1
Constructing a Method for Evaluating and Comparing the Sound of Chords

Musicians and music theorists discuss music and music theory at many different levels of abstraction. A music teacher may ask a student to play an ***mf*** E♭ on an oboe. These instructions indicate that instrumentation, dynamics and pitch are important but the octave, in which the pitch falls, may not be. The teacher could also ask someone in an orchestra to play A_4. In this case, instrumentation and dynamics (within reason) do not matter, but pitch and octave are significant. Likewise, musicians may discuss a specific A♭ dominant seventh chord in the context of a passage in a particular musical score. In this case, pitch, letter names and accidentals, dynamics, and octaves matter. At other times, musicians will discuss dominant seventh chords in general, regardless of pitch, octave, letter names and accidentals, dynamics, or particular compositions.

Different levels of abstraction can reveal or hide information about a chord. Letter names and accidentals are useful in relating a chord to a scale, but can be less useful in relating the chord to other chords in general. On the other hand, while numbers can be very useful in describing aspects of chords in general, they may not very useful in rehearsal. A music teacher will not get far asking a student to play 6 on an oboe.

Here, the challenge is to find the most useful combination of different levels of abstraction – each one not too abstract and not too specific – that provides the most useful method for evaluating and comparing the sound of chords in general.

The ideal method would be of universal application. It would provide the same results for each chord, and comparable results for every chord, regardless whether: a chord is played by a group of trumpets or a several different woodwind instruments; a chord is played quietly or loudly; a chord is pitched in a high register or a low one; the pitches in the chord are within one octave or spread across several octaves; a chord is played with different articulations, such as pizzicato on string instruments or legato on brass instruments; and so on.

The ideal method would also be as simple and user-friendly as possible, as a colour wheel is for the visual arts.

Top-down and bottom-up approaches can both be useful in constructing a method.

1.1 The top-down approach

From a top-down perspective, notes can be represented "as points… on a circle (*pitch-class space*)", which is similar to a clock face [1]. All notes can be represented as specific pitches in pitch-class space in the form of letter names with accidentals (F♯/G♭, A♯/B♭ …), irrespective of which octaves they might occur in. Ignoring information about octaves is known as octave equivalence. This approach also ignores whether a pitch is an A♯ or a B♭, what instruments might perform the notes, and which dynamic levels and articulations might be used. One can also say that this approach includes enharmonic equivalence, instrumental equivalence, dynamic equivalence, and articulation equivalence. In pitch-class space, as on a normal clock face, the distance between each point on the circle is the same. Likewise, in an equal-tempered octave, the distance between each adjacent pitch is one semitone, which is equal to 100 "cents".

Notes can also be represented as *numbered points* on a twelve-point circle (*twelve-point circular space*), which constitutes a higher level of abstraction at which it is clear that musical notes are represented in a form in which they have ceased to be notes that are performed or that have any sound. At this level, the representation of notes as numbered points on a circle is a mathematical measure of the semitone distance between specific points. Notes are represented as numbers from 0 to 11 rather than by letter names and accidentals. At this level, notes have also ceased to have any relation to frequency (kHz). Each chord is represented by a set of numbered points that begins with the number 0. Unlike pitch-class space in which a single point represents a specific pitch (such as B♭), the number 0 in twelve-point circular space may represent any pitch and other numbered points that represent other pitches in that specific chord represent the distance in semitones from pitch number 0. By replacing letter names and accidentals with numbers, the size of our method is reduced by nearly 92% without the loss of any essential information.

Traditional set class theory (Forte set classes) provides a way of capturing and grouping all possible equal-temperament chords in a twelve-point circular space. Set class theory is described in detail in other publications, so it will not be reviewed in detail here. For now, we simply provide an example: set class 3-11 includes all major and minor triads, regardless whether they are in root position, first inversion or second inversion (as those terms are traditionally used in music theory).

For our purposes, traditional set class theory seems to be too abstract to be used as a classification method because nearly all composers and listeners distinguish between the sound of major and minor triads. So, a slightly less abstract classification method would appear to be more appropriate.

The approach selected here is to use transpositional (Tn) classification to modify traditional set class theory: numbered-point sets that can be transformed into each other by transposition are members of one set-class. Transpositional set classes are numbered in increasing order: 2-1, 2-2, 2-3, and so on. These set classes do not include chords that are related by "inversion through reflection", which is defined for this purpose to mean a transformation in which every numbered point (x) of a numbered-point set is replaced by its inverse, calculated as 12 – x. Each Forte set class, which contains chords that are related by inversion through reflection, is split into two separate set classes, and the Forte set-class name is also split into two; the name of each of these transpositional set classes has the letter A or B added to it [2].

For example, set class 3-11 is divided into: 3-11A (which consists of the minor triad in root position, first inversion and second inversion); and 3-11B (which consists of the major triad in root position, first inversion and second inversion). Note that the term "first inversion" in this example has a meaning similar to that used traditionally in music theory to signify a chord where the traditional "root" (the lowest pitch below a stack of major and minor thirds) has been transposed up by at least one octave. Likewise, the term "second inversion" means a chord where the root and the next lowest pitch have both been transposed up by at least one octave. Here, the term "inversion" is applied to all chords and is not limited to chords constructed with a stack of major and minor thirds. Mathematically, these inversions are rotations in twelve-point circular space. In this transpositional twelve-point circular space, all semitone measurements are made in a clockwise direction. In this book, the words "inversion", "root position" ,"first inversion" and "second inversion" refer to rotations in twelve-point circular space; all references to "inversion through reflection" specifically use that term.

The transpositional set classes are set out in the form "TSC*-*" (for example, TSC2-1, TSC2-2, TSC2-3, TSC3-11A, and TSC3-11B).

As an example of this transpositional twelve-point circular space, every major triad in root position is abstracted to the same numbered point set (0-4-7).

In Figure 1.1 below, major triads in root position (0-4-7) are represented in transpositional twelve-point circular space. Rotating the dot pattern counter-clockwise, so that pitch 4 becomes pitch 0, shows the first inversion major triads (0-3-8). Rotating the pattern counter-clockwise again, so that pitch 7 becomes pitch 0, produces the second inversion major triads (0-5-9). These three numbered-point sets represent the three chords contained in the transpositional set class TSC3-11B. In twelve-point circular space, the number 0, when used in different sets of numbered points to represent the inversion of chords within one transpositional set class, represents a different pitch in each inversion. For example, (0-4-7), (0-3-8) and (0-5-9) may represent (C-E-G), (E-G-C) and (G-C-E). Depending on the inversion, the numbered point 0 would represent C, E or G.

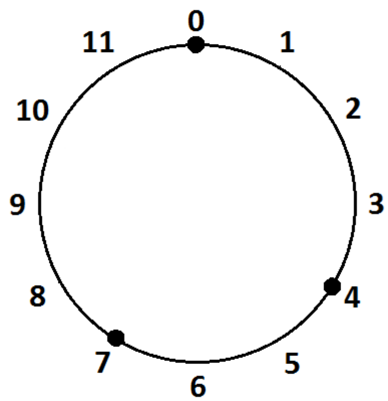

Figure 1.1 Major triads in root position (0-4-7).

In Figure 1.2 below, major triads in root position (0-4-7) are represented in transpositional twelve-point circular space as solid black dots. Numbered points 4 and 7 are reflected through the 0-6 axis, (numbered points 8 and 5 are represented as small circles), producing the second inversion minor triads (0-5-8). This is an example of inversion through reflection through the 0-6 axis.

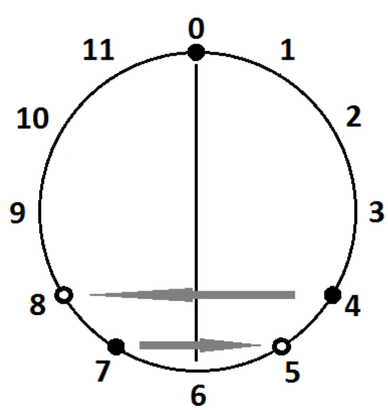

Figure 1.2 Inversion through reflection of major triads in root position (0-4-7) into second inversion minor triads (0-5-8).

In Figure 1.3 below, second inversion minor triads (0-5-8) are represented in transpositional twelve-point circular space. Rotating the dot pattern counter-clockwise, so that numbered point 5 becomes numbered point 0, shows the root position minor triads (0-3-7). Rotating the pattern counter-clockwise again, so that numbered point 8 becomes numbered point 0, produces the first inversion minor triads (0-4-9). These three numbered-point sets represent the three chords contained in the transpositional set class TSC3-11A.

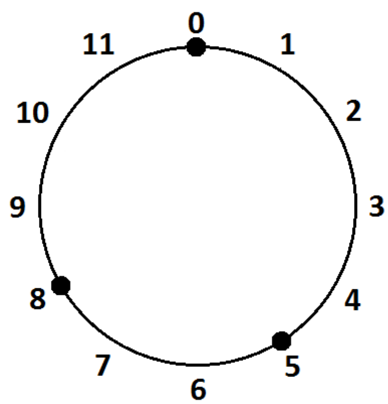

Figure 1.3 Second inversion minor triads (0-5-8).

So far, the top-down approach has provided us with a framework that can accommodate all equal-tempered chords in a transpositional twelve-point circular space. Chords are represented by the distance between numbered points, measured in semitones in a clockwise direction.

The classification of chords as transpositional set classes results in one single pitch represented in 1 TSC1 set class, 11 intervals in 6 TSC2 set classes, 55 triads and trichords in 19 TSC3 set classes, and 165 seventh chords and tetrachords in 43 TSC4 set classes.

However, this framework is completely free of sound. Musically, it is completely sterile.

It is clear that a different approach is needed to take us further. It is necessary to consider how chords are actually performed and perceived.

1.2 The bottom-up approach

From a bottom-up perspective, our approach draws on research that has been carried out by Ernst Terhardt, Dale Purves, and others.

The relationship between the sound of the human voice and the sound of musical instruments, such as strings and woodwinds, is essential to understanding the sound of chords made by singers and by instruments that produce vibrating strings or columns of air.

The following quote from Daniel Bowling and Dale Purves indicates the essential relationship between the sound of musical tones and chords, human vocalization, and the harmonic series.

> "Most natural sounds, such as those generated by forces like wind, moving water, or the movements of predators or prey, have little or no periodicity. When periodic sounds do occur in nature, they are almost always sound signals produced by animals for social communication. Although many periodic animal sounds occur in the human auditory environment, the most biologically important for our species are those produced by other humans […].
>
> Like most musical tones, vocalizations are harmonic. As with strings, vocal fold vibration generates sounds with a fundamental frequency and harmonic overtones at integer multiples of the fundamental. The presence of a harmonic series is thus characteristic of the sound signals that define human social life, attracting attention and processing by neural circuitry that responds with special efficiency. With respect to music, these facts suggest that our attraction to harmonic tones and tone combinations derives in part from their relative similarity to human vocalization. To the extent that our appreciation of tonal sounds has been shaped by the benefits of responding to conspecific vocalization, it follows that the more voice-like a tone combination is, the more we should "like" it." [3]

These relationships are integral to our approach to constructing a method for evaluating and comparing the sound of chords.

So, from a bottom-up perspective, we start with chords as they are actually performed and perceived and then abstract away as much unessential information as possible [4]. Accordingly, we modify the framework developed using the top-down approach.

Among the octaves within which humans normally speak and sing and most instrumental pitches are performed, we recognize the distinctive sound of a major triad in root position regardless of the octave or octaves in which the pitches are performed. Likewise, we recognize that sound regardless of the dynamic levels and articulations used in the performance. Using the terminology of psychoacoustics, we also recognize that distinctive sound regardless of sound pressure levels, critical bands, roughness, and pitch shifts. So, all of these variables can be safely discarded.

We also recognize the distinctive sound of a major triad in root position regardless of instrumentation or timbre. We recognize that sound whether it is played on a piano; played by a trio of string, woodwind or brass instruments; or sung by tenor, alto and soprano voices. This degree of universality suggests that voice, instrumentation and timbre can also be discarded. In all cases, some elements of timbre can be safely discarded, such as the "envelope" (for example, onset time, attack, sustain, spectral evolution, and decay). However, as we discuss further below, there is an inherent need to retain some information about voice, instrumentation and timbre (such as the frequency and amplitude of the harmonics of each performed note).

As it turns out, for our first dimension of the sound of chords – harmonic similarity of a performed chord – we can discard specific data about voice, instrumentation and timbre.

1.2.1 The harmonic similarity of a performed chord

Kamraan Gill and Dale Purves sought to determine whether the scales preferred in music worldwide and throughout history had an overall similarity to the spectral characteristics of a harmonic series [5]. They noted that, although humans can distinguish about 240 different pitches over an octave in the mid-range of hearing, they limited potential scale tones to 60 pitches to reduce the computational burden. They analysed 455,126 possible pentatonic scales, 45,057,474 possible heptatonic scales and a random sample of 10 million possible dodecatonic scales. Their research showed that the most widely used scales were comprised of intervals with the greatest overall spectral similarity to the harmonic series.

The following quote from the research of Gill and Purves indicates the relationship between human speech and the frequencies of harmonics in the harmonic series. In this approach, timbre (which includes the relative amplitudes of harmonics) is expressly ignored. Implicitly, octave equivalence and dynamic equivalence apply within the normal range of human speech.

> "[…] we favor a biologically based preference for harmonic series as the most plausible explanation for the particular scales used to make music over history and across cultures.
>
> Like any other sensory quality, the human ability to perceive tonal (i.e., periodically repeating) sound stimuli has presumably evolved because of its biological utility. In nature, such sound stimuli typically occur as harmonic series produced by objects that resonate when acted on by a force. Such resonances occur when, for example, wind or water forces air through a blowhole or some other accidental configuration, but are most commonly produced by animal species that have evolved to produce periodic sounds for social communication and ultimately reproductive success (e.g., the sounds of stridulating insects, the vibrations produced by the songbird syrinx, and the vocalizations of many mammals). Although

> all these harmonic stimuli are present in the human auditory environment, the vocalizations of other humans are presumably the most biologically relevant and frequently experienced.
>
> In humans, vocal stimuli arise in a variety of complex ways, not all of which are harmonic. Harmonic series depend on vocal fold vibrations and are characteristic of the ''voiced speech'' responsible for vowel sounds and some consonants. Although the relative amplitudes of harmonics are altered by filtering effects of the supralaryngeal vocal tract resonances to produce different vowel phones, the frequencies of harmonics remain unchanged. In consequence, the presence of a harmonic series is a salient feature of human vocalizations and essential to human speech and language. It follows that the similarity of musical intervals to harmonic series provides a plausible biological basis for the worldwide human preference for a relatively small number of musical scales defined by their overall similarity to a harmonic series." [6]

Purves, together with Deborah Ross and Jonathan Choi, indicated in another article that formant ratios in human speech are more closely related to just intonation than to equal temperament.

> "Ten of the 12 intervals generated by the analysis of either English or Mandarin vowel spectra are those used in just intonation, whereas 4 of the 12 match the Pythagorean tuning and only 1 of the 12 intervals matches those used in equal temperament." [7].

The methodology used by Gill and Purves is best described in their own words below. It is clear from their description that frequency ratios are a critical aspect of their method, which encompasses the frequencies of fundamental pitches in a scale as well as their spectral pitches and virtual pitches. They also consider each scale in isolation, unfettered by the concept of tonality, which would favour certain intervals over others; all intervals are treated equally.

> "The […] approach we take here is to quantitatively compare the harmonic structure that defines each interval in a possible scale to a harmonic series […] Accordingly our analysis does not depend on intervals and scales precisely mimicking a harmonic series, but evaluates degrees of similarity. The average similarity of all intervals in the scale is then used as a measure of the overall similarity of the scale under consideration to a harmonic series." [8]
>
> "Perceptually, the greatest common divisor of the dyad corresponds to its virtual pitch (or missing fundamental) and is used in much the same way

> as in algorithms that determine virtual pitch. Since the robustness of a virtual pitch depends on how many of the lower harmonics are present in the stimulus, this measure of similarity is both physically and perceptually relevant. For example, a dyad whose spectrum comprises 50% of the harmonic frequencies in a harmonic series would evoke a stronger virtual pitch perception than a dyad with only 10% of these frequencies. We refer to this metric as the percentage similarity of a dyad. Percentage similarity can be expressed as ((x+y-1)/(x*y))*100, where x is the numerator of the frequency ratio and y is the denominator of the ratio. For instance, a major third has a frequency ratio of 5:4; since x= 5 and y = 4, the percentage similarity is 40%.
>
> The overall conformance of a scale to a harmonic series was then determined by calculating the mean percentage similarity of the dyads in the scale in question. Using the mean as an index of similarity between a scale and a harmonic series implies that all possible dyads in the scale are equally relevant. Although in contemporary Western music any two notes in a scale can, in principle, be used together in melody or harmony, in traditional Western voice-leading and in other musical systems (e.g., classical Indian) particular tone combinations are avoided or prohibited. Nonetheless, there is no universal rule that describes which intervals might be more important in a scale than others; thus we treated all intervals equally." [9]

Gill and Purves used justly-tuned intervals in their methodology [10]. However, our top-down approach led us to a framework that accommodates all equal-tempered chords in a transpositional twelve-point circular space. Of course, music may be performed with equal-tempered intonation or just intonation. Should we modify our approach?

As Alfred Blatter notes:

> "The obvious characteristic that one hears on an even-tempered keyboard is that all perfect fifths and perfect fourths are flawed, and the thirds are of the incorrect size. […] No professional vocal or instrumental ensemble would tolerate such intonation, but the keyboardist has no choice. […] When music is performed by voices, or on nonkeyboard instruments, semitones are not all the same size. Nor are all whole tones the same. […] During performances, tuning adjustments are constantly being made." [11]

Gill and Purves observed that:

> "The use of justly tuned intervals Western music over the last few centuries has been based on equal temperament tuning, which developed as a compromise between the aesthetic value of maintaining justly tuned intervals (i.e., intervals defined by relatively small integer ratios) and the practical need to facilitate musical composition and performance in multiple keys, especially on keyboard instruments. Just intonation is generally considered the most natural tuning system […]. Moreover, just intonation is used in non-Western traditions such as classical Indian music." [12]

It is necessary to choose between equal temperament tuning and just intonation. We consider the observations above to be sufficiently compelling to justify the use of just intonation rather than equal temperament in our calculation of the harmonic similarity values of chords.

We adopt, with minor modification, the method used by Gill and Purves for analyzing scales to produce our first dimension of the sound of chords: the harmonic similarity of performed chords to the harmonic series produced by human voices and musical instruments that cause the vibration of strings or columns of air.

In a way that is very similar to the methodology used by Gill and Purves but slightly simplified, the harmonic similarity of chords in this book is analyzed with a collection of justly-tuned intervals. However, each interval has only one percentage similarity value (Figure 1.4) [13]. For example, the harmonic similarity of every major second interval that appears anywhere in a chord is calculated using the frequency ratio 9:8 and hsim 22.22, and the more remote frequency ratio 10:9 and hsim 20.00 are not used.

Transpositional twelve-point circular space with equal temperament	Equal-tempered intervals	Equal-tempered interval size (cents)	Transpositional twelve-point circular space with just intonation	Just intervals	Just interval size (cents)	Hsim of just intervals
0			0			
1	0-1	100.00	1.1944	0-1.1944	119.44	13.33
2	0-2	200.00	2.0391	0-2.0391	203.91	22.22
3	0-3	300.00	3.1564	0-3.1564	315.64	33.33
4	0-4	400.00	3.8631	0-3.8631	386.31	40.00
5	0-5	500.00	4.9804	0-4.9804	498.04	50.00
6	0-6	600.00	5.8251	0-5.8251	582.51	31.43
7	0-7	700.00	7.0196	0-7.0196	701.96	66.67
8	0-8	800.00	8.1369	0-8.1369	813.69	30.00
9	0-9	900.00	8.8436	0-8.8436	884.36	46.67
10	0-10	1000.00	10.1760	0-10.1760	1017.60	28.89
11	0-11	1100.00	10.8827	0-10.8827	1088.27	18.33

Figure 1.4 Transpositional twelve-point circular space, intervals, interval sizes, and the harmonic similarity of just intervals.

The formula for calculating the harmonic similarity (hsim) of chords, as used in this book, is set out in Appendix B (How harmonic similarity (hsim) values were calculated).

For convenience, we express the transpositional twelve-point circular space in equal-temperament form rather than just-interval form (0, 1, 2... rather than 0, 1.1944, 2.0391…). This is essentially no different than notating pitches on equally-spaced musical staves, centred on lines or between lines, with no adjustment for just intonation. Likewise, we express intervals and chords in equal-temperament form rather than justly-tuned form (a perfect fourth is 0-5 rather than 0-4.9804). But all of our calculations of harmonic similarity are based on our simplified model of just intonation.

We also treat all numbered point sets as chords, including sets that might also be used to represent scales. In other words, there is no octave doubling of any numbered points to represent the octave doubling of a tonic (or any other pitch).

For the purposes of this dimension of the sound of chords, some aspects of the framework created using the top-down approach are retained, and others are modified.

The modified framework retains all transpositional set classes. All notes continue to be represented as numbers from 0 to 11 rather than by letter names and accidentals, similar to a normal clock face. Each chord is represented by a numbered point set that begins with the number 0. Chords are represented by the distance between numbered points, measured in semitones in a clockwise direction. Individual notes do not have any relation to any specific frequencies (kHz). This framework retains octave equivalence, enharmonic equivalence, instrumental equivalence, dynamic equivalence, and articulation equivalence. It implicitly assumes that all pitches are equally and completely clear. The framework also continues to disregard the following tacitly ignored psychoacoustic concepts: sound pressure levels, critical bands, roughness, the dominant spectral region, and pitch shifts. Likewise, timbre (which includes the relative amplitudes of harmonics) is ignored.

Several aspects of the framework have been modified. The framework now uses just intonation rather than equal-temperament. All calculations of harmonic similarity are based on our simplified model of just intonation. Frequency ratios, which are a vital characteristic of this dimension, are captured in the just intervals used to calculate harmonic similarity.

Nonetheless, intervals and chords are expressed in equal-temperament form rather than justly-tuned form. Although the distance between each adjacent numbered point on the transpositional twelve-point circular space may appear to be identical, it is not. We do not adjust the distance between numbered points to account for just intonation, just as the positioning of pitches on musical staves are not adjusted for just intonation.

These modifications turn our highly-abstract top-down framework into a more musically meaningful tool that takes into account various aspects of human perception that are significant at a basic sensory level.

1.2.2 Spectral pitches and virtual root candidates

The work of Ernst Terhardt is vital to understanding the sound of chords. Research by Terhardt and his colleagues about the use of Fourier analysis and subharmonic matching leads to considerable insight into how the auditory processing system in the human brain creates the perception of chords, which is a significant advance over discussions of chordal pitches and the nature of the harmonic series, which is where many analyses end.

When a chord is performed, humans perceive the fundamental frequency of each performed note as well as spectral pitches that result from their harmonic overtones. The auditory processing system in the human brain also infers virtual pitches from those spectral pitches through a process of subharmonic matching [14]. Terhardt explained that “both spectral pitch and virtual pitch ultimately are dependent on aural Fourier analysis [but spectral] pitch is the most important carrier of auditory information, as it is an element of higher-order, Gestalt-like types of auditory precepts…” [15].

Terhardt and his colleagues expanded on the relevant auditory processing principles in relation to a specific chord:

> In the case of a major triad played on a piano, the tonal precept consists of (1) several spectral pitches which correspond to individual harmonics of the piano's complex tones, (2) several virtual pitches which correspond to the chord's musical notes, and (3) some additional virtual pitches in the frequency region below the lowest fundamental of the lowest chord note; these virtual pitches correspond to the major triad's "fundamental note" in a musical sense, i.e. the harmonic root [16].

Of course, in the case of a chord played on the piano, the tonal precept is shaped by the timbre of the piano as performed in a specific acoustic environment. We shall discuss timbre in more detail later. It should also be clear that the individual harmonics of the piano's complex tones includes the chord's musical notes.

1.2.3 Terhardt's simplified algorithm and initial revisions

With respect to the method set out in this book, Terhardt's work is important for his attempt to create a simplified algorithm that would reveal the roots of chords. He believed that the simplified algorithm could explain the psychophysical foundations of harmony.

Terhardt described his simplified algorithm as follows:

> "Fortunately (for the theory of harmony), the pitch patterns elicited by harmonic complex tones are not too heavily dependent on the shape of the Fourier spectrum, i.e., timbre. So, if one presupposes - as a kind of general convention - that musical tones throughout are to be realized as harmonic complex tones, one can indeed represent them by note symbols, at least to a reasonable approximation.
>
> Based on this convention, one can deduce from the theory of virtual pitch a root-finding algorithm that operates on the score - just as in conventional music theory […]. The principle is, that all candidates of roots must be subharmonics of the spectral pitches elicited by the actual sound, and that the prominence of any root is enhanced by "subharmonic coincidence" […]. When one translates this concept into musical notation, one gets a simple algorithm that shall be explained by the following example.

Chord notes:	c	f	a
same	C	**F**	A
- 1 fifth	**F**	B♭	D
- 1 mj 3rd	A♭	D♭	**F**
+ 1 mj 2nd	D	G	B
- 1 mj 2nd	B♭	E♭	G

The sample chord of which the root(s) are to be determined is written into the first line (2nd to 4th column, lowercase letters), i.e., c-f-a. The root candidates derived from the notes are written into the pertinent columns (uppercase letters) [.] The first candidate corresponds to both the first and second subharmonics, i.e., it is just the same note (second line), because, by definition, octave-equivalent pitches are not distinguished. The second candidate (third line), corresponding to the third subharmonic, is obtained by stepping one fifth down from the chord note, as octaves are not distinguished. The third candidate (fourth line), corresponding to the 5th subharmonic (the fourth is omitted, due to octave-equivalence) is obtained by stepping one major third down. The fourth candidate (fifth line), [corresponding] to the 7th subharmonic (the 6th is omitted, due to octave-equivalence to the third), is obtained by stepping either down by a minor seventh or stepping up one full tone (due to octave equivalence). The 8th subharmonic is omitted due to octave equivalence, such that the last candidate (6th line) corresponds to the 9th subharmonic, and it is obtained by either stepping up a 9th interval or stepping down one full tone (octave equivalence). As one can see in the table, there is one and only one candidate, namely F, that occurs in all three columns (full subharmonic match). This, by definition, indicates that F is the most prominent root of the chord.

As is apparent from this example, the determination of root candidates cannot fail for any type of chord. What may happen just is that there is no "full match", i.e., that there does not occur any candidate which is found in all columns, such that the root is less pronounced and more ambiguous than in a major triad such as above.

As another example let us consider a chord that by conventional theory has been regarded to be "at the borderline to atonality", i.e., the famous Tristan chord. Here is the corresponding table:

Tristan chord:	f	b	d♯	g♯
	F	B	D♯	G♯
	A♯	E	G♯	**C♯**
	C♯	G	B	E
	G	**C♯**	F	A♯
	D♯	A	**C♯**	F♯

> The algorithm tells us that actually there is a full match for the root C♯. Thus the chord, when considered in isolation, is far from being atonal. One can easily verify that the root C♯ indeed "makes sense", i.e., by playing it in the bass register together with the Tristan chord. So, the subharmonic matching algorithm has found out what one may as well explain in terms of the conventional theory: The Tristan chord f-b-d♯-g♯ can be said to be a major 9th chord with root C♯, of which the root note itself is missing.
>
> In summary, I believe that with the above concept of harmony, and with the explanations outlined for tone affinity and root-relationship, the psychophysical foundations of harmony are laid." [17].

Note that Terhardt stated that “the pitch patterns elicited by harmonic complex tones are not too heavily dependent on […] timbre”. This statement implies that harmonic complex tones are dependent on timbre to some degree. He also indicated that the prominence of any root is enhanced by "subharmonic coincidence". He also touched on the concept of root ambiguity. We will return to these concepts later, and show how they can be used to improve our understanding of the sound of chords.

The simplified algorithm, which Terhardt discussed above, can be set out in the form of a chart (Figure 1.5).

Spectral Pitches (Harmonics) and Virtual Root Candidates (Subharmonics)

	Intervals	Semitones													
Spectral Pitches (Harmonics)	M2 up	2 up	D	D♯ E♭	E	E♯ F	F♯ (G♭)	G	G♯ A♭	A	A♯ B♭	B	B♯ C	C♯ (D♭)	9th harmonic
	m7 up	10 up	B♭ (A♯)	B	C	C♯ D♭	D	E♭ (D♯)	E	F	F♯ G♭	G	G♯ A♭	A	7th harmonic
	M3 up	4 up	E	E♯ F	F♯ (G♭)	F♯♯ G	G♯ (A♭)	A	A♯ B♭	B	B♯ C	C♯ (D♭)	C♯♯ D	D♯ (E♭)	5th harmonic
	P5 up	7 up	G	G♯ A♭	A	A♯ B♭	B	C	C♯ D♭	D	D♯ E♭	E	E♯ F	F♯ (G♭)	3rd and 6th harmonics
	P8 up	12 up	C	C♯ D♭	D	D♯ E♭	E	F	F♯ G♭	G	G♯ A♭	A	A♯ B♭	B	2nd, 4th and 8th harmonics
Performed Chord Notes			C	C♯ D♭	D	D♯ E♭	E	F	F♯ G♭	G	G♯ A♭	A	A♯ B♭	B	
			0 - 12	**1 - 11**	**2 - 10**	**3 - 9**	**4 - 8**	**5 - 7**	**6 - 6**	**7 - 5**	**8 - 4**	**9 - 3**	**10 - 2**	**11 - 1**	1st harmonic 1st subharmonic
Virtual Root Candidates (Subharmonics)	P8 down	12 down	C	C♯ D♭	D	D♯ E♭	E	F	F♯ G♭	G	G♯ A♭	A	A♯ B♭	B	2nd, 4th and 8th subharmonics
	P5 down	7 down	F	F♯ G♭	G	G♯ A♭	A	B♭ (A♯)	B	C	C♯ D♭	D	D♯ E♭	E	3rd and 6th subharmonics
	M3 down	4 down	A♭ (G♯)	A	B♭ (A♯)	B	C	D♭ (C♯)	D	E♭ (D♯)	E	F	F♯ G♭	G	5th subharmonic
	m7 down	10 down	D	D♯ E♭	E	E♯ F	F♯ (G♭)	G	G♯ A♭	A	A♯ B♭	B	B♯ C	C♯ (D♭)	7th subharmonic
	M2 down	2 down	B♭ (A♯)	B	C	C♯ D♭	D	E♭ (D♯)	E	F	F♯ G♭	G	G♯ A♭	A	9th subharmonic

Figure 1.5 Terhardt's simplified algorithm.

Terhardt's simplified algorithm can be modified to make it function at a more abstract level by replacing the letter names and accidentals with numbers from 0 to 11 (Figure 1.6).

Spectral Pitches (Harmonics) and Virtual Root Candidates (Subharmonics)

	Intervals	Semitones													
Spectral Pitches (Harmonics)	M2 up	2 up	2	3	4	5	6	7	8	9	10	11	0	1	9th harmonic
	m7 up	10 up	10	11	0	1	2	3	4	5	6	7	8	9	7th harmonic
	M3 up	4 up	4	5	6	7	8	9	10	11	0	1	2	3	5th harmonic
	P5 up	7 up	7	8	9	10	11	0	1	2	3	4	5	6	3rd and 6th harmonics
	P8 up	12 up	0	1	2	3	4	5	6	7	8	9	10	11	2nd, 4th and 8th harmonics
Performed Chord Notes			**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	
			0 - 12	**1 - 11**	**2 - 10**	**3 - 9**	**4 - 8**	**5 - 7**	**6 - 6**	**7 - 5**	**8 - 4**	**9 - 3**	**10 - 2**	**11 - 1**	1st harmonic 1st subharmonic
Virtual Root Candidates (Subharmonics)	P8 down	12 down	0	1	2	3	4	5	6	7	8	9	10	11	2nd, 4th and 8th subharmonics
	P5 down	7 down	5	6	7	8	9	10	11	0	1	2	3	4	3rd and 6th subharmonics
	M3 down	4 down	8	9	10	11	0	1	2	3	4	5	6	7	5th subharmonic
	m7 down	10 down	2	3	4	5	6	7	8	9	10	11	0	1	7th subharmonic
	M2 down	2 down	10	11	0	1	2	3	4	5	6	7	8	9	9th subharmonic

Figure 1.6 Terhardt's simplified algorithm with letter names and accidentals replaced by numbers from 0 to 11.

In Figure 1.6, transpositional twelve-point circular space is represented as a single line of performed chord notes (0… 11), but the ends are actually connected to form a circular space (so, for example, 11 + 1 = 0). In acoustics (and psychoacoustics), the harmonic series (and spectral pitches) rise above the performed chordal pitches over more than one octave. In our method, octave equivalence applies and the spectral pitches are represented by numbers in transpositional twelve-point circular space. Likewise, octave equivalence is applied to subharmonics. Each harmonic (spectral pitch) and subharmonic level is also circular. Figure 1.7 shows the performed chord notes and their spectral pitches in transpositional twelve-point circular space.

In this method, the term "spectral pitch" is used as it clarifies the relationship between the psychoacoustic level and our mathematical formula. However, all calculations made with this formula in relation to spectral pitches produce results that are numbered points in transpositional twelve-point circular space.

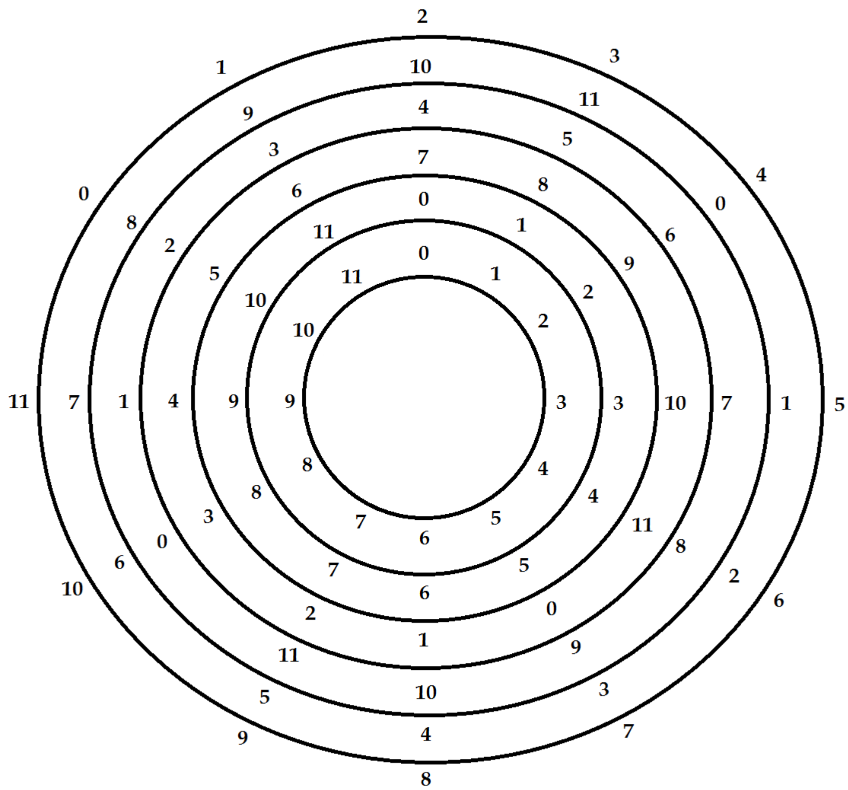

Figure 1.7 Performed chord notes circled by their spectral pitches in transpositional twelve-point circular space.

Terhardt assumed that the pitch patterns elicited by harmonic complex tones are not too heavily dependent on timbre, and concluded that it is reasonable to ignore harmonic amplitude. However, it appears that the amplitudes of the fundamental and all harmonics, and the strength of subharmonics, have more significant impact on the pitch patterns than he acknowledged.

Based on the chords that Terhardt has used as examples of the application of his simplified algorithm, a processed chord implicitly contains at least three performed chordal notes (although there does not seem to be any specific exclusion of intervals). Of course, Terhardt's simplified algorithm is not intended to identify virtual root candidates for a single pitch as there must be more than one chordal pitch to allow for the analysis of subharmonic pitch matches.

Nonetheless, for a single pitch, Terhardt's simplified algorithm can be used to identify five, equal unordered spectral pitches (0, 2, 4, 7 and 10) and five, equal unordered root candidates (0, 2, 5, 8 and 10) in transpositional twelve-point circular space. Identifying these pitches is an important first step in evaluating the sound of chords. However, these results are inconsistent with conventional music theory, which does not conclude that a single pitch has five equal roots. Treating all spectral pitches as equal, and all root candidates as equal, is inconsistent with acoustics and psychoacoustics; this approach does not reflect how people hear. Generally speaking, harmonics that are more distant from a fundamental pitch along the harmonic series tend to reverberate more quietly than less distant harmonics.

Applying Terhardt's simplified algorithm to a perfect fifth interval reveals a problem underlying an unweighted pattern matching approach.

For the perfect fifth interval (0-7), Terhardt's simplified algorithm identifies two virtual roots, 0 and 5 (Figure 1.8). Each one is selected because 2 pitches match in the subharmonic series for the chordal pitches, 0 and 7. In Figure 1.8, the root candidates 0 and 5 are equal and unordered, and 2, 3, 7, 8, 9 and 10 are another subset of equal, unordered root candidates. In other words, no prominence is implied between the two virtual roots or amongst the other root candidates, except that the two virtual roots are more prominent than the rest.

Root candidates	0	5	2	3	7	8	9	10
Number of pitches that match	2	2	None	None	None	None	None	None
Terhardt's simplified virtual root	VR	VR	-	-	-	-	-	-

Figure 1.8 Terhardt's simplified algorithm in Figure 1.6 applied to the root candidates of a perfect fifth interval (0-7).

These results are inconsistent with conventional music theory, which does not conclude that a perfect fifth has two equal roots or that 5 is the root of 0-7. These results do not appear to reflect what people hear.

Terhardt's simplified pattern-matching algorithm can also provide anomalous results for some triads.

For the A minor triad in root position (A 0-3-7), first inversion (C 0-4-9), and second inversion (E 0-5-8), Terhardt's simplified algorithm identifies one VR in each case, (5 = D; 2 = D; and 10 = D). This is inconsistent with the analysis by Terhardt and others of the virtual roots of A minor triads played on a piano, which showed that the virtual root of the A minor triad in root position (A_4, C_5, E_5) was A_4 (= 0 in transpositional twelve-point circular space) and in first inversion (C_5, E_5, A_5) was A_4 (= 9) (among other contenders). It is also inconsistent with their results for the A minor triad in second inversion (E_5, A_5, C_6), which had two ambiguous equally-weighted virtual roots (E_4 and A_3 = 0 and 5) (again, among other contenders) [18]. As well, the results produced by Terhardt's simplified algorithm for a minor triad are also inconsistent with conventional music theory and do not appear to reflect what people hear.

It appears that the Terhardt's simplified pattern-matching algorithm errs on the side of being too simplified, and provides misleading results in some cases. How can Terhardt's simplified pattern-matching algorithm be modified to be more accurate while minimizing the incorporation of acoustic or psychoacoustic variables?

1.2.4 Timbre weighting of spectral pitches and root candidates

As complex harmonic tones are dependent on timbre to some extent, a method that takes this fact into account may be able to produce more realistic and practical results than one that does not.

However, timbre is very complex. Every voice has a distinct timbre, which can change when one speaks, sings or whispers. Every type of instrument has different timbres from other instruments. Timbre changes as pitches are played in different octaves on the same instrument. Henry Brandt has concluded that there are thirteen different groups of closely-related "live" acoustic (rather than electronic, amplified or synthesized) instrumental timbres [19]. Likewise, timbre can change with how a single pitch is played, such as articulations and dynamics (including crescendos).

There is no perfect approach for determining the range of pitches to be used or the frequency of each harmonic to be used. However, the approach chosen should be reasonable, simple and workable, as the goal is to help young composers by providing them with a practical method for evaluating and comparing the sound of chords at the sensory level.

The human ear evolved along with the human voice, including singing, long before the development of sophisticated musical instruments. Furthermore, as Gill and Purves indicate, "[…] most music, even purely instrumental music, is composed within the human vocal range, and some popular instruments (e.g., the violin) bear a timbral resemblance to the human voice." [20]

Reflecting the range of the human voice to some degree would be appropriate. The range between the lowest notes of a basso cantante to the highest notes of a lyric soprano is approximately 87.31 Hz [F_2] to 1046.50 Hz [C_6].

However, this range is too wide for a single voice. Using the timbres of several voices (soprano, alto, tenor and bass) would introduce anomalies, as the frequency patterns of the harmonics for each singer's voice would differ, even when singing the same pitch. Likewise, selecting the timbres of different acoustic orchestral instruments would also introduce anomalies, whether in the same octave or different octaves.

The approach chosen for this method is to create a generic timbre that is a proxy for vocal and acoustic timbres generally and is not intended to represent the timbre of any specific note sung by any voice or played on any acoustic orchestral instrument under any circumstances. However, it is intended to be a reasonable timbre in the sense that it could represent the harmonics of many notes sung by a voice or played by an acoustic orchestral instrument (excluding percussions instruments) in the range between F_2 and C_6 [21].

To create the generic timbre used in this method, I have weighted each harmonic and subharmonic by arbitrarily dividing each value by a series of numbers from the Fibonacci series (Figure 1.9). It is important to be aware that *all* of the values derived from this method - other than those for harmonic similarity (hsim) - are directly influenced by the timbre weight percentages. Different percentages would produce different results [22]. This relationship is most easily observed in the weighting of the spectral pitches for a single pitch, which equal the percentages used in the generic timbre (Figure 1.10).

Harmonic / subharmonic	**Unweighted value**		**Fibonacci divisor**		**Percent**
2	1	/	1	=	100.00
3	1	/	2	=	50.00
5	1	/	3	=	33 1/3
7	1	/	5	=	20.00
9	1	/	8	=	12.50

Figure 1.9 Generic timbre weightings derived with Fibonacci series values.

Spectral pitches	0	5	8	2	10
Timbre-weighted spectral pitch values	100.00	50.00	33.33	20.00	12.50

Figure 1.10 Spectral pitch values for a single pitch (0) weighted by the generic timbre.

I have clarified Terhardt's pattern matching approach to focus on the harmonics and subharmonics as set out in Figure 1.11:

The 9th harmonic
The 7th harmonic
The 5th harmonic
The 3rd or 6th harmonic, whichever has the largest amplitude
The note played or the 2nd, 4th or 8th harmonic, whichever has the largest amplitude

The note played or the 2nd, 4th or 8th subharmonic, whichever has the largest amplitude
The 3rd or 6th subharmonic, whichever has the largest amplitude
The 5th subharmonic
The 7th subharmonic
The 9th subharmonic

Figure 1.11 Clarified harmonic and subharmonic frequencies.

So, the weightings from the generic timbre for the harmonics and subharmonics in this method are set out in Figure 1.12.

Harmonics and subharmonics	**Percent**
The 9th harmonic	12.50
The 7th harmonic	20.00
The 5th harmonic	33 1/3
The 3rd or 6th harmonic, whichever has the largest amplitude	50.00
The note played or the 2nd, 4th or 8th harmonic, whichever has the largest amplitude	100.00
The note played or the 2nd, 4th or 8th subharmonic, whichever has the largest amplitude	100.00
The 3rd or 6th subharmonic, whichever has the largest amplitude	50.00
The 5th subharmonic	33 1/3
The 7th subharmonic	20.00
The 9th subharmonic	12.50

Figure 1.12 The weightings from the generic timbre applied to harmonics and subharmonics.

When the preferences and percentages in Figure 1.12 are integrated into Terhardt's simplified algorithm as shown in Figure 1.6, the modified algorithm takes the form set out in Figure 1.13 below.

Spectral Pitches (Harmonics) and Root Candidates (Subharmonics)

	Percent													
Spectral Pitches (Harmonics)	12.50	2	3	4	5	6	7	8	9	10	11	0	1	9th harmonic
	20.00	10	11	0	1	2	3	4	5	6	7	8	9	7th harmonic
	33.33	4	5	6	7	8	9	10	11	0	1	2	3	5th harmonic
	50.00	7	8	9	10	11	0	1	2	3	4	5	6	The 3rd or 6th harmonic, whichever has the largest amplitude
	100.00	0	1	2	3	4	5	6	7	8	9	10	11	The note played or the 2nd, 4th or 8th harmonic, whichever has the largest amplitude
Performed Chord Notes		**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	
		0 - 12	**1 - 11**	**2 - 10**	**3 - 9**	**4 - 8**	**5 - 7**	**6 - 6**	**7 - 5**	**8 - 4**	**9 - 3**	**10 - 2**	**11 - 1**	1st harmonic 1st subharmonic
Root Candidates (Subharmonics)	100.00	0	1	2	3	4	5	6	7	8	9	10	11	The note played or the 2nd, 4th or 8th subharmonic, whichever has the largest amplitude
	50.00	5	6	7	8	9	10	11	0	1	2	3	4	The 3rd or 6th subharmonic, whichever has the largest amplitude
	33.33	8	9	10	11	0	1	2	3	4	5	6	7	5th subharmonic
	20.00	2	3	4	5	6	7	8	9	10	11	0	1	7th subharmonic
	12.50	10	11	0	1	2	3	4	5	6	7	8	9	9th subharmonic

Figure 1.13 The timbre-weighted algorithm, with weightings from the generic timbre.

For processing the spectral pitches of chords, including their weights, this timbre-weighted algorithm represents the final form of our method. It is used to calculate the spectral pitches and weights set out in Appendix A (The Method for Evaluating and Comparing the Sound of Chords).

For example, the spectral pitch series shown in the columns above the performed chordal notes 0 and 7 both contain the spectral pitch 7. In one instance, the spectral pitch weight is 100.00%; in the other, it is 50.00%. For the perfect fifth interval (0-7), the spectral pitch 7 has a spectral pitch weight of 150.00%.

However, the percentage weighting of the subharmonics changes Terhardt's modified algorithm from being a virtual pitch extraction model and turns it into a timbre-weighted root candidate model that does not identify virtual pitches. As we will show below, further modification will be required to take subharmonic coincidence into account.

Applying Terhardt's simplified algorithm to a minor triad in root position (0-3-7) identifies a false virtual root, 5 (Figure 1.14).

Root candidates	5	0	3	8	1	2	7	9	10	11
Number of pitches that match	3	2	2	2	None	None	None	None	None	None
Terhardt's simplified virtual root	VR	-	-	-	-	-	-	-	-	-

Figure 1.14 Terhardt's simplified algorithm in Figure 1.6 applied to a minor triad in root position (0-3-7).

This issue is resolved when the timbre-weighted algorithm in Figure 1.13 is applied (see Figure 1.15). The lowest pitch in the stack of thirds (0) becomes the root, consistent with conventional music theory. In addition, the prominence of the root candidates declines in a manner which reflects that harmonics, which are more distant from a fundamental pitch along the harmonic series, tend to reverberate more quietly than less distant harmonics. This spectral pattern is mirrored in the timbre-weighting of the root candidates. In this case, two equal unordered pairs of root candidates are listed, 2 and 9, and 1 and 10.

Root candidates	0	3	7	8	5	11	2	9	1	10
Timbre-weighted root-candidate values	150.00	133.33	100.00	83.33	82.50	33.33	20.00	20.00	12.50	12.50
Terhardt's simplified virtual root	-	-	-	-	VR	-	-	-	-	-

Figure 1.15 The timbre-weighted algorithm in Figure 1.13 applied to a minor triad in root position (0-3-7).

However, applying the timbre-weighted algorithm, by itself, can create problems of its own.

For example, applying Terhardt's simplified algorithm to a Tristan chord (0-3-6-10) (Figure 1.16) identifies the true virtual root, 8, as discussed by Terhardt above. For this chord, the prominence of the one virtual root is consistent with my subjective listening tests [23].

Root candidates	8	0	2	3	5	6	10	11	1	4
Number of pitches that match	4	2	2	2	2	2	2	2	None	None
Terhardt's simplified virtual root	VR	-	-	-	-	-	-	-	-	-

Figure 1.16 Terhardt's simplified algorithm in Figure 1.6 applied to a Tristan chord (0-3-6-10).

However, when the timbre-weighted algorithm in Figure 1.13 is applied, the prominence of the virtual root is lost (Figure 1.17).

Root candidates	3	6	0	8	10	11	5	2	1	4
Timbre-weighted root-candidate values	150.00	133.33	120.00	115.83	112.50	83.33	70.00	53.33	12.50	12.50
Terhardt's simplified virtual root	-	-	-	VR	-	-	-	-	-	-

Figure 1.17 The timbre-weighted algorithm in Figure 1.13 applied to a Tristan chord (0-3-6-10).

Therefore, our method needs to be revised to take into account how the brain processes subharmonic matches that are perceived as virtual roots.

1.2.5 The subharmonic enhancement multiplier

As Terhardt indicated above, the prominence of any root is enhanced by subharmonic coincidence. The approach that we use in this method to represent subharmonic coincidence is to apply a formula that increases the timbre-weighted root-candidate values of each set of pitches (*as a numbered point set*) that have subharmonic matches, with results that vary according to the number of matches for that pitch set and the subharmonic enhancement percentage applied.

Over the next two pages, the term “pitch” is used as it may clarify the relationship between subharmonic coincidences at the psychoacoustic level and our mathematical formula. However, all calculations made with this formula produce results that are numbered points in transpositional twelve-point circular space.

Let x = the subharmonic enhancement multiplier.

Let x = 1 + (y * z).

Let y = pitch multiplier for the number of matches for each pitch. If no pitches match, y = 0.

Let z = the subharmonic enhancement percentage.

For each performed pitch in a chord, polygon diagrams can be used to produce the pitch multipliers for the number of pitches among the subharmonic series that match. The first three polygons in the series are shown below (Figure 1.18). The dots represent the number of pitches that match, and the lines represent the number of matches, which are used as pitch multipliers (y).

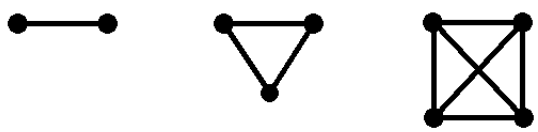

Figure 1.18 Polygon diagrams.

With one exception, the values produced by the polygon diagrams are identical to the values which were used by Gill and Purves as divisors in their formula that established the harmonic similarity of scales to the harmonic series (Figure 1.19). The exception is that the series used in this method starts with the number 2 rather than the number 1 because there must be at least two identical pitches to have a match.

It is important to be aware that this polygon model differs from the formulas set out in Terhardt's complete psychoacoustic model, which relies on the application of physics to the frequencies of performed pitches measured in Hz [24].

How many pitches match?	2	3	4	5	6	7	8	9	10	11
Pitch multiplier (y)	1	3	6	10	15	21	28	36	45	55

Figure 1.19 Pitch multipliers (y).

If 2 pitches match each other, y = 1; if 3 pitches match each other, y = 3; if 4 pitches match each other, y = 6; and so on.

These pitch multipliers are used for every set of matches that result from performing a chord. For example, if a 4-pitch chord has one root candidate with 4 pitches that match each other, and a second root candidate with 2 pitches that match each other, the pitch multiplier 6 applies to the first root candidate and the pitch multiplier 1 applies to the second root candidate.

What should the subharmonic enhancement percentage be?

My unpublished analysis of various selected chords suggests that a subharmonic enhancement percentage of 20% is close to ideal overall, but perhaps a percentage somewhat less than 20% would be best. The analysis considered alternative percentages from 10% to 70%.

My subjective listening to the same selection of chords suggests that a 20% subharmonic enhancement percentage is better than 10%, so a percentage closer to 20% than to 10% is best overall. By "better", I mean that the use of the subharmonic enhancement percentage allows the method to identify root candidates in a descending order that best reflects their sense of being the root of the chord (from most likely to be the root to least likely) when played one, two and three octaves below. Subjectively, the most likely root could occur in any one of those octaves (or in more than one octave).

First, let's consider the Tristan chord again (see Figure 1.16 and Figure 1.17).

When the subharmonic enhancement multiplier is applied to the timbre-weighted root candidates using a subharmonic enhancement percentage of 10.0%, the root candidate with the highest weight is once again the virtual root 8 (Figure 1.20). However, the root candidate value for 8 is only somewhat greater than that for 3, which suggests that there is a low degree of clarity between these two root candidates, which is not consistent with subjective listening.

Root candidates	8	3	6	0	10	11	5	2	1	4
Root-candidate values with z = 10%	185.33	165.00	146.67	132.00	123.75	91.67	77.00	58.67	12.50	12.50
Terhardt's simplified virtual root	VR	-	-	-	-	-	-	-	-	-

Figure 1.20 The timbre-weighted algorithm in Figure 1.13 and the subharmonic enhancement multiplier with a subharmonic enhancement percentage of 10.0%, applied to a Tristan chord (0-3-6-10).

When the subharmonic enhancement percentage is increased to 20.0%, there is considerably more clarity between the first two root candidates (Figure 1.21).

Root candidates	8	3	6	0	10	11	5	2	1	4
Root-candidate values with z = 20%	254.83	180.00	160.00	144.00	135.00	100.00	84.00	64.00	12.50	12.50
Terhardt's simplified virtual root	VR	-	-	-	-	-	-	-	-	-

Figure 1.21 The timbre-weighted algorithm in Figure 1.13 and the subharmonic enhancement multiplier with a subharmonic enhancement percentage of 20.0%, applied to a Tristan chord (0-3-6-10).

So, in order for the Tristan chord to have its virtual root 8 as the highest weighted and selected root, the subharmonic enhancement percentage must be at least 10%. However, as subjective listening suggests that 8 is clearly the root of this chord, a higher subharmonic enhancement percentage is required to reflect a greater degree of root clarity relative to the other root candidates. So, subjective listening suggests that a percentage closer to 20% than to 10% would be appropriate.

Second, let’s consider the minor triad in root position once more (see Figure 1.14 and Figure 1.15).

When the subharmonic enhancement multiplier is applied to the timbre-weighted root candidates using a subharmonic enhancement percentage of 10.0%, the root candidate with the highest weight remains 0, but the virtual root 5 is now weighted somewhat more highly than root candidate 7 as well as more highly than 8 but less than 0 and 3 (Figure 1.22).

Root candidates	0	3	5	7	8	11	2	9	1	10
Root-candidate values with z = 10%	165.00	146.67	107.25	100.00	91.67	33.33	20.00	20.00	12.50	12.50
Terhardt's simplified virtual root	-	-	VR	-	-	-	-	-	-	-

Figure 1.22 The timbre-weighted algorithm in Figure 1.13 and the subharmonic enhancement multiplier with a subharmonic enhancement percentage of 10.0%, applied to a minor triad in root position (0-3-7).

When the subharmonic enhancement percentage is increased to 20.0%, the virtual root 5 is now weighted significantly higher than root candidate 7 (Figure 1.23). Subjective listening suggests that this weight is too high and that a subharmonic enhancement percentage of 20.0% is too large.

Root candidates	0	3	5	7	8	11	2	9	1	10
Root-candidate values with z = 20%	180.00	160.00	132.00	100.00	100.00	33.33	20.00	20.00	12.50	12.50
Terhardt's simplified virtual root	-	-	VR	-	-	-	-	-	-	-

Figure 1.23 The timbre-weighted algorithm in Figure 1.13 and the subharmonic enhancement multiplier with a subharmonic enhancement percentage of 20.0%, applied to a minor triad in root position (0-3-7).

When the subharmonic enhancement percentage is increased to 44.6%, the virtual root 5 becomes the second most highly-weighted root. Finally, when the subharmonic enhancement percentage is increased to 69.3%, the virtual root 5 becomes the most highly-weighted root. Clearly, these results are not consistent with conventional music theory and with widely-held perceptions of the root of the minor triad in root position.

So, in order to keep the 0-3-7 minor triad from having 5 as the highest-weighted or second-highest and selected root, the subharmonic enhancement percentage must be less than 44.6%.

For the 0-3-7 minor triad, subjective listening suggests that 0 is a clearly stronger root candidate than 5, which in turn is stronger than 8. This perception is consistent with the analysis by Terhardt and others of the virtual roots of A minor triads played on a piano [25]. Subjective listening also suggests that 0 is a clearly stronger root candidate than 3, which in turn is stronger than all other root candidates; but the predominance of 7 over 5 and vice-versa depends on the octave in which those root candidates are played in. These perceptions are not consistent with the analysis by Terhardt and others, which did not report 3, 7 or 5 as being virtual root candidates. Overall, subjective listening suggests that a subharmonic enhancement percentage that is closer to 10% than to 20% would be appropriate.

A perfect subharmonic enhancement percentage does not appear to exist. The best one can do is to select a percentage that provides reasonable root-candidate values for a broad range of chords. Upward or downward shifts in the percentage may provide more reasonable values for the root candidates of some chords but less reasonable values for others. The impact of the percentage on the value of the first root candidate for a chord is relatively more important than its impact on the value of the second root candidate, which is relatively more important than its impact on the third root candidate, and so on. The optimal percentage to be used with all chords is ultimately somewhat of a compromise. It is an exercise of judgement.

Arbitrarily, I have chosen 16.66666666666667% (= 1/6) to be the subharmonic enhancement percentage (z) for the purposes of this method.

Consequently, the following subharmonic enhancement multipliers in Figure 1.24 apply. In practice, there are unlikely to be any matches that consist of large numbers of pitches.

How many pitches match?	2	3	4	5	6	7	8	9	10	11
Subharmonic enhancement multiplier (x)	1.167	1.5	2.0	2.667	3.5	4.5	5.667	7.0	8.5	10.167

Figure 1.24 Subharmonic enhancement multipliers (x).

When the subharmonic enhancement percentage is set to 16.66666666666667% and the subharmonic enhancement multipliers are applied to a minor triad in root position (0-3-

7), the virtual root 5 is now weighted somewhat more closely to root candidate 7 than it is when the subharmonic enhancement percentage is set to 20% (Figure 1.25).

Root candidates	0	3	5	7	8	11	2	9	1	10
Root-candidate values with z = 1/6	175.00	155.56	123.75	100.00	97.22	33.33	20.00	20.00	12.50	12.50
Terhardt's simplified virtual root	-	-	VR	-	-	-	-	-	-	-

Figure 1.25 The timbre-weighted algorithm in Figure 1.13 and the subharmonic enhancement multiplier with a subharmonic enhancement percentage of 16.66666666666667%, applied to a minor triad in root position (0-3-7).

The minor triad and the Tristan chord can be used again to illustrate the thought process behind the selection of 16.66666666666667% (= 1/6) to be the subharmonic enhancement percentage (z) for the purposes of this method. A somewhat lower percentage might be more appropriate to reflect the predominance of 7 and 5 as root candidates of a minor triad in root position. However, given the other constraints we have discussed, these root candidates are only vying for third and fourth place. On the other hand, this percentage clearly establishes the most-highly weighted root candidate for a Tristan chord (Figure 1.26).

Root candidates	8	3	6	0	10	11	5	2	1	4
Root-candidate values with z = 1/6	231.67	175.00	155.56	140.00	131.25	97.22	81.67	62.22	12.50	12.50
Terhardt's simplified virtual root	VR	-	-	-	-	-	-	-	-	-

Figure 1.26 The timbre-weighted algorithm in Figure 1.13 and the subharmonic enhancement multiplier with a subharmonic enhancement percentage of 16.66666666666667% (1/6), applied to a Tristan chord (0-3-6-10).

The timbre-weighted algorithm in Figure 1.13, with the root candidates enhanced by the subharmonic enhancement percentage set to 16.66666666666667%, provides the basis for calculating the root candidates of each chord, as set out in the third layer of Appendix A (The Method for Evaluating and Comparing the Sound of Chords).

The key point to understand is that the combined effect of the auditory processing of timbre and the production of subharmonic enhancement in the brain – both working together – results in the perception of the roots of chords. Either of these processes alone is incomplete and may lead to inaccurate conclusions.

1.2.6 Root clarity and root ambiguity

As discussed above, for the perfect fifth interval (0-7), Terhardt's simplified algorithm identified two virtual roots, 0 and 5 (Figure 1.8). As previously noted, these results are inconsistent with conventional music theory, which does not conclude that a perfect fifth has two equal roots or that 5 is the root of 0-7; and these results do not appear to reflect what people hear. When the timbre-weighted algorithm in Figure 1.13 and the subharmonic enhancement multiplier with a subharmonic enhancement percentage of 16.66666666666667% is applied, the error produced by the simplified algorithm is resolved (Figure 1.27). The interval has one clear root, 0, which is consistent with conventional music theory and with what people hear. The root candidate values reveal that the 5 is a false virtual root.

Root candidates	0	7	5	3	8	2	9	10
Root-candidate values with z = 1/6	175.00	100.00	72.92	33.33	33.33	20.00	20.00	12.50
Terhardt's simplified virtual root	VR	-	VR	-	-	-	-	-

Figure 1.27 The timbre-weighted algorithm in Figure 1.13 and the subharmonic enhancement multiplier with a subharmonic enhancement percentage of 16.66666666666667% (1/6), applied to a perfect fifth interval (0-7).

The clarity of the root of a chord depends primarily on the difference between the root candidate values of the first and second root candidates. In this method, root clarity is a measure of the relative difference between the root-candidate values of the first root candidate (the highest root-candidate value) and the second root candidate, where the root-candidate value of the first root candidate is greater than that of the second. The greater the relative difference between the root-candidate values of the first and second root candidates, the greater the clarity of the root. If the root-candidate values of the first and second root candidates are identical, there is complete root ambiguity and no root clarity. This ambiguity increases if more than the first two root candidates have identical root candidate values. (Perceptually, this ambiguity would also increase if the root candidate value of the third root candidate is very close to the first two values.)

We use a simple formula to measure root clarity and root ambiguity.

Let a = the root clarity value.

Let a = (b - c) / c.

Let b = the root-candidate value of the first root candidate (the highest root-candidate value).

Let c = the root-candidate value of the second root candidate.

Figure 1.28 shows the root clarity value of a perfect fifth.

Root candidates	0	7
Root-candidate values with z = 1/6	175.00	100.00
Root clarity	75.00%	

Figure 1.28 The root clarity of a perfect fifth interval (0-7).

The root clarity of a single pitch can serve as a benchmark for comparing the root clarity values of other chords (Figure 1.29).

Root candidates	0	5
Root-candidate values with z = 1/6	100.00	50.00
Root clarity	100.00%	

Figure 1.29 The root clarity of a single pitch (0).

Also consider the root candidate values of major and minor triads in root position (Figure 1.30 and Figure 1.31).

Root candidates	0	4
Root-candidate values with z = 1/6	275.00	100.00
Root clarity	175.00%	

Figure 1.30 The root clarity of major triads in root position (0-4-7).

Root candidates	0	3
Root-candidate values with z = 1/6	175.00	155.56
Root clarity	12.50%	

Figure 1.31 The root clarity of minor triads in root position (0-3-7).

Conventional music theory and widely-held perceptions lead to the conclusion that major triads have significantly clearer and more stable roots than minor triads, and that the roots of minor triads are more ambiguous. The root clarity values produced by this method for the major and minor triads in root position are consistent with those conclusions.

The root clarity value is calculated using the root-candidate value of the first root candidate even if it is not a chordal pitch. This is illustrated in Figure 1.32 with the Tristan chord.

Root candidates	8	3
Root-candidate values with z = 1/6	231.67	175.00
Root clarity	32.38%	

Figure 1.32 The root clarity of the Tristan chord (0-3-6-10).