Three Conceptions of Musical Distance

Dmitri Tymoczko
310 Woolworth Center, Princeton University, Princeton, NJ 08544.

Abstract

This paper considers three conceptions of musical distance (or inverse "similarity") that produce three different musico-geometrical spaces: the first, based on voice leading, yields a collection of continuous quotient spaces or orbifolds; the second, based on acoustics, gives rise to the Tonnetz and related "tuning lattices"; while the third, based on the total interval content of a group of notes, generates a six-dimensional "quality space" first described by Ian Quinn. I will show that although these three measures are in principle quite distinct, they are in practice surprisingly interrelated. This produces the challenge of determining which model is appropriate to a given music-theoretical circumstance. Since the different models can yield comparable results, unwary theorists could potentially find themselves using one type of structure (such as a tuning lattice) to investigate properties more perspicuously represented by another (for instance, voiceleading relationships).

1 Introduction

We begin with voice-leading spaces that make use of the log-frequency metric [1, 15, 3]. Pitches here are represented by the logarithms of their fundamental frequencies, with distance measured according to the usual metric on $\mathbb{R}$; pitches are therefore "close" if they are near each other on the piano keyboard. A point in $\mathbb{R}^n$ represents an ordered series of pitch classes. Distance in this higher-dimensional space can be interpreted as the aggregate distance moved by a collection of musical "voices" in passing from one chord to another. (We can think of this, roughly, as the aggregate physical distance traveled by the fingers on the piano keyboard.) By disregarding information–such as the octave or order of a group of notes–we "fold" $\mathbb{R}^n$ into a non-Euclidean quotient space or orbifold. (For example, imposing octave equivalence transforms $\mathbb{R}^n$ into the n-torus $\mathbb{T}^n$, while transpositional equivalence transforms $\mathbb{R}^n$ into $\mathbb{R}^{n-1}$, orthogonally projecting points onto the hyperplane whose coordinates sum to zero.) Points in the resulting orbifolds represent equivalence classes of musical objects–such as chords or set classes–while "generalized line segments" represent equivalence classes of voice leadings.[1] For example, Figure 1, from Tymoczko 2006, represents the space of two-note chords, while Figure 2, from Callender, Quinn, and Tymoczko 2008, represents the space of three-note transpositional set classes. In both spaces, the distance between two points represents the size of the smallest voice leading between the objects they represent.

Let's now turn to a very different sort of model, the *Tonnetz* [4, 5, 6] and related structures, which I will describe generically as "tuning lattices." These models are typically discrete, with adjacent points on a particular axis being separated by the same interval. The leftmost lattice in Figure 3 shows the most familiar of these structures, with the two axes representing acoustically pure perfect fifths and major thirds. (One can imagine a third axis, representing either the octave or the acoustical seventh, projecting outward from the paper.) The model asserts that the pitch G4 has an acoustic affinity to both C4 (its "underfifth") and D5 (its "overfifth"), as well as to E♭4 and B4 (its "underthird" and "overthird," respectively). The lattice thus encodes a fundamentally different notion of musical distance than the earlier voice leading models: whereas A3 and A♭3 are very close in log-frequency space, they are four steps apart our tuning lattice. Furthermore,

[1] The adjective "generalized" indicates that these "line segments" may pass through one of the space's singular points, giving rise to mathematical complications.

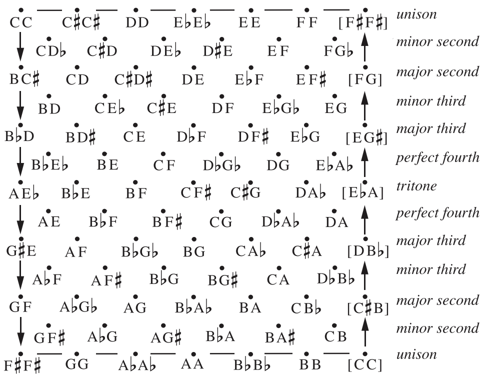

Figure 1: *The Möbius strip representing voice-leading relations among two-note chords.*

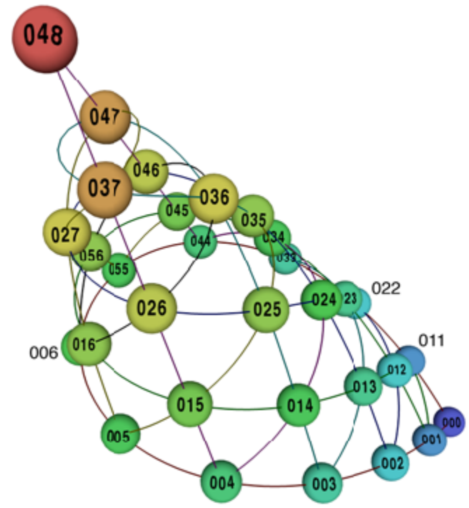

Figure 2: *The cone representing voice-leading relations among three-note transpositional set classes.*

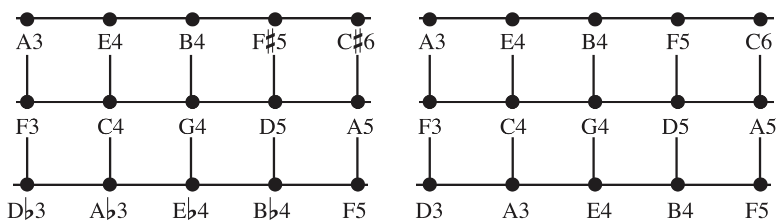

Figure 3: *Two discrete tuning lattices. On the left, the chromatic Tonnetz, where horizontally adjacent notes are linked by acoustically pure fifths, while vertically adjacent notes are linked by acoustically pure major thirds. On the right, a version of the structure that uses diatonic intervals.*

where chords (or more generally "musical objects") are represented by *points* in the voice leadings spaces, they are represented by polytopes in the lattices.

Finally, there are measures of musical distance that rely on chords' shared interval content. From this point of view, the chords C, C♯, E, F♯ and C, D♭, E♭, G resemble one another, since they are "nontrivially homometric" or "Z-related": that is, they share the same collection of pairwise distances between their notes. (For instance, both contain exactly one pair that is one semitone apart, exactly one pair that is two semitones apart, and so on.) However, these chords are *not* particularly close in either of the two models considered previously. It is not intuitively obvious that this notion of "similarity" produces any particular geometrical space. But Ian Quinn has shown that one can use the discrete Fourier transform to generate (in the familiar equal-tempered case) a six-dimensional "quality space" in which chords that share the same interval content are represented by the same point [10, 11, 12, 13, 2]. We will explore the details shortly.

Clearly, these three musical models are very different, and it would be somewhat surprising if there were to be close connections between them. But we will soon see that this is in fact this case.

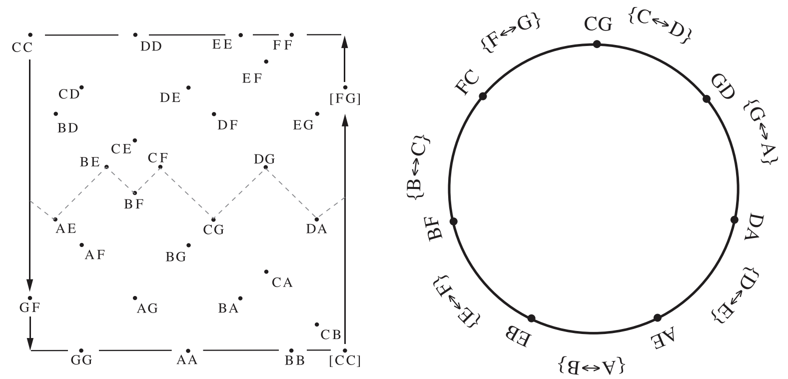

Figure 4: *Left: most efficient voice-leadings between diatonic fifths form a chain that runs through the center of the Möbius strip from Figure 1. Right: these voice leadings form an abstract circle, in which adjacent dyads are related by three-step diatonic transposition, and are linked by single-step voice leading.*

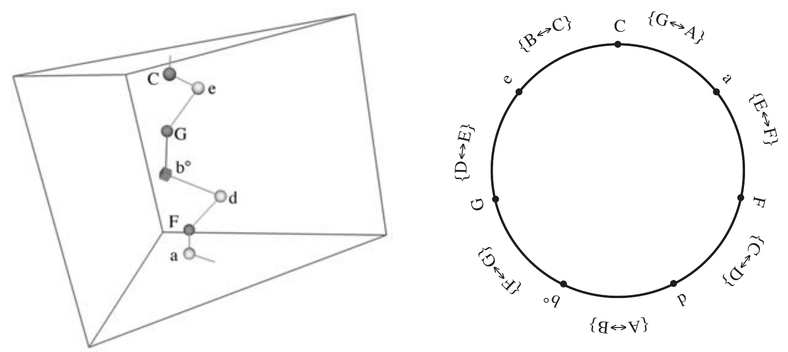

Figure 5: *Left: most efficient voice-leadings between diatonic triads form a chain that runs through the center of the orbifold representing three-note chords. Right: these voice leadings form an abstract circle, in which adjacent triads are linked by single-step voice leading. Note that here, adjacent triads are related by transposition by two diatonic steps.*

2 Voice-leading lattices and acoustic affinity

Voice-leading and acoustics seem to privilege fundamentally different conceptions of pitch distance: from a voice leading perspective, the semitone is smaller than the perfect fifth, whereas from the acoustical perspective the perfect fifth is smaller than the semitone. Intuitively, this would seem to be a fundamental gap that cannot be bridged.

Things become somewhat more complicated, however, when we consider the discrete lattices that represent voice-leading relationships among familiar diatonic or chromatic chords. For example, Figure 4 records the most efficient voice leadings among diatonic fifths–which can be represented using an irregular, one-dimensional zig-zag near the center of the Möbius strip $\mathbb{T}^2/S_2$. (The zig-zag seems to be irregular because the figure is drawn using the chromatic semitone as a unit; were we to use the diatonic step, it would be regular.) Abstractly, these voice leadings form the circle shown on the right of Figure 4. The figure demonstrates that there are purely contrapuntal reasons to associate fifth-related diatonic fifths: from this perspective {C, G} is close to {G, D}, not because of acoustics, but because the first dyad can be transformed into the second by moving the note C up by one diatonic step. One fascinating possibility–which we unfortunately cannot pursue here–is that acoustic affinities actually *derive from* voice-leading facts: it is possible that the ear asso-

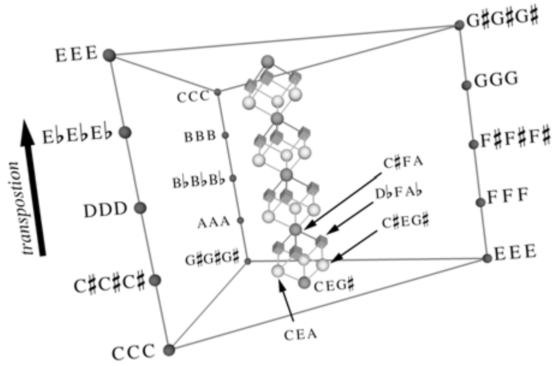

Figure 6 : *Major, minor, and augmented triads as they appear in the orbifold representing three-note chords. Here, triads are particularly close to their major-third transpositions.*

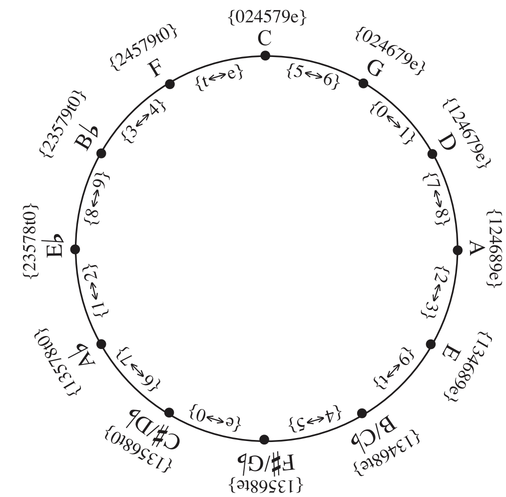

Figure 7 : *Fifth-related diatonic scales form a chain that runs through the center of the seven-dimensional orbifold representing seven-note chords. It is structurally analogous to the circles in Figures 4 and 5.*

ciates the third harmonic of a complex tone with the second harmonic of another tone a fifth above it, and the fourth harmonic of the lower note with the third of the upper, in effect tracking voice-leading relationships among the partials.

Figures 5–7 present three analogous structures: Figure 5 connects triads in the C diatonic scale by efficient voice leading, and depicts third-related triads as being particularly close; Figure 6 shows the position of major, minor, and augmented triads in three-note chromatic chord space, where major-third-related triads are close [7]; Figure 7 shows (symbolically) that fifth-related diatonic scales are close in twelve-note chromatic space. Once again, we see that there are purely contrapuntal reasons to associate fifth-related diatonic scales and third-related triads.

This observation, in turn, raises a number of theoretical questions. For instance: should we attribute the prevalence of modulations between fifth-related keys to the acoustic affinity between fifth-related pitches, or to the voice-leading relationships between fifth-related diatonic scales? One way to study this question would be to compare the frequency of modulations in classical pieces to the voice-leading distances among their associated scales. Preliminary investigations, summarized in Figure 8, suggest that voice-leading distances are in fact very closely correlated to modulation frequencies. Surprising as it may seem, the acoustic affinity of perfect fifth-related notes may be superfluous when it comes to explaining classical modulatory practice. [2]

3 Tuning lattices as approximate models of voice leading

We will now investigate the way tuning lattices like the *Tonnetz* represent voice-leading relationships among familiar sonorities. Here my argumentative strategy will by somewhat different, since it is widely recognized that the *Tonnetz* has something to do with voice leading. (This is largely due to the important work of Richard

[2] Similar points could potentially be made about the prevalence, in functionally tonal music, of root-progressions by perfect fifths. It may be that the diatonic circle of thirds shown in Figure 5 provides a more perspicuous model of functional harmony than do more traditional fifth-based representations.

		Correlation
MAJOR	Bach	.96
	Haydn	.93
	Mozart	.91
	Beethoven	.96
MINOR	Bach	.95
	Haydn	.91
	Mozart	.91
	Beethoven	.96

Figure 8 : *Correlations between modulation frequency and voice-leading distances among scales, in Bach's Well-Tempered Clavier, and the piano sonatas of Haydn, Mozart, and Beethoven. The very high correlations suggest that composers typically modulate between keys whose associated scales can be linked by efficient voice leading.*

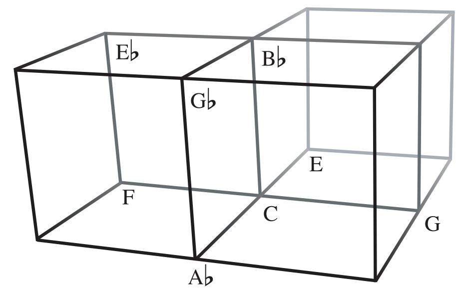

Figure 9 : *On this three-dimensional Tonnetz, the C^7 chord is represented by the tetrahedron whose vertices are C, E, G, and B♭. The $C^{ø7}$ chord is represented by the nearby tetrahedron C, E♭, G♭, B♭, which shares the C-B♭ edge.*

Cohn, who has used the *Tonnetz* to study what he calls "parsimonious" voice leading [4].) My goal will therefore be to explain why tuning lattices are only an *approximate* model of contrapuntal relationships, and only for certain chords.

The first point to note is that inversionally related chords on a tuning lattice are near each other when they share common tones.[3] For example, the *Tonnetz* represents perfect fifths by line segments; fifth-related perfect fifths, such as {C, G} and {G, D} are related by inversion around their common note, and are adjacent on the lattice (Figure 3). Similarly, major and minor triads on the *Tonnetz* are represented by triangles; inversionally related triads that share an interval, such as {C, E, G} and {C, E, A}, are joined by a common edge. (On the standard *Tonnetz*, the more common tones, the closer the chords will be: C major and A minor, which share two notes, are closer than C major and F minor, which share only one.) In the three-dimensional *Tonnetz* shown in Figure 9, where the *z* axis represents the seventh, C^7 is near its inversion $C^{ø7}$. The point is reasonably general, and does not depend on the particular structure of the *Tonnetz* or on the chords involved: on tuning lattices, inversionally related chords are close when they share common tones.[4]

The second point is that acoustically consonant chords often divide the octave relatively evenly; such chords can be linked by efficient voice leading to those inversions with which they share common notes [15, 16].[5] It follows that proximity on a tuning lattice will indicate the potential for efficient voice leading *when the chords in question are nearly even and are related by inversion*. Thus {C, G} and {G, D} can be linked by the stepwise voice leading (C, G)→(D, G), in which C moves up by two semitones. Similarly, the C major and A minor triads can be linked by the single-step voice leading (C, E, G)→(C, E, A), and C^7 can be linked to $C^{ø7}$ by the two semitone voice-leading (C, E, G, B♭)→(C, E♭, G♭, B♭). In each case the chords are

[3]This is not true of the voice leading spaces considered earlier: for example, in three-note chord space {C, D, F} is not particularly close to {F, A♭, B♭}.

[4]In the general case, the notion of "closeness" needs to be spelled out carefully, since chords can contain notes that are very far apart on the lattice. In the applications we are concerned with, chords occupy a small region of the tuning lattice, and the notion of "closeness" is fairly straightforward.

[5]The point is relatively obvious when one thinks geometrically: the two chords divide the pitch-class circle nearly evenly into the same number of pieces; hence, if any two of their notes are close, then each note of one chord is near some note of the other.

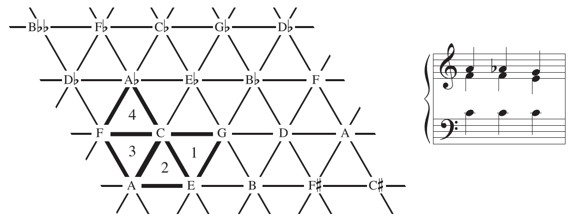

Figure 10: *On the Tonnetz, F major (Triangle 3) is closer to C major (Triangle 1) than F minor (Triangle 4) is. In actual music, however, F minor frequently appears as a passing chord between F major and C major. Note that, unlike in Figure 3, I have here used a Tonnetz in which the axes are not orthogonal; this difference is merely orthographical, however.*

also close on the relevant tuning lattice. (Interestingly, triadic distances on the diatonic *Tonnetz* in Figure 3 exactly reproduce the circle-of-thirds distances from Figure 5.) This will not be true for uneven chords: {C, E} and {E, G♯} are close on the *Tonnetz*, but cannot be linked by particularly efficient voice leading; the same holds for {C, G, A♭} and {G, A♭, D♭}. Tuning lattices are approximate models of voice-leading only when one is concerned with the nearly-even sonorities that are fundamental to Western tonality.

Furthermore, on closer inspection *Tonnetz*-distances diverge from voice-leading distances even for these chords. Some counterexamples are obvious: for instance, {C, G} and {C♯, F♯} can be linked by semitonal voice leading, but are fairly far apart on the *Tonnetz*. Slightly more subtle, but more musically pertinent, is the following example: on the *Tonnetz*, C major is two units away from F major but *three* units from F minor (Figure 10). (Here I measure distance in accordance with "neo-Riemannian" theory, which considers triangles sharing an edge to be one unit apart and which decomposes larger distances into sequences of one-unit moves.) Yet it takes only two semitones of total motion to move from C major to F minor, and three to move from C major to F major. (This is precisely why F minor often appears as a passing chord between F major and C major.) The *Tonnetz* thus depicts F major as being closer to C major than F minor is, even though contrapuntally the opposite is true. This means we cannot use the figure to explain the ubiquitous nineteenth-century IV-iv-I progression, in which the two-semitone motion $\hat{6}\to\hat{5}$ is broken into a pair of single-semitone steps $\hat{6}\flat\hat{6}\to\hat{5}$.

One way to put the point is that while adjacencies on the *Tonnetz* reflect voice-leading facts, other relationships do not. As Cohn has emphasized, two major or minor triads share an edge if they can be linked by "parsimonious" voice-leading in which a single voice moves by one or two semitones. If we are interested in this particular kind of voice leading then the *Tonnetz* provides an accurate and useful model. However, there is no analogous characterization of larger distances in the space. In other words, we do not get a recognizable notion of voice-leading distance by "decomposing" voice leadings into sequences of parsimonious moves: as we have seen, (F, A, C)→(E, G, C) can be decomposed into *two* parsimonious moves, while it takes *three* to represent (F, A♭, C)→(E, G, C); yet intuitively the first voice leading is larger than the second. The deep issue here is that it is problematic to assert that "parsimonious" voice leadings are always smaller than nonparsimonious voice-leadings: by asserting that (C, E, A)→(C, E, G) is smaller than (C, F, Af)→(C, E, G), the theorist runs afoul what Tymoczko calls "the distribution constraint," known to mathematicians as the submajorization partial order [15, 8].[6] Tymoczko argues that violations of the distribution constraint invariably produce distance measures that violate intuitions about voice leading; the problem with larger distances on the *Tonnetz* is an illustration of this general point.

Nevertheless, the fact remains that the two kinds of distance are roughly consistent: for major and minor triads, the correlation between *Tonnetz* distance and voice-leading distance is a reasonably high .79.[7]

[6] Metrics that violate the distribution constraint have counterintuitive consequences, such as preferring "crossed" voice leadings to their uncrossed alternatives. Here, the claim that A minor is closer to C major than F minor leads to the F minor/F major problem discussed in Figure 10.

[7] Here I use the L^1 or "taxicab" metric. The correlation between *Tonnetz* distances and the number of shared common tones is an even-higher .9; however, "number of shared common tones" is not interpretable as a voice-leading metric.

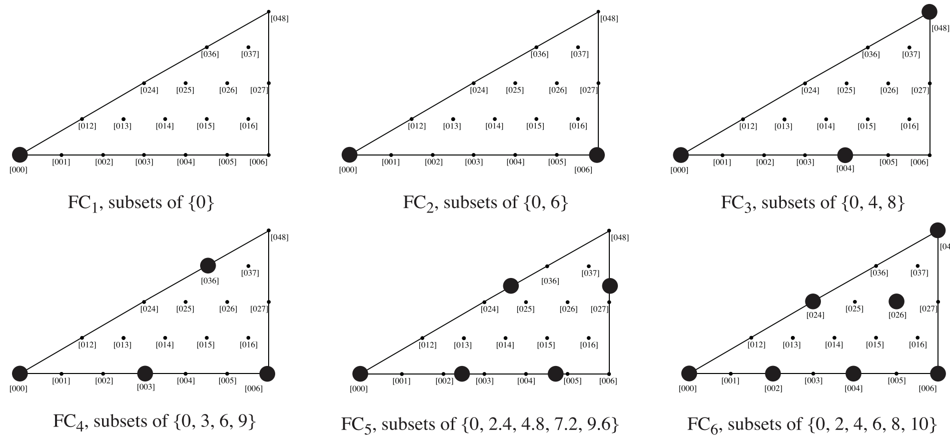

Figure 11 : *The magnitude of a set class's nth Fourier component is approximately linearly related to the size of the minimal voice leading to the nearest subset of the perfectly even n-note chord, shown here as dark spheres.*

Furthermore, since Tymoczko's "distribution constraint" is not intuitively obvious, unwary theorists might well think that they could declare the "parsimonious" voice leading (C, E, G)→(C, E, A) to be smaller than the nonparsimonious (C, E, G)→(C♯, E, G♯). (Indeed, the very meaning of the term "parsimonious" would seem to suggest that some theorists have done so.) Consequently, *Tonnetz*-distances might well appear, at first or even second blush, to reflect some reasonable notion of "voice-leading distance"; and this in turn could lead the theorist to conclude that the *Tonnetz* provides a generally applicable tool for investigating triadic voice-leading. I have argued that we should resist this conclusion: if we use the *Tonnetz* to model chromatic music, than Schubert's major-third juxtapositions will seem very different from his habit of interposing F minor between F major and C major, since the first can be readily explained using the *Tonnetz* whereas the second cannot [6]. The danger, therefore, is that we might find ourselves drawing unnecessary distinctions between these two cases–particularly if we mistakenly assume the *Tonnetz* is a fully faithful model of voice-leading relationships.

4 Voice leading, "quality space," and the Fourier transform

We conclude by investigating the relation between voice leading and the Fourier-based perspective [14, 9, 2]. The mechanics of the Fourier transform are relatively simple: for any number n from 1 to 6, and every pitch-class p in a chord, the transform assigns a two-dimensional vector whose components are

$$V_{p,n} = (\cos(2\pi pn/12), \sin(2\pi pn/12))$$

Adding these vectors together, for one particular n and all the pitch-classes p in the chord, produces a composite vector representing the chord as a whole–its "nth Fourier component." The length (or "magnitude") of this vector, Quinn observes, reveals something about the chord's harmonic character: in particular, chords saturated with (12/n)-semitone intervals, or intervals approximately equal to 12/n, tend to score highly on this index of chord quality.[8] The Fourier transform thus seems to quantify the intuitive sense that chords can be more-or-less diminished-seventh-like, perfect-fifthy, or whole-toneish. Interestingly, "Z-related" chords–or chords with the same interval content–always score identically on this measure

[8]Here I use continuous pitch-class notation where the octave always has size 12, no matter how it is divided. Thus the equal-tempered five-note scale is labeled {0, 2.4, 4.8, 7.2, 9.6}.

of chord-quality. In this sense, Fourier space (the six-dimensional hypercube whose coordinates are the Fourier magnitudes) seems to model a conception of similarity that emphasizes interval content, rather than voice leading or acoustic consonance.

However, there is again a subtle connection to voice leading: it turns out that the magnitude of a chord's *n*th Fourier component is approximately linearly related to the (Euclidean) size of the minimal voice leading to *the nearest subset* of any perfectly even *n*-note chord.[9] For instance, a chord's first Fourier component (FC_1) is approximately related to the size of the minimal voice leading to any transposition of $\{0\}$; the second Fourier component is approximately related to the size of the minimal voice leading to any transposition of either $\{0\}$ or $\{0, 6\}$; the third component is approximately related to the size of the minimal voice leading to any transposition of either $\{0\}$, $\{0, 4\}$ or $\{0, 4, 8\}$, and so on. Figure 11 shows the location of the subsets of the *n*-note perfectly even chord, as they appear in the orbifold representing three-note set-classes, for values of *n* ranging from 1 to 6 [1, 15, 3]. Associated to each graph is one of the six Fourier components. For any three-note set class, the magnitude of its *n*th Fourier component is a decreasing function of the distance to the nearest of these marked points: for instance, the magnitude of the third Fourier component (FC_3) decreases, the farther one is from the nearest of $\{0\}$, $\{0, 4\}$ and $\{0, 4, 8\}$. Thus, chords in the shaded region of Figure 12 will tend to have a relatively large FC_3, while those in the unshaded region will have a smaller FC_3. Figure 13 shows that this relationship is very-nearly linear for twelve-tone equal-tempered trichords.

Table 1 uses the Pearson correlation coefficient to estimate the relationship between the voice-leading distances and Fourier components, for twelve-tone equal-tempered multisets of various cardinalities. The strong anti-correlations indicate that one variable predicts the other with a very high degree of accuracy. Table 2 calculates the correlation coefficients for three-to-six-note chords in 48-tone equal temperament. These strong anticorrelations, very similar to those in Table 1, show that there continues to be a very close relation between Fourier magnitudes and voice-leading size in very finely quantized pitch-class space. Since 48-tone equal temperament is so finely quantized, these numbers are approximately valid for continuous, unquantized pitch-class space.[10]

Explaining these correlations, though not very difficult, is beyond the scope of this paper. From our perspective, the important question is whether we should measure chord quality using the Fourier transform or voice leading. In particular, the issue is whether the Fourier components model the musical intuitions we want to model: as we have seen, the Fourier transform requires us to measure a chord's "harmonic quality" in terms of its distance from *all* the subsets of the perfectly even *n*-note chord. But we might sometimes wish to employ a different set of harmonic prototypes. For instance, Figure 14 uses a chord's distance from the augmented triad to measure the trichordal set classes' "augmentedness." Unlike Fourier analysis, this purely voice-leading-based method does not consider the triple unison or doubled major third to be particularly "augmented-like"; hence, set classes like $\{0, 1, 4\}$ do not score particularly highly on this index of "augmentedness." This example dramatizes the fact that, when using voice leading, we are free to choose any set of harmonic prototypes, rather than accepting those the Fourier transform imposes on us.

5 Conclusion

The approximate consistency between our three models is in one sense good news: since they are closely related, it may not matter much–at least in practical terms–which we choose. We can perhaps use a tuning lattice such as the *Tonnetz* to represent voice-leading, as long as we are interested in gross contrasts ("near"

[9] Here I measure voice-leading using the Euclidean metric [1, 15, 16].

[10] It would be possible, though beyond the scope of this paper, to calculate this correlation analytically. It is also possible to use statistical methods for higher-cardinality chords. A large collection of randomly generated 24- and 100-tone chords in continuous space produced correlations of .95 and .94, respectively.

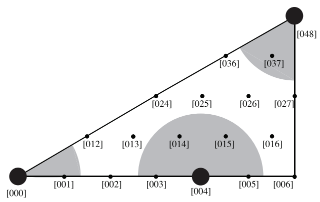

Figure 12: *Chords in the shaded region will have a large FC_3 component, since they are near subsets of {0, 4, 8}. Those in the unshaded region will have a smaller FC_3 component.*

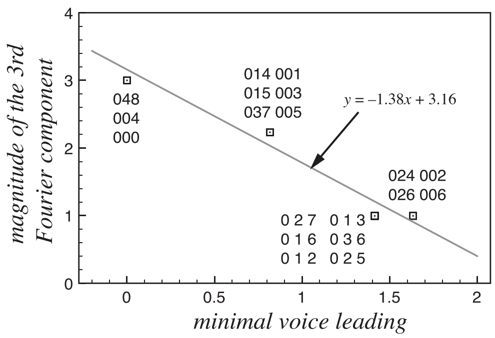

Figure 13: *For trichords, the equation* $FC_3 = -1.38VL + 3.16$ *relates the third Fourier component to the Euclidean size of the minimal voice leading to the nearest subset of* {0, 4, 8}.

Table 1: *Correlations between voice-leading distances and Fourier magnitudes.*

	FC_1	FC_2	FC_3	FC_4	FC_5	FC_6
Dyads	-.97	-.96	-.97	-1	-.97	-1*
Trichords	-.98	-.97	-.97	-.98	-.98	-1*
Tetrachords	-.96	-.96	-.95	-.98	-.96	-1*
Pentachords	-.96	-.96	-.95	-.98	-.96	-1*
Hexachords	-.96	-.96	-.95	-.96	-.96	-1*
Septachords	-.96	-.96	-.96	-.97	-.96	-1*
Octachords	-.96	-.96	-.95	-.98	-.96	-1*
Nonachords	-.96	-.96	-.96	-.98	-.96	-1*
Decachords	-.96	-.96	-.96	-.98	-.96	-1*

* Voice leading calculated using L^1 (taxicab) distance rather than L^2 (Euclidean).

Table 2: *Correlations between voice-leading distances and Fourier magnitudes in 48-tone equal temperament.*

	FC_1
Trichords	-.99
Tetrachords	-.97
Pentachords	-.97
Hexachords	-.96

vs. "far") rather than fine quantitative differences ("3 steps away" vs. "2 steps away"). Similarly, we can perhaps use voice-leading spaces to approximate the results of the Fourier analysis, as long as we are interested in modeling generic harmonic intuitions ("very fifthy" vs. "not very fifthy") rather than exploring very fine differences among Fourier magnitudes.

However, if we want to be more principled, then we need to be more careful. The resemblances among our models mean that it is possible to inadvertently use one sort of structure to discuss properties that are more directly modeled by another. And indeed, the recent history of music theory displays some fascinating (and very fruitful) imprecision about this issue. It is striking that Douthett and Steinbach, who first described several of the lattices found in the center of the voice-leading orbifolds–including Figure 6–explicitly presented their work as generalizing the familiar *Tonnetz* [7]. Their lattices, rather than depicting *parsimonious* voice leading among major and minor triads, displayed *single-semitone* voice leadings among a wider range of sonorities; and as a result of this seemingly small difference, they constucted models in which *every distance* can be interpreted as representing voice-leading size. However, this difference only became apparent after it was understood how to embed their discrete structures in the continuous geometrical figures described at the beginning of this paper. Thus one could say that the continuous voice-leading spaces evolved out of the *Tonnetz*, by way of Douthett and Steinbach's discrete lattices, even though the structures now appear to be fundamentally different. Related points could be made about Quinn's "quality space," whose connection to the voice-leading spaces took several years–and the work of several authors–to clarify.

There is, of course, nothing wrong with this: knowledge progresses slowly and fitfully. But our inves-

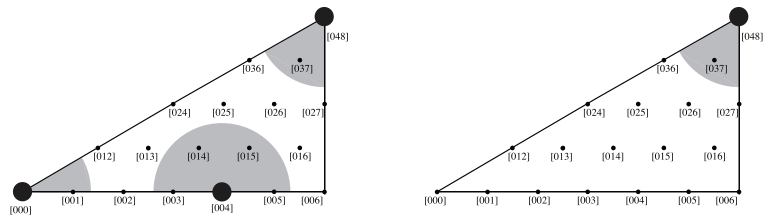

Figure 14 : *The mathematics of the Fourier transform requires that we conceive of "chord quality" in terms of the distance to all subsets of the perfectly even n-note chord (left). Purely voice-leading-based conceptions instead allow us to choose our harmonic prototypes freely (right). Thus we can use voice leading to model a chord's "augmentedness" in terms of its distance from the augmented triad, but not the tripled unison* $\{0, 0, 0\}$ *or the doubled major third* $\{0, 0, 4\}$.

tigation suggests that we may want to think carefully about which model is appropriate for which music-theoretical purpose. I have tried to show that the issues here are complicated and subtle: the mere fact that tonal pieces modulate by fifth does not, for example, require us to use a tuning lattice in which fifths are smaller than semitones. (Indeed, the "circle of fifths" C-G-D-... can be interpreted either as a one-dimensional tuning lattice incorporating octave equivalence, or as a diagram of the voice-leading relations among diatonic scales, as in Figure 7.) Likewise, there may be close connections between voice-leading spaces and the Fourier transform, even though the latter associates "Z-related" chords while the former does not. The present paper can be considered a down-payment toward a more extended inquiry, one that attempts to determine the relative strengths and weaknesses of our three different-yet-similar conceptions of musical distance.

References

[1] Clifton Callendar. Continuous transformations. *Music Theory Online*, 10(3), 2004.

[2] Clifton Callendar. Continuous harmonic spaces. Unpublished, 2007.

[3] Clifton Callendar, Ian Quinn, and Dmitri Tymoczko. Generalized voice-leading spaces. *Science*, 320:346–348, 2008.

[4] Richard Cohn. Neo-Riemannian operations, parsimonious trichords, and their 'tonnetz' representations. *Journal of Music Theory*, 41(1):1–66, 1997.

[5] Richard Cohn. Introduction to neo-Riemannian theory: A survey and a historical perspective. *Journal of Music Theory*, 42(2):167–180, 1998.

[6] Richard Cohn. As wonderful as star clusters: Instruments for gazing at tonality in Schubert. *Nineteenth-Century Music*, 22(3):213–232, 1999.

[7] Jack Douthett and Peter Steinbach. Parsiomnious graphs: a study in parsimony, contextual transformations, and modes of limited transposition. *Journal of Music Theory*, 42(2):241–263, 1998.

[8] Rachel Hall and Dmitri Tymoczko. Poverty and polyphony: a connection between music and economics. In R. Sarhangi, editor, *Bridges: Mathematical Conections in Art, Music, and Science*, 2007.

[9] Justin Hoffman. On pitch-class set cartography. Unpublished, 2007.

[10] David Lewin. Re: Intervallic relations between two collections of notes. *Journal of Music Theory*, 3:298–301, 1959.

[11] David Lewin. Special cases of the interval function between pitch-class sets X and Y. *Journal of Music Theory*, 45:1–29, 2001.

[12] Ian Quinn. General equal tempered harmony (introduction and Part I). *Perspectives of New Music*, 44(2):114–158, 2006.

[13] Ian Quinn. General equal tempered harmony (Parts II and III). *Perspectives of New Music*, 45(1):4–63, 2007.

[14] Thomas Robinson. The end of similarity? semitonal offset as similarity measure, 2006. Paper presented at the annual meeting of the Music Theory Society of New York State, Saratoga Springs, NY.

[15] Dmitri Tymoczko. The geometry of musical chords. *Science*, 313:72–74, 2006.

[16] Dmitri Tymoczko. Scale theory, serial theory, and voice leading. *Music Analysis*, 27(1):1–49, 2008.