or epiphenomenon: Evidence from experiments in the perception of melodic musical intervals. *Journal of the Acoustical Society of America*, 1978, 63, 456-468.
- Carroll, J. D., & Chang, J.-J. Analysis of individual

- differences in multidimensional scaling via an N-way generalization of Eckart-Young decomposition. *Psychometrika*, 1970, 35, 283-319.
- Charbonneau, G., & Risset, J.-C. Circularité de jugements de hauteur sonore. *Comptes Rendus Hebdomadaires des Séances de L'Academie des Sciences Paris*, 1973, 277B, 623-626.
- Chomsky, N. *Aspects of the theory of syntax*. Cambridge, Mass.: M.I.T. Press, 1965.
- Chomsky, N. *Language and mind*. New York: Harcourt, Brace & World, 1968.
- Cohen, A. *Perception of tone sequences from the Western-European chromatic scale: Tonality, transposition and the pitch set*. Unpublished doctoral dissertation, Queen's University at Kingston, Ontario, Canada, 1975.
- Coxeter, H. S. M. *Introduction to geometry*. New York: Wiley, 1961.
- de Boer, E. On the "residue" and auditory pitch perception. In W. D. Keidel & W. D. Neff (Eds.), *Handbook of sensory physiology* (Vol. 5, Pt. 3: *Clinical and special topics*). New York: Springer-Verlag, 1976, pp. 479-583.
- Deutsch, D. Octave generalization and tune recognition. *Perception & Psychophysics*, 1972, 11, 411-412.
- Deutsch, D. Octave generalization of specific interference effects in memory for tonal pitch. *Perception & Psychophysics*, 1973, 13, 271-275.
- Deutsch, D. Delayed pitch comparison and the principle of proximity. *Perception & Psychophysics*, 1978, 23, 227-230.
- Deutsch, D., & Feroe, J. The internal representation of pitch sequences in tonal music. *Psychological Review*, 1981, 88, 503-522.
- Dewar, K. M. *Context effects in recognition memory for tones*. Unpublished doctoral dissertation, Queen's University at Kingston, Ontario, Canada, 1974.
- Dewar, K. M., Cuddy, L. L., & Mewhort, D. J. K. Recognition memory for single tones with and without context. *Journal of Experimental Psychology: Human Learning and Memory*, 1977, 3, 60-67.
- Dowling, W. J. The 1215-cent octave: Convergence of Western and Nonwestern data on pitch-scaling. *Journal of the Acoustical Society of America*, 1973, 53, 373A. (Abstract) (a)
- Dowling, W. J. The perception of interleaved melodies. *Cognitive Psychology*, 1973, 5, 322-337. (b)
- Dowling, W. J. Listeners' successful search for melodies scrambled into several octaves. *Journal of the Acoustical Society of America*, 1978, 64, S146. (Abstract) (a)
- Dowling, W. J. Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, 1978, 85, 341-354. (b)
- Dowling, W. J. Musical scales and psychophysical scales: Their psychological reality. In T. Rice & R. Falck (Eds.), *Cross-cultural approaches to music*. Toronto, Ontario, Canada: University of Toronto Press, in press.
- Dowling, W. J., & Hollombe, A. W. The perception of melodies distorted by splitting into several octaves: Effects of increasing proximity and melodic contour. *Perception & Psychophysics*, 1977, 21, 60-64.
- Forte, A. *Tonal harmony in concept and practice* (3rd ed.). New York: Holt, Rinehart & Winston, 1979.
- Francès, R. *La perception de la musique*. Paris: Vrin, 1958.
- Fucks, W. Mathematical analysis of formal structure of music. *IRE Transactions on Information Theory*, 1962, 8, 225-228.
- Goldstein, H. *Classical mechanics*. Reading, Mass.: Addison-Wesley, 1950.
- Goldstein, J. L. An optimum processor theory for the central formation of the pitch of complex tones. *Journal of the Acoustical Society of America*, 1973, 54, 1496-1516.
- Greenwood, G. D. *Principles of dynamics*. Englewood Cliffs, N.J.: Prentice-Hall, 1965.
- Hahn, J., & Jones, M. R. Invariants in auditory frequency relations. *Scandinavian Journal of Psychology*, 1981, 22, 129-144.
- Hall, D. E. Quantitative evaluation of musical scale tunings. *American Journal of Psychics*, 1974, 42, 543-552.
- Helmholtz, H. von. *On the sensations of tone as a physiological basis for the theory of music*. New York: Dover, 1954. (Originally published, 1862.)
- Hilbert, D., & Cohn-Vossen, S. *Geometry and the imagination*. New York: Chelsea, 1952. (Originally published, 1932.)
- Humphreys, L. F. Generalization as a function of method of reinforcement. *Journal of Experimental Psychology*, 1939, 25, 361-372.
- Hurvich, L. M., & Jameson, D. An opponent-process theory of color vision. *Psychological Review*, 1957, 64, 384-404.
- Idson, W. L., & Massaro, D. W. A bidimensional model of pitch in the recognition of melodies. *Perception & Psychophysics*, 1978, 24, 551-565.
- Imbert, M. *L'acquisition des structures tonales chez l'enfant*. Paris: Klincksieck, 1969.
- Jones, M. R. Time, our lost dimension: Toward a new theory of perception, attention, and memory. *Psychological Review*, 1976, 83, 323-355.
- Kallman, H. J., & Massaro, D. W. Tone chroma is functional in melody recognition. *Perception & Psychophysics*, 1979, 26, 32-36.
- Kilmer, A. D., Crocker, R. L., & Brown, R. R. *Sounds from silence: Recent discoveries in ancient Near Eastern music*. Berkeley, Calif.: Bit Enki Publications, 1976.
- Koffka, K. *Principles of Gestalt psychology*. New York: Harcourt Brace, 1935.
- Köhler, W. *Gestalt psychology*. New York: Liveright, 1947.
- Krumhansl, C. L. The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, 1979, 11, 346-374.
- Krumhansl, C. L., Bharucha, J. J., & Kessler, E. J. Perceived harmonic structure of chords in three related musical keys. *Journal of Experimental Psychology: Human Perception and Performance*, 1982, 8, 24-36.
- Krumhansl, C. L., & Kessler, F. J. Tracing the dynamic changes in perceived tonal organization in a spatial representation of musical keys. *Psychological Review*, 1982, 89, 334-368.
- Krumhansl, C. L., & Shepard, R. N. Quantification of

- the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance*, 1979, 5, 579-594.
- Kubovy, M. Concurrent pitch-segregation and the theory of indispensable attributes. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization*. Hillsdale, N.J.: Erlbaum, 1981.
- Lakner, Y. A new method of representing tonal relations. *Journal of Music Theory*, 1960, 4, 194-209.
- Levelt, W. J. M., Van de Geer, J. P., & Plomp, R. Triadic comparisons of musical intervals. *British Journal of Mathematical and Statistical Psychology*, 1966, 19, 163-179.
- Lberman, A. M., Cooper, F. S., Shankweiler, D., & Studdert-Kennedy, M. Perception of the speech code. *Psychological Review*, 1967, 74, 431-461.
- Licklider, J. C. R. Basic correlates of the auditory stimulus. In S. S. Stevens (Ed.), *Handbook of experimental psychology*. New York: Wiley, 1951.
- Locke, S., & Kellar, L. Categorical perception in a non-linguistic mode. *Cortex*, 1973, 9, 355-369.
- Longuet-Higgins, H. C. Letter to a musical friend. *Music Review*, 1962, 23, 244-248.
- Longuet-Higgins, H. C. The perception of music (Review Lecture). *Proceedings of the Royal Society, London*, 1979, 205B, 307-332.
- Mathews, M. V. The digital computer as a musical instrument. *Science*, 1963, 142, 553-557.
- Mathews, M. V., & Sims, G. Perceptual discrimination of just and equal tempered tunings. *Journal of the Acoustical Society of America*, Suppl. 1, 1981, 69, 538 (Abstract).
- McAdams, S., & Bregman, A. Hearing musical streams. *Computer Music Journal*, 1979, 3, 26-43.
- Merriam, A. P. *The anthropology of music*. Evanston, Ill.: Northwestern University Press, 1964.
- Meyer, L. B. *Emotion and meaning in music*. Chicago: University of Chicago Press, 1956.
- Miller, G. A. The magic number seven, plus or minus two. *Psychological Review*, 1956, 63, 81-97.
- Miller, G. A., & Heise, G. A. The trill threshold. *Journal of the Acoustical Society of America*, 1950, 64, 637-638.
- Null, C. *Symmetry in judgments of musical pitch*. Unpublished doctoral dissertation, Michigan State University, 1974.
- O'Connell, W. Tone spaces. *Die Reihe*, 1962, 8, 34-67.
- Peterson, G. E., & Barney, H. L. Control methods used in a study of the vowels. *Journal of the Acoustical Society of America*, 1952, 24, 175-184.
- Philippot, M. *L'Arc, Beethoven*, No. 40, 1970.
- Pikler, A. G. The diatonic foundations of hearing. *Acta Psychologica*, 1955, 11, 432-445.
- Pikler, A. G. Logarithmic frequency systems. *Journal of the Acoustical Society of America*, 1966, 39, 1102-1110.
- Piston, W. *Harmony*. New York: Norton, 1941.
- Plomp, R. *Aspects of tone sensation: A psychophysical study*. New York: Academic Press, 1976.
- Plomp, R., & Levelt, W. J. M. Tonal consonance and critical band width. *Journal of the Acoustical Society of America*, 1965, 38, 548-560.
- Pressing, J. Cognitive isomorphisms in pitch and rhythm in world musics: West African, the Balkans, Thailand and Western tonality. *Ethnomusicology*, in press.
- Ratner, L. G. *Harmony: Structure and style*. New York: McGraw-Hill, 1962.
- Revesz, G. *Introduction to the psychology of music*. Norman: University of Oklahoma Press, 1954.
- Risset, J.-C. Musical acoustics. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception* (Vol. 4). New York: Academic Press, 1978, pp. 521-564.
- Ruckmick, C. A. A new classification of tonal qualities. *Psychological Review*, 1929, 36, 172-180.
- Schenker, H. *Harmony* (O. Jones, Ed. and E. M. Borgese, trans.). Cambridge, Mass.: M.I.T. Press, 1954. (Originally published, 1906).
- Shepard, R. N. Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, 1964, 1, 54-87. (a)
- Shepard, R. N. Circularity in judgments of relative pitch. *Journal of the Acoustical Society of America*, 1964, 36, 2346-2353. (b)
- Shepard, R. N. Approximation to uniform gradients of generalization by monotone transformations of scale. In D. I. Mostofsky (Ed.), *Stimulus generalization*. Stanford, Calif.: Stanford University Press, 1965, pp. 94-110.
- Shepard, R. N. Psychological representation of speech sounds. In E. E. David & P. B. Denes (Eds.), *Human communication: A unified view*. New York: McGraw-Hill, 1972, pp. 67-113.
- Shepard, R. N. Representation of structure in similarity data: Problems and prospects. *Psychometrika*, 1974, 39, 373-421.
- Shepard, R. N. The circumplex and related topological manifolds in the study of perception. In S. Shye (Ed.), *Theory construction and data analysis in the behavioral sciences*. San Francisco: Jossey-Bass, 1978, 29-80. (a)
- Shepard, R. N. Externalization of mental images and the act of creation. In B. S. Randhawa & W. E. Coffman (Eds.), *Visual learning, thinking, and communication*. New York: Academic Press, 1978. (b)
- Shepard, R. N. On the status of "direct" psychophysical measurement. In C. W. Savage (Ed.), *Minnesota studies in the philosophy of science* (Vol. 9). Minneapolis: University of Minnesota Press, 1978. (c)
- Shepard, R. N. Multidimensional scaling, tree-fitting, and clustering. *Science*, 1980, 210, 390-398.
- Shepard, R. N. Individual differences in the perception of musical pitch. In *Documentary report of the Ann Arbor symposium: Applications of psychology to the teaching and learning of music*. Reston, Va.: Music Educators National Conference, 1981. (a)
- Shepard, R. N. Psychological relations and psychophysical scales: On the status of "direct" psychophysical measurement. *Journal of Mathematical Psychology*, 1981, 24, 21-57. (b)
- Shepard, R. N. Psychophysical complementarity. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization*. Hillsdale, N.J.: Erlbaum, 1981. (c)
- Shepard, R. N. Structural representations of musical pitch. In D. Deutsch (Ed.), *Psychology of music*. New York: Academic Press, 1982.
- Shepard, R. N., & Cooper, L. A. *Mental images and their transformations*. Cambridge, Mass.: MIT Press/Bradford Books, 1982.
- Shepard, R. N., & Zajac, E. (Producers). *A pair of paradoxes*. Murray Hill, N.J.: Bell Telephone Lab-

- oratories, Technical Information Library, 1965. (Film) "Shepard's Tones." On M. V. Mathews, J.-C. Risset, et al., *The voice of the computer* (Decca Record DL 710180). Universal City, Calif.: MCA Records, Inc., 1970.
- Siegel, J. A. The nature of absolute pitch. In E. Gordon (Ed.), *Research in the psychology of music* (Vol. 8). Iowa City: University of Iowa Press, 1972.
- Siegel, J. A., & Siegel, W. Absolute identification of notes and intervals by musicians. *Perception & Psychophysics*, 1977, 21, 143-152. (a)
- Siegel, J. A., & Siegel, W. Categorical perception of tonal intervals: Musicians can't tell sharp from flat. *Perception & Psychophysics*, 1977, 21, 399-407. (b)
- Stevens, S. S. The measurement of loudness. *Journal of the Acoustical Society of America*, 1955, 27, 815-829.
- Stevens, S. S., & Volkmann, J. The relation of pitch to frequency: A revised scale. *American Journal of Psychology*, 1940, 53, 329-353.
- Stevens, S. S., Volkmann, J., & Newman, E. B. A scale for the measurement of the psychological magnitude of pitch. *Journal of the Acoustical Society of America*, 1937, 8, 185-190.
- Sundberg, J. E. F., & Lindqvist, J. Musical octaves and pitch. *Journal of the Acoustical Society of America*, 1973, 54, 922-929.
- Terhardt, E. Pitch, consonance, and harmony. *Journal of the Acoustical Society of America*, 1974, 55, 1061-1069.
- Thurlow, W. R., & Erchul, W. P. Judged similarity in pitch of octave multiples. *Perception & Psychophysics*, 1977, 22, 177-182.
- Tversky, A. Features of similarity. *Psychological Review*, 1977, 84, 327-352.
- van Noorden, L. P. A. S. *Temporal coherence in the perception of tone sequences*. Unpublished doctoral dissertation, Technische Hogeschool, Eindhoven, The Netherlands, 1975.
- Ward, W. D. Subjective musical pitch. *Journal of the Acoustical Society of America*, 1954, 26, 369-380.
- Ward, W. D. Absolute pitch. Pt. I. *Sound*, 1963, 2, 14-21. (a)
- Ward, W. D. Absolute pitch. Pt. II. *Sound*, 1963, 2, 33-41. (b)
- Ward, W. D. Musical perception. In J. V. Tobias & E. D. Hubert (Eds.), *Foundations of modern auditory theory* (Vol. 1). New York: Academic Press, 1970, pp. 407-447.
- Wightman, F. L. The pattern-transformation model of pitch. *Journal of the Acoustical Society of America*, 1973, 54, 407-416.
- Zatorre, R. S., & Halpern, A. R. Identification, discrimination, and selective adaptation of simultaneous musical intervals. *Perception & Psychophysics*, 1979, 26, 384-395.
- Zenatti, A. *Le développement génétique de la perception musicale* (Monographies Françaises de Psychologie, No. 17). Paris: Centre National de la Recherche Scientifique, 1969.
- Zuckerkandl, V. *Sound and symbol*. Princeton, N.J.: Princeton University Press, 1956.
- Zuckerkandl, V. *Man the musician*. Princeton, N.J.: Princeton University Press, 1972.

Received October 21, 1981 ■