This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001

# 10 Scales and Intervals

*Having reviewed diverse approaches to MmT and its interdisciplinary foundations, we can now return to some basic questions. Is a musical interval a frequency ratio or a pitch distance? Are musical intervals absolute and universal, or are they culturally embedded and learned? How are musical scales structured, and what is the origin of the chromatic and diatonic scales? For millennia, Western music theory has been dominated by mystical Pythagorean concepts of intervals and scale steps as ratios of small integers. Ancient and medieval theorists favored Pythagorean ratios, which were limited to factors of two and three; major 3rds were considered dissonant (64:81). Renaissance theorists such as Zarlino assumed Just tuning and intonation, in which intervals are simple ratios and major 3rds have the ratio 4:5. Musical ratio theory is an important part of cultural heritage and the history of ideas but misleading in an empirical psychological approach. Empirical studies suggest that musical intervals are inherently approximate pitch distances that deviate systematically from theoretical ratios and, like most things about music, are learned from oral-aural tradition. Their optimal tuning is a compromise among partly conflicting criteria. Whereas diatonic scales maximize the number of octave and 5th intervals between scale steps, the tuning accuracy of each interval and each scale step is limited by JNDs. This* psychocultural *approach has interesting implications for a theory of MmT.*

## Psychohistoric Origins

A musical scale is a set of tones or pitches that are used to make melody, polyphony, and harmony. It is a resource for making music. It is also an abstraction, relative to the more concrete phenomenon of melody. If a given scale can be transposed to any starting pitch, its important parameters are the intervals between the tones.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



To understand what scales are, we need to think about where they come from. For an individual listener or performer, a musical scale is learned from music. Musical interval sizes are similarly learned from music. The ideal size or intonation of a musical interval is that which is most familiar from music performances. That is part of the oral-aural tradition of music, which includes all audible details of musical performances.

Lewis Rowell (2015) offered examples from Indian music. For Western music, Anne McLucas (2010) explained:

> Oral tradition in music, then, refers to those aspects of music that are passed down by humans teaching one another the art form, whether in person or by means of recordings, radio, television, or other non-written means. It also refers to the kinds of composition that occur completely without referent to written notation, that is, the creative as well as the re-creative forms of oral tradition. (2)

> Different parameters of the musical experience may stand out when looked at from the perspective of oral tradition. The melodic and harmonic aspects so prized by Western musicologists and theorists may give way in analysis of orally transmitted music to aspects of rhythm and timbre that are less easily expressed in notation, either because of their complexity, as with timbre and the details of rhythm, or because of their all-pervasive repetitive nature, as with many danceable pieces. (157)

As the word "interval" implies, the most psychologically salient feature of a musical interval is its perceived width. In a psychoacoustical approach, that is not necessarily exactly the same as its physical width (frequency ratio or width in cents) due to pitch shifts (see chapter 6). When music students with good musical listening and performing skills hear isolated harmonic (simultaneous) intervals and are asked to label them (Plomp et al. 1973), confusions depend mainly on width: major 3rds are confused with minor 3rds more than with minor 6ths, although a minor 6th is an intervallic inversion of a major 3rd. Most of the incorrect responses in Plomp's study were a semitone away from target responses, but there were also occasional confusions between perfect 4ths and 5ths and rarer confusions between major 2nds and minor 7ths. Results were similar for pure and complex tones. Although the intervals were tuned to the equally tempered scale, the results would have been similar using Just or Pythagorean tuning.

Cazden (1954) argued that "the source of the musical scale is neither Nature nor abstract numbers, it is the history of the art of music. This is the only assumption that can account for its known variation in historical time and in different cultures. Scales are doubtless influenced, but they are not determined, by certain elementary psycho-acoustic phenomena, notably the

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



special perceptual configurations of the octave and the fifth" (301). Along similar lines, Alexander Ellis (1885) commented that "the Musical Scale is not one, not 'natural,' nor even founded necessarily on the laws of the constitution of musical sound, so beautifully worked out by Helmholtz, but very diverse, very artificial, and very capricious" (as cited in Hui 2012, 125).

In principle, a scale step can lie at any point on a frequency or pitch continuum. In practice, adjacent frequencies are far enough apart to allow them to be easily and categorically distinguished in music performance and perception. When Curt Sachs (1962) wrote somewhat condescendingly of "tumbling strains," citing Australian Aboriginal song, he seemed to be referring to melodies in which scale steps are defined relatively approximately, or in which there are relatively few scale steps per octave: "The most fascinating of the oldest melody patterns may be described as a 'tumbling strain.' Its character is wild and violent: after a leap up to the highest available note in screaming fortissimo, the voice rattles down by jumps or steps or glides to a pianissimo respite on a couple of the lowest, almost inaudible notes; then, in a mighty leap, it resumes the highest note to repeat this cascade as often as necessary" (51). But approximate tuning of scale steps is the rule—not the exception. The question is not whether the scale steps are approximate but *how* approximate they are. The inherent uncertainty in the tuning of scale steps is a musical universal that is based on the perceptual universal of JNDs and their context dependencies.

Ellis (1885) documented variations in scale tuning in different musical cultures. His conclusions undermined Western thinking about scales: "Ellis had found that pitch systems did not emerge from acoustic principles occurring naturally in the world. Instead, they were artifices fashioned diversely from place to place through direct human intervention and choice" (Stock 2007, 308). Whereas in the nineteenth century many agreed with Ellis's rejection of ratio theory, his prescient message was unfortunately not understood by all contemporary theorists of Western music, whether in humanities (e.g., Riemann) or sciences (e.g., Helmholtz). Many continued to theorize about scales and intervals using number ratios.

The inherently approximate nature of musical intervals and scale steps can be confirmed in a thought experiment. Imagine a sound recording of any music in the world that is comprised of tones that musicians or musicologists agree belong to scales. Measure the frequency of each tone. That is not as easy as it sounds—the tones have to be separated from the sonic mix,

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



and the measurement needs to be made in a temporal window that starts a certain time after the onset and has a certain duration. Since the perceived pitch depends on several harmonics (see chapter 13), inharmonicity can be a problem. In any case, the measurement should be done in a consistent way, taking advantage of the empirical methods and theories of psychoacoustics. If the musicians whose tone frequencies are being measured are free to change their intonation during performance, making small adjustments to fundamental frequencies by ear, then the measured frequencies of each tone, as it occurs in different musical contexts, will typically vary more than the physical uncertainty of the measurement procedure. Whereas that physical uncertainty might be a few cents, the musical variation (changes in the frequency of each scale step as it recurs) will typically be a few tens of cents. That uncertainty depends on two main factors that are difficult to separate: quasi-random variation (including limited ability to intone accurately) and deliberate variation. The existence of expressive intonation is evidence that scale-step frequencies or pitches are inherently imprecise.

The tuning of scales in oral-aural tradition follows different principles that can provide clues about how scales emerge and develop. One principle applies to more stable anchor tones if they span consonant intervals such as 4ths and 5ths (diatonic principle). Another principle applies to the tones between the anchors if they tend toward equal spacing (equitonic principle). If both principles apply, we may speak of a hybrid scale. The same principles show how musical scales might be hierarchically organized—similar to the hierarchical organization of MmT.

Rytis Ambrazevičius and colleagues (2015) found that "the equidistant principle of scale composition interacts with the constraint of consonances" (23): "[In] 'framework' scales, a rough, loosely intoned equitonics is anchored in a frame of consonances (mostly fifths or fourths). Equitonics can also be specific to 'dissonant' cultures, with modifications according to the psychoacoustic requirements of sonority, roughness, etc." (23–24). They offered two contrasting examples:

> Susanne Fürniss (2000) noticed that the scales of polyphonic (four-part) vocal performances of Aka pygmies are based on consonances of a polyphonic framework. They form steady trichordal reference points while the intermediate scale degrees are intoned more freely (see also Arom and Voisin 1998, 268). (23)

> The śruti system indeed exists, but it is not totally based on "simple ratios." Krishnaswamy (2003, 13) concludes that a hybrid scheme most likely appears in

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



> Carnatic instrumental music. The tuning of the strongest consonances in characteristic accords shows some resemblance to simple ratios such as 3/2 or 4/3, but this is not valid for the other tuning procedures. (32)

These arguments are examples of psychohistoric theorizing. The sizes of intervals and their musical functions are held to depend on both psychology and history. To understand what a musical interval is and how it works in a musical context, we need to combine historical and psychological approaches. Psychological theory and experimentation can shed light on how intervals are perceived and performed today, but the dependence of interval sizes on musical context may be better explained in a historical approach that considers the historical (or prehistoric) development of musical structure and tuning/intonation. In such an approach, problematic theoretical or mathematical issues of tuning disappear, and theories based on ratios become obsolete. Tuning and intonation instead become practical and aesthetic matters that depend on numerous partly contradictory constraints.

## The Pythagoreans

No account of the nature and origins of intervals and scales would be complete without reference to the Pythagoreans—the ancient Greek followers of Pythagoras, a mysterious figure about whom little is directly known. The Pythagoreans believed in the music of the spheres (or harmony of the spheres), according to which both musical intervals and the movements of the planets were determined by ratios of small integers. In some interpretations, the planets were held to produce an inaudible, spiritual form of music called *musica universalis* that represents a causal connection between the universe and the human soul. The mathematical mysticism of the Pythagoreans, like that of Empedocles (fifth century BC; Kingsley 1995), prevailed for millennia.

The Pythagoreans believed that musical intervals *are* ratios in the sense of ontology (being, existence, reality)—not merely that they *correspond* to ratios, as many believe today. They confirmed their belief empirically by subdividing a vibrating string (the monochord) into equal parts. Halving the length raised the pitch by an octave. Dividing it by three raised the pitch by an octave plus a 5th, and so on. In a Pythagorean approach, musical interval ratios are examples of Platonic forms: they represent what intervals are, or should be, ideally.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



We should be careful not to take the Pythagoreans too seriously. Thomas Britz (2022) described them as "a mystic cult of mathematicians who saw many numbers as having mystical, philosophical and even ethical significance." Along similar lines, Carl Huffman (2018) explained that:

> Pythagoras presented a cosmos that was structured according to moral principles and significant numerical relationships and may have been akin to conceptions of the cosmos found in Platonic myths . . . the planets were seen as instruments of divine vengeance ("the hounds of Persephone"), the sun and moon are the isles of the blessed where we may go, if we live a good life, while thunder functioned to frighten the souls being punished in Tartarus. The heavenly bodies also appear to have moved in accordance with the mathematical ratios that govern the concordant musical intervals in order to produce a music of the heavens, which in the later tradition developed into the harmony of the spheres.

Cazden (1954) was similarly disparaging: "Musicians are not generally aware that the 'law of Nature' doctrines, from Pythagoras to our day, while cast in a seeming scientific guise, do not result from or lead to a scientific approach to music or to Nature at all. These doctrines represent in reality a species of mysticism dealing with magic numbers. In their current forms, the Nature theories represent a retreat to mediaeval, pre-scientific modes of speculation and not a proper philosophy and method of science" (289).

The first to write extensively about the Pythagorean movement was the more down-to-earth Aristoxenus (Zhmud 2012). He tended toward what we might today call ecological materialism, with a focus on how real people create and perceive real music (although he also contributed to musical ratio theory). His proposal that the ear (not mathematics) is the ultimate judge of good or bad tuning has repeatedly been supported in modern research on intonation in music performance. Looking back, the music theory of Aristoxenus seems closer to the modern humanities and sciences than that of the Pythagoreans.

But Pythagorean thought was persistent. A millennium later, the sixth-century philosopher Boethius, guided by an immaterial but omnipresent Christian God, still regarded music of the spheres as the real music and human music as a mere approximation. Another millennium later, the scientific revolution led to the discovery of spectral analysis, after which many understood musical intervals not only as ratios of string lengths, but also as frequency ratios: ratios between the fundamental frequencies of harmonic complex tones.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



## Just versus Pythagorean

Most ratio theory in the Western history of music theory can be reduced to two approaches: Pythagorean and Just (or pure). The terminology is ambiguous: both Just and Pythagorean approaches are Pythagorean in the sense that both regard intervals as ratios of small integers. The Pythagorean approach is more characteristic of ancient Greece, whereas the Just approach was more common in the Renaissance.

Pythagorean intervals are combinations of octaves (ratio 2:1) and perfect 5ths (3:2). A Pythagorean major 3rd is four perfect 5ths minus two octaves or a ratio of $(3/2)^4/2^2 = 81/64$. That's 4.08 semitones or 408 cents—relative to 12-EDO, the tuning of modern keyboards (neglecting octave stretch for the moment). In Just tuning, major 3rds have a ratio of 5:4 or 3.86 semitones—22 cents narrower than Pythagorean. In short, Pythagorean major 3rds are slightly wider than four semitones in 12-EDO, and Just major 3rds are slightly narrower.

In music performance, it is not generally possible to distinguish between Just and Pythagorean versions of the same interval. Intonation typically varies over a range that includes both—even at high levels of musicianship (cf. Devaney et al. 2011; Duke 1985; Karrick 1998; Loosen 1993; Mason 1960; Morrison 2000; O'Keefe 1975; Sundberg 1982). Within a section of a high-level student choir (i.e., among singers singing the same note), the standard deviation in fundamental frequency under ideal conditions can be as small as 10–15 cents (Ternström and Sundberg 1988).

There is no evidence for the musical existence of two different ratios for the major-3rd interval (Just versus Pythagorean) or for a bimodal distribution in which major 3rds tend toward one or the other of the two ratios. There may, however, be a weak and inconsistent tendency to reduce the size of major 3rds in slow music with long non-vibrato tones (to reduce perceived roughness) and an opposite tendency in fast music (perhaps to clarify tonal functions or movement implications).

The empirical findings suggest that the real major 3rd is neither Just nor Pythagorean. Neither ratio is consistently preferred relative to four semitones in 12-EDO, which lies between them. Confronted by these observations, a ratio theorist might argue that idealized ratios are distorted when transformed in music performance—for reasons to do with physical constraints, musical contexts, or expression. Such distortions exist, of course—but relative to

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



which of the two ratios? The question seems to be unanswerable. So, a theory of that kind is unfalsifiable (Parncutt and Hair 2018).

Pythagorean tuning allows for every tone in a major or standard diatonic scale to be exactly defined and, on that basis, any other tone in the chromatic scale—at least in an abstract world of pure mathematics. Enharmonic equivalents are tuned differently, sharps being sharper than enharmonically equivalent flats (C♯ is sharper than D♭). Some theorists have proposed that Pythagorean tuning is ideal for Gregorian chant on the assumption that diatonic scales are created most simply by combining perfect octaves and 5ths or that Pythagorean tuning derives from ancient Greek theory (Barbour 2004; Bower 2002; Krahenbuehl and Schmidt 1962). But to my knowledge, there is no empirical evidence that chant was or is performed closer to Pythagorean than to Just.

Renaissance theorists such as Zarlino preferred Just intonation, which is limited to factors of two, three, and five: all intervals are combinations of octaves (2:1), 5ths (3:2), and major 3rds (5:4). Contrary to the Pythagorean system, accidental sharps in Just are flatter than enharmonically equivalent flats. Again, there is no empirical evidence, to my knowledge, that Renaissance polyphony was or is intoned closer to Just than to Pythagorean.

It is not mathematically possible to tune all intervals between all tones of a diatonic scale within the Just system, for example when tuning the harpsichord (Sukljan 2018). At least one interval must be tuned non-Just. The major-2nd interval has a ratio of 8:9, and the perfect 5th has a ratio of 2:3. If we create a major 6th by stacking a major 2nd on a perfect 5th, the resultant ratio is 16:27, which is not Just (a Just major 6th is 3:5). On the modern piano, if the major 2nd between C and D is tuned 8:9, and the major 6th between C and A is tuned 3:5 (the intervallic inversion of the minor 3rd, which is 5:6), then the perfect 5th between D and A is not 3:2 (Zweifel 1994). Any attempt to tune Just, or at least partially Just, results in shifting scale steps, which then sound out of tune. In that sense, there is no such thing as a Just diatonic scale in musical practice, whereas Pythagorean tuning is at least theoretically possible.

Consider a piece of music that is confined to a seven-tone diatonic scale. The music can be performed with Pythagorean tuning on a regular keyboard by tuning the white keys Pythagorean and ignoring the black keys. More complex music that modulates to any of the twenty-four major and minor tonalities can be played in Pythagorean tuning on a keyboard that has split black

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



keys with two different tunings for enharmonic equivalents (A♯ sharper than B♭). It may also be necessary to split some white keys to distinguish B♯ from C, C♭ from B, E♯ from F, and F♭ from E. Double sharps and flats cause additional problems, but their frequencies are at least clearly defined.

## Beyond Ratios

Interval ratio theory has limited usefulness in a modern, empirically based psychological approach to music theory. In practice, musical intervals may correspond approximately to ratios of small integers, but they also often deviate systematically from them. Larger intervals (not only octaves) are consistently stretched relative to both theoretical ratios and 12-EDO (Rakowski 1994; Ward 1954), whereas for small intervals (mainly minor 2nds, but also to a smaller extent major 2nds and minor 3rds), the opposite tendency has been observed (Rakowski and Miśkiewicz 1995).

Simultaneous and successive intervals between harmonic complex tones may tend under some circumstances to be tuned such that some upper partials coincide, creating simple ratios (Villegas and Cohen 2010). But that is only one of several tuning criteria. The optimal size of an interval in context also depends on the stretch of large intervals, sharp intonation of soloists relative to accompaniment, anticipation of melodic movement, and individual differences in psychoacoustic pitch shifts. Most of all, musicians learn interval sizes from music itself (Kopiez 2003). Tuning standards are passed from one generation to the next "by ear" (in oral-aural tradition). In Western culture, the tuning of the piano may be acting as a psychological standard for tuning in any tonal style.

Standard pentatonic and diatonic scales are based on perfect 5th and 4th intervals. Those intervals are learned pitch distances—not ratios, although they are certainly close to 3:2 and 4:3, respectively. If listeners are sensitive to (and prefer) the consonance of perfect intervals (a culturally arbitrary criterion), the standard scales have few alternatives. These simple observations answer the question "Why are standard pentatonic and diatonic scales the most common in Western music?" without the need to consider ratios (e.g., Zweifel 1996).

The circle of 5ths is a music-theoretical structure with a long history. Ancient Greek music theorists, regarding musical intervals as simple ratios and music theory as a branch of mathematics, realized that one of the twelve

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



5ths must be mistuned to complete the circle, because $(3/2)^{12}$ is not exactly equal to an integer power of two. A whole number of ideal 5ths can never equal a whole number of ideal octaves; $(3/2)^{12} = 129.7$, but $2^7 = 128$.

That may be an interesting mathematical point, but its practical musical and psychological significance is limited. In musical practice, scales are collections of pitch categories—not exact pitches. The tuning of all intervals, including octaves and 5ths, is inherently flexible. The adjustment that piano tuners make to the perfect 5th interval to solve the problem (reducing its size from an exact 2:3 ratio of 702 cents to a tempered 5th of 700 cents) is imperceptible in musical contexts and, in that sense, musically irrelevant (Ward and Martin 1961). Joos Vos (1986) found that under musically atypical ideal conditions, purity ratings for two simultaneous harmonic complex tones spanning a perfect 5th were only affected if mistuning exceeded 2–10 cents. In real music, the variation in interval size is usually much larger. In a psychological approach that puts the experience of listeners (including performers, improvisers, and composers) in the foreground, the pitch-class circle and the chromatic scale are effectively closed and complete.

Music is not the only phenomenon that has been explained mathematically since ancient times. A well-known theory of visual aesthetics invokes the golden ratio or divine proportion ($\phi = 1.618$). The golden ratio plays an important role in certain geometric figures and is related to the Fibonacci sequence. The idea fascinated the Pythagoreans, Euclid, da Vinci, Kepler, and Euler. But architecture, paintings, and human bodies are not more beautiful if their dimensions correspond to the golden ratio (Stieger and Swami 2015). Moreover, and contrary to rumor, ancient Greek builders and architects were seldom guided by the golden ratio (Foutakis 2014). Incidentally, the ratio of the sides of DIN-A4 paper is $1{:}\sqrt{2} = 1{:}1.41$, which is much smaller than the golden ratio—but that does not make A4 paper any less beautiful, and US letter is not even less beautiful because it lies even further from the golden ratio (1:1.29).

## Stretched Octaves

The width of an octave is not only inherently approximate, but also stretched or enlarged. If it is not stretched, it may sound out of tune. If all octaves are played exactly 2:1 in a piece of music, very low tones sound sharp, and very high tones sound flat. In the violin concerto repertoire, high natural

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



harmonics sound flat, but when physically measured, they tend to be sharp relative to 2:1 (Guettler 2002). In listening experiments, octaves are preferred if they are 15–20 cents too wide (Dobbins and Cuddy 1982; Fransson et al. 1974; Houtsma et al. 1987, Demonstration 15).

If an octave is not a simple number ratio, no interval is a simple number ratio. But ratio theory continues to thrive, despite the clear empirical counterevidence. In one modern approach, Timothy Hubbard (2022) proposed that the Pythagorean comma can explain the stretched musical octave in the following way. The Pythagorean comma is obtained by moving around the circle of 5ths, adding together twelve 5ths (by multiplying their ratios) and subtracting seven octaves: $(3/2)^{12}/2^7 = 3^{12}/2^{19} = 531441/524288 \approx 24$ cent $\approx 1/4$ semitone. That corresponds approximately to empirical measures of octave stretch in music (Dobbins and Cuddy 1982), suggesting a link between the two. Hubbard explained the link in terms of "a desire to return to the tonic or other central pitch" or be "consistent with the idea of movement through pitch space" (677). He implied that when listeners perceive an octave interval in music, they imagine (or somehow cognitively process) a twelve-step (!) voyage around the circle of 5ths, in which each 5th is tuned to within 2 cents of 3:2, and each octave is tuned to within 2 cents of 2:1.

If Hubbard was wrong, how else can octave stretch be explained? We need to think about the origin of the musical octave from an ecological-evolutionary or psychophysiological viewpoint. An important function of hearing is speech processing. The auditory system is constantly responding to voiced speech sounds—harmonic complex tones. The octave is the perceived distance between the 1st and 2nd harmonics (or the 2nd and 4th, or the 3rd and 6th). Perception of octaves within harmonic complex tones is enabled by the basilar membrane in the inner ear, but the basilar membrane is not a perfect spectral analyzer. The information it sends down the auditory nerve is slightly distorted due to the enormous frequency and amplitude range of hearing. As a result, the perceived octave between the 1st and 2nd harmonics of a typical harmonic complex tone tends to be slightly stretched (Terhardt et al. 1982b). The stretch can be explained in either a spectral or a temporal approach (Hartmann 1993); in both cases, it involves the neurophysiology of the auditory periphery.

An octave is a musical phenomenon—not a mathematical one. In a circular operational definition, an octave is the pitch distance or frequency ratio between musical tones that are perceived to lie an octave apart. That being

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



the case, we can establish the size of an octave by measuring musical sound. Empirical findings suggest that the perceived octave between the 1st and 2nd harmonics of a single complex tone is stretched relative to a 2:1 ratio due to pitch shifts by roughly the same amount as a typical performed octave in music—suggesting the two have a common origin, or the perceived octave represents the origin of the performed octave.

There is also a physical (acoustical) reason for musical octave stretch. Octaves on the piano (or any instrument with freely vibrating strings such as the guitar) are physically stretched due to the stretching of intervals between partials. The effective length of a piano string is different for each partial, because the string is not perfectly elastic: higher partials cause the string to bend more at the bridge. When piano tuners line up the 2nd partial of a lower tone with the 1st partial of a higher tone, the result is a slightly stretched octave (Giordano 2015). In such cases, a stretch of 20 cents is typical, but the stretch is smaller in the central range and larger at low and high frequencies. A piano whose octaves are not stretched in this way sounds clearly out of tune (D. W. Martin and Ward 1961).

## 12-EDO

The standard against which tuning in Western music is most appropriately measured in practice is neither Just nor Pythagorean. It is a division of the octave into twelve equal semitones (12-EDO), such that each semitone corresponds approximately to a ratio of the twelfth root of two ($1:\sqrt[12]{2}$) or about 1:1.06 (a 6-percent change in frequency). In addition, the octave should be slightly stretched, causing all the semitones to be stretched by the same amount, but that stretch is not normally included in the definition of 12-EDO. Based on what we know about intonation in real music, 12-EDO is the most realistic and convenient reference for measuring the size of intervals.

In music performance, the size of an interval can vary considerably. Consider the most common interval between successive tones: the major 2nd. Ancient theorists argued over whether its ratio was 8:9 or 9:10, but in performance—even by the best modern ensembles with the most sensitive tuning—tuning varies over a range that includes both these ratios. Moreover, nothing special happens, psychologically or musically, in the immediate vicinity of the ratios themselves. Like all intervals, major 2nds can be relatively large or relatively small—nothing more. In everyday musical contexts,

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



the interval of a minor 2nd can vary across a relatively wide range—say, 0.5–1.5 semitones—and still be recognized as a minor 2nd (cf. Devaney et al. 2011). Given the inherent uncertainty in the tuning and intonation of all scales, Western music being no exception, it is appropriate to define the chromatic scale as approximate 12-EDO.

Why is the piano keyboard based on 12-EDO and not 13-EDO, 11-EDO, or another number? That is one of the few questions that ratio theory can clearly and plausibly answer. In separate studies, Ramon Fuller (1991), Dirk de Klerk (1979), and Sebastian Von Hoerner (1976) showed mathematically that certain equal divisions of the octave come closer to Just intervals than others. The most promising divisions have five, seven, twelve, nineteen, thirty-one, forty-three, or fifty-three equal parts. To understand why MmT is based on 12-EDO and not one of the other numbers in this list, we need to think about how real music works:

- Two musical tones that are a semitone (100 cents) apart are usually perceived as clearly and categorically different in pitch—even in relatively complex musical contexts and considering typical limitations on the intonation of the voice and melodic instruments. In the psychology laboratory, under ideal conditions and using trained listeners, the JND for the frequency of successive pure tones in the central range can be as small as 3 cents, but the smallest microtonal intervals in musical use are in the range 33–50 cents (for a literature review, see Parncutt and Cohen 1995).
- The number twelve evidently approaches the largest the number of pitches per octave that an expert musician can hold in imagination in typical musical contexts—a kind of upper limit of human cognitive processing capacity. There is no clear evidence for this assertion, but it is interesting to compare George Miller (1956).
- The difference between Just tuning and 12-EDO for musically important intervals such as major 3rds is small enough to lie within normal ranges of intonation in vocal or instrumental performance.
- The twelve-tone scale enables a large number of scalar subsets to be created and enumerated in pitch-class set theory (Forte 1973). Those subsets are usually unequally spaced, which helps listeners to get their bearings relative to tonal references or tonics (Trehub et al. 1999).

If 12-EDO is an appropriate standard against which tuning and temperament can be measured, it follows that Just and Pythagorean tuning are not.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



That casts a shadow of doubt over the word "temperament" itself. The word suggests mistuning relative to simple frequency ratios, on the assumption that those simple ratios are standard or ideal. But neither Just nor Pythagorean tuning can be considered standard or ideal if there is no evidence that musicians consistently prefer one over the other.

The idea of measuring interval sizes relative to 12-EDO was introduced to the discipline of comparative musicology (an approach to musicology based on intercultural comparison—in particular, comparison to Western music) by Alexander Ellis in the colonialist nineteenth century. Ellis proposed measuring musical intervals in cents, or hundredths of a semitone. His project to measure the scales of different cultures systematically was promising but only partly successful. Today, despite the efforts of Ellis and his followers, we cannot draw upon a systematic collection of common non-Western tunings because of the wide range of tunings of any given interval in any given cultural context. Western music is typical in that regard: operatic singers depart from 12-EDO by a wide margin in both vibrato (often plus/minus a semitone) and deviations between the center of the vibrato (perceived as the intended pitch) and the intonation of the accompaniment (Prame 1997).

In an innovative attempt to address this issue, Simha Arom and Frédéric Voisin (1998) had musicians in Central Africa and Java tune an electronic keyboard with natural-sounding timbres to find out what intonations they intended or preferred—what interval sizes were stored in their minds. To my knowledge, this promising technique has not yet been applied systematically to musical styles in different geographical regions—possibility due to ethical considerations or simply intonational variability. Arom nevertheless claimed to have measured interval sizes to within 10–20 cents.

Today, most Western music is played in or near 12-EDO. That includes music written long before the idea of 12-EDO crystallized in the early eighteenth century with the first book of J. S. Bach's *Wohltemperiertes Klavier*. Even then, Bach was not promoting exactly equal semitones; *wohltemperiert* is not the same as *gleichstufig temperiert* (Rasch 2011). Anyway, the idea was not new in the eighteenth century: much sixteenth-century lute music was performed in something approaching 12-EDO (Ozmo 2016). The extent to which Bach's clavichord might have been (almost) equally tempered in the modern sense is unclear, but we do know from the music itself that all twenty-four major and minor keys were eminently usable, even if there were slight tuning variations from one key to another (Rasch 2011). Small tuning variations of that

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



kind are not relevant for the theory presented here. Quarter-tone tuning in Middle Eastern musical traditions (Persian *dastgah*; Arabic *maqam*) can be similarly approximate.

One can argue that the chromatic scale, defined as approximate 12-EDO, is much older than Bach. The division of the perfect 4th into two whole steps and one half step in ancient Greek theory is consistent with 12-EDO. During the next few centuries, Chinese tuning theory independently developed something close to 12-EDO based on the circle of 5ths (McClain 1979). The ancient Greek tetrachord was four tones, spanning a perfect 4th interval or 3:4 ratio. In the diatonic genus, and starting from the higher tone of the 4th and moving downward, the tetrachord comprised two whole-tone intervals (understood as 8:9 or 204 cents each) and one semitone (243:256 or 90 cents). These three intervals multiply to 3:4 and add to 498 cents (cf. Barbera 1977; Gollin 2004). Although the 243:256 ratio was smaller than that of the semitone in today's 12-EDO and corresponds to Pythagorean tuning (producing slightly sharp leading tones), it still lies well within the usual range of tuning in today's music performance. From that, we may guess that a lot of ancient Greek music practice conformed approximately to 12-EDO—similar to MmT today.

## Octave Equivalence

The music-theoretical principle of octave equivalence says that tones an octave apart have the same harmonic function—at least from the point of view of Western tonality. But octave equivalence is not a musical universal, as is often claimed:

- In Indonesian gamelan tunings, even if octave equivalence seems to apply, the octaves may be stretched more than Western octaves (Carterette and Kendall 1994)—possibly due to the familiarity of gamelan players with the nonharmonic spectra of gamelan gongs (McLachlan, Marco, and Wilson 2013).
- Some Australian indigenous melodies seem to lack octaves (Will 1997), but that does not mean the performers or listeners are insensitive to them. Instead, octaves may simply be unimportant or irrelevant in this musical style. Another possibility is that the category boundaries for the perception of musical pitches and intervals are different in Australian indigenous

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



and Western music, causing musically trained Western listeners to get the impression that the music lacks octaves.

- Nori Jacoby and colleagues (2019) remarked that "sung correlates of octave equivalence are undetectable in Amazonians, suggesting effects of culture-specific experience . . . octave equivalence may be culturally contingent, plausibly dependent on pitch representations that develop from experience with particular musical systems" (3229). Specifically, "it is possible that perceptual octave equivalence only emerges in the presence of an octave-based musical system with a large melodic range" (3241).

Even if octave equivalence is not a musical universal, sensitivity to octaves can be considered a perceptual universal, given the importance of octaves in the harmonic series that every human perceives within voiced speech sounds when tracking the fundamental frequency of speech. Laurent Demany and Françoise Armand (1984) found on the basis of a careful empirical study that "two pure tones forming an octave interval have some degree of perceptual equivalence for three-month-old human infants" (63). However, their finding could be accounted for by familiarity either with the harmonic series within voiced speech sounds or with Western music, both of which can be heard both before and after birth. The perceptual octave effect can be quite weak; Howard Kallman (1982) presented successive pure tones spanning chromatic intervals to musicians and asked them to rate similarity. He observed almost no effect at the octave. Responses depended more on interval size.

Despite the ambiguity of the empirical findings, it is reasonable to assume that everyone in the world is capable of perceiving the similarity of successive harmonic complex tones spanning octave intervals, in which the audible harmonics of the first tone line up with those of the next, by comparison to mistuned octaves—even if octave equivalence plays no role in their music. Regardless of cultural background, every human has a basilar membrane that enables a running spectral analysis. Similarly, everyone in the world is capable of perceiving the consonance (harmonicity and/or smoothness) of simultaneous octaves by comparison to slightly larger or smaller intervals (cf. Plomp and Levelt 1965). The ability to recognize harmonicity in running spectra logically includes the ability to recognize octaves (frequency ratios of approximately 2:1) as the most common and hence most important indicator of spectral harmonicity. All of that is presumably true across cultures, even if octave relationships play no role in the music of a given culture or if they play a different role.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



Octave equivalence in Western and other musics may be considered psychohistoric in the sense that neither psychology nor history alone can account for it, but only a combination of the two. Octave equivalence depends not only on perceptual properties of tones and intervals as we perceive them (the psychological aspect), but also on the musical contexts in which octave intervals occur and the history of perceiving and performing octaves in diverse musical contexts. Octave equivalence may be learned from exposure to octave intervals within the spectra of individual harmonic complex tones in speech, music, and everyday sounds, in parallel with harmonic fusion (Demany et al. 2021).

In Western music theory, tones an octave apart are equivalent, but tones a 5th apart are categorically different. The distinction may be clearer for musicians who can easily recognize pitch patterns due to their musical training, and their understanding of music theory and music notation in which octave equivalence plays an important role. Experiments on similarity judgments of musical tones at different intervals (e.g., Parncutt 1989) do not imply a categorical difference between octaves and 5ths for nonmusicians.

## Perfect Intervals

In the history of Western music theory, octaves, 5ths, and 4ths between simultaneous and successive tones are *perfect consonances*. But for simultaneous tones, the 4th is paradoxically often regarded as a dissonance. When the 4th occurs within a chord and the lower tone of the 4th is in the bass, the chord is dissonant and requires special treatment or resolution (e.g., a major or minor triad in 2nd inversion).

Despite the ambiguous C/D of the simultaneous perfect 4th interval, melodic perfect 4ths and 5ths evidently played an important role in the prehistory of standard pentatonic and diatonic scales. In oral tradition, intonation is learned from performance and passed from one musician or generation to the next. Quasi-diatonic scales may have emerged when consonant melodic intervals (octaves, 5ths, 4ths) came to be preferred between nonadjacent scale tones (Gauldin 1983).

A possible explanation involves pitch commonality. The pitch of a musical tone is ambiguous (Terhardt 1974a), and tones are heard to go well with each other if they have pitches in common (Parncutt 1989). Pitch commonality involves both spectral pitches (audible partials) and virtual pitches

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



(possible fundamentals; see chapter 13). Successive tones at perfect octave, 5th, and 4th intervals have both spectral and virtual pitches in common (see figures 13.1 and 13.2). The perceptual coherence of a scale may depend on the degree to which its tones are related by perfect intervals.

The idea that pitch commonality drove the prehistoric emergence of scales in oral tradition can hardly be tested. We have no recordings of ancient music; we only have recordings of traditional music that we guess was not influenced by MmT. But if the idea is valid—and it is certainly straightforward—we can apply it in an attempt to explain why some scales, understood as sets of pitch categories, are more common than other scales—at least within Western music. Familiar scales from music theory (church modes, the major scale, variants of the minor scale, the blues scale) tend to maximize the number of perfect 4th or 5th relationships. After the octave, that is the strongest harmonic relationship—the one with the highest pitch commonality between harmonic complex tones.

The principle applies in particular to standard pentatonic and diatonic scales, whose steps correspond to adjacent points on the circle of 5ths. Starting from C, the first five tones on the circle (C, G, D, A, E) make up a standard pentatonic scale, as do the five tones starting with the second element (G, D, A, E, B), and so on. The first seven tones (C, G, D, A, E, B) make a standard diatonic scale, as do the seven tones starting with the 2nd element (G, D, A, E, B, F♯), and so on. In a psychohistoric account of the origin of such scales, pentatonic and diatonic scales developed as 5ths were favored between non-adjacent tones due to their pitch commonality.

In an ecological approach, the process did not rely on musicians and listeners recognizing scales as such. Instead, they recognized melodies, which were preferred if they conformed to scales. Scales may have been recognized later at a more abstract level (perhaps by trained musicians). Recognizing a scale meant somehow understanding the pitch pattern as a whole; it also meant recognizing the more common (usually smaller) intervals within that context. The process did not necessarily involve the development of a cognitive representation of the circle of 5ths.

## Asymmetry and Evenness

The intervals of the seven-tone diatonic scale, considered as a subset of 12-EDO, are unequal (separated by whole tones and semitones). The scale cannot be transposed onto itself without changing one of the tones unless the

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



transposition is through an octave. The scale's asymmetry makes it easier for listeners to get their bearings and hear the music relative to tonal references or tonics. Whereas Sandra Trehub and colleagues (1999) suggested that humans might have an "inherent processing bias favoring unequal-step scales," one could also say in an ecological approach that it is inherently easier for any perceiving machine or organism to get oriented relative to an unequally spaced scale because the unequal interval sizes allow one to locate the tonic or other pitch reference unambiguously. For similar reasons, scalar asymmetry facilitates the learning of new melodies (Pelofi and Farbood 2021).

A related property of diatonic scales is evenness. The whole tones and semitones are distributed almost equally around the pitch-class circle (Clough and Douthett 1991). As a result, the size of most intervals corresponds roughly to the number of scale steps, give or take a semitone: intervals of a diatonic 2nd are one or two semitones, 3rds are three or four, 4ths are five or six, 5ths are six or seven, 6ths are eight or nine, and 7ths are ten or eleven (see table 3.1). Each diatonic interval has two common variants a semitone apart, which limits variation and simplifies the mapping of diatonic onto chromatic. If one interval between neighboring tones is smaller than average, the intervals on each side of it are usually larger than average, restoring evenness. Within such a diatonic scale, when pitch patterns are transposed to different starting pitches, the exact sizes of individual intervals may change, but the relationship between larger and smaller intervals stays the same.

The diatonic scale can be constructed by stacking six perfect 5th intervals and transposing back down by octaves. The result is a scale with seven tones per octave: CDEFGAB. It is interesting to ask why the scale has seven tones and not six or eight. A scale of six tones, CDEFGA, would be inconsistent: two of its minor 3rds (DF and EG) would be split (by E and F), whereas AC would not be split. It is more logical to split either all minor 3rds (producing a standard diatonic scale) or none of them (producing a standard pentatonic). Similarly, a scale of eight tones, such as CDEFGAB♭B, would be inconsistent because the minor 3rd AC would be split into three, whereas DF and EG would be split into two.

## Performed Intonation

Tuning is frequency adjustment before performance, such as the tuning of a piano or a guitar. Intonation is frequency adjustment during performance—in real time, from one tone to the next, as in a choir or string quartet. Musical

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



intonation strays in both quasi-random and systematic-intended ways from theoretically correct frequencies or pitches: various studies of Western performance suggested that intonation tends to be closer, on average, to Pythagorean or 12-EDO than to Just (e.g., Mason 1960).

Good intonation is about honing in on the familiar average sizes of musical intervals. It also involves keeping the pitch of scale degrees constant when they occur at different times and in different musical contexts, but it may mean changing those pitches if the musical style includes rule-based microtonal variations (Ahlbäck 2018). In ensemble performance, good intonation also involves responding appropriately and sensitively to intentional and unintentional changes in intonation by other performers. Good intonation is inseparable from other musical skills and other aspects of familiarity with relevant musical styles.

During the course of a performance on instruments with variable intonation (including voices), there is no exact standard of pitch or frequency. Instead, the tuning of more stable tones such as tonics tends to be more exact, and the tuning of less stable tones such as leading tones tends to be more variable (Ambrazevičius 2005). The intonation of a given scale step varies over a range that seems smaller to the listener than to the acoustician making physical measurements (cf. Devaney et al. 2011). The range of variation is greater in everyday singing by nonmusicians, in which aesthetic judgments can be remarkably independent of intonation; in that case, "interval tuning is imprecise in both production and perception" (Pfordresher and Brown 2016, 11).

Musicians and listeners intuitively apply several conflicting criteria to decide whether a pitch or an interval is "in tune" in the context of a musical performance. The criteria include:

- Context-dependent learned interval sizes. Interval sizes are not only learned from music, they can also depend on musical context (Ahlbäck 2018).
- Interval stretch (Rakowski 1994). Different people prefer different amounts of stretch, and upright pianos have more stretch than grands (Railsback 1938).
- The relationship between foreground and background (solo versus accompaniment). Solos are often intoned sharp relative to accompaniment (Kantorski 1986).

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



- Anticipation of pitch movement. Rising leading tones are intoned sharp, and falling leading tones are intoned flat. In both cases, intonation is more variable than for other scale steps (Fyk 1997). Rytis Ambrazevičius (pers. commun.) found a similar principle in Lithuanian traditional singing: Scale degrees adjacent to tonal anchors tended to gravitate toward them, an interval of about two semitones shifting toward about one-and-a-half semitones.
- Emotional expression. This can affect intonation in many ways, perhaps by imitating emotional speech, but there are also typically musical gestures, such as sliding toward target pitches (P. Johnson 2004).
- Style. Slow, sustained music without vibrato may tend in the direction of Just intonation if that makes beats less audible. Singers who think Renaissance polyphony should be sung with Just intonation may deliberately reduce the size of their major 3rds.
- Pitch shifts. In principle, all pitches in a piece of music are slightly shifted due to perceptual distortions, for which performers may compensate. Pitch shifts depend on sound level or masking and are smaller for complex than for pure tones (Terhardt 1975).
- Idiosynchronies of musical instruments. High natural harmonics on string instruments may sound flat because octaves are not stretched (or not stretched enough; Guettler 2002).
- Individual psychological and physiological differences. These apply for instance to the size of pitch shifts.

Good intonation is generally a compromise among such conflicting criteria. Consistent with that idea, Anders Friberg and colleagues (2006) developed, adjusted, and tested a set of performance rules for computer-controlled expressive performance. The rules included:

- *High sharp*: Stretch all intervals in proportion to size.
- *Melodic intonation*: Intonate according to melodic context.
- *Harmonic intonation*: Intonate according to harmonic context.
- *Mixed intonation*: Intonate using a combination of melodic and harmonic intonation.

In melodic intonation, semitones are performed smaller than in 12-EDO. In harmonic intonation, beating in chords is minimized, which sometimes means tuning in the opposite direction (toward Just rather than Pythagorean), leading to questions of mixed intonation.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



In addition to the listed criteria, intonation in musical performance is inherently uncertain. Musicians (including singers) have limited control over the pitch of their instruments, leading to quasi-random variation around approximately defined target frequencies. Eric Prame (1997) analyzed twenty-five tones sung by sopranos in each of ten commercially available recordings of Franz Schubert's *Ave Maria*. Whereas the pitch of the tones was heard near the middle of the vibrato (as confirmed by Shonle and Horan 1976), the mean extent of vibrato according to Prame was 0.7 of a semitone in both directions. So, it spanned 1.4 semitones. Larger vibratos were measured for Verdi opera arias than for Schubert. The difference between relatively sharp and relatively flat singers was remarkable: "The mean difference across artists between sharpest and flattest intonation was 54 cent" (Prame 1997, 618). Moreover, "the intonation of tones with greater vibrato extent tended to be sharper" (619).

Although the JND for the frequency of two successive pure tones, when measured in a psychoacoustic experiment and under ideal conditions, is much smaller than a semitone (typically 5 cents, with large individual differences), it is larger in complex musical contexts, especially when those contexts feature audible mistunings (often of the order of a semitone) due to the technical demands of musical sensorimotor coordination (e.g., choral singing). Because our ears are used to variation around target pitches in music, we are surprisingly tolerant of it. That can explain why the tuning variations that are revealed in physical measurements of music performance tend to be greater than the mistunings that we think we are hearing. Musicians who are courageous enough to have their own intonation measured may be surprised at how poor it is from a physical viewpoint.

But it is not the physical viewpoint that counts. What counts is perception. Besides, musicians may deliberately and audibly change the pitch of a tone during the course of the tone for expressive reasons, starting a little flat and gradually rising to a target pitch. The tonal references relative to which intonation is being adjusted may also change, from melodic intonation (where tonal references are successive) to harmonic intonation (where tonal references are simultaneous). If such adjustments are performed effectively, no one will describe the deviations as "mistuning" or complain about poor intonation.

The importance of beating (amplitude modulation from almost coinciding partials) for tuning in music performance is smaller than many musicians believe. The sound of beating is familiar from the "wa-wa" sound of a single

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



note on a badly out-of-tune or "honky-tonk" piano. If the beats are too fast to be heard individually (more than twenty per second) and slower than about 250 per second (Terhardt 1968), the sound is rough. One criterion for good intonation in music performance is to minimize or avoid audible beats and roughness, but that roughness or beating may be only perceptible if the tones are long and sustained and lack vibrato. In music, it is common either for tones to die away quickly (like piano tones do) or for their fundamental frequencies to change slightly during the course of the tone (as often happens in vocal, woodwind, brass, or bowed-string music).

## Categorical Perception

Perception of MmT involves psychological pitch and time categories that correspond roughly to notated pitches and durations. Given the variability and context dependence of intonation in music performance, the psychological categories that make up a Western musical scale are typically about a semitone wide (i.e., a central pitch plus or minus a quarter tone).

The German word *zurechthören* means to hear a tone correctly, even though it was not played correctly, or to guess which tone a musician intended to play and hear the sound as if that tone had been played. We do that all the time when we hear music that is sung or played on instruments with variable intonation. Within the range of *zurechthören*, both the performer and the audience perceive a tone as in tune; professional performers are usually more sensitive to mistuning than their audiences.

Melodies tend to focus on a small set of approximately defined pitches called scale steps, such that musical pitch is perceived in categories. They do not usually slide up and down like speech does (portamento). The ultimate origin of psychological pitch categories in music is unclear. Various possible psychohistoric scenarios might have led to their emergence. All of them are speculative, since we have no way of going back in time and performing psychological experiments:

- Musical pitch categories may enable and facilitate melody transmission in oral tradition. They may allow melodies to be memorized and passed from one generation to the next with a minimum of information content (even if the music is not written down). A modern analogy is the MIDI file, which takes up much less space on a computer hard disc than a corresponding MP3 or .WAV file. But the idea is problematic. People in oral traditions also

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a portion of the eBook at doi:10.7551/mitpress/15050.001.0001



remember and reproduce subtle variations in intonation (McLucas 2010): they remember much more than just the "notes."

- Musical pitch categories may be a consequence of the physics of musical instruments, including the finger holes of flutes or techniques to tune string instruments that involve avoiding beating. Musical instruments therefore tend to produce specific frequencies rather than pitch slides. According to this idea, pitch categories emerged in prehistory when early singers imitated instruments (E. Schubert and Wolfe 2013).
- Pitch categories may be related to category boundaries for the mistuning of harmonics within complex tones. Modern psychoacoustic experiments have shown that the spectral pitches of partials within complex tones are perceived categorically: a partial is perceived either as a harmonic of a given fundamental or as a separate sound, depending on the degree of mistuning. The ear is surprisingly tolerant of mistunings of partials within individual tones—perhaps because many tones in the physical world have almost-harmonic partials. Bells are an example: usually, some partials in the spectrum correspond to an incomplete, mistuned harmonic series, and others do not. Another example is the piano: the partials of piano tones are stretched relative to the harmonic series, the octave between the 1st and 2nd partials being perceptibly bigger than 2:1, the perfect 5th between the 2nd and 3rd bigger than 2:3, and so on. Moreover, the ear cannot hear pitch very exactly in very short sounds (due to the uncertainty principle—a hard physical limitation; see chapter 5). A harmonic partial within a harmonic complex tone can therefore be heard as part of the tone, even if it is noticeably mistuned, and the maximum mistuning typically lies between a quarter tone and a semitone (B. C. J. Moore et al. 1986). The categorical pitch perception of partials in complex tones could explain why the pitches of tones in melodies are also perceived categorically.
- We tend to categorize everything that we perceive in different ways. In everyday life, we categorize the distal stimulus (people talking, dogs, bells, trucks, the wind in the trees, rivers flowing). In music, we categorize the proximal stimulus or musical surface (the pitch, timbre, loudness, and perceived duration of voices and instruments). The two perceptual processes are evidently related. A focus on specific pitches—held and repeated during a melody—attracts our attention to the proximal stimulus and is typical of music listening.

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

This is a section of doi:10.7551/mitpress/15050.001.0001

# Psychoacoustic Foundations of Major-Minor Tonality

## By: Richard Parncutt

## Citation:

*Psychoacoustic Foundations of Major-Minor Tonality*
**By:** Richard Parncutt
**DOI:** 10.7551/mitpress/15050.001.0001
**ISBN (electronic):** 9780262377362
**Publisher:** The MIT Press
**Published:** 2024

The open access edition of this book was made possible by generous funding and support from MIT Press Direct to Open

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024

MIT Press Direct

© 2024 Massachusetts Institute of Technology

This work is subject to a Creative Commons CC-BY-ND-NC license.
Subject to such license, all rights are reserved.

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Stone Serif and Stone Sans by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data
Names: Parncutt, Richard, 1957– author.
Title: Psychoacoustic foundations of major-minor tonality / Richard Parncutt.
Description: Cambridge, Massachusetts : The MIT Press, 2024. | Includes bibliographical references and index.
Identifiers: LCCN 2023007827 (print) | LCCN 2023007828 (ebook) | ISBN 9780262547352 (paperback) | ISBN 9780262377379 (epub) | ISBN 9780262377362 (pdf)
Subjects: LCSH: Tonality. | Psychoacoustics. | Chords (Music) | Musical intervals and scales.
Classification: LCC ML3811 .P37 2024 (print) | LCC ML3811 (ebook) | DDC 781.2/6—dc23/eng/20230606
LC record available at https://lccn.loc.gov/2023007827
LC ebook record available at https://lccn.loc.gov/2023007828

Downloaded from http://direct.mit.edu/books/oa-monograph/chapter-pdf/2333018/c008300_9780262377362.pdf by guest on 08 September 2024