4

Musical Scales

The individual parts of a melody reach the ear in succession. We cannot perceive them all at once. We cannot perceive backwards and forwards at pleasure. Hence for a clear and sure measurement of the change of pitch, no means was left but progression by determinate degrees. This series of degrees is laid down in the musical scale.

(Helmholtz, 1877/1954, p. 252)

INTRODUCTION

Of the perceptual dimensions used in music, pitch is unique in having a scale dividing it fairly rigidly into discrete steps. Steps of loudness are designated in musical scores with labels such as *p, mp, mf,* and so forth, but these are relative indicators whose meaning changes with context. The temporal continuum is divided into discrete beats, but the duration of the beat changes with tempo. Timbre is multidimensional, and all its dimensions admit of more or less continuous variation. Pitch alone is organized into discrete steps. This is true in almost all the cultures of the world. We say almost all because there are some cultures that often use chants on two pitch levels for which it is difficult to define a scale. For example, the Hawaiian *oli* chant (described by Roberts, 1967) is primarily monotonic, with wide latitude in the pitch separation of a lower secondary tone. Like the Hawaiian, several cultures in widely separated parts of the world use two-pitch chants, and it seems a matter of semantics whether we call the two pitches with their single discrete step a scale or not. Further, there is some evidence that as such a chant becomes more intense and excited, the single melodic interval becomes expanded, perhaps in continuous fashion (Sachs, 1965). Nevertheless, most of the world's music is based on scales having stable, discrete steps. The sizes of the steps vary from culture to culture, though virtually all use the octave as a basic interval.

In our discussion, we first consider a set of cognitive constraints on scale construction that seem to operate through much of the world. We show how the application of those constraints in various combinations leads the various possible forms of scale in different cultures. Following that, we turn to a discussion of alternative ways psychologists and musicians have looked at pitch scales, leading to a discussion of contemporary multidimensional-scaling approaches. We conclude the chapter with an overview of the use of scales in a variety of cultures.

It is a puzzle for psychology why the music of the world uses discrete steps from pitch to pitch rather than the continuous series of all possible pitches. In his discussion of this problem, Helmholtz (1877/1954, pp. 250ff.) suggests that a scale of discrete pitch levels provides a psychological standard by which the listener can measure melodic motion. If melodies consisted of continuous changes of pitch like the wailing of a siren, the listener would still have to learn a scale in order to comprehend the amount of those changes. In that case, the listener would have a very difficult time learning the scale, since the actual notes of the scale would seldom occur (and seldom be marked) in the perceived music. The cognitive tasks of musician and listener are immensely simplified by restricting the set of pitches to the graduated degrees of a scale. Knowing a musical scale gives the listener an immediate basis for comparing the sizes of pitch intervals and judging the extent of melodic motion. (Helmholtz's argument is a special case of a general argument concerning cognitive frameworks made by Kant, 1787/1933, pp. B xxxix f., footnote.)

Beyond providing a measure of melodic motion, the scale provides a cognitive framework that facilitates the remembering of the pitches of a melody. This is especially important in nonliterate cultures where the human memory is the only vehicle by which melodies are preserved. Without culturally established scales, the reproduction of tunes and their transmission from generation to generation is a haphazard affair. We present evidence below that people's memory for familiar tunes using familiar scales is indeed very good. Some researchers go so far as to argue that because the scale is a mutually shared category system among performers and listeners, the listener "hears one and the same music, in different rooms, played on various instruments or sung, transposed, recorded with various equipment with such and such distortion, etc." (Francès, 1958, pp. 34–35).

The pitch categories of a musical scale serve the same kind of psychological function as do discrete categories used in the recoding of many types of messages communicated across noisy channels. Languages use this kind of discrete categorization. There is a certain range of physical sounds which, in English, are heard as *p* (Lane, 1965; Liberman, Harris, Kinney, & Lane, 1961). Certain physical parameters can be changed along a continuum, and up to a point, the sound will remain *p*. However, at a certain point the sound will become a *b* for the listener. Speakers are given considerable latitude in the way they produce *p*'s. They can generate physical sounds anywhere along that part of the continuum mutually recognized as belonging to the *p* category and their listeners will understand them. Further, the listeners will be able to economize what they must store in memory in order to remember the message. Rather than remember all the acoustic parameters of the waveform the speaker actually produced, all the listener must remember is the perceptual category into which those sounds fall. As in language, the categorization of both stimuli and responses afforded by musical scales reduces ambiguity and memory load in hearing, remembering, and producing music. As in most human communication systems, some indeterminacy within categories (intonation error) is allowed in order to achieve greater stability and clarity at the broader level of category identification.

WESTERN SCALES AND EQUAL TEMPERAMENT

Constraints on Scale Construction

The structure of scales in Western music provides a convenient starting place to begin the description of musical scales of pitch. After describing Western scales, we go on to describe certain features of the scale systems of other cultures. Figure 4.1 shows a typical Western tonal scale (C major) along with the fundamental frequencies of the notes. There are several psychological constraints on tonal scale construction found all over the world. First, the pitches of a scale must be discriminable from one another when played in succession. In Western music, the smallest pitch interval between notes, called a semitone, represents a frequency difference of about 5.9%. Such an interval can be found between E and F and between B and C in Figure 4.1. Since humans are capable of discriminat-

Figure 4.1 Standard fundamental frequencies of the Western C-major scale.

ing frequency differences of the order of 1.0% in this frequency range (Shower & Biddulph, 1931), intervals of a semitone would seem to be safely discriminable. The composer Harry Partch (1974), a long-time proponent of microtonal scales, which have intervals smaller than the semitone, notes that the principal limitations on the smallness of intervals are not given by their psychophysical discriminability but by the cognitive systems of the listener—a set of constraints to which we return below.

A second constraint is that tones whose fundamental frequencies stand in a 2:1 ratio (or nearly so; we discuss complications below) are treated as very similar to each other. The interval between frequencies in a 2:1 ratio is called an octave. The two notes labeled *C* in Figure 4.1 are an octave apart. Tones an octave apart are perceived as similar not only by humans but also by white rats (Blackwell & Schlosberg, 1943). In cultures having labels for the pitches of the scale, tones an octave apart are given the same name. In Western music, for example, tones with fundamental frequencies of 32.75, 65.5, 131, 262, 524, 1048, 2096, and 4192 Hz are all called C. Further, in cultures with functional harmony like the western European tones of the same name have the same harmonic functions when combined simultaneously with other tones. The overwhelming majority of cultures in the world make use of the equivalence of tones an octave apart. The only exceptions we have found are certain groups of Australian aborigines. In their cultures, melodic imitations at roughly octave intervals do not always use the same logarithmic scale intervals, and when men and women sing together, they do so in unison and not in octaves (Ellis, 1965).

A third constraint is that when the octave is filled in with the intervals of the scale, there should only be a moderate number of different pitches, say five or seven. This constraint arises from the cognitive limitation on the number of different values along a psychological dimension people can handle without confusion. Miller (1956) argues that across various sensory modalities, the number of stimuli along a given dimension people can categorize consistently is typically 7 ± 2. With more than seven or so different pitches, people begin to hear two or more pitches as falling in the same category and so defeat the purpose Helmholtz saw for the scale. Most cultures in the world use five to seven pitches in their tonal scales.

A fourth constraint that operates in a few cultures of the world, among them the Western and the Chinese, is that the octave should be divided into a series of minimal intervals, all equal in size, which are added together to construct all the intervals used in melodic scales. Western tuning divides the octave into 12 semitones (a chromatic scale). One advantage of such a system is that a given melody can be transposed so as to start on any note of the tuning system and be reproduced without

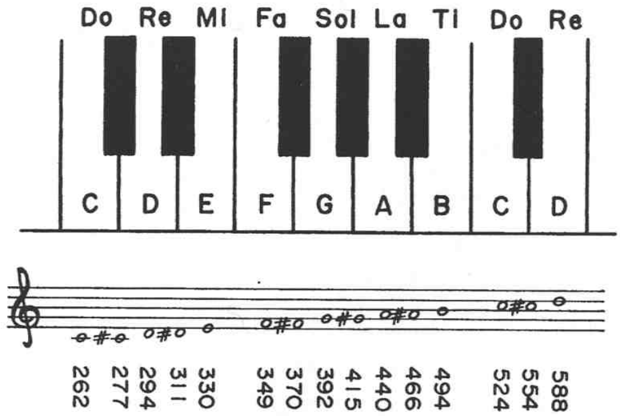

Figure 4.2 The frequencies of the chromatic scale on the piano keyboard.

distortion. Each pitch interval in the transposition will contain the same number of semitones as the corresponding interval in the original. Figure 4.2 shows the octave, beginning on middle C, divided into semitones, with the frequencies of the notes and their arrangement on the piano keyboard. Note that the major scale made up of the white keys (the C-major scale of Figure 4.1) has intervals between successive notes of either 1 or 2 semitones. You could construct a major scale beginning on any one of the keys in Figure 4.2 simply by preserving the same sequence of intervals as in the white-key scale—that is, 1-semitone intervals between the third and fourth and between the seventh and eighth notes of the scale and 2-semitone intervals elsewhere. And each melody with which you are familiar will remain the same melody whichever scale you use. In fact, unless you have absolute pitch, every time you sing ''Happy Birthday'' you are likely to sing it with a different set of pitches than the time before (that is, in a different key) without noticing the difference. You automatically reproduce the same intervals (measured in semitones) as the time before.

The system of tuning in which all intervals are constructed by adding semitones is called equal temperament. It is important to realize that the notion that all scale intervals should be derivable by the combination of some minimal modular interval (the semitone) arose as a rationalization of the structure of tonal scales already in use. Melodies had been sung using the major scale for centuries (if not millennia) before people thought of the possibility of expressing its intervals in semitones of equal size. While practical approximations to equal temperament were developed over a long period of time both in China and in the West (Kuttner, 1975; McClain, 1979), the exact mathematics of the system were derived in China around 1580 by the scholar Chu Tsai-Yü (Needham, 1962). This discovery made its way to Europe by 1630, and over the next century, came into more and more common use. Bach wrote his *Well-Tempered Clavier* in the 1720s and 1730s as a tour de force demonstration of its usefulness, systematically using major and minor scales beginning on all the 12 possible notes. In what follows, we describe how the frequencies of the semitone intervals are derived.

There is one more constraint that applies in conjunction with the fourth constraint, though it is usually just implicitly assumed. That is that when we apply the various constraints on scale construction that we have just discussed, then the scales we construct should consist of the intervals of scales and melodies already traditionally in use.

Before proceeding to the algebraic derivation of the frequencies of the equal-tempered chromatic scale, let us pause for a brief overview of the musical and psychological requirements such a scale is designed to satisfy:

1. discriminability of intervals,
2. octave equivalence,
3. a moderate number of pitches within the octave (usually about seven), and
4. the use of a uniform modular pitch interval (the semitone) with which to construct approximations of all the intervals of scales traditionally in use.

Equal Temperament

In deriving the equal-tempered chromatic scale of pitches, we rely mainly on the requirements that all octaves (2:1 frequency ratios) are of psychologically equal size and that each octave is divided into 12 equal pitch intervals. One consequence of taking tones separated by equal frequency ratios (e.g., 2:1) and placing them at psychologically equal intervals along the pitch scale (e.g., at octave intervals) is that the resulting psychophysical scale relating pitch to frequency will be logarithmic. Remember from Chapter 2 the definition of logarithm—if

$$a^b = c$$

then

$$\log_a c = b.$$

That is, the logarithm (b) of a number (c) is the power to which one must raise some base (a) to get that number. In Chapter 2, we used 10 as the base. Here we use 2, because of the 2:1 ratio of the octave. Consider the following series of frequencies of pitches we call C:

$$
\begin{aligned}
1 \times 32.75 \text{ Hz} &= 32.75 \text{ Hz} \\
2 \times 32.75 \text{ Hz} &= 65.5 \text{ Hz} \\
4 \times 32.75 \text{ Hz} &= 131 \text{ Hz} \\
8 \times 32.75 \text{ Hz} &= 262 \text{ Hz} \\
16 \times 32.75 \text{ Hz} &= 524 \text{ Hz} \\
&\vdots
\end{aligned}
$$

The series of numbers 1, 2, 4, 8, 16, and so on are all powers of 2:

$$
\begin{aligned}
2^0 &= 1 \\
2^1 &= 2 \\
2^2 &= 4 \\
2^3 &= 8 \\
&\vdots
\end{aligned}
$$

Note that the exponent gives the number of octaves a given C lies above the lowest C at 32.75 Hz. These exponents can be written as logarithms using 2 as a base:

$$
\begin{aligned}
\log_2 1 &= 0 \\
\log_2 2 &= 1 \\
\log_2 4 &= 2 \\
&\vdots
\end{aligned}
$$

Hence,

$$
\begin{aligned}
\log_2 (32.75/32.75) &= \log_2 1 = 0 \\
\log_2 (65.5/32.75) &= \log_2 2 = 1 \\
\log_2 (131/32.75) &= \log_2 4 = 2 \\
&\vdots
\end{aligned}
$$

In general, the number of octaves between two pitches is given by the binary logarithm (that is, the logarithm with base 2) of the ratio of their frequencies. The pitch interval P in octaves between two frequencies f_1 and f_2 is given by

$$P = \log_2(f_2/f_1).$$

Many graphs use a logarithmic coordinates for frequency, so that equal distances on the graph correspond to equal frequency *ratios* (or equal numbers of octaves) rather than equal frequency *differences*. Figure 4.3 presents pitch in octaves as a function of frequency, first with a linear frequency coordinate (A) and then a logarithmic frequency coordinate (B). The piano keyboard resembles (B)—equal distances along the keyboard represent equal frequency ratios, no matter where along the keyboard one is.

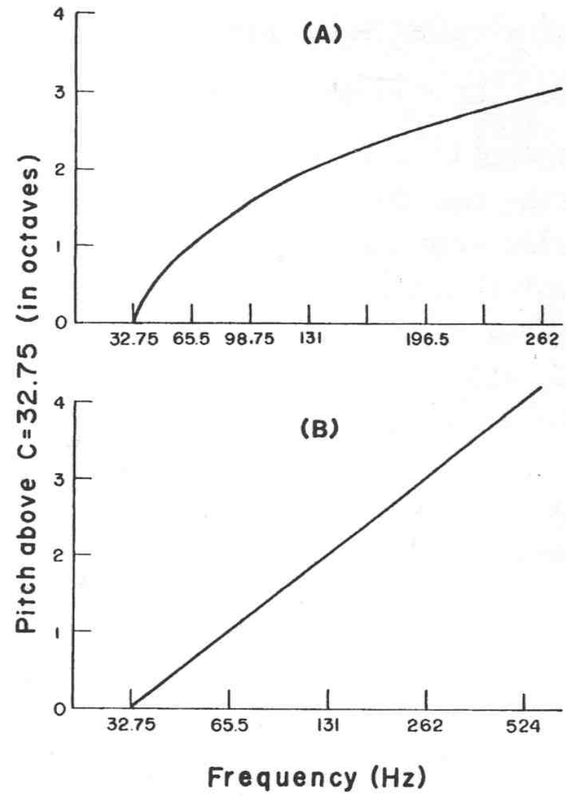

Figure 4.3 The pitch scale on linear (A) and logarithmic (B) coordinates.

Since all the octaves represent equal ratios of frequency, then all the semitones must represent equal ratios as well. This is because each octave contains 12 equal semitones. If we start with an octave C–C′ and move up 1 semitone to the octave C♯–C♯′, we still must have a 2:1 frequency ratio. Here we added a semitone at the top of the octave. In order to preserve the 2:1 ratio of the C♯–C♯′ octave, the semitone we added must have twice as large a frequency difference as the one we took away, as can be seen in Figure 4.2. Both semitones represent a constant 5.9% increase in frequency—to increase pitch by 1 semitone we multiply frequency by 1.059. The number 1.059 arises because it is the twelfth root of 2:

$$2^{\frac{1}{12}} = 1.0594631\ldots$$

If we start on any pitch and go up to 12 semitones, we reach a pitch one octave higher, a 2:1 frequency ratio. This comes out even if for each semitone increase, we multiply the frequency by $2^{\frac{1}{12}}$, since $(2^{\frac{1}{12}})^{12}$ is just 2. For example, start with middle C (262 Hz) and multiply by $2^{\frac{1}{12}}$ to get C♯:

$$f_{C} = 262\text{ Hz} \times 2^{\frac{1}{12}} = 277\text{ Hz}.$$

To get D, multiply the frequency of C by $2^{\frac{1}{12}}$:

$$f_D = f_{C\sharp} \times 2^{\frac{1}{12}} = 262\text{ Hz} \times 2^{\frac{1}{12}} \times 2^{\frac{1}{12}} = 262\text{ Hz} \times (2^{\frac{1}{12}})^2 = 294\text{ Hz},$$

and so on. We reach C′ after doing this twelve times, so

$$f_{C'} = 262 \text{ Hz} \times (2^{1/12})^{12} = 262 \text{ Hz} \times 2 = 524 \text{ Hz}.$$

You can work all this out on your handy pocket calculator, and we recommend playing with the calculation of various frequencies of pitches and intervals to familiarize yourself with the scale patterns involved.

Most of the cultures of the world do not use equal-tempered tuning as a basis of their tonal scale systems; that is, they do not use the fourth constraint. They do, however, base their scales on the octave, and that is sufficient to produce a logarithmic scale of frequency. In the above argument concerning the overlapping octaves C–C′ and C♯–C♯′, we could replace C with any arbitrarily chosen pitch X between C and C′ and get the same result. That is, the frequency ratio $f_{X'}/f_{C'}$ has to equal the ratio f_X/f_C. This follows since the X's being one octave apart, fall in a 2:1 frequency ratio. Thus, on a logarithmic scale of frequency, the interval X–X' will equal the interval C–C′, and the interval C–X will equal the interval C′–X'. In most cultures, the note X will be at some interval to C that cannot be expressed as a whole number of semitones.

There are further constraints that often operate on the selection of pitches in tonal scales whether the constraint of equal temperament is operating or not. (In what follows, we use the term *scale* to refer to the sort of tonal scale used in melodies, as distinct from the chromatic scale of all semitones, for example.) First, we might want a scale to contain a variety of intervals. Scales with only one logarithmic interval size between successive pitches do not afford as much possibility for melodic variation as do scales with more different interval sizes. This is apparent in even a simple tune like "Three Blind Mice" (Figure 4.4). Note that the second pair of phrases repeats the contour of the first pair, but with a subtle change of logarithmic interval size (Figure 4.4a). If we were to make all the intervals of uniform size, it would destroy some of the interest of the tune (Figure 4.4b). In Western music, the scale in which all intervals are 2 semitones is called the *whole-tone* scale. Some composers

Figure 4.4 "Three Blind Mice" with contour and semitone intervals: (A) familiar version; (B) version using whole tone scale, in which all intervals are equal to 2 semitones.

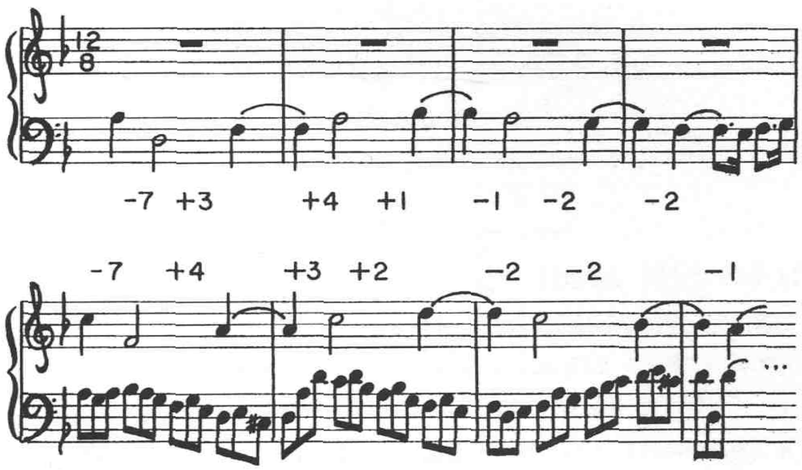

Figure 4.5 The opening of J. S. Bach's *Contrapunctus XIV* from *The Art of the Fugue.*

(e.g., Debussy in the 1890s) have experimented with its use, but the consensus among musicians is that it fails to offer enough intervallic variety to qualify as anything more than a novelty.

The use of the variety of intervals that arise when a thematic contour is moved up and down the scale is particularly apparent in the relationship of a tonal answer to the initial subject in a fugue. The start of a fugue is constructed like the start of a canon or "round" (e.g., "Frère Jacques"), with one part starting out alone with the subject and then continuing with an accompanying line while a second part enters with the subject. Sometimes this second part presents a literal transposition of the subject to a new key, preserving all its logarithmic interval sizes. But in other fugues, the second presentation of the subject is translated along the tonal scale in the same key as the initial presentation, preserving its contour but resulting in a change of its logarithmic interval sizes. Figure 4.5 shows the beginning of "Contrapunctus XIV" from Bach's *The Art of Fugue,* in which the second appearance of the subject varies from the first in its intervallic detail, providing a certain amount of melodic interest. (In fact, the answer changes the mode from minor to major, a contrast Bach plays upon throughout the rest of the fugue.)

Another aspect of the requirement of variety—one that appears only in an equal-tempered tuning system—is what we might call intervallic completeness. Consider the pentatonic scale shown in Figure 4.6. (The interval pattern is the same as that of the black notes on the piano, and it is very similar to the scales used in Chinese, Tibetan, American Indian, and Celtic folksongs.) Under the scale, the possible intervals are diagrammed. Note that only 8 interval sizes (measured in semitones) smaller than an octave occur, out of the 11 intervals that are possible when all the pitches of the chromatic scale are used. There are intervals available in the tem-

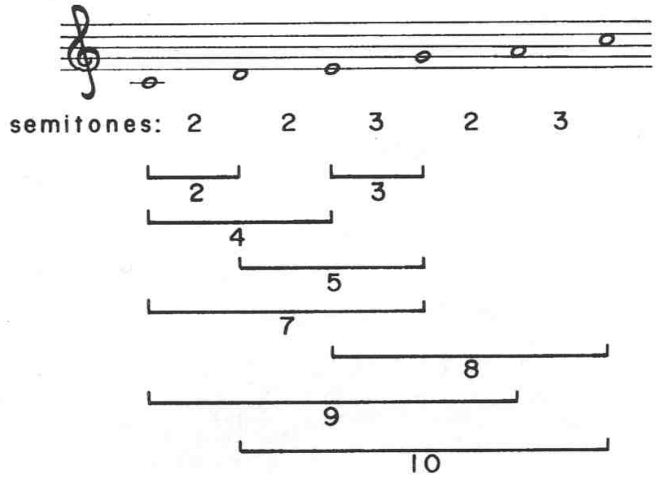

Figure 4.6 A pentatonic scale, with instances of its possible interval sizes within the octave.

pered tuning system of semitones that the pentatonic scale does not use (namely, intervals of 1, 6, and 11 semitones). Thus the pentatonic scale, though it provides considerable variety of intervals, does not provide the greatest variety possible in the semitone system. Balzano (1980) has shown that the smallest number of pitches that provide all of the possible intervals is seven. Figure 4.7 shows that, in fact, the major scale of seven pitches contains all the intervals.

One possibly desirable property of a melodic scale that applies to all systems, equal-tempered or not, is what Balzano (1980) calls coherence. Note that for both the pentatonic and the major scale, all intervals of two scale steps (as measured in semitones) are larger than any interval of one scale step, all three-step intervals are larger than any two-step interval, and so on. Such scales are called coherent. In scales not based on semi-

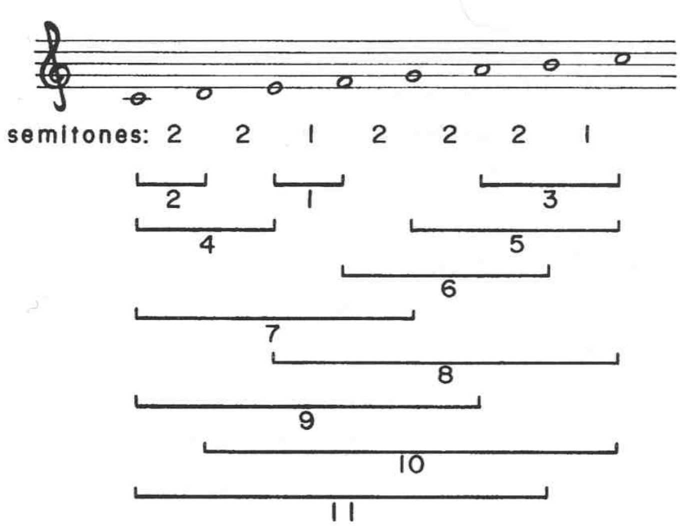

Figure 4.7 The C-major scale, with instances of its possible interval sizes.

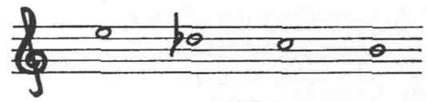

Figure 4.8 The ancient Greek chromatic mode.

tone intervals, coherence requires that logarithmic interval size increase with scale-step interval size with no reversals (namely, small logarithmic intervals containing more scale steps than larger logarithmic intervals). Most of the scales in the world exhibit coherence in this sense. Exceptions occur in lyre tunings from ancient Greece (Boring, 1929; Grout, 1960). Figure 4.8 shows the chromatic tuning system described by Aristoxenus (a pupil of Aristotle in the fourth century B.C.). Note that the scale is not coherent in that there is a 3-semitone interval containing one scale step, and a 2-semitone interval containing two scale steps. Coherence seems to be desirable in terms of Helmholtz's "mental measurement of pitch changes" approach, but it is clearly not so essential as to be universally present.

At this point, we can add two more optional constraints to our list of requirements that often figure in the determination of scale structure:

5. maximizing intervallic variety (completeness), and
6. preserving coherence of large and small interval sizes.

CROSS-CULTURAL EVIDENCE

The account given above of the relationship of the psychological dimension of pitch to the physical dimension of frequency is the one that, after weighing the evidence, we prefer to the various alternatives. The main reason for preferring this logarithmic system is its basis in the octave. People are very precise at making octave judgments of successive tones. Further, both musicians and nonmusicians can transpose logarithmic intervals along the scale without distortion, so long as the intervals are part of a familiar tune such as the NBC chimes (Attneave and Olson, 1971). This precision in using octaves and other logarithmic intervals seems nearly universal to all the cultures of the world, even to the extent of agreement on particular kinds of deviation from exact 2:1 frequency ratios for the octave. The evidence comes from both laboratory experiments and instrument tunings.

When adjusting successive tones to a subjective octave, listeners produce a "stretch" in the size of the octave, amounting to about 0.15 semitone in the midrange, or a ratio of about 2.009:1. The amount of stretching increases markedly in the higher registers. In Ward's (1954) study, although listeners differed somewhat among themselves, each individual was quite self-consistent, and Ward was able to develop a coherent (and essentially logarithmic) pitch scale based on their stretched octaves. (All we need to do is to use a semitone based on the twelfth root of 2.009). Figure 4.9 combines data from Ward (1954) and Walliser (1969) to show cumulative deviations of octaves from a 2:1 ratio over a four-octave range. The best confirmation Ward got for this stretched-octave scale was from one of his listeners who had absolute pitch. That listener's scale, based on her octave matches, agreed almost perfectly with her scale based on adjustments of isolated tones to match her internalized scale of note names.

Pianos are tuned with stretched octaves, but the stretch is only about half that found for subjective tuning of successive pitches. The stretch in piano tunings is due to the fact that piano strings, being somewhat stiff and not ideally flexible, have upper partials that are sharp (high in frequency) in comparison with integer-ratio harmonics. As would be expected from the discussion of tonal consonance in Chapter 3, simultaneous tone combinations on the piano sound best when the fundamentals of upper notes are tuned to coincide with the slightly sharp partials of

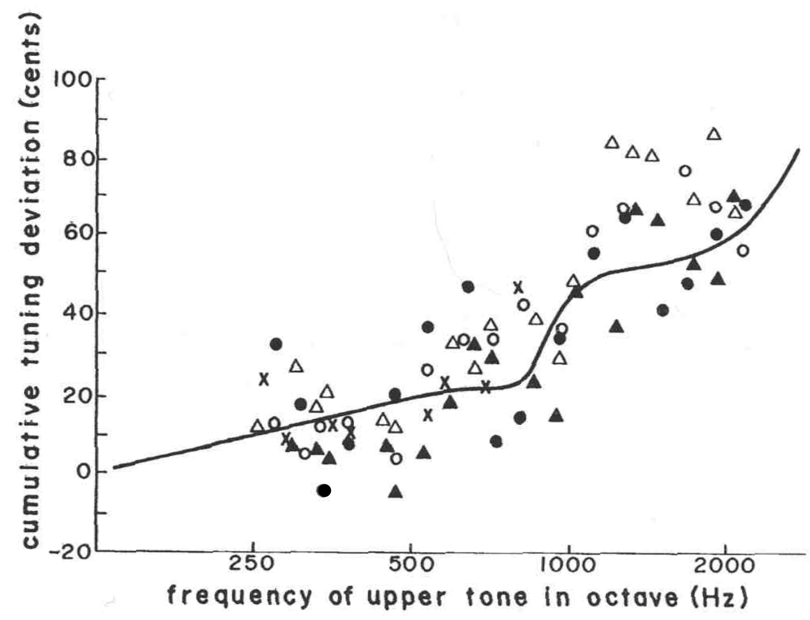

Figure 4.9 Data from octave judgments of Western listeners (solid line: from Ward, 1954, & Walliser, 1969) plotted with tuning measurements from non-Western instruments. X, Burmese harp (Williamson, 1968); circles, *gamelan sléndro,* triangles, *gamelan pélog;* open symbols from Surjodiningrat, Sudarjana, and Susanto, 1969; filled symbols from Hood, 1966. The abscissa represents deviations in cents (hundredths of a semitone) from octaves with 2:1 frequency ratios. (After Dowling, 1978.)

lower notes. This stretch is most pronounced where the relative stiffness of the strings is greatest, due to relative large ratios of diameter to length, namely, in the highest and lowest octaves (Martin & Ward, 1961; Young, 1952). That Western musicians adhere to the larger stretch of the subjective octave described above, rather than gravitate toward the smaller stretch of the familiar piano, strongly supports the argument for the inherent nature of the stretched octave for successive tones.

The adjustment of successive tones to stretched octaves has been replicated with non-Western listeners (Burns, 1974) and is reflected in numerous non-Western patterns of musical instrument tuning. Figure 4.9 shows data from all the precise measurements of tunings of instruments in actual use that we could find. The data definitely cluster around Ward's (1954) octave-judgment curve rather than the baseline. (The baseline represents octaves tuned to an exact 2:1 ratio.) It is essential that data like these be drawn from instruments in current use. Much of the literature on non-Western tunings derives from measurements on museum instruments long out of use. It is obvious that the tunings of stringed instruments would deteriorate rapidly over time. What is not so obvious is that seemingly more robust instruments such as the marimba- and xylophone-like instruments of the Indonesian *gamelan* also lose their tuning over time. Bronze keys need continual tuning adjustment (by filing) as the molecular structure of the metal gradually changes. Wooden keys can be tuned by filing or adding lumps of clay (adding and subtracting mass; see Chapter 2), with the obvious difficulty that the clay might fall off and disrupt the tuning. The resonator tubes of marimba-type (*gender*) instruments are tuned by changing the diameter of the hole in the top of the tube and by filling them with pebbles or sand to the appropriate length, an adjustment that is lost if the instrument is upended.

Most of the data in Figure 4.9 are from Indonesian *gamelan* tunings. It should not be supposed that the variability around Ward's curve represents a more or less ineffectual attempt by the Indonesian musician to match the stretched-octave pitch scale. The variability around the curve is intentional and results from the unique tunings of the various sets of instruments involving the stretching and compression of certain intervals in certain octaves, an effect we discuss below. It is clear from the account of vibrato-producing beats in the paired tunings of Balinese *gamelans* that the instrument makers are capable of achieving considerable precision in tuning. What is clear from Figure 4.9 is that, whatever tuning deviations are introduced, they occur as deviations from a stretched-octave curve and not from 2:1 octaves. (These deviations are small relative to the intervals between pitches in the tonal scales and so do not violate the principle of repetition of scale pattern within each octave.)

Another difficulty with the literature on instrument tunings is a bias on the part of many early investigators to find octaves with 2:1 frequency ratios. This, coupled with the imprecision of measuring instruments, led to more reports of 2:1 octaves than we would expect just from the normal variability of tuning and measurement. Ellis' appendixes to his English translation of Helmholtz (1877/1954) are full of such observations. This bias is very similar to the bias we are about to discuss concerning the establishment of tuning systems on patterns of small-integer frequency ratios.

ALTERNATIVE ACCOUNTS

Along with our preferred logarithmic scale, we need to present two common alternative accounts of pitch scaling and our reasons for rejecting them. One of these accounts comes from the psychophysical scaling tradition of psychology and claims that pitch is related to frequency, not by a logarithmic, but by a power function. The second comes from a numerological tradition in musicology and claims that all the intervals of a tonal scale are (or should be) derivable from small-integer ratios.

Psychophysical Scales

We are claiming that the relationship between pitch and frequency is logarithmic. The most serious alternative relationship, proposed by Stevens (Stevens & Volkmann, 1940), is that of a power function. While the logarithmic function describes correspondences between *additive* pitch increments (semitones and octaves) and frequency *ratios* (Figure 4.3), the power function describes correspondences between pitch *ratios* and frequency *ratios*. In constructing his pitch scale, Stevens used scaling methods similar to those described in Chapter 2 for the sone scale of loudness. For example, he had listeners perform magnitude estimation tasks in which they judged pitch by estimating how many times higher one note sounded than another. (That is, a pitch twice as high should receive a number twice as large.) Stevens calls his unit of pitch the mel, and his mel scale is shown in Figure 4.10, in which both pitch and frequency coordinates are logarithmic. If a power function were really appropriate, then Figure 4.10 would show a straight line relating pitch ratios to frequency ratios. The function is more or less straight in the midrange but breaks down especially in the high register. But the fact that the mel scale is not a power function should not bother us unduly. It might have been a useful function even if it curved (as it does). However, there are more serious

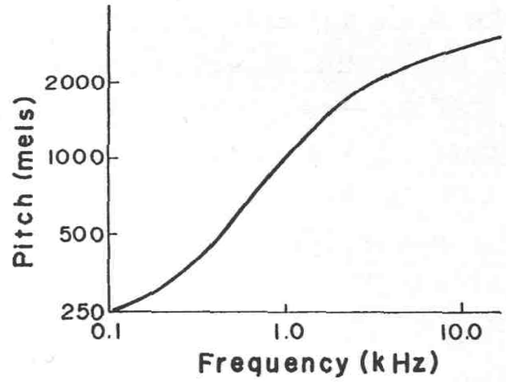

Figure 4.10 Stevens' mel scale for pitch.

reasons for rejecting the mel scale in favor of the octave scale as a psychophysical scale for pitch.

The main reason we prefer the logarithmic octave scale is that people are very precise at making octave judgments and very imprecise at the kinds of judgments required by magnitude estimation. As Ward (1970, p. 412) comments, "Measuring pitch in mels . . . is analogous to pacing off a room for wall to wall carpeting when a steel measuring tape is handy." Both musicians and nonmusicians are precise in making octave judgments for tones presented successively. Further, even with methods very similar to those used by Stevens, results emerge that support the logarithmic pitch scale, at least below 2000 Hz. Null (1974) used both the method of bisection (in which listeners adjust a comparison tone so that it falls midway in pitch between two standards) and a modified magnitude-estimation task. With both methods and both for musicians and for nonmusicians, the logarithmic scale gave a good fit to the data. (But for a dissenting view, see Schneider, Parker, & Upenieks, 1982).

Integer Ratios

A second source of doubts about the logarithmic pitch scale presented here is the musicologist who believes that the intervals of the scale derive from ratios of small whole numbers. For example, Bernstein (1976), in an otherwise admirable book, carries this approach to extremes. We reject this view for several reasons. First, musicians' intonation in practice is best described as an approximation of the tempered tuning system we have been describing (Ward, 1970). The exception to this rule is when small groups play or sing slow passages and have the opportunity to adjust their intonation to small-integer ratios, producing the effects of consonance discussed in Chapter 3. But that practice arose (in the West, at least) only after the development of polyphony and, hence, long after the development of the scale systems. Further, expectations of melodic direction can override the requirements of integer-ratio tuning, as in the case of a sharpened leading tone (seventh scale degree) resolving to a tonic. Insisting on integer ratios (just intonation) that run counter to such dynamic tendencies can be musically disastrous. Rosen (1972, p. 28) notes: "I once heard a quartet play [a certain passage] in just intonation with horrible effect. This is not to say that string players play, or should play, in strict equal temperament: pitch is always subtly altered, but for expressive reasons which have little to do with just intonation."

A second reason to reject integer ratios as a basis of melodic scales is that the supposedly simplest ratio, the octave, does not occur in its pure 2:1 ratio in listeners' adjustments of successive tones. People all over the world stretch the octave by about the same amount, suggesting that the stretched, and not the 2:1, octave is what is built into the human auditory system. Third, it is clear from the history of Western music since 1600 that musicians have found that the advantages for musical practice afforded by tempered tuning far outweigh the disadvantages. This would not be true if integer-ratio tuning had the overriding importance sometimes claimed for it.

It is important to realize that the approaches of both equal-temperament and small-integer ratios arise from attempts to rationalize existing traditional melodic scales. Small-integer systems were the earlier of the two to arise, and equal temperament arose in turn as an answer to musical problems raised by small-integer systems. The importance of small-integer ratios was realized more than 2000 years ago in both China and Greece (Boring, 1929; McClain, 1979), but, of course, the ratios that were measured were not ratios of frequencies (since the ancients did not know about frequency) but rather ratios of the lengths of vibrating bodies: strings (in Greece) or air columns (in China). Musicians in those cultures noticed the practical value of tuning the pitches of the scale so that certain simultaneous combinations of them sounded especially consonant, as discussed in Chapter 3. They noticed also that these consonant intervals arose from strings and pitch pipes having small-integer ratios of length (e.g., the fifth produced by a 3:2 ratio). They then tried to extend the system to derive the pitches of other keys than the one they started with. Such a project was important in second century B.C. China, for example, since each of the 12 keys had come to be associated with a month of the year (Nakaseko, 1957).

Here the early acousticians ran into a problem. A pure 3:2 ratio for the fifth and a pure 2:1 ratio for the octave cannot exist in the same scale. If we generate a cycle of fifths using 3:2 ratios, we use all 12 pitches without repeating: C–G–D–A–E–B–F♯–C♯–G♯–D♯–A♯–E♯(close to F). If we carry that one step further, we should arrive on a C seven octaves above the one we started with. But unfortunately, the twelfth power of 3:2 is larger than the seventh power of 2, and so the C arrived at by the cycle of fifths will have a different pitch from the C obtained by a cycle of octaves. The difference becomes gradually more pronounced as we progress around the cycle of fifths, so that intonation differences between instruments tuned to play in one key and those tuned to play in another become more noticeable. (Imagine an instrument with a scale based on the A♯ in the above cycle of fifths playing a C simultaneously with an instrument tuned to the original C.) The compromise represented by equal temperament can be seen as an agreement to hold firmly to the 2:1 ratio for the octave, making the frequency ratio for the fifth somewhat smaller than 3:2, namely, $2^{7/12}$.

In summary, the tempered logarithmic tuning system used in Western music is the result of compromises among a variety of constraints arising from the human auditory system, people's cognitive capabilities, and the requirements of musical interest. We turn now to elaborations of the system that reflect further features of human cognition.

MULTIDIMENSIONAL APPROACHES

So far we have been considering the construction of a psychophysical scale for pitch in terms of the two dimensions of pitch and frequency. But there is good reason to believe that pitch is not well represented by just one subjective dimension. Miller's (1956) argument to which we alluded above under the third constraint suggests that along any psychological dimension, the greatest number of categories we can use reliably is about seven. But a musician working within a tonal framework can use many more pitch categories than that. The musician can label the notes of the tonal scale (do, re, mi, etc.) and also tell which octave they came from. Across five octaves, that amounts to using about 36 categories consistently. Thus, if Miller's magic number of 7 ± 2 holds for pitch, there must be more than one cognitive dimension involved.

The intuitions of musicians also suggest that there are at least two dimensions to musical pitch. Pitches can be similar in two principle ways: They can be near each other in pitch, falling in the same general frequency range, and they can be similar in occupying the same place in their respective octaves. Two tones an octave apart (two C's, for example) are similar and are said to have the same chroma. A C and a D may lie next to each other in frequency and may be said to have similar *pitch height*. Figure 4.11 represents these two kinds of similarity by means of a spiral. Tones

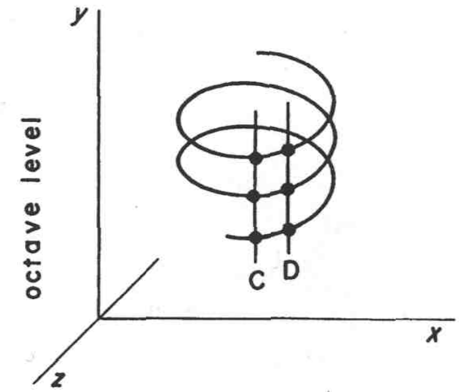

Figure 4.11 Helical model of the pitch scale.

of similar chroma are arranged in the same vertical column, and each turn of the spiral represents an increase in pitch of one octave. Tones similar in pitch height are adjacent to each other along the spiral and project onto adjacent points on the spiral's axis.

The independence of chroma and pitch height can be readily demonstrated with a piano. Start by playing the four highest C's on the piano. Then move down almost an octave and play four D's. Continue dropping by sevenths, playing the notes of an ascending C-major scale. You can hear both the descending series of pitch heights and the ascending series of chromas. Shepard (1964) constructed a more formal demonstration of this independence by creating a pattern in which chroma continuously changes while pitch height remains the same. Shepard's pattern (Figure
4a 4.12 and Example 4a) consists of 10 sine-wave components spaced at octave intervals. The components in the middle of the frequency range are more intense than those on the ends—the top and bottom components would be barely audible if presented alone. Following presentation of 10

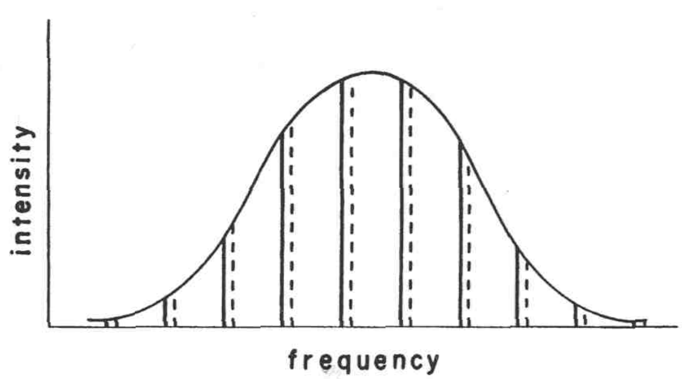

Figure 4.12 Shepard's auditory barber pole, illustrating a change of chroma without a change of tone height. Dashed lines indicate an upward shift in frequency of all components. (After Shepard, 1964.)

C's, the pattern shifts to 10 C♯'s (dotted lines in the figure). As the components shift up, those below the middle of the frequency range increase in intensity and those above the middle decrease, leaving the overall intensity envelope centered on the same frequency. The pattern sounds as though it has shifted up in pitch, though only the chromas have changed. The distribution of intensities across the frequency range has not changed. A succession of such changes sounds like a continuously rising pitch, which can be continued indefinitely—a sort of auditory barber pole. The listener can follow the rising pitch only so far, of course, after which it jumps down before starting up again. Burns (1981) has shown that this demonstration works even when the components are spaced at intervals somewhat larger or smaller than the octave. Burns's finding does not necessarily mean that tones an octave apart are not equivalent, however, but only that in such cases, pitch height is the overriding dimension.

In his opera *Wozzeck* (act 3, scene 4), Alban Berg used an effect similar to Shepard's—a continuously rising scale orchestrated so that as the upper instruments reached the top of their compass and dropped out, lower instruments entered at the bottom with the rising pattern. This creates an eery effect appropriate to the nightmare quality of the scene—Wozzeck searching the dark pond for the murder weapon. Berg's effect is not simply tacked on, but grows organically out of the musical structure of the scene. Ascending scales play an important thematic role in this part of the opera, leading to their striking use at the end of the scene.

The spiral in Figure 4.11 represents two aspects that are important in people's understanding of pitch: the close relationship of tones adjacent in the scale and the similarity of tones an octave apart. In accomplishing that representation, we have allowed the physical dimension of frequency to sink into the background. In fact, we could eliminate the frequency dimension entirely, and Figure 4.11 would still be a good representation of the psychological relationships among pitches. It would no longer be a psycho*physical* representation, but rather a multidimensional representation of mental structure. It could represent, for example, the pattern of judgments that listeners produce when judging the similarity of pairs of tones. The similarity judgments could be used to generate a model of the listener's cognitive structure of pitch relationships by means of multidimensional scaling techniques (discussed in Chapter 3). A computer program would take the similarity judgments and try to find a spatial arrangement of the pitches such that more similar pitches would lie closer together and less similar pitches farther apart. The program follows a trial-and-error procedure to search for a pattern that does the least violence to the similarity judgments in the data. The theoretical question is, Will the results of multidimensional scaling correspond to our intuitions concerning the mental structure as shown in Figure 4.11?

Krumhansl (1979) presents some evidence bearing on this issue. Krumhansl played listeners a major triad, establishing the framework of a tonality. Then she played them a pair of notes, and they judged the similarity of the two notes within the tonal framework. The pattern of similarity judgments led to a multidimensional scaling solution in which the pitches were grouped in three more or less concentric rings. The two inner rings contained the pitches of the major scale arranged in order of chroma around a half circle, and the outer ring contained the pitches not included in the major scale arranged so each was closest to its nearest neighbor in the scale. Similarity judgments thus lend partial support to the notion that one component of a multidimensional representation of pitch might be the ordered set of tone chromas arranged around a circle.

Krumhansl and Shepard (1979) collected data of a somewhat different sort. They presented listeners with a major scale minus its last note—the tonic at either the bottom or top. Instead of the tonic, they ended the scale with another pitch in the octave beyond the end of the scale. The listener was asked to judge how good that pitch was as a completion of the scale. This data, too, could be analyzed by multidimensional scaling techniques. Listeners differed in the degree to which chroma (vs. tone height) influenced their judgments, with musically more experienced persons basing their judgments more on chroma. In this study, too, the circle of pitch chromas compatible with the helix of Figure 4.11 appeared. Shepard (1982b) derived a multidimensional scaling solution from this data in which (viewed from one direction) the pitches are arranged chromatically around a circle and in which the tone-height difference between the top and bottom notes is represented by a gap where the circle fails to close. The same solution viewed from another direction shows the pitches arranged as a circle of fifths (C, G, D, A, E, B, F♯, C♯, G♯, D♯, A♯, F, C′) with the C's an octave apart occupying the same position. The circle of fifths represents some important structural properties in the organization of musical pitch. One reason for its importance arises from the central importance of the interval of the fifth itself within the scale, approximating as it does a 3:2 frequency ratio. Another feature of the circle of fifths is that by slicing it through with a straight line, one can separate the seven pitches of a given major scale from the other five pitches. Thus it can be used to represent the close relationship of the pitches of a scale versus the extraneous pitches.

Shepard (1982a, 1982b) took such considerations as the foregoing into account in constructing an idealized model of the psychological structure of musical pitch incorporating two intertwined spirals. This double helix model is shown in Figure 4.13A. In the double helix, each link along the curve of a spiral represents a whole step (2 semitones). Each link across

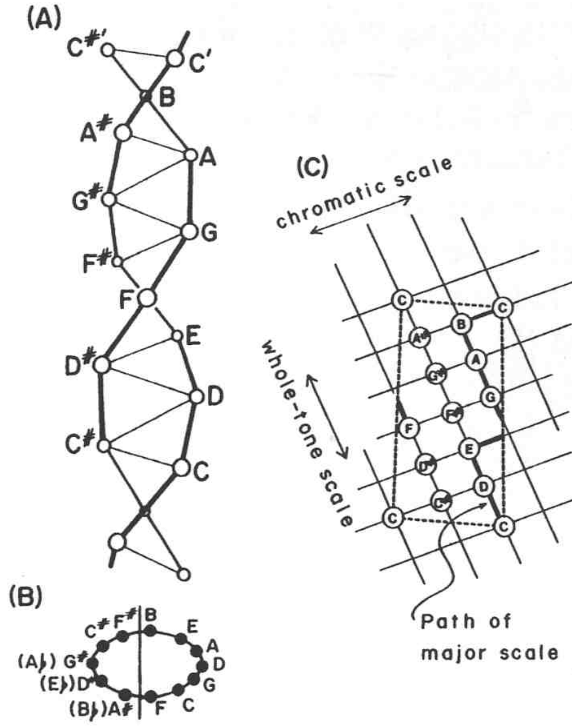

Figure 4.13 (A) Shepard's double helix representation of the cognitive structure of musical pitch. Note that the pattern projects downward onto the circle of fifths (B). (C) The double helix unwrapped onto a plane. (After Shepard, 1982.)

the model from one spiral to the other represents a half step (1 semitone). The vertical dimension in the model is simply tone height, and one could climb an ascending chromatic scale by tracing the sequence of links back and forth between the spirals. Imagine that the model is standing on its end (just as shown) in the middle of the floor and we dropped a ribbon from each node to the floor, with the label of the node at the bottom. Each ribbon extended upward to the ceiling would pass through a series of tones having the same label and lying at octave intervals from each other. The pattern of node labels on the floor is the circle of fifths (Figure 4.13B). Just as noted above, we could draw a line through the circle of fifths dividing the seven notes of a major scale from the others. In Figure 4.13C this has been done for the C-major scale. And the plane rising vertically from that line into the double helix (among the ribbons) would then slice off the notes of that scale. Such a plane also would separate out the pitches of the commonest sort of pentatonic scale as well.

The circle of fifths is important for another reason that we explore more fully in Chapter 5: It expresses the psychological difference between musical keys, or key distance. If we take the plane that divides the notes of a given scale from the other notes and rotate it around the vertical axis by one notch (adding one note and deleting one note), we get a scale in a new key that is very similar to the one we started with. That is, the new key shares all but one of its pitches with the old key. However, if we rotate the plane several notches around the circle of fifths, we get a key that shares fewer notes with the first. Such a key is said to be distant from the first key. Krumhansl, Bharucha, and Castellano (1982) have demonstrated the psychological reality of key distance using listeners' judgments of chordal patterns. In Chapter 5, we present evidence that key distance is a very important factor in judgments of melodic similarity.

Now imagine that the double helix is one of those cookie rollers that imprints a repeating pattern on a flat surface of cookie dough. As we roll it across the surface, the pattern of pitches is "unwrapped" onto the plane as shown in Figure 4.13b. Now the major scale that we sliced off with the vertical plane becomes a zig-zag path. The chromatic scale of half steps is a straight line, as is the whole-tone scale. In addition to capturing the significance of the octave, the circle of fifths, and the major scale, Shepard's model also captures the fact that in context, the different-sized intervals of the major scale are heard as psychologically equal. This is represented by equal distances for both half and whole steps, with the contextual difference represented by difference of direction. The beauty of Shepard's double helix lies in its success in capturing succinctly several important facts of pitch perception in one coherent structure.

We want to emphasize that the structure in Figure 4.13 is a model designed to express certain relationships that listeners perceive among pitches. This does not mean the structure need literally be stored in the brain but only that the relationships it expresses have analogs in brain structure. Note also that a structure like this has no intrinsic implications for psychological processes such as perceiving, judging, and remembering. Different aspects of the structure will become dominant, depending on the task we give the listener in an experiment. We return in Chapter 5 to the effect of task on the listener's perception and behavior.

Before we leave the multidimensional scaling of pitch, we quote Krumhansl and Shepard (1979) in a statement that expresses some significant truths concerning research with musical phenomena. They suggest that multidimensional scaling approaches can be used to disclose individual differences among listeners as well as differences due to the auditory context in which judgments are made.

> To the extent that tones differing in frequency are not interpreted as musical stimuli—because they are presented in isolation from a musical context, because the tones themselves are stripped of harmonic content, or because they are played to musically unsophisticated listeners—the most potent factor governing their perceived relations is simply their separation along a one-dimensional continuum of pitch-height. To the extent that tones are interpreted musically—because they are embedded in a musical context, because they are rich in overtones, and because

> they are played to musically sophisticated listeners—simple physical separation in log frequency gives way to structurally more complex factors, including octave equivalence or its psychological counterpart, tone chroma, and a hierarchy of tonal functions specific to the tonality induced by the context. (p. 529)

SCALES IN OTHER CULTURES

The set of pitches in Shepard's model we have just discussed is specific to Western music. We cannot expect the details of its double helix to apply directly to non-Western scale systems. For example, the circle of fifths, a very strong structural element in Western music and a central component of the model, does not play nearly so great a role in non-Western musical systems. However, the more basic structural properties that went into the construction of the logarithmic psychophysical scale for pitch (Figure 4.3) and the single helix expressing octave equivalence (Figure 4.11) seem to us to be quite generally applicable to most of the musical systems throughout the world. This is because nearly all cultures use logarithmic pitch scales in which octave equivalence plays a central role. Those cultures that cannot be said to use such a model do not use an alternative model, but rather use such restricted melodic patterns that the form of their scales beyond the range of one octave is moot. Thus, for nearly all the cultures of the world, the helix of Figure 4.11 represents the underlying psychophysical organization of pitch material. Pitches an octave apart are treated as functionally equivalent in melodies, though each culture fills in the pattern of pitches within the octave in its own way. (In what follows, we discuss scale structure in terms of sets of pitches. For many purposes, the equivalent description in terms of intervals would be more appropriate, since in most cases the anchoring of a pitch label to a specific frequency (as in A = 440 Hz) is irrelevant to both the psychological structure and the culture being discussed. We simply use the pitch-set description for convenience.)

There are certain regularities in the way a culture fills in the octave and in the way it uses pitches in melodies that are best disclosed by considering melodic scales to be constructed out of the underlying tonal material through a process involving several levels of psychological analysis (Dowling, 1978b, 1982c). These levels, shown in Figure 4.14, are (1) the underlying *psychophysical pitch function* that assigns pitches to frequencies, (2) the *tonal material* consisting of all the possible pitches that could be used in melodies, (3) the *tuning system* consisting of a subset of the pitches in the tonal material that could be used as the basis of a variety of modal scales, and (4) the *modal scale* in which the pitches of a tuning

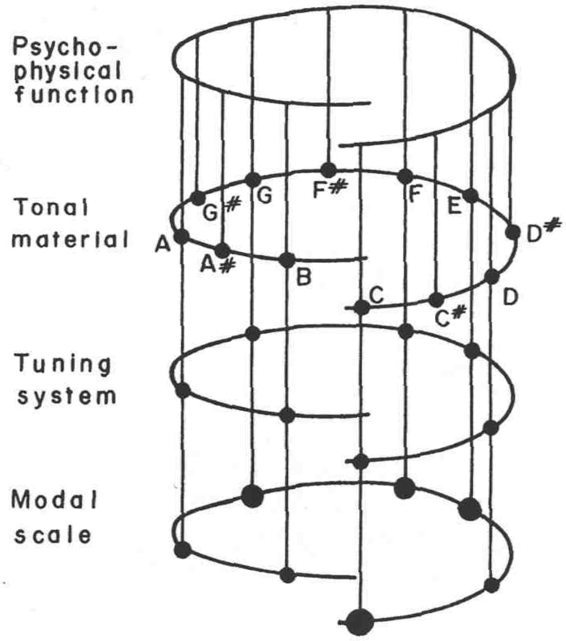

Figure 4.14 Levels of analysis of musical scales, using Western pitch labels.

system are hierarchically organized with a tonal center (tonic) and are used in actual melodies. (Hood, 1971, presents a similar approach that differs from this one in some details.)

Each of these four levels of analysis is formed by making a selection of pitches out of the next higher level or by imposing some constraint on them. The levels progress from the highly abstract psychophysical scale, containing all possible pitches, to the very concrete modal scale, in which the pitches are the actual pitches of melodies. There is good evidence for the psychological reality of these levels, both from the analyses of music theorists in a number of cultures and from laboratory experiments in our own culture. We review these laboratory experiments in Chapter 5. Here we present the theoretical bases of this outline.

Figure 4.14 is constructed in terms of the Western scale system, since that is the most familiar for us and is likely to be familiar to our readers. In it the tonal material consists of all the notes on the piano—the chromatic scale of 12 steps to the octave. The tuning system consists of a selection of 7 out of those 12 pitches—the white notes on the piano that can form the basis of various modal scales. The tuning system is more abstract than any of the modal scales formed from it in that it contains a cyclic pattern of intervals without any specified point of origin. Selecting a point of origin in the tuning system amounts to choosing a specific modal scale (C major in Figure 4.14). In some cultures, the construction of a modal scale out of the tuning system involves the omission of certain pitches as well. Modal scales are the most concrete level in the scheme and involve not only the selection of the pitches of melodies, but also the imposition of a tonal hierarchy that shapes the listener's expectations.

Tonal Material

In many cultures, the number of pitches that are potentially available in all melodies exceeds the number that might actually appear in a single melody. This larger set we call tonal material. Tonal material is the level at which the undifferentiated continuum of the psychophysical scale is divided into discrete categories. In Western music the tonal material consists of the pitches of the chromatic scale, as discussed above. North Indian music provides a good illustration of the use of a level for tonal material. The North Indian octave can be divided into 12 more or less equal steps, and the tonal material divides into a set of two pitches a perfect fifth apart that function as the first and fifth degrees of every modal scale, plus a set of five pitches each of which can appear in either a higher or a lower variant (Jairazbhoy, 1971). The pitches of the tonal material are shown in quasi-Western notation in Figure 4.15. C is used as a point of origin here, but that should not be taken as fixing the pitch level; nor should the Western pitches be taken as anything more than rough approximations of the actual pitches used.

Tuning System

In Western music, we use basically one tuning system corresponding to the interval pattern of the white notes on the piano (see Figure 4.2). This tuning system should not be thought of as having a fixed pitch, however, but rather as capable of sliding up and down. That is, we could recreate the same interval pattern by starting with any pitch on the piano. And we could also create it by singing the same interval pattern beginning with any pitch whatsoever, whether on the piano or not. The interval pattern between successive notes is, in semitones, 2, 2, 1, 2, 2, 2, 1 (see Figure 4.7).

In North Indian music, too, we should conceive of the tuning system as an interval set rather than as a series of fixed pitches. The five varying

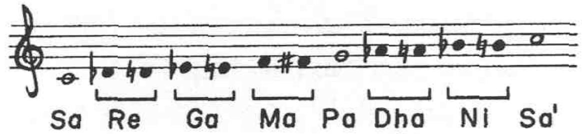

Figure 4.15 The tonal material of North Indian music. *Sa, Re, Ga* . . . corresponds to do, re, mi. . . . Brackets denote pairs of pitches of which only one member typically appears in a scale.

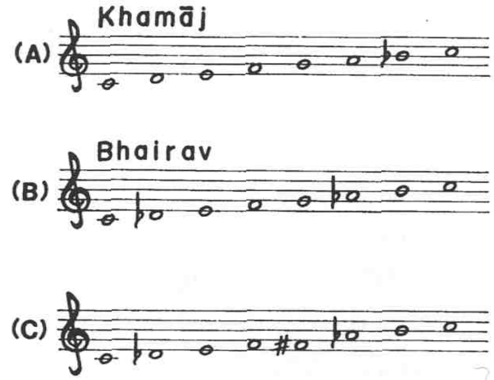

Figure 4.16 Three North Indian tuning systems generated by making different selections from the tonal material.

pitch categories from the tonal material (*re, ga, ma, dha, ni*), each of which can take one of two values, generate $2^5 = 32$ *possible* scale frameworks or tuning systems. In addition, there are three patterns in which the F—*ma*—can appear in both natural and sharp forms, giving a total of 35 possible tuning systems. Of these 35, 10 are in common use, and about 10 more are used occasionally. Figure 4.16 gives some examples of tuning system scales (Sanskrit *tḥat,* "framework," pronounced "taht") generated from the tonal material of Figure 4.15. The level of tuning system is especially useful here, since the pitch series of a *tḥat* does not typically appear intact as a modal scale.

Modal Scales

A modal scale is created from a tuning system by imposing a structural hierarchy on the set of pitches. In its simplest form, this structure designates some tones as more important than others (some may drop out entirely) and establishes dynamic patterns of expectation concerning where pitches might not lead in a melodic sequence. At present, Western music relies primarily on two modal scales: major and minor. These are generated from the intervals of the tuning system by taking different starting points in the cycle of intervals (Figure 4.17). Any of the 12 pitches in the tonal material can be taken as a starting point for either mode, giving 24 major and minor keys.

A complication arises in the case of the minor mode as it has been used since the seventeenth century. Melodies in the minor often use a different set of pitches for ascending and descending (Figure 4.18). The descending series is the same as in Figure 4.17, but the ascending series is more like the major mode starting on the same pitch in that the sixth and seventh degrees of the scale are raised. An intermediate form, in which only the seventh degree is raised both ascending and descending, also occurs. In

Figure 4.17 Major (A) and minor (B) modal scales in Western music, with intervals shown in semitones. (C) illustrates the generation of the minor mode starting on the pitch C using the interval sequence of (C). The flat (♭) lowers the pitch by 1 semitone.

terms of the conceptual scheme of Figure 4.14, it seems most convenient to introduce these alternate pitches by dipping directly into the tonal material, that is, not by including them in the tuning system.

In Western music through the early 1600s, there were four additional modes available besides the major and minor: those beginning on the second, third, fourth, and fifth steps of the white-note pattern (Figure 4.19). Of these modes, the Dorian and Phrygian are most likely to be encountered today. A common version of the tune "Greensleeves" (Figure 4.20A) is largely in the Dorian, though this tune is often assimilated to the modern minor mode by lowering the sixth degree of the scale (B♭ for B in Figure 4.20A). Phrygian tunes are rarer. One of the more familiar is the old hymn "Oh Haupt voll Blut und Wunden," which appeared as a popular song in the 1960s. Bach used this tune extensively, but by the early eighteenth century, the Phrygian was a very old-fashioned mode. Of five harmonizations by Bach given in Riemenschneider (1941), only one takes the Phrygian tonic (E in Figure 4.20B) as the tonic for the harmonization. Three take the sixth degree (C) as the tonic, letting the melody end on the third degree of the final chord. The remaining harmonization is a hybrid of those two approaches.

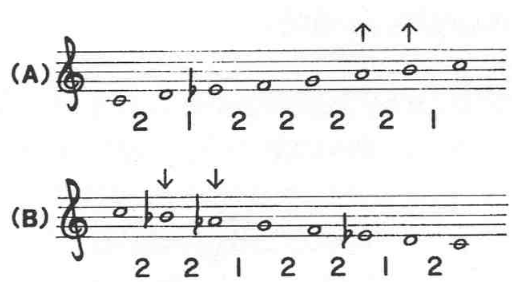

Figure 4.18 Ascending (A) and descending (B) forms of the melodic minor mode. Arrows indicate chromatic alterations.

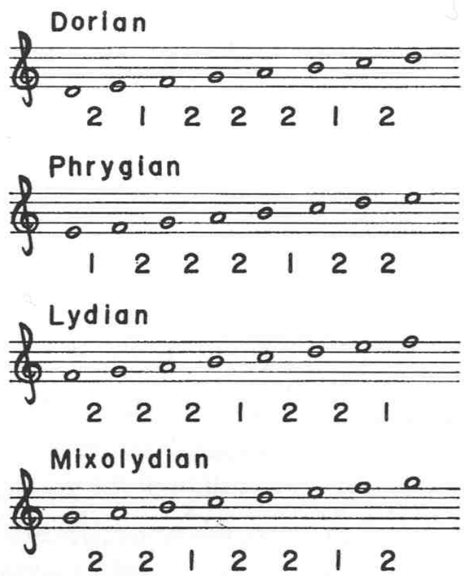

Figure 4.19 Four medieval Western modes.

In turning to a consideration of modal scales in North Indian music, we encounter several complications that are not unusual in non-Western music. First, a mode (Sanskrit *rāg*) includes more than just the pitch series of a modal scale. A *rāg* is associated with a particular aesthetic–affective quality (*rasa*), includes particular melodic phrases and expectations (especially cadence formulas), and is customarily performed at a particular time of day. Second, not only does the modal scale impose a structural hierarchy on the pitches, but modal scales are often "gapped," using only five or six of the pitches available in the *ṭhat* tuning system. Further, the

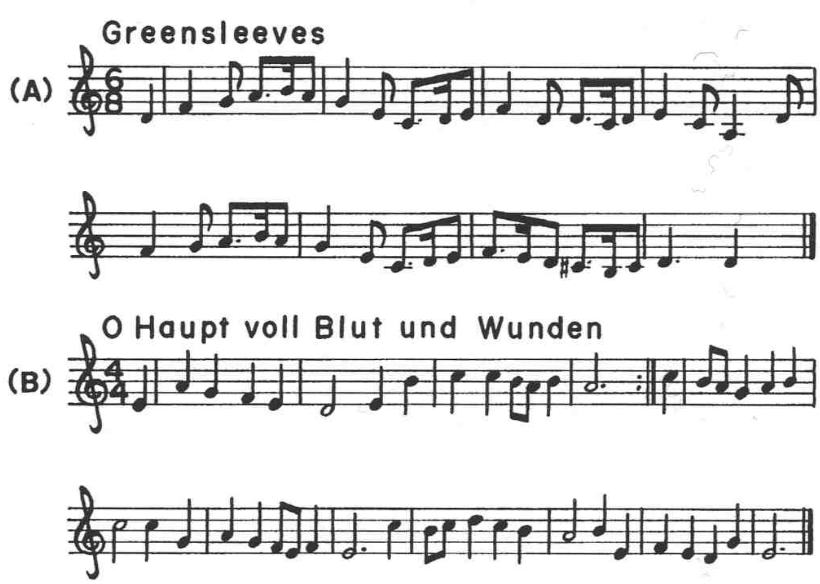

Figure 4.20 Melodies in the keys of (A) D Dorian ("Greensleeves") and (B) E Phrygian ("O Haupt voll Blut und Wuden").

Figure 4.21 Two North Indian scale patterns, based on Figure 4.16A and B, respectively.

type of pattern used in the Western minor mode, where ascending and descending melodic lines use different sets of pitches, is common in North Indian music. In such cases, the descending line is usually taken as a more basic, being involved as it is in cadence patterns at the ends of phrases. Figure 4.21 gives two examples of modal scales, with characteristic directional patterns of melodic lines in the *rāgs* indicated in a sort of shorthand notation, more important notes being given longer time values. *Rāg Des* (Figure 4.21A) is based on *that* of Figure 4.16A. Note that the descending pattern uses seven pitches and contains irregularities in which the first and fourth steps are approached by skips. The ascending pattern is pentatonic (gapped), and the seventh degree is raised. More elaborate examples of chromatic alteration can be seen in *Rāg Ramkali* (Figure 4.21b), based on Figure 4.16b. The ascent is six-toned, while the descent has elaborate turns of phrase involving chromatic alteration of the fourth and seventh degrees.

Indonesian music illustrates some of the same principles as Indian music in the relationship of tuning system and mode, though it possesses unique characteristics. Classical Javanese music has two tuning systems: *pélog* (heptatonic) and *sléndro* (pentatonic). Different sets of instruments in the same *gamelan* are tuned to the two systems, which typically share only one pitch. (The common pitch is useful in occasional modulations from one system to the other in contemporary performances.) Modes in the *pélog* system are created by selecting gapped pentatonic scales from the seven pitches in the system. The term closest to mode is *pathet,* but like the Indian *rāg,* the Javanese *pathet* is as much characterized by particular melodic phrases as by the pitches of its scale (Becker, 1977). In fact, of the three *pathet* in the *pélog* system, two share the same modal scale. All three *pathet* of the *sléndro* system simply use the pentatonic scale of the tuning system pitches. In those cases, characteristic melodic contours are the main distinguishing features of the modes. Unlike the Indian *rāg,* a *pathet* is not associated with a specific *rasa* (affective tone)—there are not enough *pathet* for them to serve that function.

Rather, the classification of *pathet* (at least for the *sléndro* modes) seems to have arisen from the association of particular songs with particular phases of dramatic performances they accompany, in turn associated with particular time periods in the course of an all night performance.

Not only are the two tuning systems of a *gamelan* different from each other, but (as we mentioned in connection with Figure 4.10) each *gamelan* uses different tuning systems from the next, and even within the same *gamelan,* the interval patterns of successive octaves differ in detail. These tuning variations are associated with specific aesthetic qualities. For example, when the interval between the first and second degrees of the tuning system is especially large, the effect is characterized as bright and cheerful, whereas if it is small, the effect is thought to be soft and gentle (Becker, 1977). A particular *gamelan* might thus be thought more apt for playing certain *pathet* than others. Many *gamelans* have proper names that emphasize these affinities. for example, the *gamelan* at the University of Michigan is called Khyai Telaga Madu (Venerable Lake of Honey) because of its "sweet" tuning.

Ornament Tones

There is one remaining class of pitches from the level of tonal material that we have not yet mentioned. Members of this class are what we might call ornament tones. In the music of India especially, there are clusters of small intervals around the notes of the modal scale that can be used to ornament the principle scale pitches with trills and grace notes. The use of these tones, illustrated in Example 4b, has led some to think of Indian music as using microtonal intervals in musical scales. However, as we have seen, the North Indian system of modal scales uses at most seven pitches to the octave, and the additional pitches (15 more in some analyses) appear in a strictly subsidiary role.

Ornament tones also appear in the music of Indonesia. Most instruments of the *gamelan* have a fixed set of five or six or seven pitches to the octave. Vocal soloists, however, are not so restricted. The vocalist can use, as passing tones, pitches from the tuning system omitted from a gapped pentatonic scale by the instruments, as well as introduce microtonal intervals by way of ornament (*surupan*). Players of the bowed two-string spike-fiddle (*rebab*) also follow this practice.

The above outline of levels, going from the abstract and general level of the psychophysical relationship of pitch and frequency to the very specific and concrete level of the selection of actual pitches that occur in actual melodies is found in enough musical cultures to justify according it a certain amount of psychological reality. Additional evidence bearing on this issue appears in Chapter 5. But before leaving pitch scales, we must attend to one more topic.

ABSOLUTE PITCH

No chapter on pitch perception in music would be complete without mention of the phenomenon of absolute pitch—the ability of some individuals to identify the notes of the chromatic scale by name when presented in isolation from other pitches. This ability is a relatively recent development, specific to Western culture, since it is only in the last few hundred years that anything like standard pitches (such as an A of 440 Hz) have come into general use. Ellis (Helmholtz, 1877/1954, Appendix 20, Section H) presents evidence for a fairly wide range of pitch standards in Europe even in the nineteenth century. (Of course pitch standards were in use throughout the history of ancient China, but the standards were customarily changed with the ascent of each new emperor (Needham, 1962). The phenomena associated with absolute pitch cast additional light on some of the main themes of discussion in this chapter: the importance of the octave, the importance of an internalized framework of the scale, and the hierarchical nature of that framework.

The phenomena of absolute pitch have been the focus of heated debate over whether the ability is innate or acquired through musical experience. As Cuddy (1968) points out, the tendency to put the question in all-or-nothing nature–nurture terms gravitates toward the scientifically meaningless, since no matter what experiences a person has, it is impossible to prove that the emergence of absolute pitch was not due to maturation in favorable circumstances rather than to more ordinary processes of learning. In any case, "one can always choose to define absolute pitch so as to exclude all cases where some kind of formal training can be detected" (Cuddy, 1968, p. 1069). It is clear that some factors involved in pitch judgment, both absolute and relative, are very likely innate. (*Relative* pitch involves proficiency at *interval* recognition and production.) The evidence on the importance and pervasiveness of octave generalization suggests such innateness for the primacy of octaves in pitch scales. However, some aspects of absolute pitch abilities cannot be innate, for example, the pairing of note names with frequencies. Although it is clear from the excellent reviews of the subject that are available that absolute pitch is more easily acquired at an early age (Shuter-Dyson & Gabriel, 1981; Ward, 1963), evidence has been accumulating on adults' success in acquiring the ability. This evidence also bears on the question of the psychological structure of pitch scales.

It seems unlikely to us that the ability for absolute pitch is bimodally distributed in the population, that is, that some have it and others do not. It seems more consonant with our experience that people possess the ability in varying degrees and that whether the ability shows up depends on the particular task demands the person faces. Certainly, some of the component abilities necessary to the naming of pitches are present in greater or lesser degree throughout the population (Hurni-Schlegel & Lang, 1978). Terhardt and Ward (1982) found that musicians could discriminate alterations in key of 5-sec excerpts from Bach's *Well-Tempered Clavier*. Their listeners performed well above chance in telling whether each excerpt had been transposed up or down from the original key shown in a score printed on the answer sheet. And nearly everyone possesses a temporally local absolute pitch. That is, after the age of 5 or 6 people are generally able to sing familiar tunes maintaining the same tonic reference pitch throughout; and when the pitch shifts, it is generally to a new consistent frame of reference (Dowling, 1982b). Choir members learn to remember the pitch they just finished singing through all sorts of confusing context in order to enter on that pitch again 20 or 30 sec later. And persons possessing long-term absolute pitch differ in their labeling accuracy, the frequency range over which they can succeed in the labeling task, and the variety of musical instruments for which they can identify pitches. Individuals with good absolute pitch seem to have "stored a limited number of points along the frequency continuum in long-term memory and . . . use this information for classifying current pitch inputs" (Siegel, 1972, p. 86).

It has been clear for many years that training can improve note-naming performance (Ward, 1963), and since the 1960s, more dramatic results have been obtained in adults' acquisition of absolute pitch. Brady (1970) reports his experience teaching himself absolute pitch at the age of 32. He programmed his laboratory computer to produce tapes consisting of a succession of sine waves tuned to different pitches in the chromatic scale. The earlier tapes in the sequence contained high proportions of the pitch C, providing a stable reference point. Brady practiced naming the notes on the tape a half hour a day for two months, receiving feedback after each trial. He avoided trying to solve the task by figuring the relative pitch intervals between successive notes. His conscious task throughout was to retain just the one pitch, the C. After two months, Brady's error rate was negligible and his experience was that "the 'chroma' dimension so overwhelmingly overrode the high–low dimension that most of the tasks, with practice, became very easy" (Brady, 1970, p. 884). Sounds in the environment began to take on codable pitch qualities—the B refrigerator, the child's pull-toy in A. This is not to say that absolute pitch is not easier to learn in infancy, just that it *can* be acquired by some adults.

In his learning of absolute pitch, Brady took advantage of a method explored by Cuddy (1968, 1970, 1971) in a series of experiments. Cuddy found that training methods that emphasized the acquisition of a framework of reference tones among all the notes of the scale led to better pitch identification than training methods in which all the scale notes received equal attention. This was true even when persons of moderate musical experience were learning to label pitch sets defined by intervals of equal numbers of mels rather than by equal musical intervals (Cuddy, 1970). And when the framework for which labels are learned is already a familiar pattern—for example, the F-major triad in the case of music students—performance was especially good (Cuddy, 1971). All of this gives further support to the notion that the scale schemata that are internalized are well characterized as frameworks with definite focal points, hierarchically arranged. This is a theme we return to in Chapter 5.

SUMMARY

Musical pitch is uniquely represented by scales with discrete steps, which function as perceptual categories for hearing, remembering, and producing music. These musical scales appear to be governed by a few major constraints: (1) discriminability of intervals, (2) octave equivalence, and (3) a moderate number of pitches within the octave. This psychological structuring of the pitch dimension is related to the physical dimension of frequency logarithmically, and we have presented theoretical and experimental evidence to support this view. We have also argued that the tempered logarithmic tuning system used in Western music is a product of compromises of these constraints and of performance practice.

We described recent studies that indicate that pitch is more than a unidimensional subjective experience. Pitch adjacency and pitch height (or chroma) can be represented by an idealized model incorporating a double helix, suggesting important relationships for musical cognitive processing. An examination of the Western tempered tuning and scale systems from other cultures supports the hypothesis that people construct scales through several levels of processing: (1) the underlying psychophysical pitch function, (2) the tonal material, (3) the tuning system, and (4) the modal scale. The psychological reality of this anslysis draws face validity from the many traditions of musical pitch systems it describes and is the basis for the study of melodic organization processes in the next chapter.