Journal of Mathematics and Music, 2013
Vol. 7, No. 2, 127–144, http://dx.doi.org/10.1080/17459737.2013.818724

Geometry and the quest for theoretical generality

Dmitri Tymoczko*

Music, Princeton University, 310 Woolworth Center, Princeton, NJ 08544, USA

(*Received 5 February 2013; final version received 19 June 2013*)

This paper reconsiders scale-theoretical ideas from the standpoint of voice leading. I begin by showing that Clough and Douthett's definition of 'maximal evenness' makes covert reference to voice-leading distances, albeit disguised by the form of their equations. I then argue that theirs is a special case of the broader problem of quantizing continuous chords to a scale, deriving some new results in this more general setting – including an analogue of the 'cardinality equals variety' property. In the second part of the paper, I show how Clough and Douthett's 'J-function' began as a device for generating maximally even collections, only later evolving into a tool for studying the voice leadings between them. Particularly important here are Julian Hook's 'signature transformations', which I generalize to a wider range of collections. I conclude with a few remarks about history and methodology in music theory.

Keywords: voice leading; geometry; maximal evenness; signature transformations; transformational theory

In what follows, I want to retrace the history of thinking about 'maximally even' sets from a modern perspective, detailing how the investigation makes covert reference to voice leading at several points. Along the way, I will show that recent ideas allow us to extend familiar results, opening up new questions about quantization and the generalized signature transformation. This historical reconstruction is meant as a case study in the generalizing power of geometry, one that could in principle be replicated for many other areas of theory; indeed, I have already made analogous arguments in [1] (which uses geometry to analyse key signatures), [2] (using voice leading to model twentieth-century modulation), [3] (treating K nets), [4] (treating K nets, contour, and set-class similarity), [5] (dealing with set-class similarity and the Fourier transform), and [6] (generalizing the traditional *Tonnetz*). Taken together, these papers form a larger argument that voice leading should play a central role in the music theorist's toolkit.

1. Quantization, voice leading, and the generalized J-function

Intuitively, a 'maximally even' n-note chord, residing in scale s, is a chord in s that is *as close as possible* to the nearest of the perfectly even chords that divide the octave into n precisely equal pieces.[1] (These perfectly even chords will not in general lie within the scale.[2]) In exploring maximal evenness, we might therefore begin by (a) specifying the notion of distance implicit in

*Email: dmitri@princeton.edu

the phrase ‘as close as possible’; and (b) formulating the *quantization function* that sends any set of pitch classes (whether in the scale or not) to the nearest scale-chord.[3] With these definitions in hand, we could then catalogue the results of quantizing the perfectly even chords. This, however, is not the strategy pursued in the important scale-theoretical work of John Clough and Jack Douthett. Rather than defining a general distance-measure, they instead provide a metaphor about electrons in a circular space [7, p. 96], along with a formal definition of ‘maximal evenness’ that does not obviously connect with any antecedent intuitions [7, Def. 1.7].[4] The result is a subtle logical or metalogical gap between the underlying mathematical formalism (unobjectionable in its own terms) and the intuitions that formalism is meant to represent.

Let us therefore pursue an alternative path to the maximally even collections. We begin by assigning every fundamental frequency f a pitch label p, as follows:

$$p = r + c\log_2\left(\frac{f}{440}\right). \tag{1}$$

Here c is an integer that determines the size of the octave, r is an arbitrary real number that determines the value of A440, and every conceivable pitch is given a label. We can define *scale tones* as those pitches with integer labels. (In more general applications we could consider non-equal tempered scales that are piecewise combinations of equal-tempered scales, joined at scale tones, but we will not need that complication here. These non-equal-tempered scales define metrics on pitch space according to which their steps are size 1.) If we adopt the absolute-value metric on $\mathbb{R}$, Equation (1) yields a familiar measure of musical distance (‘pitch-space distance’) whose units are ‘semitones’ or ‘scale steps’. This in turn allows us to write down the ‘quantization function’ that sends a continuous pitch p to the nearest integer value q:

$$q = \lfloor p + 0.5 \rfloor. \tag{2}$$

The additive factor 0.5 ensures that notes move by the smallest possible distance: those with fractional part ≥ 0.5 are rounded up while others are rounded down. This factor will be familiar to anyone who has ever programmed a computer to convert a floating-point number into an integer.

A *voice-leading metric* extends this distance measure to n-tuples of pitches. Suppose, for example, instrument 1 plays note n_1 at time t_1 and n_2 at t_2, while instrument 2 plays m_1 at time t_1 and m_2 at t_2, so that the two instruments articulate the musical lines $n_1 \to n_2$ and $m_1 \to m_2$. These lines imply two distances, $d_1 = \mathcal{D}(n_1, n_2)$ and $d_2 = \mathcal{D}(m_1, m_2)$. We might wonder whether the pair of distances $\{d_1, d_2\}$ is collectively larger or smaller than some other pair $\{d_3, d_4\}$. Measures of *voice-leading size* allow us to answer this question. Formally, a measure of voice-leading size is a *strongly isotone function*, a function from n-tuples of non-negative numbers $(d_1, d_2, \ldots, d_n)$ to real numbers, subject to the following constraints:

(P1) It is completely symmetric in all arguments;
(P2) It is nondecreasing in each argument; and
(P3) It is consistent with the submajorization partial order.

(For details see [8,9], and [10], which treats the more general case where we have a *partial order* defined on n-tuples of real numbers.) With these definitions in hand, we can write down the ‘generalized quantization function’ that sends any continuous chord $P = \{p_i\}$ to the nearest scalar chord $Q = \{q_i\}$

$$Q = \{\lfloor p_i + 0.5 \rfloor\}. \tag{3}$$

Since this moves each chord tone to the nearest chromatic scale-tone, property P2 assures us that it sends the chord as a whole (represented by a point in a higher-dimensional configuration space) to the scalar chord that is as close as possible.

Given an octave of size c, we therefore quantize the perfectly even n-note chord $P = \{0c/n, 1c/n, 2c/n, \ldots\}$ as follows:

$$q_i = \left\lfloor \frac{ic}{n} + m + 0.5 \right\rfloor, \ 0 \leq i < n. \tag{4}$$

Here m is a 'mode index' that allows us to transpose the chord arbitrarily. To explore the maximally even chords in scale s, we simply need to consider all the chords resulting from an equation such as (4). Fortunately, Equation (4) is equivalent to Clough and Douthett's 'J-function' but for two trivial differences: the additive constant 0.5 and a scaling factor applied to m.[5] Their paper thus completes our derivation of maximally even sets by describing the collections that can result from the quantization function (4).

We have therefore filled the logical hole in Clough and Douthett's paper, answering the question: "according to *what notion of distance* are the 'maximally even' chords as close as possible to perfectly even chords?" Our response is: according to any notion of distance that respects properties P1 and P2 above, and *a fortiori* any notion of voice-leading distance. We have also seen that the derivation of maximally even sets is just a special case of the important musical problem of *quantization*, in which we are asked to find the scalar chord X that is *as close as possible* to some continuous chord Y.[6] This in turn is a special case of an even more general problem, that of finding the smallest voice leading between two chords or set classes. We conclude that voice leading and maximal evenness are intimately related: the notion of 'maximal evenness' implicitly involves reference to some notion of musical distance among chords, and the form of Clough and Douthett's 'J-function' indicates that the relevant notion includes all reasonable measures of voice-leading distance; indeed, the 'J-function' (modulo the additive factor of 0.5) is precisely the quantization function appropriate for all reasonable voice leading metrics. Unfortunately, these connections are not widely appreciated among mathematical music theorists. In large part, this is because Clough and Douthett absorb the factor of 0.5 in their versions of Equations (2)–(4), thus hiding the link between the J-function and the intuitive notion of 'moving to the nearest scale tone'.[7] A likely secondary factor is the reluctance to consider how discrete equal-tempered pitches are inherently situated within continuous, perceptually real log-frequency space.[8]

With these preliminaries out of the way, we can explore how quantization converts an underlying process of *continuous transposition* (represented, for example, by varying the parameter m in the J-function [Equation (4)]) into a series of single-step voice leadings in discrete space. Imagine that we have some continuous chord $C = \{c_i\}$, and consider the fractional values $f_i = c_i - \lfloor c_i \rfloor$. As Figure 1 shows, quantization sends fractional values in the range [0, 0.5) downward to the next-lowest pitch and those in the range [0.5, 1) upward to the next highest pitch. Now imagine transposing our chord continuously downward: for very small transpositions, nothing will happen to our quantized chord (assuming no chord tone has a fractional value of exactly 0.5). Eventually, however, some fractional value will cross from the range [0.5, 1) to the range [0, 0.5), at which point its quantized value will shift downward by semitone. This process will repeat for each voice until the entire chord has moved down by semitone. Quantization therefore transforms continuous descending transposition (applied to an unquantized chord-in-continuous-space) into a series of *discrete voice leadings* semitonally lowering each note of the quantized chord in

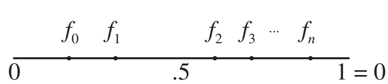

Figure 1. The fractional parts of a continuous chord. Those in the range [0, 0.5) move downward to the next-lower scale tone, while those in the range [0.5, 1) move upward to the next-higher scale tone. As the continuous chord is transposed downward, the quantized form of each note descends semitonally when its fractional part crosses 0.5.

(a)

{0, 2, 4, 6, 7, 9, 10}	for t_f in [0.09, 0.46)	(*acoustic scale*)
{0, 3, 4, 6, 7, 9, 10}	for t_f in [0.46, 0.48)	(*octatonic subset*)
{0, 3, 4, 6, 8, 9, 10}	for t_f in [0.48, 0.50)	
{0, 2, 3, 5, 7, 8, 9}	for t_f in [0.50, 0.64)	
{0, 2, 4, 5, 7, 8, 9}	for t_f in [0.64, 0.81)	
{0, 2, 4, 5, 7, 8, 10}	for t_f in [0.81, 0.99)	(*acoustic scale, second mode*)
{0, 2, 4, 6, 7, 8, 10}	for t_f in [0.99, 1) or [0, 0.09)	(*whole-tone plus one*)

(b)

{0, 4, 7}	for t_f in [0, .48) or [.64, 0)	(*major triad*)
{0, 4, 8}	for t_f in [.48, .5)	(*augmented triad*)
{0, 3, 7}	for t_f in [.5, .64)	(*minor triad*)

Figure 2. (a) The 12-tone equal-tempered quantizations of the first seven pitch classes in the harmonic series $\{0.00, 2.04, 3.86, 5.51, 7.02, 8.41, 9.69\}$, expressed as a function of the transposition's fractional value t_f. (b) The 12-tone equal-tempered quantizations of the just major triad $\{0, 3.86, 7.02\}$. Only 2% of the transpositions quantize to the augmented triad, whereas 84% quantize to the major triad.

turn.[9] (Observe that the ordering of the fractional parts will determine the order of the voices' descent, a point that will be relevant momentarily.) Historically, this is important because rigorous definitions of voice leading arrived only recently; for a time, these sorts of ad hoc models – featuring quantization and the J-function – provided important tools for investigating certain kinds of voice leading.[10]

Musically, this means that it is wrong to ask for 'the' quantization of a set class. (For instance: 'what is *the* twelve-tone equal-tempered set class corresponding to the first seven pitch classes in the harmonic series [or the Pythagorean major scale, etc.]?') In general, the continuous transpositions of an n-note chord will quantize to n different set classes, though this number can be smaller if the chord is symmetrical or nearly so, or if any of the its notes share the same fractional values (Figure 2).[11] (As we will see, this fact is closely related to the scale-theoretical result that n-note diatonic set classes come in n chromatic forms.) Instead we should ask *what range of transpositions* quantizes to a particular chromatic chord-type. In Figure 2(b), for example, a wide range of transpositions quantize to the equal-tempered major triad, with only a very small range quantizing to the augmented triad.[12] Both intuitively and geometrically, the unquantized chord is 'closer' to a quantized chord if the associated range of transpositions is wider.[13]

To understand this geometrically, we consider transpositional set-class space or 'chord-type space': this is the n-torus $\mathbb{T}^n$, representing ordered tuples of pitch classes, modulo translation and reordering, or $\mathbb{T}^{n-1}/\mathcal{S}_n$ (see [4]). Every continuous set class lies within a simplicial region M, which I will call the 'minimal equal-tempered simplex'; this, as the name suggests, is the smallest simplex whose vertices are all equal-tempered set classes (Figure 3).[14] For any set class C in this region, and any of the vertices of the minimal equal-tempered simplex M, we can find a transposition of C that quantizes to that vertex; quantizing a continuous descending series of transpositions sends the set class to each of the vertices of M in turn. To see this, start with an equal-tempered transpositional set class $C = \{c_i\}$. Now consider the simplicial region M formed by adding to the coordinates c_i fractional values f_i satisfying the inequalities[15]

$$f_a \leq f_b \leq f_c \leq \cdots \leq f_n \leq f_a + 1, \quad \sum f_i = 0. \tag{5}$$

These fractional parts are arbitrarily ordered, so that f_0 (the fractional part to be added to note c_0) may be larger or smaller than f_1. As we require, this simplex's boundary is traversed by lowering each note of the set class semitonally in turn; this is because the sum of the consecutive distances $f_b - f_a, f_c - f_b, \ldots, f_a + 1 - f_n$ must be equal to 1, with each fractional part at most one semitone above its predecessor. The vertices therefore have the form

$$f_a = f_b = \cdots = f_i = f_j - 1 = f_k - 1 = \cdots = f_n - 1. \tag{6}$$

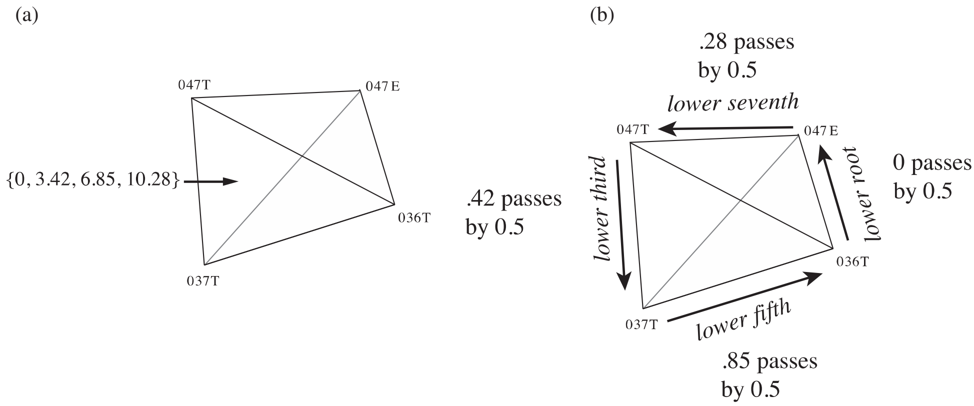

Figure 3. (a) The minimal equal-tempered simplex in four-note set-class space, $\mathbb{T}^3/\mathcal{S}_4$, containing the transpositional set class $\{0, 3.42, 6.85, 10.28\}$. ('T' and 'E' abbreviate 10 and 11.) Quantizing the set $\{0, 3.42, 6.85, 10.28\}$ gives us $\{0, 3, 7, 10\}$ or 037T. Ordering the notes by fractional distance above 0.5, we obtain (6.85, 0, 10.28, 3.42). (b) As we transpose downward, we lower voices in this order, producing 036T, 047E, 047T, 037T, returning to the original set class.

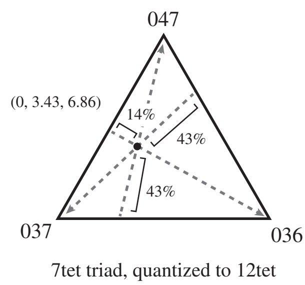

Figure 4. The 7-tone equal-tempered triad $\{0, 3.43, 6.86\}$. Its barycentric coordinates are (0.43, 0.43, 0.14), representing the fraction of the distance from each side to the opposite vertex. These numbers are equal to the fraction of transpositions quantizing to that vertex.

(Recall also that we are in set class space, where we discard translations.) Turning back to Figure 1, we see that Equation (6) simply expresses the effect of quantizing some fractional parts 'downward' and others 'upward' ($f_a, f_b, \ldots, f_i$ and $f_j, f_k, \ldots, f_n$ respectively). As discussed previously, continuous transposition will move the boundary between notes that quantize downward and notes that go upward, so that it cycles through each pair of adjacent fractional parts. This implies that quantization sends the continuous set class to each vertex of the minimal equal-tempered simplex.

In general, each equal-tempered set class of size n is a vertex of $n!$ such 'minimal equal-tempered simplexes', corresponding to the $n!$ ways in which its notes can be sequentially lowered. The closer a continuous chord is to a vertex, the more of its transpositions will quantize to it; in fact, the set class's barycentric coordinates record the proportion of transpositions quantizing to each vertex (Figure 4).[16] In the special case where the continuous chord divides the octave perfectly evenly, it will lie within a simplex (or subsimplex) whose vertices are transpositionally related. (This follows from the Chinese Remainder Theorem.[17]) The three possibilities are shown in Figure 5. In the first case, where the size of the chord divides the size of the scale, the chord is found on a vertex (0-dimensional subsimplex) of the minimal equal-tempered simplex; this means that the J-function (upon continuously varying the parameter m) defines a series of 'purely

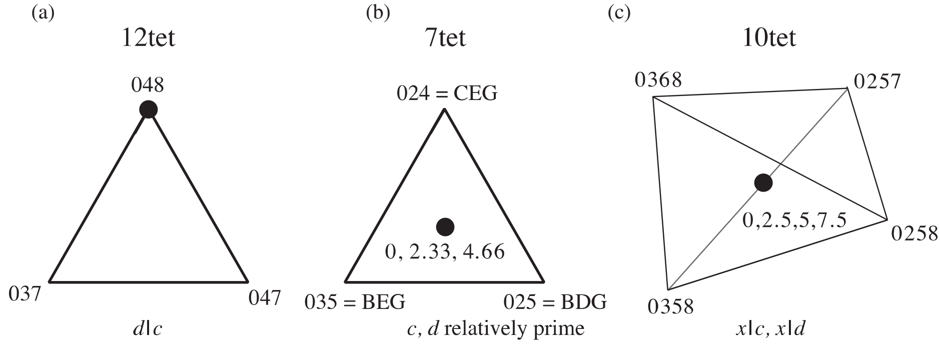

Figure 5. The perfectly even chord can either be found on the vertex of a minimal equal-tempered simplex (a), in the centre of the simplex (b), or in the centre of an edge (or more generally, facet) (c). These correspond to the cases where the size of the chord divides the size of the scale, where the two numbers are relatively prime, and where they share a common factor, respectively.

transpositional' voice leadings that move all notes in the chord by semitone. In the second, where the two sizes are relatively prime, the chord is found in the very centre of the minimal equal-tempered simplex. Here, the J-function voice leadings produce a 'generalized circle of fifths' (or 'signature transform') linking the chord's transpositions by single-semitone voice leadings.[18] And in the third, where the size of the chord shares a common factor with the size of the scale, the chord is found in the centre of one of the minimal equal-tempered simplex's subsimplexes (facets, edges, etc.). In this case, the J-function produces a transpositionally symmetrical version of the 'generalized circle of fifths' where multiple voices, forming a completely even chord, move semitonally at each step.[19] In all three cases, the resulting voice leadings move down the centre of chord space, connecting the scalar chords that are as even as possible.

Two final points. First, it is natural to wonder how to extend these results to the case where we quantize to a scale that is not perfectly even. (NB: when we use the metric defined by the scale itself, this question is trivial, since every scale is perfectly even according to its own metric; the interesting question arises when we are quantizing to one scale but measuring distance with another, as when we quantize to the diatonic but measure intervals using semitones.) The basic principle is that the more uneven the scale is, the more various will be the quantized forms of a continuous set class; rather than cycling around the vertices of a single simplex, quantization will send the chord through a wider region of set-class space. Finding a general theory here is an open problem, as is understanding repeated quantizations to multiple scales (cf. Douthett and Plotkin's work on 'filtered point-symmetry', to be discussed shortly). It seems plausible that one could derive some interesting quantitative relationships between the evenness of the quantizing scale and the range of set classes that can result from quantization.

Second, recall that Clough, Douthett, and Myerson were motivated by the desire to find *generalized versions* of the 7-in-12 diatonic scale. (Indeed 'maximal evenness' is supposed to be a generalization of diatonicity.) In previous work, I have extended this project by providing analogues of the *melodic* and *harmonic minor* scales whenever the size of our chord (or 'generalized diatonic scale') is relatively prime to the size of the chromatic scale containing it. The basic idea here, discussed in Chapters 3, 4, and 7 of [9], is to scramble the voice leadings on the generalized circle of fifths, as shown in Figure 6 (reprinted from [9]). Thus if one is concerned to identify generalized analogues of familiar scales, we can move beyond the maximally even collections to a broader range of *nearly* or *relatively even* collections, all linked by structurally analogous voice leadings.[8,9]

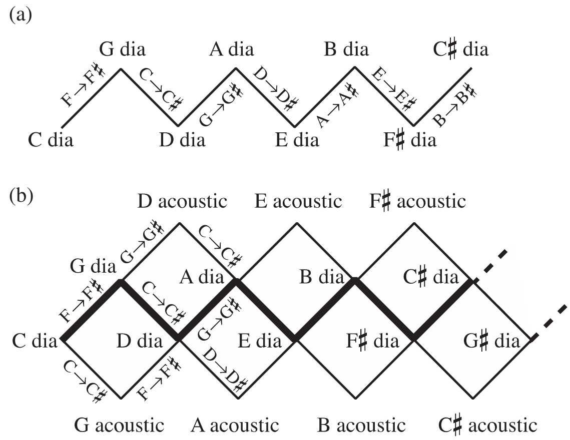

Figure 6. (a) The circle of fifths connects diatonic scales by single-semitone voice leading. (b) We can generate the acoustic scale (or ascending melodic minor) by reversing the order of every pair of voice leadings, for instance by letting C → C♯ operate on C diatonic before F → F♯. We can create analogues of the acoustic scale whenever the size of our collection is relatively prime to the size of the chromatic scale containing it. To generate analogues of the harmonic major and minor scales we scramble the order of any *three* consecutive voice-leading motions.

2. From the *J*-function to scalar transpositions

I now want to trace the evolution of Clough and Douthett's '*J*-function' from a device for tabulating maximally even sets into a model for voice leading more generally, focusing on the role played by Julian Hook's 'signature transformations'. In 2003, Hook defined 'transformations' that change diatonic passages by adding one sharp or flat to their key signatures.[12] (We assume that the music contains no additional accidentals beyond those in the signature.) Though it was not clear at the time – and certainly not part of his early presentations on the subject – this process can be understood as the repeated application of a particular *voice leading schema* (i.e. 'raise the fourth degree by semitone') to the diatonic collection defined by the key signature.[20] Thus Hook's work represents an early instance of the more general claim that modulation involves voice leading *applied at the level of the underlying scale rather than chords*.[9,24]

From my perspective, Hook's 'signature transformations' are a set of voice leadings generated by the voice leading $(C, D, E, F, G, A, B) \rightarrow (C, D, E, F\sharp, G, A, B)$, its transpositions, and their inversions (or retrogrades). This is also the collection of voice leadings generated by Clough and Douthett's *J*-function, for seven-note chords in a 12-note scale; geometrically, these voice leadings move down by step along the grid at the centre of chord space. One of Hook's insights was that this single voice-leading schema (or 'signature transformation') generates *all the 'scalar' and 'chromatic' transpositions* of the diatonic scale. That is, by applying seven signature transformations in a row, one transposes each note of the diatonic scale upward by semitone. (This semitonal transposition, along with its inverse, in turn generates chromatic transpositions by any number of semitones.) Similarly, by applying 12 signature transpositions in a row one transposes each note of the scale up by one diatonic scale step (an operation which, with its inverse, generates scalar transpositions by any number of scale steps). Thus, rather than thinking of the signature transformation as a combination of scalar and chromatic transpositions – down four steps from ionian to lydian, and up seven semitones from C diatonic to G diatonic – we can view the signature transformation as the more fundamental operation producing these other transpositions.

When I first heard Hook speak about this topic, at the 2003 national conference of the Society for Music Theory, I was interested in moving scale theory away from what seemed like a narrow-minded fixation on maximally even collections. For in actual music one finds a range of chords and scales that are nearly but not *maximally* even, including melodic and harmonic minor scales, major triads in chromatic space, fourth chords in diatonic space, and so on. (The more general point here is that 'satisficing' is often more important than maximizing; it is more valuable to be *even enough* than *as even as one could possibly be.*) It seemed then (and seems now) that music theory often fixates on mathematically elegant special cases at the expense of the full range of common musical practice. (In [9] I therefore focus on the broader class of 'fairly even' collections.) So as I listened to Hook talk about signature transformations, I began to think about how to generalize the idea to any possible scale. It seemed to me that the natural extension was simply to consider all the voice leadings that could be generated by combining scalar and chromatic transposition. (This approach emphasizes the converse of Hook's insight, namely that the signature transformation can always be represented as a combination of scalar and chromatic transpositions.) This led me to propose the *scalar* and *interscalar* interval matrices as devices for tracking the combination of these two operations,[24] devices that in turn proved useful for interpreting the geometry of voice-leading spaces. For it turns out that the *interscalar transpositions* between any two collections are just the bijective, strongly crossing-free voice leadings connecting them – that is, those bijective voice leadings that will never have crossings no matter what octave the voices are in, represented geometrically by line segments that avoid the singular boundary of the orbifold $\mathbb{T}^n/\mathcal{S}_n$.[21] By property P3 above, this collection of voice leadings must contain the smallest possible (bijective) voice leading between those two collections [8–10]. In many practical contexts – for instance if we are interested in minimal (or efficient) voice leadings – we can therefore restrict our attention to these interscalar transpositions.

Hook's central result, then, is that, in the case of the diatonic scale, the bijective, strongly crossing-free voice leadings are generated not by *two* radically distinct voice leadings (one chromatic and one scalar transposition), but by *one* voice leading and its inverse (the 'signature transformation'). Perhaps even more surprising is that this generalizes to any chord whose size is relatively prime to the size of the enclosing scale. (Hook hints at this idea in his work on 'spelled heptachords',[13] but the general statement is new to this paper.) Consider, for example, the case of seven-note chords in the 12-note chromatic scale. The diatonic signature transformation $(\mathrm{C}, \mathrm{D}, \mathrm{E}, \mathrm{F}, \mathrm{G}, \mathrm{A}, \mathrm{B}) \rightarrow (\mathrm{C}, \mathrm{D}, \mathrm{E}, \mathrm{F}\sharp, \mathrm{G}, \mathrm{A}, \mathrm{B})$ combines chromatic transposition up by seven semitones (from C major to G major) with scalar transposition down by four steps (from ionian mode down to lydian). Now consider some other seven-note scale such as the acoustic. If we transpose the collection $(\mathrm{C}, \mathrm{D}, \mathrm{E}, \mathrm{F}\sharp, \mathrm{G}, \mathrm{A}, \mathrm{B}\flat)$ up by seven semitones and down by four scale steps, we obtain the voice leading $(\mathrm{C}, \mathrm{D}, \mathrm{E}, \mathrm{F}\sharp, \mathrm{G}, \mathrm{A}, \mathrm{B}\flat) \rightarrow (\mathrm{C}\sharp, \mathrm{D}, \mathrm{E}, \mathrm{F}, \mathrm{G}, \mathrm{A}, \mathrm{B})$. Iterating this voice leading (so that we always transpose up by seven semitones and down by four steps) we can generate *any* bijective strongly crossing-free voice leading between acoustic scales.[22] This, then, is the 'signature transformation' for the acoustic collection.

The underlying mathematics here is very simple. Since the size of the chord n is relatively prime to the size c of the scale, transpositionally related pitch-class sets, residing in that scale, cannot have the same sum mod c.[23] Since chromatic transposition down by semitone subtracts n from the sum of a chord's pitch classes, and since one-step scalar transposition adds c to them, the Chinese Remainder Theorem tells us that there is a combination of scalar and chromatic transpositions that subtracts 1 from the sum of the chord's pitch classes. This combination is *independent of the internal structure of the chord,* depending only on the numbers n and c. From the general structure of chord space, we know that any strongly crossing-free voice leading, connecting transpositionally related chords with paths summing to ± 1, will generate all the scalar and chromatic transpositions involving that chord type.[24] It follows that we can model these voice leadings using a one-dimensional circular graph, no matter how large our chord, and no matter how many dimensions in the ambient space.

There is however an important difference between Hook's original 'signature transformation' and the more general version we are considering. When a chord is maximally even, the signature transformation (or 'generalized circle of fifths') involves a single-step voice leading at every turn. When composing these voice leadings, we therefore move progressively farther from our starting point, so that if we start at the C major scale, it takes progressively larger voice leadings to get to G major, D major, A major, etc. (Indeed, each step on the generalized circle of fifths adds one semitone of overall motion to the voice leading.) When our original chord is not maximally even, this is no longer the case. Thus we can *generate* all strongly crossing-free acoustic-scale voice leadings by iterating the schema

$$(\mathrm{C}, \mathrm{D}, \mathrm{E}, \mathrm{F}\sharp, \mathrm{G}, \mathrm{A}, \mathrm{B}\flat) \rightarrow (\mathrm{C}\sharp, \mathrm{D}, \mathrm{E}, \mathrm{F}, \mathrm{G}, \mathrm{A}, \mathrm{B}) \rightarrow \ldots$$

but it is not the case that the generating schema is itself the minimal voice leading between equal-tempered acoustic scales. In fact, if we iterate the schema twice, we obtain

$$(\mathrm{C}, \mathrm{D}, \mathrm{E}, \mathrm{F}\sharp, \mathrm{G}, \mathrm{A}, \mathrm{B}\flat) \rightarrow (\mathrm{C}, \mathrm{D}, \mathrm{E}, \mathrm{F}\sharp, \mathrm{G}\sharp, \mathrm{A}, \mathrm{B})$$

which involves two semitones of total motion rather than the three semitones of the generating voice leading. The issue here is that the iterated pair of voice leadings moves C up to C♯ and then back down to C, with the two one-semitone moves cancelling out. Clearly, this happens because the generating voice leading involves both ascending and descending motion, which will always be the case whenever our chord is not maximally even.

In any case, the important point is that, even though the acoustic scale is found in seven-dimensional space, we can depict all the bijective, strongly crossing-free voice-leading possibilities between acoustic scales using a simple circular graph. (Remember that for many theoretical and analytical purposes we can limit our attention to these interscalar transpositions.) We can include other scales on this graph simply by decomposing our three-semitone schema into three single-semitone voice leadings:

$$\begin{aligned}(\mathrm{C}, \mathrm{D}, \mathrm{E}, \mathrm{F}\sharp, \mathrm{G}, \mathrm{A}, \mathrm{B}\flat) &\rightarrow (\mathrm{C}\sharp, \mathrm{D}, \mathrm{E}, \mathrm{F}\sharp, \mathrm{G}, \mathrm{A}, \mathrm{B}\flat) \rightarrow (\mathrm{C}\sharp, \mathrm{D}, \mathrm{E}, \mathrm{F}, \mathrm{G}, \mathrm{A}, \mathrm{B}\flat)\\ &\rightarrow (\mathrm{C}\sharp, \mathrm{D}, \mathrm{E}, \mathrm{F}, \mathrm{G}, \mathrm{A}, \mathrm{B}) \rightarrow\end{aligned}$$

Figure 7 (adapted from [9]) shows all the strongly crossing-free voice leadings between acoustic, harmonic minor, and harmonic major collections. It is, I think, rather remarkable that we can capture so many musically important voice leadings using such a simple one-dimensional graph. And while distance along this circle does not correspond perfectly to voice leading distance, it does provide a reasonable approximation. The more uneven the chord becomes, the greater will be the divergences between 'graph distance' (number of steps on the circle) and the size of the voice leadings it depicts.

This result is interesting enough to warrant a few more examples. First, consider the case of three-note chords in a seven-tone scale (which we can take to be the regular diatonic). The 'generalized circle of fifths' combines transposition up by triadic step (that is, one-step transposition along the three-note 'scale' defined by the triad itself) with chromatic transposition down by third, producing

$$(\mathrm{C}, \mathrm{E}, \mathrm{G}) \rightarrow (\mathrm{C}, \mathrm{E}, \mathrm{A}) \rightarrow (\mathrm{C}, \mathrm{F}, \mathrm{A}) \rightarrow \ldots$$

The analogue for the fourth chord (C, F, G) is therefore

$$(\mathrm{C}, \mathrm{F}, \mathrm{G}) \rightarrow (\mathrm{D}, \mathrm{E}, \mathrm{A}) \rightarrow (\mathrm{C}, \mathrm{F}, \mathrm{B}) \rightarrow (\mathrm{D}, \mathrm{G}, \mathrm{A}) \rightarrow \ldots$$

Notice that, as in the acoustic-scale signature transform, the generating voice leading involves three semitones of total motion, two ascending and one ascending. (This is always true of the

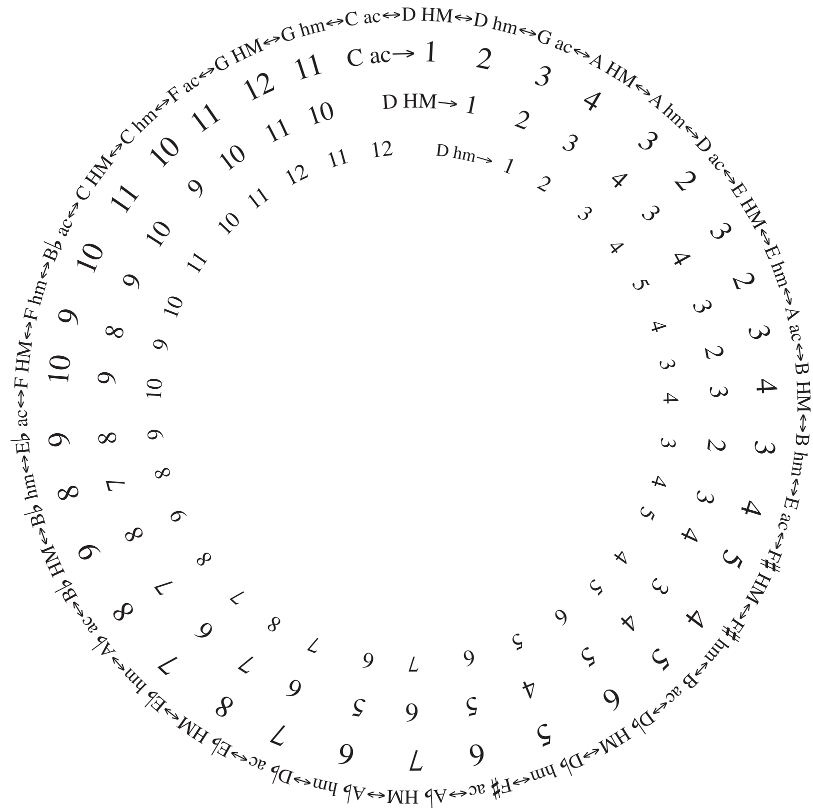

Figure 7. An expanded 'circle of fifths' showing all the bijective, strongly crossing-free voice leadings between acoustic, harmonic minor, and harmonic major scales. The numbers in the circle show the total semitonal motion involved in the voice leading from the chord at the beginning of the series of numbers; observe that although each sequence is different, they ultimately settle into a pattern of the form (5, 6, 5, 6, 7, 6, …).

second most-even chord's signature transform.) Similarly, chords that are two steps apart on the circle (here {C, F, G} and {C, F, B}) are separated by two steps of total motion, as in the case of the acoustic scale. (Again, this is always true of the second most-even chord's signature transform.) For the five-note chromatic cluster (C, C♯, D, D♯, E) we modify the circle of pentatonic scales (C, D, E, G, A) → (B, D, E, G, A) → (B, D, E, F♯, A) → … which combines chromatic transposition up by fifth with scalar transposition down by three steps. This produces the voice leading (C, C♯, D, D♯, E) → (A, A♯, B, G, G♯) → … Figure 8 compares the sizes of the voice leadings in the pentatonic, 'dominant ninth', and chromatic signature transforms; note that for each transposition the maximally even chord moves its notes by the smallest total number of semitones, a number which can be attained but never surpassed by the other chords.

We can generalize this line of thinking even further, to the case where the size of the chord is not relatively prime to the size of the scale. Suppose we have an n-note chord X contained in an equal tempered c-note scale, and suppose that n and c share a common factor f (which might be equal to n). In that case, the scale will contain f different transpositions of the chord X that sum to the same value modulo c (assuming chord X is not itself transpositionally symmetric). This is because we can add c/f to each note of the chord, so we change the chord's sum by ic, where i is an integer and ic is some number of octaves. In this case, the bijective, strongly crossing-free voice leadings will be generated by very different kinds of voice leadings, for example one that connects same-sum chords, and another that connects different-sum chords.

(C, D, E, G, A) →		**(C, D, E, F♯, A) →**		**(C, C♯, D, D♯, E) →**	
→(C, D, E, G, A)	0	→(C, D, E, F♯, A)	0	→(C, C♯, D, D♯, E)	0
→(B, D, E, G, A)	1	→(B, C♯, E, G, A)	3	→(A, A♯, B, G, G♯)	17
→(B, D, E, F♯, A)	2	→(B, D, E, F♯, G♯)	2	→(F♯, D, D♯, E, F)	10
→(B, C♯, E, F♯, A)	3	→(B, C♯, D♯, F♯, A)	3	→(B♭, B, C, C♯, A)	13
→(B, C♯, E, F♯, G♯)	4	→(A♯, C♯, E, F♯, G♯)	4	→(G, G♯, E, F, F♯)	16
→(B, C♯, D♯, F♯, G♯)	5	→(B, C♯, D♯, E♯, G♯)	5	→(B, C, C♯, D, D♯)	5
(etc.)		(etc.)		(etc.)	

Figure 8. The generalized circle of fifths (or signature transform) for three progressively less-even five-note chords, with the total number of semitones moved (= L^1 size of the voice leading) from the chord at the top of the column. As the chord becomes more uneven, we see larger and larger divergences between graph distance and voice-leading size.

Consider 013 trichords in the familiar chromatic scale: since the pitch classes in {C, D♭, E♭}, {E, F, G}, and {G♯, A, B} sum to the same value, the chords can be linked by 'purely contrary' voice leading in which the amount of ascending voice leading balances the amount of descending voice-leading, such as (C, D♭, E♭) → (G, E, F). To link chords with different sums we can use the 'transpositional' voice leading (C, D♭, E♭) → (C♯, D, E). Together, these two voice-leading schemas (along with their inverses) do the job of the signature transformation, composing to generate all the strongly crossing-free voice leadings between the chord's transpositions. Or consider the C 'double diatonic' collection (C, D, E, F, F♯, G, A, B). Here, minor-third related double-diatonic collections share the same sum so that C, E♭, F♯, and A double diatonic can all be linked by purely contrary motion. (Indeed, we can generate all the strongly crossing-free purely contrary voice leadings from (C, D, E, F, F♯, G, A, B) → (C♯, D, D♯, E, F♯, G♯, A, B).) To generate the remaining voice leadings we need to supplement these 'pure contrary' voice leadings with a linearly independent voice leading such as (C, D, E, F, F♯, G, A, B) → (C, D♭, E♭, F, G♭, G, A♭, B♭).[25] The upshot is that in these cases, the graph of strongly crossing-free voice leadings will be *inherently two dimensional*: in place of the circular graph generated by the signature transformation, where motions can be specified using a single number, we now have a two-dimensional toroidal graph generated by two different kinds of voice leadings, which for simplicity we can take to be *scalar* and *chromatic* transpositions.[26] This is a mathematical reformulation of my initial response to Hook's work – namely, that in the general case one needs to replace the one-dimensional 'signature transformation' with the two-dimensional combination of scalar and chromatic transpositions.

The geometry underlying these relationships can be nicely illustrated in the two-dimensional case. In two-note pitch space, $\mathbb{R}^2$, an equal-tempered scale produces a square lattice, rotated 45° relative to the direction of transposition. A quotient of this lattice appears in two-note chord space ($\mathbb{T}^2/\mathcal{S}_2$, a Mobius strip with singular boundary), shown in Figure 9. When the perfectly even chord passes through the vertices of the central squares, transposition by half-octave acts as a reflection fixing the perfectly even chords; it follows that the other two vertices of the central squares are related by transposition, and share the same sum. (To see why, consider the transposition that fixes the square as a whole.) The voice-leadings connecting these transpositionally related chords will be linearly independent of the purely transpositional voice leadings, so it takes two different kinds of voice leadings to generate all the voice leadings between A and its transpositions, as shown on the figure. But when the perfectly even chord passes through the centre of two sides, as shown on the bottom of the figure, then there is no transposition that fixes the square; since each transposition has a distinct sum, we now have a signature transformation, a voice leading $A \to \mathbf{T}_x(A)$ that can be iterated to link all of the transpositions of A. These relationships generalize to all the two-note chords in the space, no matter how far from the centre, and analogous facts hold in all higher dimensions. Figure 9 may therefore help readers visualize the non-intuitive

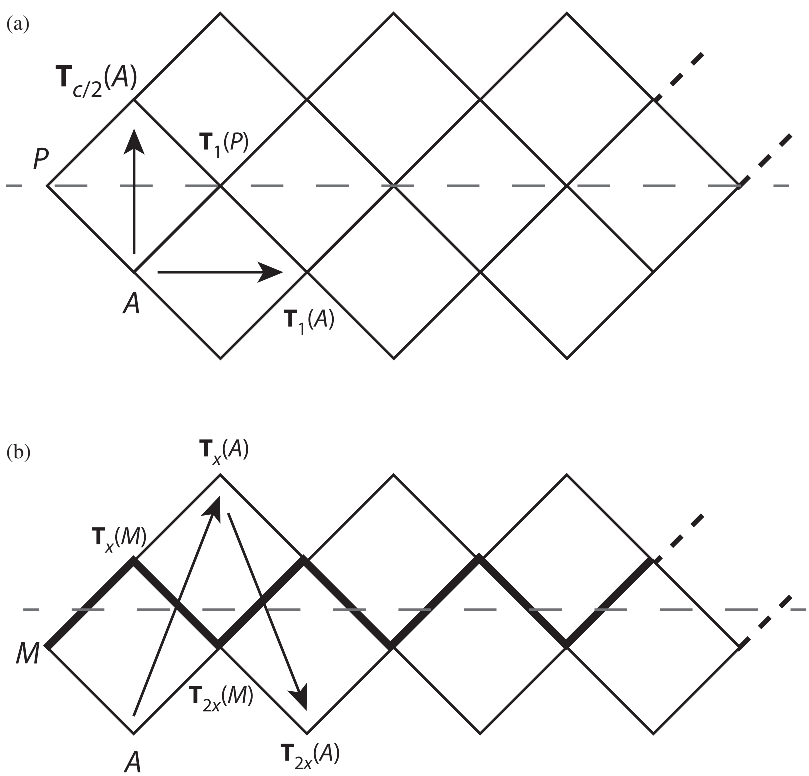

Figure 9. The generators of the strongly crossing-free voice leadings as a function of the relative primality of the size of the chord (here, 2) and the size of the scale. (a) When the cardinality of the scale is even, the perfectly even chord (dashed line) runs through the central vertices of the central squares; transposition by half-octave acts as a reflection (or in odd dimensions, rotation) that fixes these squares, and we need two types of voice leading to generate the strongly crossing-free voice leadings connecting the transpositions of A. (b) When the scale is odd, no transposition fixes the central squares; in this case we have a signature transformation that links all of the transpositions of A, essentially because the transpositional voice leading is linearly dependent on the voice leadings shown by the arrows. This linear dependence is obvious from the picture, since composing the two arrows yields a horizontal voice leading.

connection between the generators of the strongly crossing-free voice leadings and the relative primality of chord size and scale size.

Another important feature of Hook's work is the application of signature transformations to *subsets* of the diatonic scale. As Hook noticed, signature transformations will generally transform subsets by shifting notes semitonally. Consider, for example, the flatward signature transformations of the B diminished triad.

$$(\mathrm{B}, \mathrm{D}, \mathrm{F}) \to (\mathrm{B}\flat, \mathrm{D}, \mathrm{F}) \to (\mathrm{B}\flat, \mathrm{D}\flat, \mathrm{F}) \to (\mathrm{B}\flat, \mathrm{D}\flat, \mathrm{F}\flat) \to \ldots$$

This semitonal descent is strongly reminiscent of – indeed identical to – the stepwise descents that result from quantization (Figure 1), for reasons we will discuss in a moment. In the more general case, where our set is not maximally even, its subsets will generate a similar circle, though notes will not always descend one at a time and there may be more than n chromatic forms of each n-note subset. For example, consider the effect of the acoustic scale's 'signature transform' on its triadic subsets:

$$(\mathrm{B}\flat, \mathrm{D}, \mathrm{F}\sharp) \to (\mathrm{B}, \mathrm{D}, \mathrm{F}) \to (\mathrm{B}, \mathrm{D}, \mathrm{F}\sharp) \to (\mathrm{B}, \mathrm{D}\sharp, \mathrm{F}\sharp) \to \ldots$$

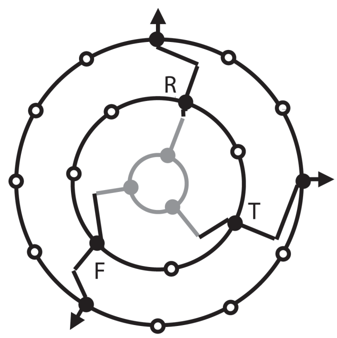

Figure 10. Douthett's pictorial representation of the J-function depicts a series of perfectly even inner beacons (corresponding to the term *ic/n* in Equation (4)) being quantized to a perfectly even scale (as in Equation (2)). Here we see a multistage quantization where the completely even three-note chord is quantized to a completely even seven-note chord, and then a completely even 12-note scale. This could be used to model music in which triads live in a diatonic scale which itself moves through chromatic space.

Here we find four different types of triad (major, minor, diminished, and augmented) with augmented triads linked to diminished triads by two-semitone motion. Indeed, as long as our original set is not maximally even, we can *always* find n-note subsets that, under the signature transform, are linked by more than one semitone of motion, producing more than n chromatic forms in all.

Hook's 'signature transformations' represent one natural extension of the 'J-function'. A second originates in Douthett's (and later, Plotkin's) notion of 'filtered point-symmetry'.[14,15] In its simplest form, filtered point-symmetry is simply a graphical depiction of the J-function, adding no new mathematical content. The force of the idea lies in the use of continuous transposition and multiple layers of quantization to model as wide a range of voice leadings as possible. (This is facilitated by transposing both the continuous chord and the quantizing scales in complex ways.) But while this approach can encompass an impressively wide range of voice leadings, it suffers to my mind from important conceptual limitations. First, scales play an ambiguous role, sometimes modelling actual features of a musical context (as in Figure 10, which could depict triads in a diatonic scale in a chromatic universe), and sometimes appearing purely as technical tools for ensuring that multiple quantizations deliver the desired voice leadings. Thus the approach cannot model major triads directly existing in chromatic space, for example, but must embed these triads within seven-note scalar structures that often have no analytical significance. Second, the approach is not well-suited to scales that are not maximally even, including the ubiquitous melodic and harmonic minor collections, nor to situations where these scales alternate rapidly – as in traditional minor-mode music. Third, the model is problematic when the space of strongly crossing-free voice leadings has two degrees of freedom. For consider any continuous chord along with some set of quantizing scales, all subjected to arbitrary continuous time-dependent transpositions; this configuration will generate voice leadings forming a *one-dimensional circular graph* (whose parameter is time), as in the original signature transformation. Neither Douthett nor Plotkin has proposed general models in which scalar and chromatic transposition are clearly represented.[27] For this reason, filtered point-symmetry seems more appropriate to the study of *sequential patterns* than voice leading more generally (cf. Yust's contribution to this volume). This conclusion is further reinforced by the fact that Douthett and Plotkin cannot model voice leadings with crossings and by the fact that they treat non-bijective voice leading only awkwardly.[28] Fourth, as we saw in the preceding section, the geometrical theory of voice leading provides a principled explanation of just *how* quantization and transposition lead to the voice leadings they do, while it is much harder to run the explanation in the other direction – that is, to build a general theory of voice leading using the materials of filtered point-symmetry. In this sense, the one theory seems more fundamental than the other. Fifth, I suspect that composers are more likely to think about connecting familiar chords by way of efficient voice leading rather than *quantizing perfectly even chords into multiple scales*. (It also seems much more likely that listeners respond directly to voice-leading distances than that they hypothesize the existence of perfectly even continuous chords only very imperfectly realized in equal temperament.) Sixth and finally, in Douthett and Plotkin's work, the connection to quantization and voice leading is constantly obscured by the universal absorbtion of the additive constant 0.5 into their version of Equations (2)–(4). (Graphically, this means that their lines always move *counterclockwise* even when there are closer clockwise destinations available.) This small notational issue has important theoretical implications, effectively obscuring the conceptual meaning of the entire approach. (Indeed neither the word nor the concept 'quantization' appears in Douthett and Plotkin's writings – though it is, of course, implicit in the J-function itself.) This unclarity is a minor issue in Clough and Douthett's original work, where the J-function appears merely as a tool for tabulating maximally even collections; it becomes more worrisome when the J-function is itself presented as an object worthy of music-theoretical study – for if we do not understand what the J-function *is*, then how can it help us understand particular pieces?

In personal conversations, and in their response in this volume, Plotkin and Douthett attempt to justify their model on the basis of 'analytic utility'. The power of this response largely depends how we understand the nature and purpose of the analytical enterprise. On one view, promulgated by David Lewin among others, analyses are to be evaluated rather like works of art: to say that an analysis is good is simply to say that we enjoy thinking about music in the way it describes. I have criticized this approach, since virtually *any* claim about music – or indeed anything else – can be justified with a simple statement that it is enjoyable.[16] (Perhaps it is also enjoyable to practice phrenology or worship the Ancient Greek Gods.) In my own view, analysis generally seeks to uncover *general principles* that (consciously or unconsciously) guide the creation or perception of works of art. Thus one might seek to explain why Debussy's *L'Isle Joyeuse* moves between whole-tone, acoustic, and diatonic scales, or why Chopin uses a particular sequence of chords in the E minor prelude, or even why there are so few iii chords in Mozart's G major piano sonata. From this perspective, models can be expected to have a robust analytical utility only insofar they perspicuously represent principles that – consciously or unconsciously – guide composers' choices or listeners' responses. (The deeper methodological issue here is that we should demand some *causal explanation* that connects mathematical relationships to the activities of composing or listening; and any such causal link will almost certainly involve the brain.) But there is very little reason to think that quantization and continuous transposition, represented graphically by the concentric circles of filtered point-symmetry, fit the bill. What is more likely: that composers were motivated by general principles from the theory of voice leading (such as 'move to the nearest chord-tone'), or by unconsciously applying complex combinations of quantizations and continuous transpositions?

3. Conclusion

In this paper we have encountered two general music-theoretical facts. First, the 'generalized cardinality-equals-variety (CV) property', which states that the continuous transpositions of a generic n-note set class (not too symmetrical, with fractional parts all distinct) quantize to n set-classes, located on the vertices of the minimal equal-tempered simplex containing it; these quantized forms are connected by a circle of single-semitone voice leadings. Second, the 'generalized signature transform': whenever the size of a chord is relatively prime to the size of its surrounding scale, we can find a 'signature transform' (or 'generalized circle of fifths'), defined as a voice-leading schema that (together with its inverse) generates all the bijective, strongly crossing-free voice leadings between the chord's transpositions. Neither of these facts has anything to do with evenness, maximal or otherwise. Rather, what is unique to the maximally even case is the connection between them: *only when we have a maximally even chord can the signature transform be represented as the effect of quantization.*[29] (Here we connect our generalized CV property to traditional scale theory's special case.) This gives us a much clearer understanding of the logic underlying this whole conceptual territory, allowing us to differentiate three topics – quantization, signature transforms, and maximal evenness – that have often blurred together.[30]

In constructing this revisionist history of maximally even chords, we have found voice leading playing a role at several points, from Clough and Douthett's formal definition of 'maximally even' to Douthett and Hook's eventual interpretation of the 'J-function' as a generalized circle of fifths.[31] Indeed, this concern with voice leading becomes more explicit over time: whereas Clough, Douthett, and Myerson originally aspired just to catalogue the maximally even collections, Hook, Douthett, and Plotkin are avowedly focused on the voice leadings between them. Absent a fully general theory of voice leading, and an understanding of the many ways it figures into this story, one can see how these papers would have engendered a significant sense of mystery, as if they were portents of deep but obscure relationships.

Once the full geometrical perspective is at hand, however, the hints lose some of their potency. The maximally even chords are simply the chords nearest the centre of whatever chord space we are talking about, lying on the lattice-quotient determined by the scale in question. The 'J-function' converts a continuous process of descending transposition into a series of descending voice leadings near the centre of this lattice(-quotient), lowering notes in turn. This is entirely in keeping with what happens in the general case (Equation (3), Figure 3), with the only complication being that, if we start with completely even chords, the quantized results are always transpositionally related. This is again a relatively straightforward consequence of the geometry of set-class space: in the generic interior of the space, the minimal equal-tempered simplex will contain n different set-classes on its vertices; near the perfectly even chord, however, the vertices of the simplex will lie in different fundamental domains, and will be transpositionally related (see [9, Figure 3.8.6]).

Some readers may find this generality threatening, others liberating. The mathematician Groethendieck spoke of mathematical problems not as nuts to be cracked, but as rocks to be submerged and dissolved by a larger theory that goes beyond the desired results. In this case, older ideas about maximally even collections can be subsumed in a much more general geometrical perspective that allows us to consider every conceivable chord, whether inside a scale or not. We can apply traditional concepts – including 'cardinality equals variety' and Hook's signature transformation – to a much broader range of collections. This can in turn break our fixation on maximally even sets, allowing us to cultivate more flexible habits of mind appropriate to a broader range of analytical and compositional circumstances.

Acknowledgements

Conversations with Andrew Jones, Jason Yust, and Marek Žabka have been influential on this paper, as have the comments by the anonymous referees and the other participants in the 2012 John Clough Memorial Conference. (Thanks to Rick Cohn for organizing and supporting it.) Yust's recent work touches on many of the same topics in a more analytical register.

Notes

1. If this intuition seems controversial, consider an economic analogy. Suppose we want to divide \$4 fairly among three people. Intuitively we are trying to get as close as possible to the ideal division where each person has \$1 1/3: if we have only dollar bills, one person will get \$2 and the others \$1; if we have lots of pennies, one person gets \$1.34 and the others \$1.33. Here, as in music, the notion of 'as close as possible' involves principles P1–3 described below.

2. Terminological note: I will generally use the term 'chord' to refer to some subset of a larger collection, or 'scale', so that 'maximally even chords' are contained within (perfectly even) chromatic scales. In Part II, I will sometimes use the term 'scalar transposition' to refer to (bijective) strongly crossing-free voice leadings from any chord to itself; when the chord under discussion is the diatonic scale, these correspond to the familiar 'scalar transpositions' that turn ionian into dorian, etc.; when applied to a major triad, these scalar transpositions send root position to first inversion, etc. Context should clarify my meaning here.
3. By 'continuous chord' I mean any multiset of pitch classes whatsoever – the analogue to \$1 1/3 in the previous footnote. A 'scale chord', naturally enough, is a chord in the scale, analogous to \$1 or \$1.33.
4. Block and Douthett [17] consider other formal measures of 'evenness', again without connecting them to more general notions of musical distance; in [11], Dothett and Krantz consider measures of evenness that rely on chord lengths in two-dimensional Euclidean space, whose musical meaningfulness is doubtful.
5. The 'J-function' first appears near the end of [18], though without its name. Jack Douthett tells me that the 'J' stands for both 'Jack' and 'John [Clough]'.
6. Along the way we have noticed that similar problems can arise in economics, where we might like to use an analogue of voice-leading distance.[10]
7. Clough and Douthett's 'J-function' sends the chord $\{0.99, 4.99, 8.99\}$ not to the nearby $\{1, 5, 9\}$, from which it is aurally indistinguishable, but to the more distant $\{0, 4, 8\}$. Perhaps for this reason, the connection between the 'J-function' and quantization (or 'moving to the closest note') has not previously appeared in print.
8. In responding to this paper, Jack Douthett emphasized the unfamiliarity of the continuous perspective, noting that he and Clough thought exclusively in discrete terms. But continuous log-frequency space exists whether we think about it or not, just as quantities like 'a third of a dollar' can arise even if we only have dollar bills in our wallet.
9. Of course, if the chord has multiple notes with the same fractional value then they will descend at the same time.
10. In most musical styles, there is an important syntactical difference between motion by ascending semitone and motion by descending major seventh. Earlier discussions of voice leading [19–22] erase this distinction by adopting traditional pitch-class intervals; as a result, they arguably fail to capture (or formalize) the musical intuitions basic to the very notion of voice leading.
11. Cf. Figure 2(a), where the nearly even harmonic series quantizes to two separate modes of the acoustic scale, or Figure 5, where the augmented triad quantizes to three separate modes of the equiheptatonic triad. A similar thing can occur if a chord is inversionally symmetrical or nearly so.
12. We can, of course, ask 'what quantization moves notes by the smallest overall distance', though this in general requires choosing a metric of voice leading size. Furthermore, it may be more important that a chord is roughly equidistant from several different quantizations than that it happens to be *slightly* closer to one of these.
13. A common mistake, when quantizing set classes, is to arbitrarily fix some note in the chord. For instance, if we require that the harmonic series start on an equal-tempered note, then quantization of its first seven pitch classes will *not* produce the acoustic scale. However, if we consider all possible transpositions then the acoustic scale *is* the most common quantization of the start of the harmonic series.
14. If the fractional parts of two notes are equal, then the chord will sit on a boundary of the minimal equal-tempered simplex; in this case, multiple notes will descend at the same time.
15. These regions are congruent to the cross section of n-note chord space, which is given by the equation: $x_0 \leq x_1 \leq x_2 \leq \cdots \leq x_{n-1} \leq x_0 + 12, \sum x_i = c$.
16. This follows from considerations in [9, Appendix B]. Start with the simplex whose coordinates represent the notes' fractional parts, as defined by Equation (5). We obtain the standard simplex by applying an affine transformation that records the intervals between adjacent notes (and from the last note to the note a semitone above the first). These intervals sum to 1, as required. But as is obvious from Figure 1, these intervals also represent the range of transpositions quantizing to each individual chord type.
17. Clough and Douthett [7] provide a lengthy proof, which Demaine et al. [23] shorten considerably. For the outline of an even shorter proof note that the statement is trivial when the size of the chord n divides the size of the scale c. (Here the chord lies on a vertex, or zero-dimensional subsimplex, of an n-dimensional minimal equal-tempered simplex, as in Figure 5(a).) When the numbers are relatively prime, then, by the Chinese Remainder Theorem the integer multiples of c cycle through all possible values mod n. When $ic \equiv \pm 1 \pmod{n}$, the spectrum of the ith scalar interval is a 'near interval cycle': a circular graph of pitch classes with all but one of the notes linked to its neighbours by the same interval, and with the unusual interval being just a semitone larger or smaller. From this it follows that the chord can be linked to two of its transpositions by single-semitone voice leading [9, Figure 3.11.4], and hence that there exists a minimal equal-tempered simplex in set class space whose vertices are all forms of this set class; this necessarily contains the perfectly even chord (since the spectrum of the ith scalar interval encloses the interval containing *ic/n*). This is shown in Figure 5(b). Finally, the case where n and c share a common factor f reduces to the previous case when one considers the chord of size n/f in the 'reduced octave' of size c/f.
18. Note that Clough and Myerson [18] use the term 'generalized circle of fifths' to refer to a 'near interval cycle' as defined in the previous footnote. By contrast, I use the term to refer to a circular sequence of single-semitone *voice leadings* connecting transpositionally related chords.
19. For instance, the three octatonic scales are linked by the voice leadings $(0, 1, 3, 4, 6, 7, 9, 10) \rightarrow (0, 2, 3, 5, 6, 8, 9, 11) \rightarrow (1, 2, 4, 5, 7, 8, 10, 11)$.
20. I am using the term 'voice-leading schema' to refer to what I have elsewhere called 'uniformly T-related voice leadings'.[9] Suppose we have a voice leading $A \rightarrow B$, with chord $\mathbf{T}_x(A)$ transpositionally related to A. We can say that the (uniformly T-related) voice leadings $A \rightarrow B$ and $\mathbf{T}_x(A \rightarrow B)$ *instantiate the same voice-leading schema.* The new terminology is meant to capture the difference between analysis and composition: where the language of

'uniform T-relationships' suggests recognizing relationships out there in the world, the term 'voice leading schema' is appropriate when we are wondering how to 'do $A \to B$' to the chord $\mathbf{T}_x(A)$.

21. A voice leading, as I define it in [9], is a multiset of ordered pairs (p, r) where p is a pitch class and r is a real number representing a path in pitch class space; the pitch class p is sent to $q(= p + r)$ along the path r. A voice leading between two pitch-class multisets $\{p_i\}$ and $\{q_i\}$ is bijective if each element in one multiset is paired with exactly one element in the other. The voice leading $(\mathrm{C}, \mathrm{C}, \mathrm{E}, \mathrm{G}) \to (\mathrm{A}, \mathrm{C}, \mathrm{F}, \mathrm{A})$ (notes are assumed to move along the shortest path when not specified) is bijective between {C, C, E, G} and {F, A, A, C} but non-bijective between {C, E, G} and {F, A, C}. The orbifold $\mathbb{T}^n/\mathcal{S}_n$ is the n-torus modulo the symmetric group $\mathcal{S}_n$, or the space of unordered pitch classes. It is often called n-note *chord space*. Every point in this space represents an n-note chord, and every line segment represents a voice leading between its endpoints. For more, see [4,8,9].
22. We can think of these voice leadings as the composite $\mathrm{A}^{-1}\mathrm{B}^n\mathrm{A}$, where A is a strongly crossing-free voice-leading schema that takes the acoustic scale to the diatonic, B^n is the n-fold application of the diatonic scale's signature transformation, and A^{-1} is the inverse of A. Thanks here to Jason Yust.
23. The sum of a pitch-class set is simply the sum of its pitch classes modulo the size of the octave. Chords with the same sum live on the same 'cross section' of chord space.[9]
24. Scalar transpositions compose 'nicely', by way of vector addition; this is easily seen by considering a chord without pitch-class duplications, in the interior of the space.
25. Note that we cannot use the 8-semitone purely transpositional voice leading between C and D♭ double diatonic, because then we would never be able to generate the one-step scalar transposition. Instead, we use the four-semitone voice leading in the main text, which combines one-semitone chromatic transposition with one-step scalar transposition.
26. An important caveat: in speaking of the dimension of these spaces, I am adopting an abstract, graph-theoretical perspective. In the actual geometry of chord space the chords summing to the same value will lie on the vertices of a simplex whose precise shape is determined by the intervallic structure of the chord. There is a combination of scalar and chromatic transpositions that cycles through the vertices of this simplex, but this generating voice leading is not geometrically privileged. However, for many analytical purposes it is useful to disregard the higher-dimensional geometry in favour of the simpler graph-theoretical approach.
27. In the 3-in-7-in-12 case, one can, for certain configurations, rotate the middle and outer beacons to produce these two independent degrees of freedom, but that is not true in general.
28. Plotkin [15] allows the inner ring to undergo a continuous transition between maximally even chords of different cardinalities, isomorphic to voice leadings. But it is unclear why we should limit ourselves to just these possibilities. Why not allow the inner ring to articulate any voice leading whatsoever? And if we draw on the full theory of voice leading in this way, what is gained by the additional apparatus of quantization, since we can already model all possible voice leadings directly? Note by contrast that more general theories of voice leading can encompass non-bijectivity; see [8], supplementary section 8, [24], Section IV, and [9], §2.9, §3.12, and §8.2.
29. Note that it can happen that a continuous and not perfectly even chord X quantizes in some scales but not others to the maximally even collection. But we will always be able to find scales in which X quantizes to a chord that is not maximally even. If X *always* quantizes to a maximally even chord, in every possible equal-tempered scale, then X is perfectly even.
30. This poses a problem for the filtered point-symmetry approach: if only the maximally even chord's signature transformation can be directly generated by quantization and continuous transposition, then other chords will require superfluous layers of quantizing scales.
31. See note 18.

References

[1] Tymoczko D. Voice leadings as generalized key signatures. Music Theory Online. 2005;11(4).
[2] Tymoczko D. Scale networks and Debussy. J Music Theory. 2004;48(1):219–294.
[3] Tymoczko D. Recasting K-nets. Music Theory Online. 2007;13(3).
[4] Callender C, Quinn I, Tymoczko D. Generalized voice-leading spaces. Science. 2008;320(5874):346–348.
[5] Tymoczko D. Set-class similarity, voice leading, and the Fourier transform. J Music Theory. 2008;52(2):251–272.
[6] Tymoczko D. The generalized *Tonnetz*. J Music Theory. 2012;56(1):1–52.
[7] Clough J, Douthett J. Maximally even sets. J Music Theory. 1991;35(1–2):93–173.
[8] Tymoczko D. The geometry of musical chords. Science. 2006;313(5783):72–74.
[9] Tymoczko D. A geometry of music: harmony and counterpoint in the extended common practice. New York: Oxford University Press; 2011.
[10] Hall R, Tymoczko D. Submajorization and the geometry of unordered collections. Am Math Mon. 2012;119(4):263–283.
[11] Douthett J, Krantz R. Maximally even sets and configurations. J Comb Optim. 2007;14(4):385–410.
[12] Hook J. Signature transformations. In: Douthett J, Hyde M, Smith C, editors. Music theory and mathematics: chords, collections, and transformations. Rochester, NY: University of Rochester Press; 2008. p. 137–160.
[13] Hook J. Spelled heptachords. In: Agon C, Amiot E, Andreatta M, Assayag G, Bresson J, Mandereau J, editors. Mathematics and computation in music – MCM. Proceedings; 2011. Heidelberg: Springer; 2011. p. 84–97.

[14] Douthett J. Filtered point-symmetry and dynamical voice leading. In: Douthett J, Hyde M, Smith C, editors. Music theory and mathematics: chords, collections, and transformations. Rochester, NY: University of Rochester Press; 2008. p. 72–106.
[15] Plotkin R. Cardinality Transformations in diatonic space. In: Agon C, Amiot E, Andreatta M, Assayag G, Bresson J, Mandereau J, editors. Mathematics and computation in music – MCM. Proceedings; 2011. Heidelberg: Springer; 2011. p. 207–219.
[16] Tymoczko D. Basket cases. Music Theory Spectr. 2012;34(1):153–154.
[17] Block S, Douthett J. Vector products and intervallic weighting. J Music Theory. 1994;38(1):21–41.
[18] Clough J, Myerson G. Variety and multiplicity in diatonic systems. J Music Theory. 1985;29(2):249–270.
[19] Roeder J. A geometric representation of pitch-class series. Perspect New Music. 1986;25(1–2):362–409.
[20] Lewin, D. Some ideas about voice-leading between pcsets. J Music Theory. 1998;42(1):15–72.
[21] Morris, R. Voice-leading spaces. Music Theory Spectr. 1998;20(2):175–208.
[22] Straus, J. Uniformity, balance, and smoothness in atonal voice leading. Music Theory Spectr. 2003;25(2):305–352.
[23] Demaine E, Gomez-Martin F, Meijer H, Rappaport D, Taslakian P, Toussaint G, Winograd T, Wood D. The distance geometry of music. Comput Geom. 2009;42(5):429–454.
[24] Tymoczko D. Scale theory, serial theory, and voice leading. Music Anal. 2008;27(1):1–49.