Music Theory Online

The Online Journal of the

Society for Music Theory

Volume 10, Number 3, September 2004

Clifton Callender*

Continuous Transformations

KEYWORDS: transformation, voice leading, interpolation, Saariaho, Nancarrow, Ligeti

ABSTRACT: Since musical systems are generally limited to a discrete number of points in both the frequency and time domains, music theorists have properly concentrated on techniques for modeling discrete transformations of musical objects. However, a number of contemporary composers have explored continuous transformations of musical objects in the frequency, time, and other domains. Such processes include continuous transformations of pitch simultaneities, rhythmic patterns, and tempos. This paper discusses techniques for modeling infinitesimal motions in continuous musical spaces (as well as discrete motion in atypical spaces such as non-tempered microtonal progressions) with analytic application to works by Kaija Saariaho, Conlon Nancarrow, and György Ligeti.

Submission received July 2004

---

1 Introduction

[1.1] The point of departure for this paper is Kaija Saariaho's *Vers le blanc*, a tape piece written using IRCAM's CHANT vocal synthesis program.[1] Figure 1 summarizes the pitch structure, which consists of a single large-scale interpolation, achieved by three independent and continuous glissandos, between two chords over a duration of 15 minutes. (Saariaho 1987) About this work the composer writes, "The harmony is a continuous stream and cannot be heard as a series of changing chords. One only notices from time to time that the harmonic situation has changed." To which any analyst of the piece must respond, "What are these harmonic 'situations'?"[2]

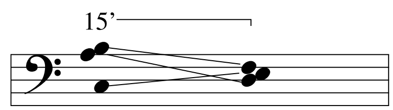


Figure 1: *Harmonic interpolation in* Vers le blanc.

[1.2] The harmonic interpolation in [1.1] is a continuous transformation from the ordered pitch set $\langle 0, 9, 11\rangle$ to $\langle 4, 2, 5\rangle$ with $C_3 = 0$. Order position defines each voice so that voice 1 moves from 0 to 4, voice 2 moves from 9 to 2, and voice 3 moves from 11 to 5. In this case each voice is a continuously changing stream of pitch. While the process described above is a clear example of infinitesimal voice leading, the possibilities are not limited to interpolations between pitch sets. For example, the same ordered sets could be reinterpreted as ordered beat-class sets in a meter of twelve pulses, as shown in Figure 2. Suppose we were to take 101 samples over the course of the interpolation. Voice 1 would begin at beat 0 in the first measure, beat 0.04 in the second measure, beat 0.08 in the third measure, and so forth. The perception of this very slight change in beat position is that voice 1 gradually migrates from beat 0 to beat 4. Now the question has shifted slightly to "What are the intermediate rhythmic 'situations'?"

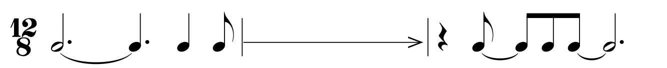


Figure 2: *Hypothetical rhythmic interpolation*

[1.3] Or suppose the sets are interpreted as values for tempos of the form $2^{\frac{x}{12}}c$ beats per minute, so that (with $c = 60$) one tempo accelerates from $2^{\frac{0}{12}} \cdot 60 = 60$ to $2^{\frac{4}{12}} \cdot 60 \approx 76$, while the other

[1] The present paper is a development of papers presented at the Society for Music Theory 2002 Annual Conference, the American Mathematical Society 2003 Southeastern Sectional, a seminar on "Modeling Musical Systems" at the University of Chicago, and a workshop on "Atonal Voice Leading" at the 2003 Mannes Institute for Advanced Studies in Music Theory. Comments and suggestions from members of these audiences as well as discussions with my colleagues at Florida State University, particularly Michael Buchler, have been very beneficial and are greatly appreciated. Special thanks to Robert Peck, Judith Baxter, Richard Cohn, and Joseph Straus for inviting me to present at these venues. I would also like to thank John Roeder and an anonymous reader for their helpful criticisms of an earlier draft. Discussions with Ian Quinn greatly sharpened my understanding of the geometric concepts underlying Section 3.

[2] There are many other processes unfolding simultaneously—rhythmic interpolations, timbral interpolations, and so forth. Only the pitch interpolation will be discussed here.

two tempos decelerate from $2^{\frac{9}{12}} \cdot 60 \approx 101$ to $2^{\frac{7}{12}} \cdot 60 \approx 67$ and $2^{\frac{11}{12}} \cdot 60 \approx 113$ to $2^{\frac{5}{12}} \cdot 60 \approx 80$, respectively. What tempo relationships will result and how will these evolve over time? The sets could also be interpreted as ordered dynamics (e.g., 0 = ***ppppp***, 1 = ***pppp***, ..., 11 = ***fffff***) or gradations within a timbral space. For the latter, consider 0 to be *molto sul tasto*, 11 to be *molto sul ponticello*, and the values in between to represent median bow positions for a string section. Gradual changes of timbre could be achieved by staggering changes of bow position throughout the section, as in Ligeti's *Lontano* and other works.[3]

[1.4] The most general form of these interpolations is a transformation that acts on one musical object and yields a second object, $S \xrightarrow{f(t)} T$. However, instead of achieving the motion instantaneously, the transformation is smeared over time so that we hear the (potentially) infinite musical states that lie between S and T. Our formidable theoretical techniques are of use at exactly two moments in time for each of the above examples—the commencement and completion of the gradual processes. A random sample of any moment within the interpolations is unlikely to conform to existing musical categories, even taken in their most general form. Such notions as n-tone equal temperament, additive rhythms based on a unit pulse, or hemiolas of the form $x : y$ do not apply. Worse, any sample gives no information about the immediately preceding or succeeding moments.

[1.5] The solution offered in this paper is to consider these continuous transformations as trajectories through a space constructed according to a particular distance metric. The primary focus will be on these trajectories and their analytical and compositional applications, though it will be necessary to discuss the basics of the space in which they travel.[4] Before proceeding, the reader is encouraged to listen to Example 1 [modem] [broadband]—an audio simulation of the pitch interpolation in *Vers le blanc*.[5]

2 Preliminaries

[2.1] We will be discussing *ordered* sets almost exclusively. Ordered sets will designated by angled brackets while unordered sets will be designated by curly brackets. For instance, $\langle a, b\rangle$ and $\langle b, a\rangle$ are both possible orderings of the unordered set $\{a, b\}$. Unless indicated otherwise, A, B, etc., will be ordered sets of n voices such that $A = \langle a_1, \dots, a_n\rangle$, $B = \langle b_1, \dots, b_n\rangle$, and so forth, where $a_i, b_i \in \mathbb{R}$. For general cases, voice 1, voice 2, ..., voice n will be identified as v_1, v_2, ..., v_n. For instance, the solution set for the equation $v_2 = v_3$ contains all sets whose second and third voices are equal.

[2.2] The set of ordered sets equivalent to A under transposition will be designated as $/A/_T$, the transposition class or T-class of A. The set class of A—the set of all sets equivalent to A under transpositional, permutational, inversional, and modular (octave) equivalence—will be designated as $/A/$.[6] Since the members of A are real numbers, the order of both $/A/_T$ and $/A/$ is infinite.

[2.3] If one set can be transformed into another by moving a single voice a distance of h, then the

[3] Presumably, this one dimensional timbral space would be a vector within a multi-dimensional space such as Grey 1977.

[4] See Callender 2004 for a more thorough examination of the geometry of these spaces.

[5] This simulation compresses the fifteen-minute interpolation into two and a half minutes.

[6] If $A' = T_x(A), x \in \mathbb{R}$, then $A' \in /A/_T$. For $A' \in /A/_T$, if A'' is a permutation of $\{a'_1 + c_1 m, \dots, a'_n + c_n m\}$, where $c_i \in \mathbb{Z}$ and $m \in \mathbb{R}$ is the modulus, then $A'', I(A'') \in /A/$, where I is inversion about 0.

sets are Δ^h-related. That is, if $\langle |b_1 - a_1|, \ldots, |b_n - a_n| \rangle$ is a permutation of $\langle 0^{n-1}, h \rangle$ (i.e., an ordered set containing $(n-1)$ 0s and 1 h), then A and B are Δ^h-related, written $A \Delta^h B$.[7] Furthermore, the classes to which A and B belong are also Δ^h-related. Thus, $/A/_T \; \Delta^h /B/_T$ and $/A/ \; \Delta^h /B/$. For example, $\langle -1, \pi, e \rangle$ and $\langle -3.5, \pi, e \rangle$ are $\Delta^{2.5}$-related, as are the equivalence classes to which they belong.

[2.4] The transformations we will be considering are continuous functions of time, conventionally notated $f(t)$, $g(t)$, and so forth.[8] If the value of voice i at any time is $f_i(t)$, then the transformation is given by $f(t) = \langle f_1(t), \ldots, f_n(t) \rangle$. The notation $A \rightarrow B$ implies that $f(0) = A$, $f(1) = B$, and the motion from a_i to b_i is linear; i.e., $f_i(t) = (b_i - a_i)t + a_i$.[9] For example, the continuous transformation in *Vers le blanc* is $\langle 0, 9, 11 \rangle \rightarrow \langle 4, 2, 5 \rangle$. Since voice one begins at 0 and ascends four semitones over the course of the entire piece,

$$f_1(t) = 4t.$$

Similarly, since voice two begins at 9 and descends seven semitones,

$$f_2(t) = -7t + 9.$$

Finally, since voice three begins at 11 and descends six semitones,

$$f_3(t) = -6t + 11.$$

Having defined the function for each voice, the ordered pitch set of *Vers le blanc* at time t is

$$f(t) = \langle f_1(t), f_2(t), f_3(t) \rangle. \tag{1}$$

3 T-class space and regions

3.1 Ordered set space

We begin with two assumptions about the space in which these transformations take place.

Assumption 1. The space should be Euclidean.

[3.1.1] Euclidean space of n dimensions is the set $\mathbb{R}^n$, where the distance between any two points in the space, $X = (x_1, \ldots, x_n)$ and $Y = (y_1, \ldots, y_n)$, is defined as

$$d(X, Y) = \sqrt{\sum (y_i - x_i)^2}. \tag{2}$$

[7] This is a generalization of relations variously defined as P^1 in Callender 1998, P_1 in Childs 1998, maximal smoothness in Cohn 1996, $P_{1,0}$ in Douthett and Steinbach 1998, and DOUTH1 in Lewin 1996.

[8] Intuitively, a continuous function is one that can be drawn without having to lift your pencil from the paper. Mathematically, a transformation $f(t)$ is continuous if, for every t_0 in the domain of f, $\lim_{t \to t_0} f(t) = f(t_0)$ (read the limit of $f(t)$ as t approaches t_0 equals $f(t_0)$). If, for any $\epsilon > 0$, there exists a $\delta > 0$ such that $|f(t) - c)| < \epsilon$ whenever $0 < |t - t_0| < \delta$, then $\lim_{t \to t_0} f(t) = c$. (Weisstein) Perceptually, a transformation is continuous as long any discontinuities are small enough to remain undetected by the listener.

[9] This does not necessarily imply that $f(t)$ is defined only for $t \in [0, 1]$. For $t > 1$, $f(t)$ is a continuation of the motion from A to B. For $t < 0$, $f(t)$ is a continuation of the motion from B to A.

This metric taps into our typical notions of distance in physical space and is well suited to the algebraic manipulations of Sections 6 and 7. An example of a common metric that is not Euclidean is the "city-block metric:"

$$\sum |y_i - x_i| .^{10} \tag{3}$$

Equation 3 is often used for "voice-leading distance," where the distance between two ordered sets is the sum of the unordered intervals traversed by each voice.[11,12] (Cohn 1998, Lewin 1998, Roeder 1987, Straus 2003)

Assumption 2. If one set can be transformed into another by moving a single voice, the distance between the two sets should be the distance this voice moves. That is, if $A\,\Delta^h B$, then $d(A,B) = h$.

[3.1.2] Consider a texture of a single voice that can vary continuously in pitch. The position of this voice in pitch space can be modeled by points on a real number line with some pitch arbitrarily assigned to 0. If the voice moves from x to $x \pm h$, then the distance the voice has moved is h. Assumption 2 states that if we add any number of stationary voices to the texture, the distance between the resulting sets will still be h. For example, if $A = \langle -2, 6.5, 10 \rangle$ and $B = \langle -2, 5.25, 10 \rangle$, then the distance between A and B should be 1.25, since voice two moves from 6.5 to 5.25 while the other voices remain the same. However, based on these two assumptions we cannot (yet) make any claims concerning the distance between sets that differ by more than a single voice.

[3.1.3] For ordered sets of n voices the simplest way to satisfy the above assumptions is to map each voice onto an axis in n-dimensional Euclidean space, using the Cartesian coordinate system. It is easy enough to verify that Assumption 2 is satisfied. Moreover, we can measure the distance between any two sets of an equal number of voices using equation 2. For $n = 3$ the distance between A and B is

$$d(A,B) = \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2 + (b_3 - a_3)^2}).^{13} \tag{4}$$

[3.1.4] The beginning and ending sets from *Vers le blanc* are located at $\alpha = \langle 0, 9, 11 \rangle$ and $\omega = \langle 4, 2, 5 \rangle$, respectively. The interpolation, f, consists of a straight line connecting these two points. The rate of change from α to ω remains constant so that $t_1 = \frac{1}{2}$ (7'30") corresponds to the midpoint of f, $t_2 = \frac{1}{9}$ (1'40") corresponds to the point lying one ninth of the way from α to ω, and so forth. Points corresponding to 12-tone equal tempered sets form a cubic lattice. We can get a sense of the harmonic "situation" at any moment, t_i, by finding those familar tempered sets that lie nearest $f(t_i)$. For example, $f(t_1) = \langle 2, 5\frac{1}{2}, 8 \rangle$ lies precisely halfway in between $\langle 2, 5, 8 \rangle$ and $\langle 2, 6, 8 \rangle$. $f(t_2) = \langle \frac{4}{9}, 8\frac{2}{9}, 10\frac{3}{9} \rangle$ lies within the cube whose vertices are the ordered sets such that voice one is 0 or 1, voice two is 8 or 9, and voice three is 10 or 11.

[3.1.5] However, ordered set space is of limited value. Simply transposing *Vers le blanc* would yield an entirely different path in the same space. Likewise, labeling the voices differently, so that the

[10] Equations 2 and 3 are examples of p-norm distances, with p equal to 2 and 1, respectively. p-norm distance is defined as $d_p(X,Y) = (\sum |y_i - x_i|^p)^{\frac{1}{p}}$.

[11] In addition to satisfying Assumption 1, I believe equation 2 yields more intuitive results than equation 3 in certain voice-leading situations. To take an extreme example, consider a 20-note chord. Now move each voice up or down by $\frac{1}{20}$ of a semitone, or 5 cents. According to equation 3 the distance between the two chords, which are barely distinguishable, is 1—exactly the same as if a single voice had moved by a semitone. In contrast, equation 2 yields a distance of approximately 0.22.

[12] Polansky 1996 offers a nice discussion of the differences between Euclidean and city-block distance.

[13] Mapping voices onto axes of an oblique coordinate system would yield a different distance function between ordered sets, though Assumptions 1 and 2 would still be satisfied.

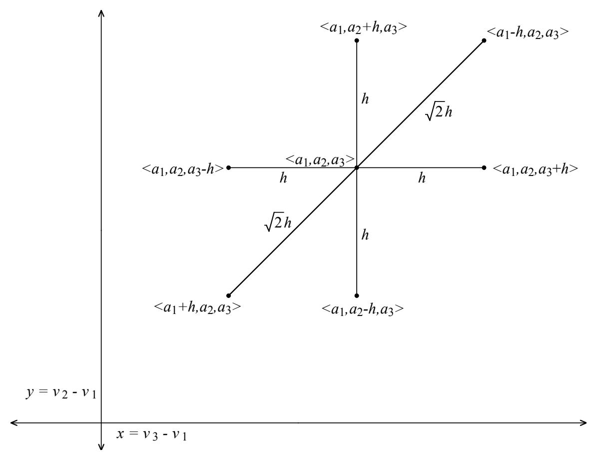


Figure 3: *Problematic mapping of T-classes using Cartesian coordinates.*

interpolation is $\langle 9, 0, 11\rangle \rightarrow \langle 2, 4, 5\rangle$, would yield a separate path. An inversion of *Vers le blanc*—for example, $\langle 11, 2, 0\rangle \rightarrow \langle 7, 9, 6\rangle$ or $T_{11}I(f(t))$—would be yet another path. We would like to have a space in which transposition, permutation, and inversion of voices do not yield separate paths. In short, we would like a space that allows transpositional, permutational, inversional, and even modular (octave) equivalence to be taken into account.

3.2 Transpositional equivalence and T-class space

[3.2.1] We begin by factoring out transposition. The class representative of $/A/_T$ is taken to be $\langle 0, a_2 - a_1, \ldots, a_n - a_1\rangle$, which we will abbreviate as $\langle a_2 - a_1, \ldots, a_n - a_1\rangle_T$.[14] In the case of three voices, the class representative is $\langle a_2 - a_1, a_3 - a_1\rangle_T$. Since three voices have been reduced to two intervals, we can use a two-dimensional space for T-classes of three-voice sets. First, it is necessary to update Assumption 2 for distances between T-classes, designated as ρ instead of d.

Assumption 2a. If a member of one T-class can be transformed into a member of another T-class by moving a single voice, the distance between the two T-classes should be the distance this voice moves. That is, if $/A/_T\ \Delta^h /B/_T$, then $\rho(A, B) = h$.[15]

For example, the distance between $/\langle 2, 6, 10\rangle/_T$ and $/\langle 1, 4, 9\rangle/_T$ should be 1, since $\langle 0, 4, 8\rangle \in /\langle 2, 6, 10\rangle/_T$ can be transformed into $\langle 0, 3, 8\rangle \in /\langle 1, 4, 9\rangle/_T$ by moving voice two a distance of one semitone.

[3.2.2] A simple way to map T-classes onto the Euclidean plane is to set $x = v_3 - v_1$ and $y = v_2 - v_1$, where x and y are the standard orthogonal axes. However, this mapping runs afoul of Assumption 2a. Figure 3 plots the set $A = \langle a_1, a_2, a_3\rangle$ and the six sets to which A is Δ^h-related. Moving either voice two or voice three by $\pm h$ yields a distance of h between A and $\langle a_1, a_2, a_3 \pm h\rangle$ or $\langle a_1, a_2 \pm h, a_3\rangle$. However, moving voice one by $\pm h$ yields a distance of $\sqrt{2}h$ from A to $\langle a_1 \pm h, a_2, a_3\rangle$.

[14] The class representative of $/A/_T$ is the same for all members of the class, since for $A' = \langle a_1 + x, \ldots, a_n + x\rangle$, $\langle (a_2 + x) - (a_1 + x), \ldots, (a_n + x) - (a_1 + x)\rangle_T = \langle a_2 - a_1, \ldots, a_n - a_1\rangle_T$.

[15] The arguments for the function ρ are the T-classes $/A/_T$ and $/B/_T$. Since ρ is not defined on the *sets* A and B, we can use the shorthand $\rho(A, B)$ for the more explicit but cumbersome notation $\rho(/A/_T, /B/_T)$.

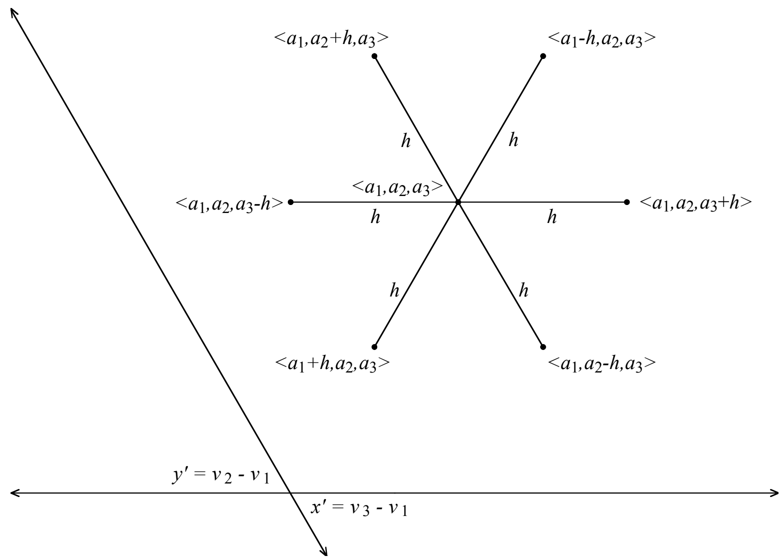


Figure 4: *Mapping of T-classes using oblique coordinates.*

[3.2.3] We can correct this situation by using an oblique coordinate system in which the axes, x' and y', are in a 120° relation rather than 90°. (The axes in this oblique coordinate system are labeled x' and y' to avoid confusion with the orthogonal axes x and y.) We now set $x' = v_3 - v_1$ and $y' = v_2 - v_1$. Figure 4 plots the sets from Figure 3 in this oblique coordinate system. The distance between A and each of the other sets is h. Since this holds for any value of h, Assumption 2a is satisfied. In order to use the standard Euclidean metric (equation 2) to measure the distance between T-classes, it is necessary to map this oblique system onto the normal Cartesian system with orthogonal axes. Any location in the oblique system, (x', y'), will be located at $(x' - \frac{1}{2}y', \frac{\sqrt{3}}{2}y')$ in the Cartesian system. Thus, in the Cartesian system T-classes are located at $\left(v_3 - v_1 - \frac{1}{2}(v_2 - v_1), \frac{\sqrt{3}}{2}(v_2 - v_1)\right)$. The distance between T-classes $/A/_T$ and $/B/_T$ is

$$\rho(A,B) = \sqrt{\left(b_3 - a_3 - \frac{1}{2}\left((b_2 - a_2) + (b_1 - a_1)\right)\right)^2 + \left(\frac{\sqrt{3}}{2}\left((b_2 - a_2) - (b_1 - a_1)\right)\right)^2}. \quad (5)$$

By expanding and recombining terms, we can rewrite equation 5 as

$$\rho(A,B) = \sqrt{\frac{3}{2}\left(\sum_{i=1}^{3}(b_i - a_i)^2 - \frac{\left(\sum_{i=1}^{3}(b_i - a_i)\right)^2}{3}\right)}, \quad (6)$$

which anticipates the form of the general equation for n voices derived in Section 7.[16] The space defined by this mapping and distance function will be refered to as the “T-class space” for ordered sets of three voices.

[16] If $A\,\Delta^h B$, then $\rho(A,B) = \sqrt{\frac{3}{2}\left(h^2 - \frac{h^2}{3}\right)} = h$, satisfying Assumption 2a.

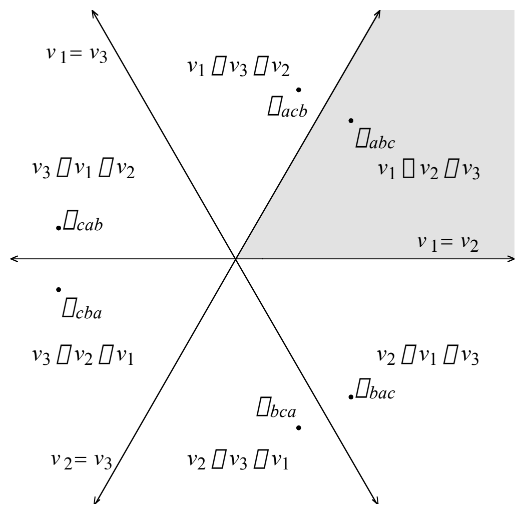


Figure 5: *Permutation regions of T-class space.*

3.3 Permutational equivalence and the normal region

[3.3.1] There are a total of six permutations of voices for any trichord, $\langle a,b,c\rangle$, namely $\langle a,b,c\rangle$, $\langle a,c,b\rangle$, $\langle b,a,c\rangle$, $\langle b,c,a\rangle$, $\langle c,a,b\rangle$, and $\langle c,b,a\rangle$. Figure 5 shows the location in T-class space of the six permutations of α: $\alpha_{abc} = \langle 0,9,11\rangle$, $\alpha_{acb} = \langle 0,11,9\rangle$, $\alpha_{bac} = \langle 9,0,11\rangle$, $\alpha_{bca} = \langle 9,11,0\rangle$, $\alpha_{cab} = \langle 11,0,9\rangle$, and $\alpha_{cba} = \langle 11,9,0\rangle$. Each permutation lies in one of six regions of T-class space bounded by the lines $v_1 = v_2$, $v_2 = v_3$, and $v_1 = v_3$. Each region consists of all sets satisfying one of the following inequalities: $v_1 \le v_2 \le v_3$, $v_1 \le v_3 \le v_2$, $v_2 \le v_1 \le v_3$, $v_2 \le v_3 \le v_1$, $v_3 \le v_1 \le v_2$, or $v_3 \le v_2 \le v_1$. We will take the region defined by the inequality $v_1 \le v_2 \le v_3$ to be the representative permutational region, or the *normal region*, shaded in Figure 5. Every point in T-class space is

[3.3.2] The lines bounding the permutational regions act as mirrors in T-class space. That is, points related by reflection about one of these lines are also related by some permutation. For example α_{abc} and α_{bac} are related by reflection about the line $v_1 = v_2$ and the permutation that exchanges the order position of voices one and two. Reflection about the line $v_i = v_j$ is equivalent to the permutation that exchanges the order position of voices i and j.

3.4 Inversional equivalence and the normal half region

[3.4.1] Figure 6 plots the location of α, ω, $I(\alpha)$ (the inversion of α about 0), and $I(\omega)$ in the normal region. (By the location of A in the normal region, we will intend the location of A' in T-class space, where A' is the permutation of A such that $a'_1 \le a'_2 \le a'_3$.) The normal region can be divided in half by the line $v_3 - v_2 = v_2 - v_1$. What is the musical significance of this line? Any trichord in pitch space is inversionally invariant (I-invariant) under some T_xI if the ordered interval from the bottom to the middle voice is the same as the ordered interval from the middle to the top voice. Thus all three-voice sets in the normal region that are I-invariant lie on the line $v_3 - v_2 = v_2 - v_1$. As can be readily observed in Figure 6, α and ω are congruent to their inversions through reflection about this line.

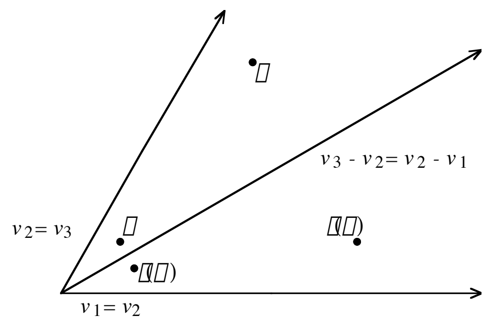


Figure 6: *Location of α, ω, $I(\alpha)$, and $I(\omega)$ in the normal region of T-class space.*

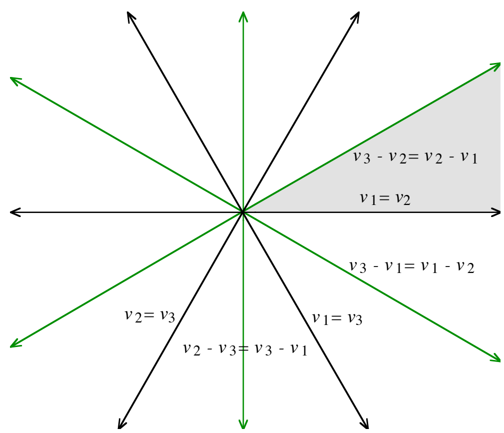


Figure 7: *Half regions in T-class space.*

[3.4.2] Each permutational region is bisected into two half regions by a line corresponding to I-invariant sets as shown in Figure 7. (In Figure 7 boundaries of permutational regions are shown in black, while the bisecting lines are shown in green.) The normal and southwest regions are bisected by $v_3-v_2 = v_2-v_1$, $v_2-v_3 = v_3-v_1$ bisects the north and south regions, and $v_3-v_1 = v_1-v_2$ bisects the southeast and northwest regions. These lines also act as mirrors in T-class space for three-voice sets—points opposite these lines are related by inversion. The lower half region (east northeast) of the normal region contains all sets satisfying the inequality $v_3 - v_2 \geq v_2 - v_1$ (in addition to the inequalities defining the normal region), while the upper half region (north northeast) contains all sets satisfying $v_3 - v_2 \leq v_2 - v_1$. We will take the lower half region to be the representative inversional region, or the *normal half region*, shaded in Figure 7. Every point in T-class space is congruent to some point in the normal half region under permutational *and* inversional equivalence, and no two points within the normal half region are congruent.

3.5 Modular equivalence and the fundamental region

[3.5.1] The last equivalence to consider is modular (or octave) equivalence. Some might wonder if invoking octave equivalence is helpful in understanding continuous transformations like *Vers le*

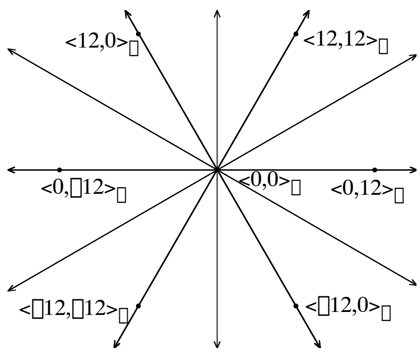


Figure 8: *Points congruent to the origin mod 12 in T-class space.*

blanc. After all, it would seem that a gradual change of octave is not possible. While the objection is certainly reasonable, there are several reasons for taking this last step:

> 1) Consider a variation of *Vers le blanc* in which the bottom voice is raised by an octave; i.e., $\langle 12, 9, 11\rangle \to \langle 16, 2, 5\rangle$. There are noticable differences from the original including the lack of a convergence on a unison and motion from a relatively close spacing in pitch space to a relatively open spacing, which is the opposite of *Vers le blanc*. However, the convergence on an octave instead of a unison is still a very salient event, the relative motion among the voices remains the same, and the resulting harmonic trajectory is similar enough to warrant investigating how these two versions of *Vers le blanc* compare.
>
> 2) If the motion is continuous (or at least nearly so) and the rate of change is slow, a gradual change of pitch class *can* be heard as continuously evolving despite the lack of a consistent register. Example 2 provides a demonstration. The example is a gradual change in pitch class with random changes of octave at periodic time intervals.
>
> 3) It *is* possible to change the perceived octave of a pitch gradually by making use of Shepard tones. These are auditory illusions first created by Roger Shepard in which a continuously descending glissando does not descend in register. (Shepard 1964) (The same illusion is possible for ascending glissandos.) The trick is that as the perceived pitch descends it gradually decreases in volume while the pitch an octave above gradually increases in volume. (Typically, multiple contiguous octaves of the pc are used in order to smooth out the transition.) If executed properly, the perceived pitch rises to the upper octave without the listener being able to identify when the switch occurs.[17] Listen to Example 3 [modem] [broadband] to hear this illusion.[18]

[3.5.2] In asserting modular equivalence, we state that for $P = \langle p_1 \ldots, p_n\rangle$ and $P' = \langle p_1 + c_1 m, \ldots, p_n + c_n m\rangle$, where m is a real number and c_i is an integer, P' is congruent to P mod m. Figure 8 shows a portion of T-class space and those points that are equivalent to the origin, $\langle 0, 0\rangle_T$,

[17] Jean-Claude Risset has explored this illusion in numerous works and has extended it so that the register of a gradually descending pitch ascends or vice-versa.

[18] Example 3 is available on the internet at http://www.crowncity.net/ratcave/Audio/Shepard_Tones.mp3. Similar demonstrations are available at http://asa.aip.org/demo27.html, which is maintained by the Acoustical Society of America.

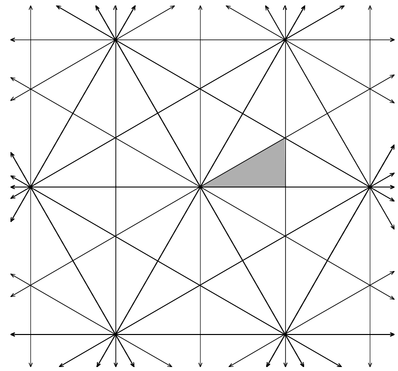


Figure 9: *T-class space tiled by the fundamental region.*

with $m = 12$. Each of these points can be viewed as a translation of the origin. For instance, we reach $\langle 0, 12\rangle_T$ by moving 12 units along the x'-axis. Since $\langle 0, 0\rangle_T$ and $\langle 0, 12\rangle_T$ are equivalent, the boundaries of the half regions that intersect at the origin are equivalent to their translation 12 units along the x'-axis, where they intersect at $\langle 0, 12\rangle_T$. For example, the line $v_2 = v_3$ is congruent to the line $v_2 = v_3 - 12$. Figure 9 shows the same portion of T-class space with half region boundaries translated to intersect at those points that are equivalent to the origin. These lines carve T-class space into equivalent triangular regions that are some combination of translation and/or rotation and/or reflection of the shaded region. (This is the triangular region that lies in the normal half region and contains the origin.) The infinite plane of T-class space is tiled by this region without overlaps or gaps. Thus, this region is the *fundamental region.*[19] The fundamental region contains all possible trichordal set classes—that is, every ordered set of three *real* numbers is equivalent to some point in the fundamental region under transpositional, permutational, inversional, and modular equivalence—and no two points in the fundamental region belong to the same set class. We will use the symbol Π^3 to refer to the fundamental region of T-class space for three-voice sets.

[3.5.3] Figure 10 plots the twelve familiar trichordal set classes, as well as the seven multiset classes along the x'-axis, of 12-tone equal temperament in the fundamental region.[20] The upper and lower boundaries of Π^3 have already been discussed. The right boundary, $v_2 + 12 - v_3 = v_3 - v_1$ or $v_1 + v_2 + 12 = 2v_3$, is a translation of the vertical line bisecting the north and sorth permutational regions corresponding to the I-invariant sets in those regions. Since this boundary is a translation of a line of I-invariance, points opposite this line are equivalent under pc inversion.

[3.5.4] Distances between points in Π^3 represent the minimal distance between members of the respective set classes. For instance, the distance between $\langle 2, 5\rangle_T$ ([0,2,5]) and $\langle 3, 7\rangle_T$ ([0,3,7]) is $\sqrt{3} \approx 1.73$. There are no two members of $[0, 2, 5]$ and $[0, 3, 7]$ that are closer than $\sqrt{3}$ in T-class space. We can define ρ for set classes as

$$\rho(/A/, /B/) = \min\left(\rho(A', B')\right), \tag{7}$$

for all $A' \in /A/$ and $B' \in /B/$.

[19]See Coxeter 1973, pp. 63, 76-87, for a definition and discussion of fundamental regions.

[20]Multisets are unordered sets in which multiple occurrences of an element are counted separately. For example, $\langle a, b, a\rangle$ is an ordering of the multiset $\{a, a, b\}$, which is *not* the same as $\{a, b\}$.

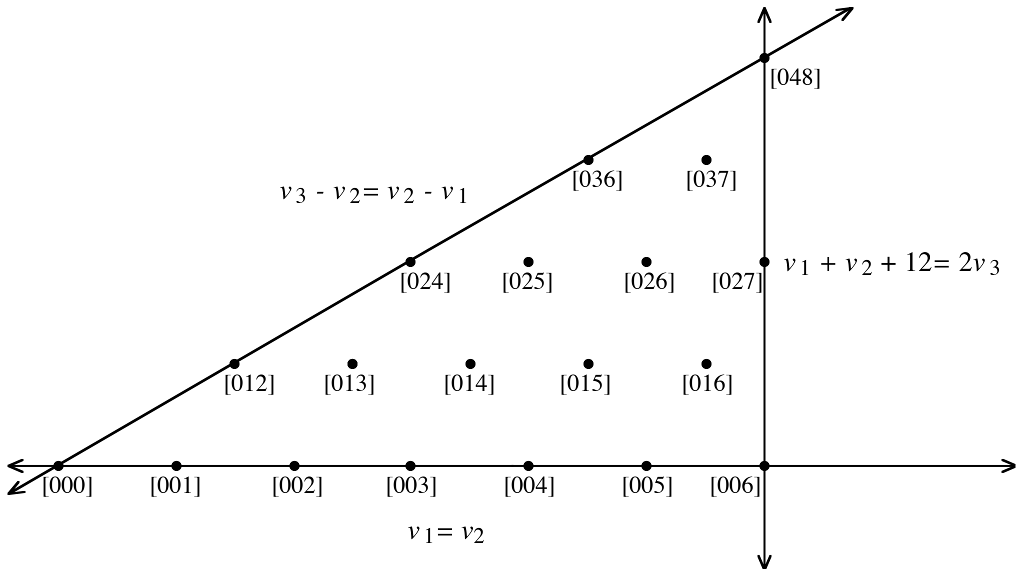


Figure 10: *Trichord set-classes in 12-tone equal temperament in the fundamental region.*

3.6 Comparison with Roeder 1987

[3.6.1] The construction of T-class space and its various regions is very similar to the geometric approach in Roeder 1987, and a brief comparison of the two is in order.[21] Roeder constructs an "ordered interval space" (OI-space) for ordered pcsets, where the trichord A is located in a 12 x 12 region of the Cartesian plane at $(a_2 - a_1, a_3 - a_2)$, and the arithmetic is mod 12. (The model is extendable to higher dimensions for pcsets of greater cardinality.) Since it is limited to pcsets where the pitch classes are integers, OI-space for trichords is a toroidal lattice. Roeder provides algebraic and geometric descriptions of pcset inclusion, exclusion, and inversion, and divides ordered interval space into regions of permutational and inversional equivalence. (These equations can be translated into corresponding equations for T-class space.) These regions contain members of each T_nI-type without redundancy, similar to Π^3.[22]

[3.6.2] There are several differences between OI-space and T-class space: 1) the latter is Euclidean (modular equivalence carves T-class space into regions as opposed to being assumed in toroidal OI-space); 2) the latter is continuous (though, since Roeder's space is generalizable to any modulus, m, it can be made continuous by allowing m to approach infinity); and 3) the axes in T-class space are oblique rather than orthogonal. The first difference allows the use of the Euclidean metric. The third difference accounts for the fact that Roeder's T_nI regions are not congruent, in contrast with Π^3, which is a fundamental region. (Skewing OI-space by placing the axes in a 60° relation yields the same triangular regions as in T-class space.)

[3.6.3] The biggest difference between the two spaces for the purposes of this paper is that distance in T-class space measures how far the *voices* move, whereas OI-space measures how much the

[21] My thanks to Robert Morris for initially pointing out the similarities and directing me to Roeder's work.

[22] There are numerous other details that the interested reader is encouraged to study in both Roeder 1987 and 1984.

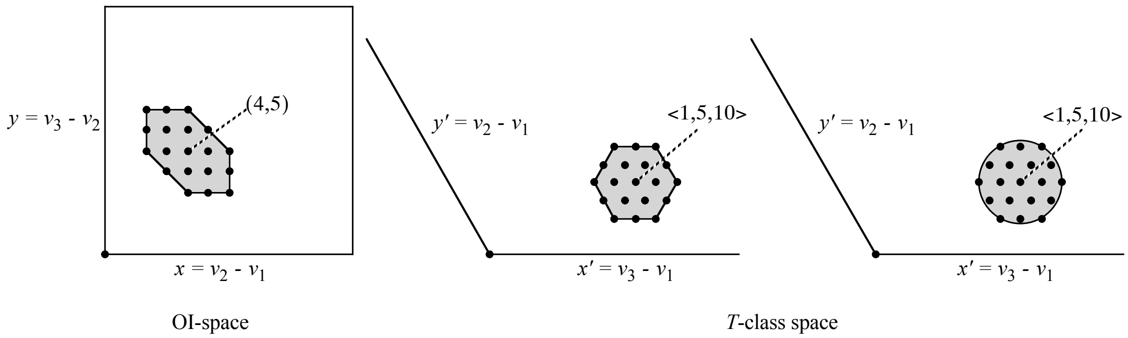


Figure 11: *Semitonal voice-leading regions in OI and T-class space.*

intervals change. Certainly, there is a strong relation between the two, but they are not identical. The distance between two three-voice sets A and B in T-class space is given by equation 6. In OI-space the distance between A and B is the "absolute sum of the differences, considered as interval classes, of the intervals" of each voice in A and B respectively. If $A = \langle 1, 5, 10\rangle$ and $B = \langle 1, 6, 10\rangle$, the distance between them in T-class space is 1, since a single voice moves a distance of 1. In OI-space the distance is two, because the ordered interval series changes from $\langle\langle 4, 5\rangle\rangle$ to $\langle\langle 5, 4\rangle\rangle$, using double brackets to distinguish ordered interval series from ordered sets. Depending on the analytical focus, either space may be more suitable than the other. For another example that demonstrates these differences, consider two ordered sets of three voices, Q and R, where each voice moves at most a distance of 1: $|r_i - q_i| \leq 1$. For $Q = A$, the region satisfying these inequalities in OI-space is the region on the left side of Figure 11. (See Roeder 1987, Example 15) In T-class space the same system of inequalities leads to the hexagonal region in the center of the figure. Generalizing the situation, the same equal-tempered sets are contained in the circular region on the right, where Q is the center and the radius is two. This is the region of all sets, R, such that $\rho(Q, R) \leq 2$. (Similar regions obtain for analogous situations in higher dimensions.)

4 Saariaho, Vers le blanc

[4.1] Figure 12 shows the path taken by *Vers le blanc* in T-class space. T-classes of 12-tone equal tempered sets that lie near this path are included in Figure 12 as reference points. As discussed previously, one advantage of plotting f in T-class space versus the ordered set space of Section 3.1 is that all transpositions of *Vers le blanc* result in the same path. Since sets of integers form a cubic lattice in ordered set space and a triangular lattice in T-class space, another advantage is that the number of neighboring 12-tone equal tempered reference points is reduced from eight to three. For instance, $f(\frac{1}{9})$ lies in the shaded triangle whose vertices are the T-classes $\langle 8, 10\rangle_T$, $\langle 7, 9\rangle_T$, and $\langle 7, 10\rangle_T$. Additionally, the convergence of voices one and two on a unison is explicitly represented by the intersection of f and the x'-axis (where $y' = v_2 - v_1 = 0$).

[4.2] A few cautionary notes are in order. Not only do constant transpositions of f trace the same path in T-class space, but variable transpositions in which the contour of the voices is distorted also trace the same path. For instance, consider a variation of *Vers le blanc* in which α is untransposed but ω is transposed by 30 semitones—i.e., $\langle 0, 9, 11\rangle \to \langle 34, 32, 35\rangle$. We could call this variation

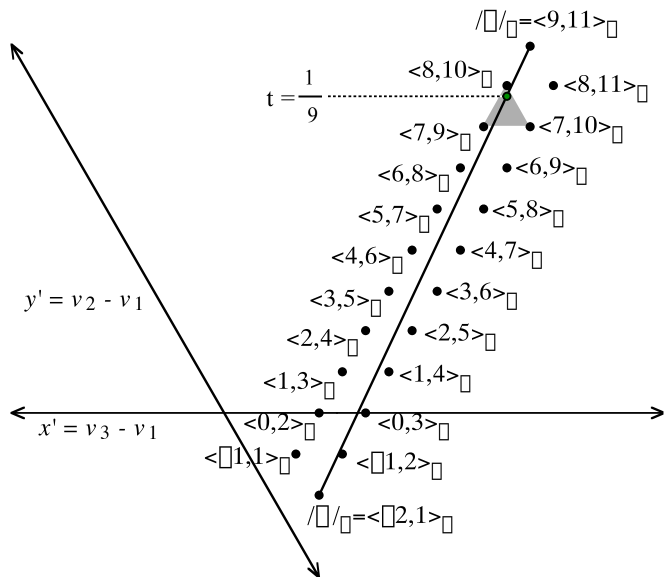


Figure 12: *Trajectory of* Vers le blanc *in T-class space.*

$T_{30t}(f)$. Given the fast rate of change (relative to the original) and uniform direction of all voices, $T_{30t}(f)$ would seem to be fairly different from f, yet it traces the same path in T-class space. For a more drastic variation, consider a transposition of f that changes randomly and abruptly every 200 milliseconds. Since transposition has been factored out, this more drastic variation also traces the same path in T-class space. The path in Figure 12 thus represents a family of interpolations that move linearly from $/\alpha/_T$ to $/\omega/_T$, *Vers le blanc* being just one member of this family. Nonetheless, as long as we remember the specifics of a given continuous transformation, we can take advantage of the perspectives offered by Figure 12 and the next few figures.[23]

[4.3] Figure 13 shows the six possible paths that *Vers le blanc* can take depending on how we order the voices. f_{abc} extends from α_{abc} to ω_{abc} $(= \langle 4, 2, 5\rangle)$, f_{bac} extends from α_{bac} to ω_{bac}, and so forth. It is easy to see in Figure 13 that f_{abc} is congruent to f_{bac} through reflection about the line $v_1 = v_2$. f_{abc} is also congruent to f_{acb} through reflection about the line $v_2 = v_3$ and f_{cba} through reflection about the line $v_1 = v_3$. f_{abc} begins in the normal region and moves out of this region as it crosses the boundary $v_1 = v_2$. Likewise, f_{bac} begins out of the normal region and moves into this region as it crosses $v_1 = v_2$, intersecting this line at the same point at which it intersects f_{abc}. Limiting our perspective to just the normal region, we can trace the path of *Vers le blanc* by taking only the portion of f_{abc} and f_{bac} lying in the normal region, shown as a blue line in Figure 13. The voice crossing in *Vers le blanc* is now shown by the reflection of f off of the boundary $v_1 = v_2$. The perceptual effect of this and other reflections will be discussed in greater detail below.

[4.4] In Figure 14 the portion of the blue line in Figure 13 that lies outside of the normal half region

[23] The tradeoffs between abstract categories and specific instances of these categories are of course very familiar. In particular, the present discussion is closely related to graphs of minimal voice leading between trichordal and tetrachordal set-classes in Cohn 2003 and Straus 2003. These graphs represent the *potential* for minimal voice-leading between members of connected set classes. Likewise, points that lie within an arbitrarily small neighborhood in T-class space represent the *potential* for infinitesimal transformations connecting members of the respective T-classes.

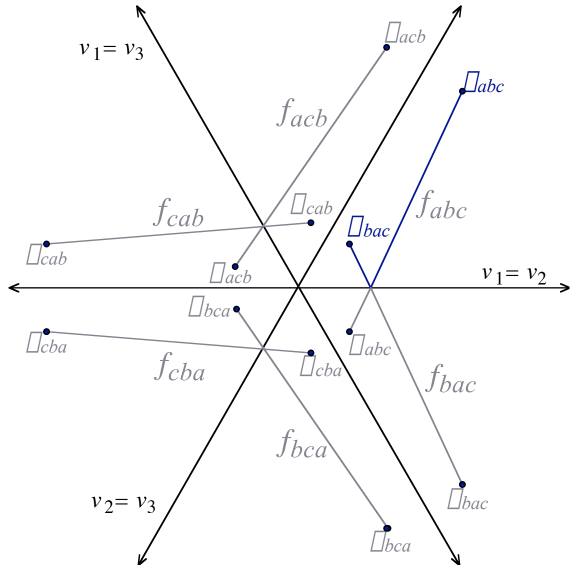


Figure 13: *Permutations of* Vers le blanc *and its trajectory through the normal region.*

is shown in grey. Reflecting these grey segments about the line of I-invariant pitch sets yields the path of *Vers le blanc* in the normal half region, shown in blue. Beginning at α on the far right, the path moves to the left, toward the line of I-invariance. Reflecting off of this boundary, it moves toward the permutation boundary and reflects back toward the inversion boundary. Reflecting for a second time off of the inversion boundary, *Vers le blanc* moves to its conclusion at ω.

[4.5] Similarly, the portion of Figure 14 that lies outside of the fundamental region is the grey segment on the right side of Figure 15, AB. Reflecting this segment about the right boundary of Π^3 yields the segment BD. A portion of this reflected segmented also lies outside of Π^3. Reflecting CD about the upper boundary yields $C\omega$ and completes the path of *Vers le blanc* in the fundamental region. This is the graph of the set-class of f, or $/f/$.

[4.6] What does this tell us about the large-scale organization of the piece? The composer suggests that as the harmony resulting from the multiple glissandos approaches a more traditional sonority, listeners tend to notice that the harmonic situation has changed. (Saariaho 1987) So we can get a better sense of the harmonic trajectory by inspecting its relation to set classes in 12-tone equal temperament, which serve as cognitive reference points.[24] (See Figure 16.) Beginning with [013], f initially moves toward [024]. It then moves in the direction of [025], [026], and [027] and back again, with increasing distance from these familiar sonorities. The next section consists of a motion to a dyad somewhat less than a minor third, followed by a final motion toward, away from, and then to its starting point, [013].[25] It is important to remember that ricochets in the graph of f do

[24] This is another strong motivation for invoking modular (octave) equivalence for a continuous transformation in *pitch* space—limiting the number of reference points to the usual 12 trichordal set classes (and 7 multiset classes) instead of the infinite reference points in an unbounded region.

[25] In personal correspondence, Dmitri Tymoczko has raised interesting questions about the conditions under which a linear interpolation will cross an equal-tempered set class. For an interpolation, $A \to B$, suppose that A is an equal-tempered T-class; i.e. $(a_i - a_1) \in \mathbb{Z}$. If there exists a c such that $\frac{b_i - a_i}{c} \in \mathbb{Z}$, where the greatest common factor of $\frac{b_i - a_i}{c}$ for all i is 1, then for all $t' = \frac{k}{c}, k \in \mathbb{Z}$, $f(t')$ is an equal-tempered T-class; $(f_i(t') - f_1(t')) \in \mathbb{Z}$. Otherwise,

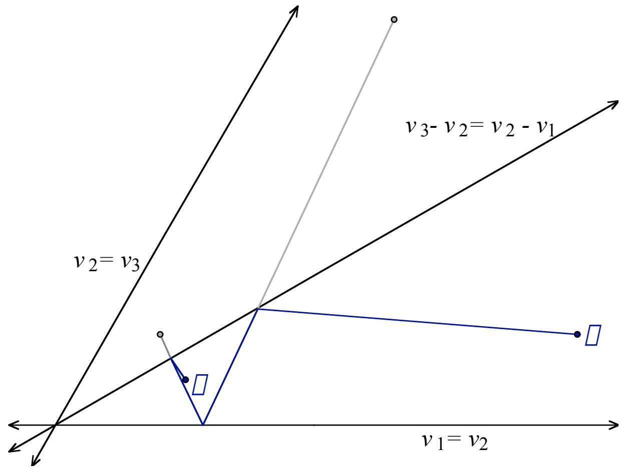


Figure 14: *Trajectory of* Vers le blanc *through the normal half region.*

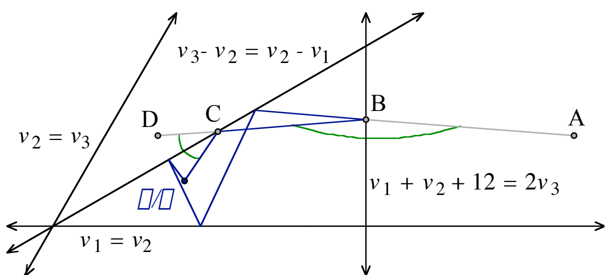


Figure 15: *Trajectory of* Vers le blanc *through the fundamental region.*

not represent sharp changes in the harmonic interpolation. They are, rather, gradual convergences with the boundaries of Π^3.

[4.7] The perceptual effect of these convergences differs depending on the boundary involved. (Figure 17 highlights three segments of f and labels 12 time points, t_0 through t_{11}, including every reflection against a boundary of Π^3. The table beneath the figure lists each timepoint and its corresponding set class and timings in *Vers le blanc* and the simulation in Example 1 [modem] [broadband].) The reflection at t_8 is clearly audible as the moment when voices one and two converge on a unison. The moments when f converges on the other boundaries are not as salient, since it is difficult to hear the precise moment when f becomes an I-invariant set. However, one can hear these reflections indirectly, especially those occuring at t_3 and t_{10}, where f retraces nearly the same path in Π^3. The passage of *Vers le blanc* from t_2 to t_4 remains in a small region of Π^3 where the set classes are highly similar. The distance between $f(t_2)$ and $f(t_4)$ is only $\rho = \sqrt{\frac{1}{147}} \approx 0.082$. For the passage between t_9 and t_{11}, the situtation is identical. (This is due to the fact that f "strikes" both boundaries at the same angle of incidence.) Compare these passages with an equal stretch of time from t_6 to t_7, but with no intervening reflection. The distance from $f(t_6)$ to $f(t_7)$

there are no values of $t \neq 0$ such that $f(t)$ is an equal-tempered T-class.

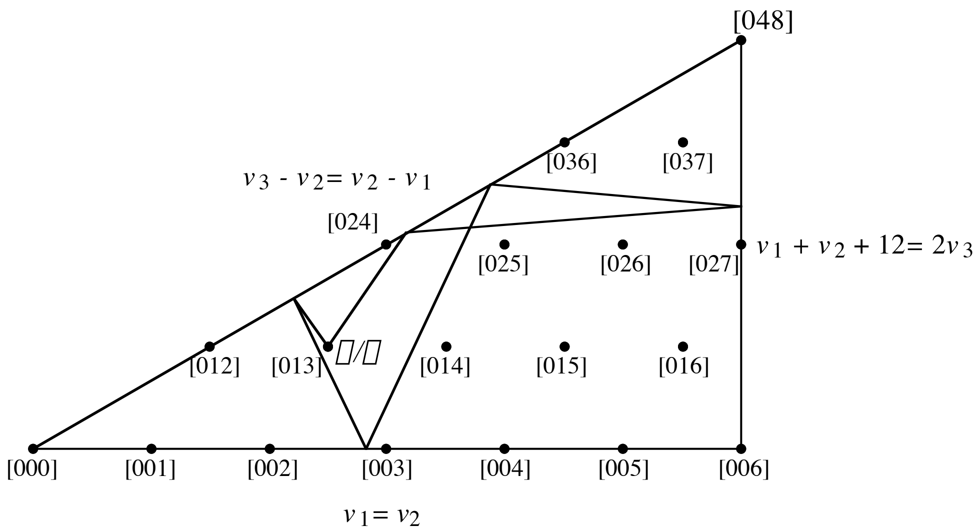


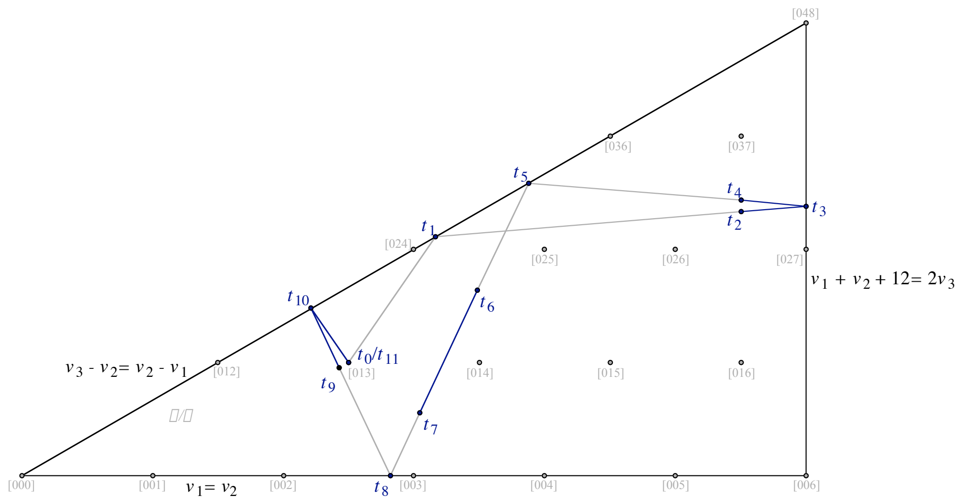
Figure 16: */f/ with 12-tone equal tempered set classes as reference points.*

is $\rho = \sqrt{\frac{148}{147}} \approx 1.003$. This is, of course, an immediate consequence of the fact that the boundaries act as mirrors. The closer the angle of incidence formed by f and the boundary is to 90°, the more pronounced the effect is.

[4.8] In considering the relation of f to familiar reference points from 12-tone equal temperament, it is useful to note that some set classes have a greater cognitive pull than others.[26] For example, f does not need to be as close to [037] in order to make the association as it does to [036]. To find the point at which f is closest to [037] (or some other consonant-triad-like sonority), we draw a perpendicular from the set class to f. (Figure 18) The distance from f to [037] is minimized at $t = \frac{16}{37} \approx$ 6'25" ($\approx$ 1'05" through the simulation in Example 1). Prior to [037], f is closest to [027] when $t = \frac{14}{37} \approx$ 5'51" ($\approx$ 57" into Example 1). Very gradual voice leading from [027] to [037] has the potential, exploited in *Vers le blanc*, to sound like a 4-3 suspension. (The potential is realized here since the second chord is a *major* triad and the rate of change for each voice, particularly the bass, is very slow.) The aural experience is somewhat akin to a distorted rendering of a V^{4-3} in a major key. Of course, I would not want to take the analogy too far, but the gradual emergence and submergence of a motion even vaguely similar to tonal voice leading is likely to be latched on to by a listener at sea in very unfamiliar aural waters.

[4.9] Other reference points are suggested by the convergence of one of the intervals with pure thirds, fourths, and fifths. These convergences are marked by interference beats before and after the pure interval, itself marked by the complete cessation of beats. By fixing two pitches in an interval class of a pure fourth/fifth or pure major third and allowing the third pitch to vary independently, two closed loops are formed in Π^3. (In Figure 19, green lines indicate sets containing a pure major third and red lines indicate sets containing a pure fourth/fifth.) Each loop corresponds to all possible trichords containing the respective pure intervals. Intersections between f and the two loops occur with regularity over the opening three-fourths of the piece, as summarized in Figure 19, including

[26]This line of thinking was sparked by a comment from Nancy Rogers.



t_i	$/f(t_i)/$	time in *Vers le blanc*	time in simulation
t_0	$[0, 1, 3]$	0’	$\approx$ 0’
t_1	$\left[0, 2\frac{1}{9}, 4\frac{2}{9}\right]$	$\frac{1}{9} \approx$ 1’55”	$\approx$ 12”
t_2	$\left[0, 2\frac{1}{3}, 6\frac{2}{3}\right]$	$\frac{7}{21} =$ 5’	50”
t_3	$\left[0, 4\frac{17}{21}, 7\frac{4}{21}\right]$	$\frac{8}{21} \approx$ 5’43”	$\approx$ 57”
t_4	$\left[0, 2\frac{3}{7}, 6\frac{5}{7}\right]$	$\frac{9}{21} \approx$ 6’26”	$\approx$ 1’04”
t_5	$\left[0, 2\frac{7}{12}, 5\frac{1}{6}\right]$	$\frac{7}{12} \approx$ 8’45”	$\approx$ 1’27”
t_6	$\left[0, 1\frac{2}{3}, 4\frac{1}{3}\right]$	$\frac{14}{21} =$ 10’	1’40”
t_7	$\left[0, \frac{13}{21}, 3\frac{8}{21}\right]$	$\frac{16}{21} \approx$ 11’26”	$\approx$ 1’54”
t_8	$\left[0, 0, 2\frac{9}{11}\right]$	$\frac{9}{11} \approx$ 13’34”	$\approx$ 2’3”
t_9	$\left[0, \frac{20}{21}, 2\frac{19}{21}\right]$	$\frac{19}{21} \approx$ 13’34”	$\approx$ 2’16”
t_{10}	$\left[0, 1\frac{10}{21}, 2\frac{20}{21}\right]$	$\frac{20}{21} \approx$ 14’17”	$\approx$ 2’23”
t_{11}	$[0, 1, 3]$	15’	2’30”

Figure 17: *Reflections of f off of the boundaries of* Π^3.

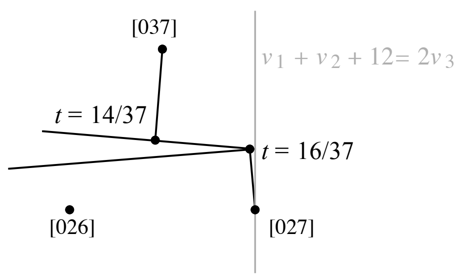


Figure 18: V^{4-3} *in* Vers le blanc.

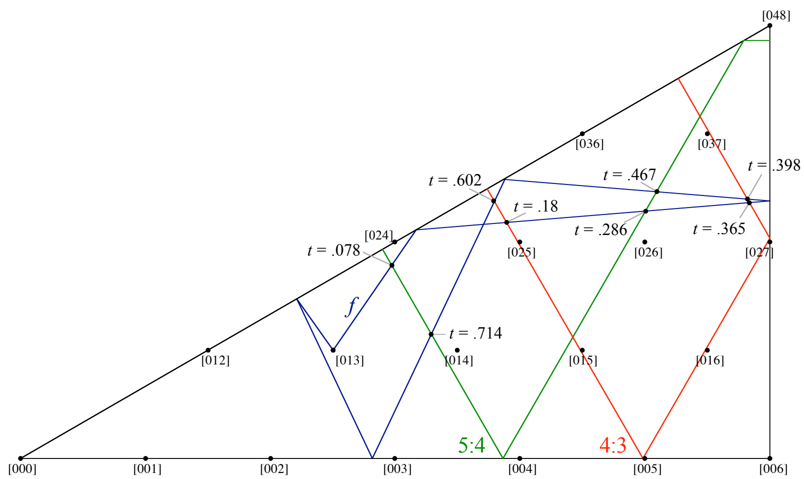


Figure 19: *Intersection of f with pure 4^{th}s/5^{th}s and pure major 3^{rd}s.*

the pure fourth and pure fifth in rapid succession near the vertical boundary. (This is another consequence of the effect discussed in [4.7].) Pure harmonic intervals are noticeably lacking in the closing fourth of the piece.

[4.10] We can also characterize the space in terms of relative evenness and unevenness of set classes. (Clough and Douthett 1991) Points in the space that lie closer to maximally even sets, [048] for trichords, represent set classes with a more even distribution of pitch classes than points that lie farther away.[27] For example, [037] is closer to [048] than [025], and is thus more even. The harmonic progression of *Vers le blanc* is a large-scale motion toward and then away from greater evenness.

[4.11] Finally, the intersection of f with the x'-axis is a very prominent formal event, occurring $\frac{9}{11}$ or 12'16" into the piece. The moments just before and after the convergence with a flat minor third contain the highest degree of dissonance in the entire piece. Indeed, associating tension with dissonance, *Vers le blanc* is weighted heavily toward the end, with the highest levels of tension coming in the final third of the work. This yields a dramatic profile not easily intuited from the information provided in Figure 1.

5 Nancarrow, Study No. 22

[5.1] For an example of how these ideas apply to continuous transformations in a domain other than pitch, we turn to an acceleration canon from the player-piano works of Conlon Nancarrow. The opening section of Study No. 22 is a canon in three voices in which each voice accelerates at a different rate. The specifics of this acceleration, taken from Gann 1995, are summarized in figure 20. The lowest voice enters first, beginning at a tempo of approximately 45 beats per minute, and accelerates at a rate of 1% to a tempo of approximately 116 beats per minute. The highest voice enters approximately 14 seconds later at approximately 45 beats per minute, but accelerates at a

[27] For another approach to determining the relative evenness of a set class see Bloch and Douthett 1994.

lowest voice: M.M. $\approx$ 45 to 116 (acceleration = 1%)
highest voice: M.M. $\approx$ 45 to 185 (acceleration = 1.5%)
middle voice: M.M. $\approx$ 45 to 373 (acceleration = 2.25%)

Figure 20: *Accelerations in the opening of Nancarrow's Study No. 22.*

rate of 1.5% to a final tempo of approximately 185 beats per minute. The final voice to enter, the middle voice, begins about 29 seconds after the highest voice and accelerates from approximately 45 to almost 373 beats per minute at a rate of 2.25%.[28] The entrances of each voice are timed such that the final note of the canon line is attacked simultaneously in all voices approximately 1.37 minutes or 82 seconds into the piece.

[5.2] While the information in figure 20 provides us with a starting point in understanding the passage, it does not show the tempo relationships between the voices and how these relationships change over time. Figure 21, which graphs the acceleration of each voice over time, is an improvement. With this graph we can note certain prominent tempo relationships between the voices, such as various moments when tempos are equivalent and which voices are the fastest or slowest at any moment. But we are still lacking a detailed representation of how these accelerandos interact.

[5.3] It would be nice if we could use the T-class space already constructed for *Vers le blanc*, but there are important obstacles. The primary problem is that intervals between pitches are measured by their differences while intervals between tempos are measured by their ratios. This leads to numerous differences between pitch and tempo spaces. For instance, the sets $\langle 3,4,5\rangle$ and $\langle 4,5,6\rangle$ are equivalent (under transposition) as pitch sets, but not as tempo sets. The sets $\langle 3,4,5\rangle$ and $\langle 6,8,10\rangle$ are equivalent as tempo sets—the second is twice as fast as the first—but not as pitch sets. However, the problem lies in the default level of description for pitch and tempo. If we normally described pitch in terms of frequency, our notion of interval would carry over naturally from one domain to the other.

[5.4] To make this more clear, let's consider a dyad expressed as the frequencies f_1 and f_2 measured in beats per second (Hertz). The interval between the two frequencies is given by the ratio $\frac{f_2}{f_1}$. We wish to convert this *frequency ratio* to a *pitch interval*. Setting $f_1 = 2^{\frac{v_1}{c}}$ and $f_2 = 2^{\frac{v_2}{c}}$, where c is the number of equal divisions of the octave, the interval is

$$\frac{f_2}{f_1} = \frac{2^{\frac{v_2}{c}}}{2^{\frac{v_1}{c}}} = 2^{\frac{v_2-v_1}{c}}. \tag{8}$$

[28] Gann refers to this type of acceleration as a geometric acceleration. For an acceleration of x%, each duration is $1+\frac{x}{100}$ as long as the following duration. Expressing this acceleration as a continuous function with respect to time we have

$$f(t) = -\frac{(r-1)^2}{(d+(c-t)(r-1))^2 \ln r},$$

where d is the initial duration, $r = \left(1+\frac{x}{100}\right)^{-1}$, c is the time at which the acceleration begins, t is measured in minutes, and $f(t)$ yields the number of elapsed beats at time t. The corresponding tempo function, used for the graphs in Figure 21, is

$$\frac{d}{dt} f(t) = \frac{(r-1)}{(d+(c-t)(r-1)) \ln r}.$$

For all three voices in this example the initial duration is $\frac{1}{45}$ minutes or $1\frac{1}{3}$ seconds. For more on this type of acceleration and the derivation of these functions, see Callender 2001.

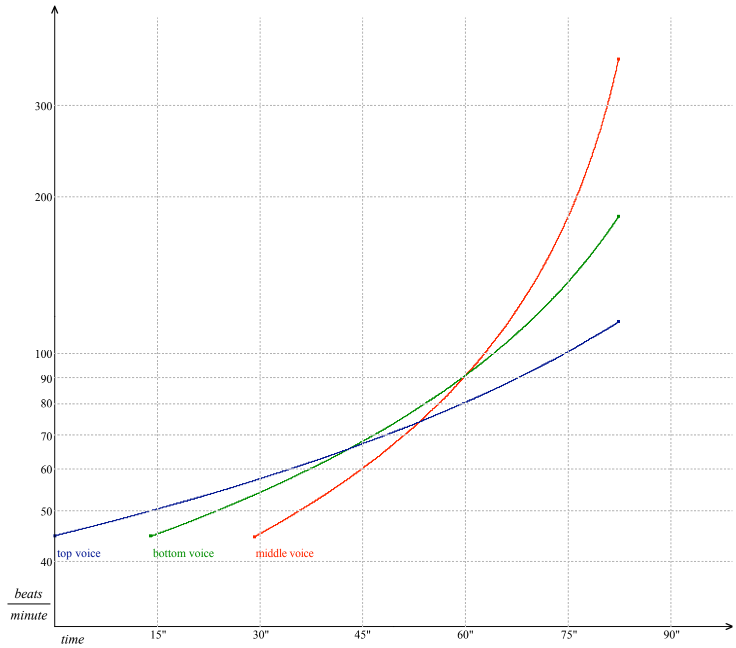


Figure 21: *Tempo graph for opening of Nancarrow's Study No. 22.*

Solving for $v_2 - v_1$ we have

$$v_2 - v_1 = c \log_2 \frac{f_2}{f_1}, \tag{9}$$

which, for $c = 12$, gives the *directed pitch interval* of the dyad in terms of semitones.

[5.5] For example, let $A_f = \langle f_1, f_2, f_3 \rangle$ be an ordered set of frequencies (measured in Hertz) with $f_1{:}f_2{:}f_3 = 3{:}4{:}5$. A_f is a just-intonation triad in six-four position. Let i_1 be the directed pitch interval from f_1 to f_2 and i_2 be the directed pitch interval from f_1 to f_3. We can then plot A_f in T-class space with $x' = i_2$ and $y' = i_1$. According to equation 9, $i_1 = 12 \log_2 \frac{4}{3} \approx 4.98$ semitones and $i_2 = 12 \log_2 \frac{5}{3} \approx 8.84$ semitones. If $A = \langle a_1, a_2, a_3 \rangle$ is the set of pitches corresponding to f_1, f_2, and f_3, then A_f and A are located at the same point in T-class space. Scaling a frequency set by a constant is the equivalent of transposing a pitch set: $zA_f = \langle zf_1, zf_2, zf_3 \rangle = T_{c \log_2 z}(A)$. The *scalar class* of A_f is $\left\{ A'_f \mid A'_f = zA_f, z \in \mathbb{R} \right\}$. We will write the scalar class of A_f as $\langle f_1{:}f_2{:}f_3 \rangle$. The isomorphism between T-classes and scalar classes is

$$\begin{aligned} \langle f_1{:}f_2{:}f_3 \rangle &= /\langle 0, c \log_2 \tfrac{f_2}{f_1}, c \log_2 \tfrac{f_3}{f_1} \rangle/_T, \text{and} \\ /\langle a_1, a_2, a_3 \rangle/_T &= \langle 2^{\frac{a_1}{c}}{:}2^{\frac{a_2}{c}}{:}2^{\frac{a_3}{c}} \rangle. \end{aligned} \tag{10}$$

Thus we can write $\langle 3{:}4{:}5 \rangle = /\langle 0, 12 \log_2 \frac{4}{3}, 12 \log_2 \frac{5}{3} \rangle/_T$. (Hereafter, we will refer to scalar classes as T-classes.)

[5.6] Suppose the frequencies in the above example were not measured in Hertz but in beats per *minute*. Then f_1, f_2, and f_3 will be perceived as *tempos* (assuming their values fall within the

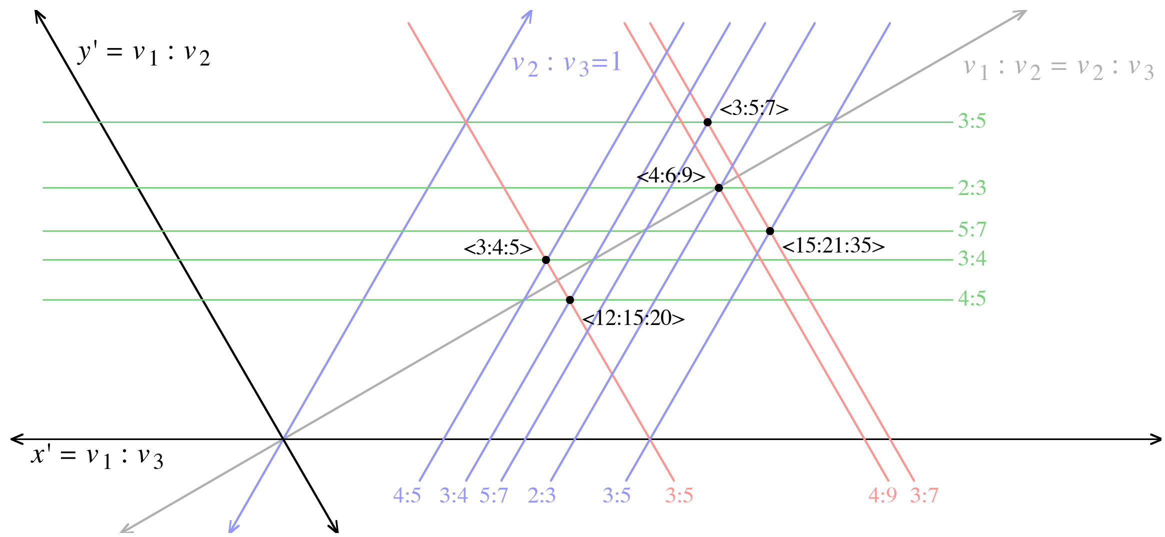


Figure 22: *Tempo sets plotted in T-class space.*

appropriate range). However, with respect to T-class space the absolute values of f_1, f_2, and f_3 are irrelevant. What matters is that their ratios are identical. Thus A_f is located at the same point in T-class space regardless of whether its elements are perceived as pitches or tempos.[29] (This assumes that $c = 12$. In fact, it is more intuitive to set $c = 1$ for tempos, using the "octave" or 1:2 ratio as the basic unit.) Figure 22 plots tempo sets with the ratios ⟨3:4:5⟩, ⟨3:5:7⟩, ⟨4:6:9⟩, ⟨12:15:20⟩, and ⟨15:21:35⟩ in T-class space.[30] The green, red, and blue lines correspond to various ratios. The green lines represent the ratio between voices 1 and 2 and are parallel to the x'-axis, itself corresponding to the ratio 1:1. Red lines represent the ratio between voices 1 and 3 and are parallel to the y'-axis, also corresponding to the ratio 1:1. Blue lines represent the ratio between voices 2 and 3 and are parallel to the line forming a 120° angle with both the x'- and y'-axes. (N.B.—A red line labeled by the ratio f_1:f_2 is the line $x' = \log_2 \frac{f_2}{f_1}$ <u>not</u> $x' = \frac{f_2}{f_1}$. The grey line will be discussed below.) Since we are more comfortable discussing tempos as ratios, these lines are a more convenient reference than equal divisions of the "octave."

[5.7] To graph the acceleration canon as a trajectory we follow the same steps as for the pitch interpolation of *Vers le blanc*. Figure 23 graphs the acceleration canon in T-class space. (In Figure 23 the x'- and y'-axes are labeled as the lines $v_1 : v_2 = 1$ and $v_1 : v_3 = 1$, respectively. Also, the figure does not take equivalence classes other than transposition into account. We will do so shortly.) The graph begins with the entrance of the middle voice (where $t = 0.484$ minutes and $g(0.484) = \alpha$) when the tempo relationships are approximately 1.21:1:1.3, which is very close to 17:14:18. (NB—17:14:18, which indicates that voice 1 is faster than voice 2, is *not* the same as 14:17:18, which indicates that voice 2 is faster than voice 1.) The trajectory, designated by the function g, curves around the origin and appears to flatten out near the conclusion (where $t = 1.372$ and $g(1.372) = \omega$).[31] We can also see the tempo convergences where g crosses black lines

[29] Locating tempo sets in T-class space is essentially the same step taken in Lewin 1987 (chapter 4), where tempos from a portion of Elliot Carter's String Quartet no. 1 are represented as pitches.

[30] Nancarrow's Study No. 46 contains ostinatos in 3:4:5 tempo ratios; Studies Nos. 17 and 19 are canons in 12:15:20; Ligeti's Piano Etude, No. 6, opens with tempo ratios 15:21:35; and 4:6:9 are the tempo ratios for the first movement of the author's own *Clarinet Concerto*.

[31] In general, since the accelerations are non-linear, g will not be a straight line. However, if the acceleration were allowed to continue beyond the end of the canon, the fastest voice (voice 2) would increase at such a high rate that

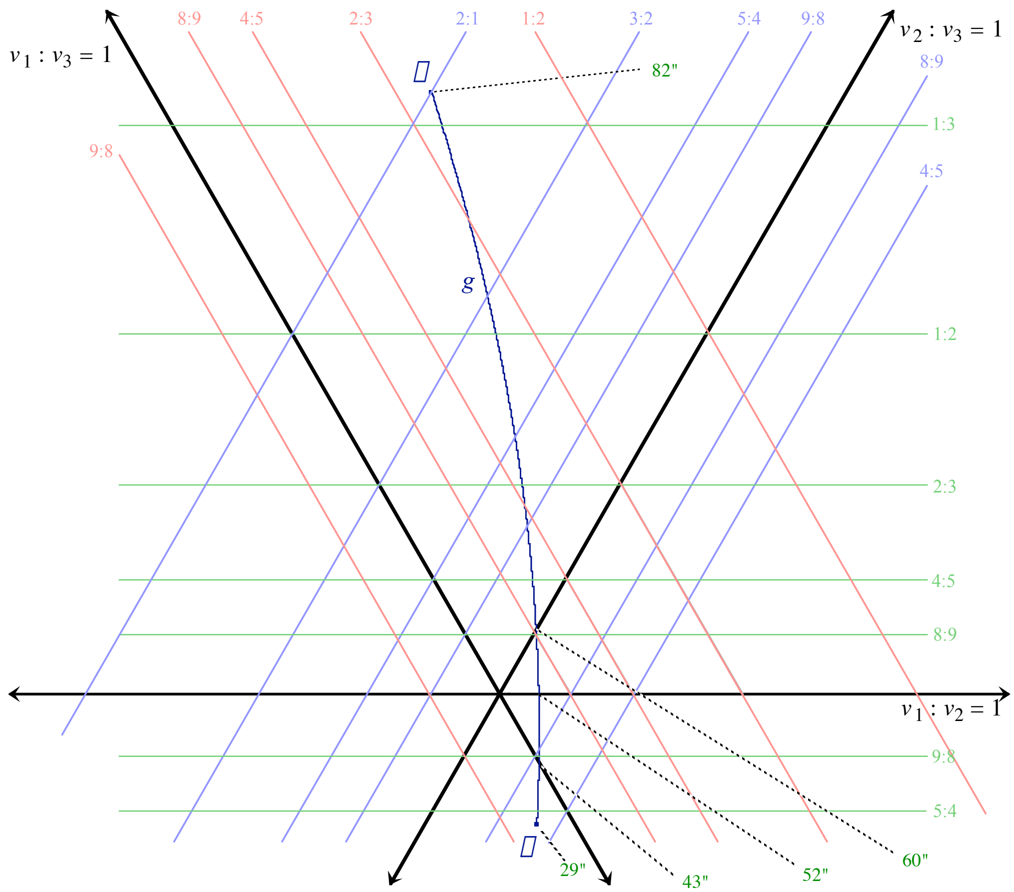


Figure 23: *Tempo trajectory for the opening of Nancarrow's Study No. 22.*

corresponding to 1:1 ratios.

[5.8] As with *Vers le blanc*, we will proceed to consider equivalence classes relevant to the musical situation at hand. Since the timbre of all three strata in Study No. 22 is homogenous and their relative registers are somewhat obscured, it makes sense to assert permutational equivalence and limit our perspective to the normal region. In other words, we want to focus on the relationships between tempos ordered from slowest to fastest, rather than ordered by register. Figure 24 graphs g in the normal region. Beginning with the entrance of the middle voice all tempo ratios contract toward a convergence of the fastest two tempos, where g intersects the line $v_2 : v_3 = 1$. At the first tempo convergence, which occurs at $t \approx 0.717$ minutes (≈ 43 seconds), the ratio between the slowest tempo and the other tempos is approximately 7:8. The second convergence occurs at $t \approx 0.887$ minutes (≈ 52 seconds) when the slowest two tempos are equal, indicated by the horizontal line $v_1 : v_2 = 1$, and form a ratio of approximately 16:17 with the faster tempo. After backtracking nearly the same area of T-class space to the third tempo convergence at $t \approx 1.002$ minutes (≈ 60 seconds)—the second convergence at $v_2 : v_3 = 1$—the tempos quickly and progressively diverge for the remainder of the opening canon.

[5.9] The next equivalence to consider is inversion. Is there a meaningful sense in which two tempo

g would approach a straight line running parallel to the y'-axis.

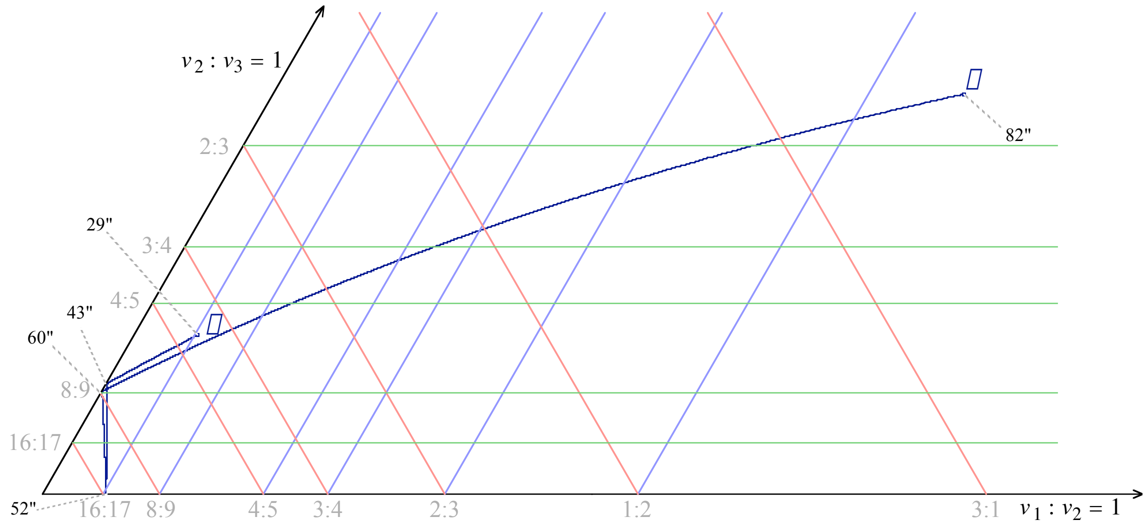


Figure 24: *g in the normal region.*

sets could be related by inversion? Let's briefly review the effects of inversion on the ordered interval series for a pitch set, limiting our focus to ordered pitch sets in the normal region. The ordered interval series of A is $\langle\langle a_2 - a_1, a_3 - a_2\rangle\rangle$, again using double brackets to eliminate confusion with ordered sets. The inversion of a pitch set results in the retrograde of its ordered interval series, so that the ordered interval series of $I(A)$ is $\langle\langle a_3 - a_2, a_2 - a_1\rangle\rangle$.[32] As discussed in Section 3.4, for sets in the normal region this results in a reflection about the line $v_3 - v_2 = v_2 - v_1$. Let f_1, f_2, and f_3 be the frequencies of a_1, a_2, and a_3. Then the ordered interval series of A can also be written as $\langle\langle f_1{:}f_2, f_2{:}f_3\rangle\rangle$, and the ordered interval series of $I(A)$ can be written as $\langle\langle f_2{:}f_3, f_1{:}f_2\rangle\rangle$.

[5.10] For example, let's return to A_f, which has an ordered interval series of $\langle\langle 3{:}4, 4{:}5\rangle\rangle$. Its inversion will yield the ordered interval series $\langle\langle 4{:}5, 3{:}4\rangle\rangle$, so the T-class of $I(A_f)$ is permutationally equivalent to $\langle 12{:}15{:}20\rangle$. Figure 25 interprets A_f and $I(A_f)$ as tempo sets. (In figure 25, the inversion maps the tempo of the middle voice onto itself.) The pulse of the middle voice stays at M.M. = 60 throughout. The pulse of the bottom voice switches from M.M. = 45 in the first system to M.M. = 48 in the second system, while the top voice switches from M.M. = 75 to M.M. = 80. The ordered interval series, progressing from the bottom to top staff is $\langle\langle 3{:}4, 4{:}5\rangle\rangle$ in the first system and $\langle\langle 4{:}5, 3{:}4\rangle\rangle$ in the second system. There are certainly differences between the two systems: 1) pulses in system one coincide every four beats, whereas pulses in system two coincide every 15 beats; and 2) each rhythmic layer in system one is likely to be heard as a division of a longer time span (equivalent to one bar), whereas rhythmic layers in system two are likely to be heard as multiples of shorter pulse (equivalent to a sixteenth note). But there is also a strong similarity between the tempo sets of the two systems due to their identical interval content—both sets contain the ratios 3:4, 4:5, and 3:5. In Figure 22 these two tempo sets are readily observed to be related by reflection about the line of I-invariant tempo sets, $v_1 : v_2 = v_2 : v_3$. The I-related pair $\langle 3{:}5{:}7\rangle$ and $\langle 15{:}21{:}35\rangle$ are also related by reflection about this line,[33] while$\langle 4{:}6{:}9\rangle$ lies on the

[32] $I(A) = \langle -a_1, -a_2, -a_3\rangle$, which is permutationally equivalent to $\langle -a_3, -a_2, -a_1\rangle$. $\langle\langle -a_2 - (-a_3), -a_1 - (-a_2)\rangle\rangle = \langle\langle a_3 - a_2, a_2 - a_1\rangle\rangle$.

[33] In the opening of Ligeti's Piano Etude, No. 6, three different tempos are achieved by pulse streams articulated every 3rd, 5th, or 7th sixteenth. The resulting set of tempos belongs to the T-class $\langle 15{:}21{:}35\rangle$, which is the inversion of

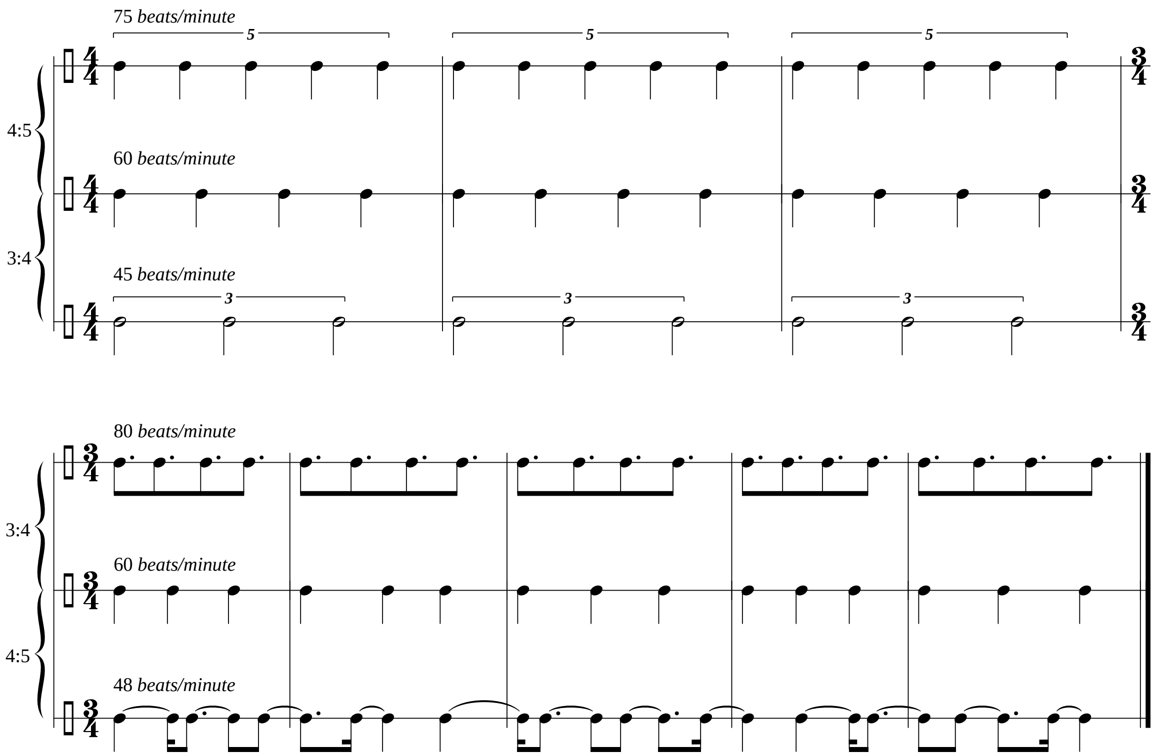



Figure 25: *I-related tempo sets.*

line of I-invariant sets.

[5.11] Modular equivalence for tempos is context dependent in a way that octave equivalence for pitches is not. In certain contexts tempos in a 1:2 ratio could be considered equivalent. However, in other contexts, such as compound meter, a 1:3 ratio could be considered equivalent. Since there is no discernible meter in the opening of Study No. 22 anyway, we will not take modular equivalence into consideration.

[5.12] Figure 26 graphs g in the normal half region, again beginning with the entrance of the middle voice. The portions of the graph in Figure 24 that lie above the normal half region are reflected about its upper boundary. All three tempo convergences correspond to reflections off of the line $v_1 : v_2 = 1$. In addition, Figure 26 shows the three convergences on I-invariant tempo sets corresponding to reflections off of the line $v_1 : v_2 = v_2 : v_3$. The first of these occurs at $t = 0.837$' (or ≈50"), the second occurs at $t = 0.931$' (or ≈56"), and the third occurs at $t = 1.268$ (or ≈76"). Figure 26 shows even more clearly how the canon retraces its own path (again, due to reflections off of the boundaries of the normal half region) prior to the divergence of tempos at the end of the canon. Though it appears as if this divergence takes up the majority of the passage, this is not the case. (Time is not a dimension of the graph.) In fact, only the final $\frac{2}{5}$ of the graph in Figure 26 occurs after the third tempo convergence, with most of the canon characterized by very close tempo

the duration ratios. Conversely, pulse streams articulated every 15th, 21st, and 35th sixteenth would yield the T-class $\langle 3{:}5{:}7 \rangle$. (See Taylor 1997 for a detailed analysis of this Etude.)

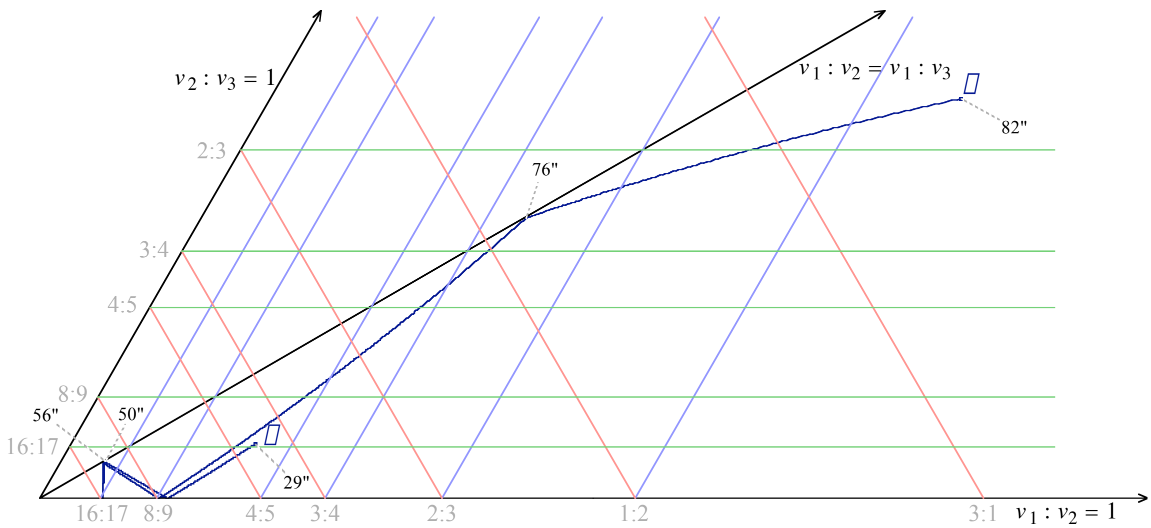


Figure 26: *g in the normal half region.*

ratios. Even in this final $\frac{2}{5}$ most of the divergence occurs near the very end of the canon—note how much distance in T-class space is covered in the final six seconds.[34] This sheds light on my own perception of the canon. When tempos are close, but not equivalent, the result is a kind of tempo dissonance, in which it is difficult to distinguish and separate the individual tempos. As the tempos diverge and the ratios become less close, it is easier to experience the tempos as separate polyphonic layers, particularly when they approach the simpler ratios of 2:3, 1:2, 1:3, and so forth.

6 Ligeti, Hamburgisches Konzert

[6.1] Figure 27 is an annotated score of the concluding chorale (marked *Choral* in the original score) from the second movement of György Ligeti's *Hamburgisches Konzert*, for solo horn, four obbligato natural horns, and chamber orchestra. The chorale is given almost entirely to the four natural horns (pitched in F, E, $E\flat$, and D), who are instructed to play throughout the concerto in natural tuning; i.e., the pitch is not to be "corrected" by hand to coincide with equal temperament.[35] Figure 28 gives the typical spelling of the first thirteen partials of an E_1 fundamental with deviations from 12-tone equal temperament indicated in cents above the staff. For instance, the seventh partial is approximately 31 cents below the equal tempered D_4. (In Figure 27, pitches that are not octave related to the fundamental of each horn are identified by partial number the first time they occur.) While the chorale is obviously discrete in the pitch domain, it is necessary to use the techniques

[34] Allowing the canon to continue beyond 82 seconds, as t approaches 1.494 minutes (approximately 90 seconds), $g_2(t)$ goes to infinity, so $g(t)$ would pass through T-class space at an infinite rate.

[35] The score is not explicit as to the precise intervallic relationship between the fundamentals of the natural horns. In the absence of any indications specifying particular frequency ratios, such as 19:18 for semitone-related fundamentals, I am assuming the fundamentals belong to the same equal-tempered universe as the rest of the ensemble. This is in keeping with Ligeti's usual flexible approach to microtonal structures, taking advantage of "mistuned" natural harmonics and overtones on strings and brass, employing instruments like ocarinas and slide whistles in which the tuning is precarious, and using instructions such as "just a little bit lower (less than a quarter tone)" to achieve a very complex non-tempered sound world, rather than erecting an alternative tuning system based on just intonation, quarter tones, or the like.

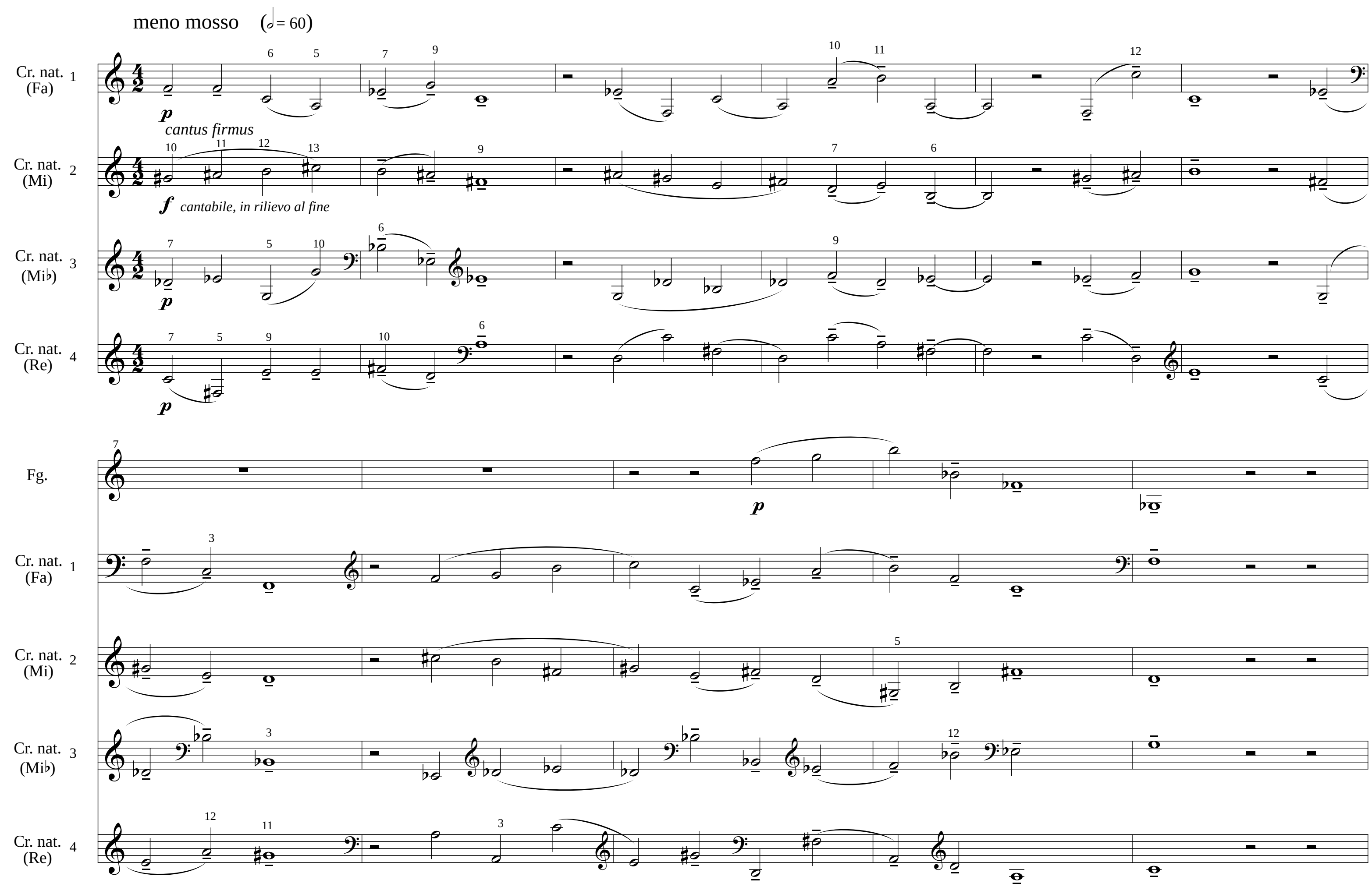



Figure 27: Choral *from György Ligeti's* Hamburgisches Konzert, *second movement*

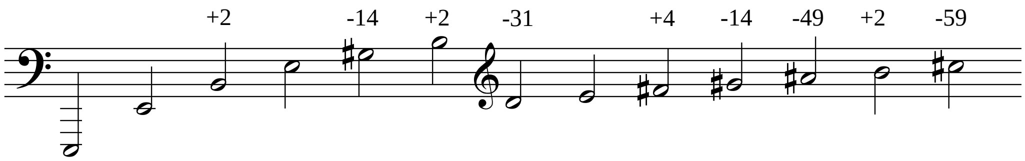


Figure 28: *The first 13 partials of* E_1 *and their deviations from twelve-tone equal temperament, measured in cents.*

developed for continuous transformations in the foregoing in order to answer questions about the harmonic structure. (Example 4 is a MIDI realization of the chorale, minus the bassoon.)[36]

[6.2] For example, chords 1.1, 1.3, 2.1, 2.2, 3.2, 3.3, and 4.1 (reckoning chords by measure.beat) are all fairly similar to [0158], the set class of major seventh chords. ("Fairly similar" in the sense that if each pitch were moved to its nearest neighbor in 12-tone equal temperament, the resulting chord would be a member of [0158].) But are some more similar than others? Chord 1.3, with a single deviation of 14 cents from equal temperament, is likely to be heard as more similar to [0158] than chord 2.1. But is the latter chord, with two deviations, one of 31 cents and one of 14 cents, necessarily closer to [0158] than chord 1.1, with three deviations, two of 31 cents and one of 14 cents? For those chords that are further removed from [0158], to which other set classes are they most similar? And does rounding each pitch of a chord to the nearest equal-tempered pitch always yield the equal-tempered set class closest to the chord?

[6.3] To answer these questions, we need to find the 12-tone equal tempered set classes that are most similar to each chord in the chorale. First we will find the 12-tone equal tempered T-classes that are closest to each chord in the T-class space for ordered sets of four voices. It is possible to construct a three-dimensional Euclidean space analogous to that developed in Section 3, determining distances geometrically. However, we will proceed with an algebraic approach that generalizes nicely to T-class spaces of any dimension. (As long as Assumptions 1 and 2a are satisfied, distances determined geometrically or algebraically are identical.)[37]

[6.4] Let's begin by considering the distance between chord 1.1, $P = \langle -0.31, 0.69, 7.86, 5\rangle$, and its closest neighbor in 12-tone equal temperament, $Q = \langle 0, 1, 8, 5\rangle$ ($C_4 = 0$). The mapping of pitches from one chord to the other, along with the corresponding distances, is

$P \to Q$	d_i
$5 \to 5$	$d_4 = 0$
$7.86 \to 8$	$d_3 = 0.14$
$0.69 \to 1$	$d_2 = 0.31$
$-0.31 \to 0$	$d_1 = 0.31$

[36] The score is notated in C. In the original score, Ligeti uses three different types of arrows to indicate the microtonal deviations of the 5^{th} (and 10^{th}), 7^{th}, and 11^{th} harmonics. The arrow for the 13^{th} harmonic is the same as that for the 7^{th} harmonic, though the respective deviations are not equal. The slight deviations of the 3^{rd} (6^{th} and 12^{th}) and 9^{th} harmonics are not indicated. Thus, the original score is only an approximation of the sounding pitches—we will use the actual microtonal deviations (rounded to the nearest cent), as shown in Figure 28, in the analysis that follows.

[37] What follows is an extension of Lewin 1998, Sections 7 and 8.

where d_i is the distance the i^{th} voice moves. The total distance from P to Q is

$$d(P,Q) = \sqrt{d_1^2 + d_2^2 + d_3^2 + d_4^2} \approx 0.46. \tag{11}$$

[6.5] However, since we are interested in the distance between the T-classes of P and Q, a better fit may be found by allowing Q to vary continuously. Transposing Q by a continuous variable, x, the mapping and corresponding distance from P to $T_x(Q)$ is:

$P \to T_x(Q)$	d_i
$5 \to 5 + x$	$d_4 = x$
$7.86 \to 8 + x$	$d_3 = x + 0.14$
$0.69 \to 1 + x$	$d_2 = x + 0.31$
$-0.31 \to x$	$d_1 = x + 0.31$

The distance between P and $T_x(Q)$ varies as a function of x:

$$\begin{aligned} f(x) = d(P, T_x(Q))^2 &= x^2 + (x+0.14)^2 + 2(x+0.31)^2 \\ &= 4x^2 + 1.52x + 0.2118. \end{aligned} \tag{12}$$

In order to find the value for x that minimizes f, we could simply inspect the graph of f. However, to be precise, we will take the derivative of $f(x)$, $\frac{d}{dx}f(x)$, and solve $\frac{d}{dx}f(x) = 0$ for x:[38]

$$\begin{aligned} \tfrac{d}{dx}f(x) &= 8x + 1.52 = 0 \\ x &= -\tfrac{1.52}{8} = -0.19. \end{aligned} \tag{13}$$

In other words, we can find the best fit by transposing Q down 0.19 semitones. Substituting -0.19 for x in equation 12, the distance between P and $T_{-0.19}(Q)$ is

$$d(P, T_{-0.19}(Q)) = \sqrt{4 \cdot (-0.19)^2 - 1.52 \cdot 0.19 + 0.2118} \approx 0.26. \tag{14}$$

[6.6] According to this metric, the distance between Δ^h-related T-classes is $\frac{\sqrt{3}}{2}h$ rather than h, which contradicts Assumption 2a. This seems somewhat counterintuitive and is worth a brief discussion. Consider the two pitch sets $A = \langle 0, 4, 7, 11\rangle$ and $B = \langle 0, 4, 7, 10\rangle$. A and B are Δ^1-related as are $/A/_T$ and $/B/_T$. The distance between A and B is equal to 1, but if we allow B to be transposed continuously, we can find a closer fit. Specifically, $d(A, T_x(B))$ is minimized when $x = 1/4$, or when B is raised by one eighth-tone. The resulting distance is $d\left(A, T_{\frac{1}{4}}(B)\right) = \frac{\sqrt{3}}{2}$. Since $/A/_T \; \Delta^1 /B/_T$, $\rho(A,B)$ should be equal to 1. Therefore it is necessary to scale $d\left(A, T_{\frac{1}{4}}(B)\right)$ by $\left(\frac{\sqrt{3}}{2}\right)^{-1}$, so that

$$\rho(A,B) = d\left(A, T_{\frac{1}{4}}(B)\right) \cdot \tfrac{2}{\sqrt{3}} = 1.$$[39]

[38] For readers who may be anxious at the mention of calculus, only the most elementary techniques are needed for our purposes. Here is a brief explanation: The graph of $f(x)$ is a "U"-shaped curve. We wish to find the bottom of this "U" (the point at which the distance is minimized). For any point x_0 on this curve we can draw the tangent—a straight line that touches the curve at x_0 and no other point. The slope of this tangent is the derivative of $f(x)$, written $\frac{d}{dx}f(x)$. For a function of the form $g(x) = ax^2 + bx + c$, the derivative is $\frac{d}{dx}g(x) = 2ax + b$. The tangent to the lowermost point of $f(x)$ has a slope of 0, which is why we wish to find the point where $\frac{d}{dx}f(x) = 0$.

[39] The scaling factor for n-voice textures is derived in the following section.

Accordingly, the distance between $/P/_T$ and $/Q/_T$ is

$$\rho(P, Q) = d(P, T_{-0.19}(B)) \cdot \frac{2}{\sqrt{3}} \approx 0.3, \tag{15}$$

or nearly one third the distance of a semitone deviation. Since $Q \in [0158]$ and no other member of [0158] is closer to P in T-class space, then by equation 7, $\rho(/P/, /Q/) \approx 0.3$.

[6.7] Figure 29 gives the respective distances (ρ_1, ρ_2, ρ_3) between the set class of each chord in the chorale (X) and the three T-classes that are closest to it (Y_1, Y_2, Y_3). (Horizontal lines have been inserted at the end of each phrase. $\rho_i = \rho(/X/, /Y_i/)$.) We can now answer the questions posed

X	$/Y_1/$	ρ_1	$/Y_2/$	ρ_2	$/Y_3/$	ρ_3
1.1	[0158]	$\rho \approx 0.300$	[0157]	$\rho \approx 0.764$	[0258]	$\rho \approx 0.877$
1.2	[0237]	$\rho \approx 0.424$	[0236]	$\rho \approx 0.707$	[0137]	$\rho \approx 0.744$
1.3	[0158]	$\rho \approx 0.168$	[0157]	$\rho \approx 0.834$	[0157]	$\rho \approx 0.939$
1.4	[0358]	$\rho \approx 0.519$	[0258]	$\rho \approx 0.537$	[0248]	$\rho \approx 0.793$
2.1	[0158]	$\rho \approx 0.315$	[0148]	$\rho \approx 0.739$	[0147]	$\rho \approx 0.879$
2.2	[0157]	$\rho \approx 0.498$	[0158]	$\rho \approx 0.505$	[0157]	$\rho \approx 0.921$
2.3	[0369]	$\rho \approx 0.033$	[0258]	$\rho \approx 0.974$	[0258]	$\rho \approx 0.974$
3.2	[0158]	$\rho \approx 0.424$	[0157]	$\rho \approx 0.707$	[0147]	$\rho \approx 0.744$
3.3	[0158]	$\rho \approx 0.300$	[0157]	$\rho \approx 0.764$	[0258]	$\rho \approx 0.877$
3.4	[0268]	$\rho \approx 0.154$	[0157]	$\rho \approx 0.847$	[0258]	$\rho \approx 0.951$
4.1	[0158]	$\rho \approx 0.317$	[0157]	$\rho \approx 0.849$	[0268]	$\rho \approx 0.883$
4.2	[0358]	$\rho \approx 0.334$	[0258]	$\rho \approx 0.725$	[0248]	$\rho \approx 0.867$
4.3	[0247]	$\rho \approx 0.496$	[0147]	$\rho \approx 0.678$	[0137]	$\rho \approx 0.697$
4.4	[0258]	$\rho \approx 0.174$	[0369]	$\rho \approx 0.896$	[0358]	$\rho \approx 0.911$
5.3	[0358]	$\rho \approx 0.295$	[0148]	$\rho \approx 0.887$	[0258]	$\rho \approx 0.887$
5.4	[0358]	$\rho \approx 0.491$	[0247]	$\rho \approx 0.511$	[0148]	$\rho \approx 0.928$
6.1	[0158]	$\rho \approx 0.168$	[0157]	$\rho \approx 0.834$	[0157]	$\rho \approx 0.939$
6.4	[0147]	$\rho \approx 0.334$	[0137]	$\rho \approx 0.725$	[0158]	$\rho \approx 0.867$
7.1	[0347]	$\rho \approx 0.317$	[0147]	$\rho \approx 0.849$	[0237]	$\rho \approx 0.883$
7.2	[0137]	$\rho \approx 0.020$	[0136]	$\rho \approx 0.980$	[0237]	$\rho \approx 0.994$
7.3	[0258]	$\rho \approx 0.496$	[0258]	$\rho \approx 0.697$	[0358]	$\rho \approx 0.678$
8.2	[0258]	$\rho \approx 0.404$	[0248]	$\rho \approx 0.597$	[0358]	$\rho \approx 0.936$
8.3	[0246]	$\rho \approx 0.337$	[0135]	$\rho \approx 0.664$	[0135]	$\rho \approx 0.924$
8.4	[0147]	$\rho \approx 0.508$	[0258]	$\rho \approx 0.677$	[0158]	$\rho \approx 0.686$
9.1	[0148]	$\rho \approx 0.326$	[0137]	$\rho \approx 0.860$	[0158]	$\rho \approx 0.872$
9.2	[0258]	$\rho \approx 0.497$	[0248]	$\rho \approx 0.504$	[0358]	$\rho \approx 0.938$
9.3	[0148]	$\rho \approx 0.332$	[0048]	$\rho \approx 0.671$	[0158]	$\rho \approx 0.915$
9.4	[0147]	$\rho \approx 0.254$	[0258]	$\rho \approx 0.794$	[0247]	$\rho \approx 0.819$
10.1	[0347]	$\rho \approx 0.491$	[0236]	$\rho \approx 0.560$	[0246]	$\rho \approx 0.809$
10.2	[0147]	$\rho \approx 0.023$	[0247]	$\rho \approx 0.987$	[0137]	$\rho \approx 0.987$
10.3	[0369]	$\rho \approx 0.033$	[0258]	$\rho \approx 0.974$	[0258]	$\rho \approx 0.974$
11.1	[0257]	$\rho \approx 0.300$	[0157]	$\rho \approx 0.764$	[0258]	$\rho \approx 0.877$

Figure 29: *Set-class distances in* Choral.

in [6.2]. Chord 1.3 is indeed the closest chord to [0158], not only in the first phrase, but in the entire chorale.[40] Though one might be led to believe that chord 2.1 is closer to [0158] than chord 1.1 based on the number of deviations from equal temperament, this is not the case. In addition, chords 1.1 and 2.1 deviate from [0158] in different ways, with the first chord leaning to [0157] as the second closest set class and the second chord leaning to [0148]. And though rounding each pitch of chord 2.2 to the nearest equal tempered pitch yields a member of [0158], the chord as a whole is actually slightly closer to [0157].

[6.8] A few other points can be gathered from studying Figure 29. The first three phrases end on relatively "in tune" chords, while the fourth phrase ends on a decidedly "out of tune" sonority. This unsettled phrase ending ushers in the extended final phrase, which begins with the least tempered stretch in the entire chorale. The first five notes of the final phrase average a distance of $\rho_1 \approx 0.414$ from equal tempered set classes—the largest of any consecutive five chords. One can also readily see those chords that are likely to be the most ambiguous to listeners conditioned to 12-tone equal temperament. Chords 1.4, 2.2, 3.2, 4.3, 5.4, 7.3, 8.2, 8.4, 9.2, and 10.1 are all quite distant from their nearest 12-tone equal tempered neighbors. Chords 1.4 and 8.4 are the most extreme in this regard.

7 General distance metric for T-class space of n voices

[7.1] In this section we wish to generalize the distance metric of equation 5 for ordered sets of n voices adopting the approach of Section 6. Let $P = \langle p_1, \ldots, p_n \rangle$ and $Q = \langle q_1, \ldots, q_n \rangle$ be points in n-dimensional Euclidean space. The distance between P and Q is

$$d(P,Q)^2 = \sum (q_i - p_i)^2. \tag{16}$$

We factor out transposition by allowing Q to be transposed by the continuous variable x. The distance squared between P and $T_x(Q)$ is

$$\begin{aligned} f(x) = d(P, T_x(Q))^2 &= \sum (q_i + x - p_i)^2 \\ &= nx^2 + 2x \sum (q_i - p_i) + \sum (q_i - p_i)^2. \end{aligned} \tag{17}$$

In order to find the transposition of Q closest to P, we take the derivative of $f(x)$ and find the value of x for which $\frac{d}{dx} f(x) = 0$. The derivative of $f(x)$ is

$$\tfrac{d}{dx} f(x) = 2nx + 2 \sum (q_i - p_i). \tag{18}$$

Setting $\frac{d}{dx} f(\xi) = 0$,

$$\xi = -\frac{\sum (q_i - p_i)}{n}. \tag{19}$$

[40] Given that each horn is limited to the pitches drawn from the overtone series, many of which are quite "out of tune," one might question if a chord exists under these constraints that is closer to [0158]. The answer is yes—for instance, the non-tempered set $\langle C, E, F, A \rangle$, ordered from horn one to four, is ≈ 0.033 from [0158].

Thus, $T_\xi(Q)$ minimizes the distance between P and $T_x(Q)$. Substituting ξ for x in equation 17, we have

$$\begin{aligned} f(\xi) = d(P, T_\xi(Q))^2 &= n\left(-\frac{\sum(q_i - p_i)}{n}\right)^2 - 2\sum(q_i - p_i)\frac{\sum(q_i - p_i)}{n} + \sum(q_i - p_i)^2 \\ &= \frac{\left(\sum(q_i - p_i)\right)^2}{n} - 2\frac{\left(\sum(q_i - p_i)\right)^2}{n} + \sum(q_i - p_i)^2 \\ d(P, T_\xi(Q)) &= \sqrt{\sum(q_i - p_i)^2 - \frac{\left(\sum(q_i - p_i)\right)^2}{n}}. \end{aligned} \tag{20}$$

[7.2] It is necessary to scale equation 20 so that Assumption 2a is not violated. If $A\,\Delta^h B$, then

$$d(A, T_\xi(B)) = h\sqrt{\frac{n-1}{n}}. \tag{21}$$

Since, by Assumption 2a, $\rho(A, B) = h$, we must scale equation 21 by $\sqrt{\frac{n}{n-1}}$ so that $\rho(A, B) = \sqrt{\frac{n}{n-1}}d(A, B) = h$. Scaling equation 20 similarly we have,

$$\rho(P, Q) = \sqrt{\frac{n}{n-1}}d(P, T_\xi(Q)) = \sqrt{\frac{n}{n-1}\left(\sum(q_i - p_i)^2 - \frac{\left(\sum(q_i - p_i)\right)^2}{n}\right)}, \tag{22}$$

which gives the distance between T-classes of ordered sets of n voices. (For $n = 3$ equation 22 is equivalent to equation 6.)[41]

8 Two higher-dimensional examples

[8.0] As we have seen, continuous transformations in three-voice texture can be represented in two-dimensional (2D) space. The non-tempered chords of Section 6 could have been represented in 3D space, and textures of n voices can be represented in $(n-1)$D space. Obviously, there are enormous difficulties in intuiting such higher-dimensional spaces. In some situations it is possible to reduce the number of dimensions and work with a more easily intuited space. In other situations it is necessary to observe the space "indirectly." This section features one example of each type of situation.

[41] If P and Q are interpreted as pitch sets, the foregoing dovetails nicely with recent work on voice-leading distance and "transposition-like" voice leading (Lewin 1998, Quinn 1996, Straus 2003). The basic notion is that the motion from P to Q may be likened to some transposition with a certain amount of deviation, or offset. The metric used for the offset from a "T_x-like" motion in both Lewin and Straus is $\sum|q_i - (p_i + x)|$. For example, if V is the motion from $\langle B\flat_3, C_4, E_4\rangle$ to $\langle E_4, F_4, A_4\rangle$, then V is "T_5-like" with an offset of 1, since two voices move by T_5 and the remaining voice moves by T_6. Furthermore, T_5 is the transposition that best approximates V, since any T_x with $x \neq 5$ will yield an offset greater than 1. In the context of the present paper, we can say that the motion from P to Q is "$T_{-\xi}$-like", where ξ is defined as in equation 19, with an offset of $\rho(P, Q)$. A comparison with these approaches (in particular, Lewin 1998, Section 8, which explores voice leading in a continuous pc space) would be very relevant, but is beyond the scope of the present paper.

8.1 Spiral: continuous transpositional combination

[8.1.1] Consider the following harmonic interpolation, designated $f(t)$ as t varies from 0 to 1:

$$\begin{aligned} F_5 &\to A_4 \\ C\sharp_5 &\to F\sharp_4 \\ A_4 &\to D_4 \\ F\sharp_4 &\to B_4 \\ D_4 &\to G\sharp_4 \\ B\flat_3 &\to E_4 \end{aligned}$$

The top three voices descend continuously from an augmented triad to a major triad, while the bottom three voices ascend continuously from an augmented triad to a major triad. Though the space for hexachords is 5D, this particular transformation can be represented in lower dimensions.

[8.1.2] We begin by massaging $f(t)$ into a more usable form. Let $Q \parallel R = \langle q_1, \ldots, q_n, r_1, \ldots, r_n \rangle$ be the *concatenation* of Q and R. Expressing $f(t)$ explicitly in terms of a concatenation of transpositionally equivalent trichords, we have $f(t) = \alpha \to \omega$, where $\alpha = \langle B\flat_3, D_4, F\sharp_4 \rangle \parallel T_{11}(\langle B\flat_3, D_4, F\sharp_4 \rangle)$ and $\omega = \langle E_4, G\sharp_4, B_4 \rangle \parallel T_{-2}(\langle E_4, G\sharp_4, B_4 \rangle)$. Recalling the discussion in [4.2], the path of $f(t)$ in T-class space is indistinguishable from any interpolation of the form $T_x(\alpha) \to T_y(\omega)$. In particular, $g(t) = T_2(\alpha) \to T_{-4}(\omega)$ is identical to $f(t)$ in T-class space. We can write $g(t)$ as the concatenation of two functions, $g(t) = g_1(t) \parallel g_2(t)$, setting $g_1(t) = \langle C_4, E_4, G\sharp_4 \rangle \to \langle C_4, E_4, G_4 \rangle$ and $g_2(t) = T_{11}(\langle C_4, E_4, G\sharp_4 \rangle) \to T_{-2}(\langle C_4, E_4, G_4 \rangle) = T_x(g_1(t))$, where $x = 11 - 13t$. Let $P_y = \langle 0, 4 \rangle \parallel \langle y \rangle$. Since $g_1(t)$ always contains 0 and 4 as its first two members ($C_4 = 0$), $g_1(t) = P_y$, where $y = 8 - t$. Finally, let $P_y * x = P_y \parallel T_x(P_y)$. Thus, $g(t) = P_y * x$.[42]

[8.1.3] Allowing x and y to vary independently over the reals generates a 2D slice of the 5D T-class space for ordered sets of six voices. The region containing the hexachordal T-classes of the form $/P_y * x/_T$ with $x, y \in [0, 12)$ can be represented as a rectangular region of the Euclidean plane, shown in Figure 30. For the remainder of this section, we will assume sets to be ordered pcsets. Taking the values of x and y mod 12, we should imagine the region curled so that arrow heads are coincident with their tails, forming the familiar torus. (For the purpose of visual clarity, the torus and subsequent cross sections will be shown "unwrapped" as a region of the plane.) A few hexachords are plotted in Figure 30. (In the figure and for the remainder of this section, pc letter names followed by a "$_+$" indicate augmented triads ordered root, third, fifth; pc letter names not followed by a "$_+$" indicate major triads ordered similarly.) The distance between the T-classes of two ordered pcsets, Q and R, is defined as

$$\rho_{\text{pc}}(Q, R) = \min \rho(Q, R'), \tag{23}$$

for all R' such that $r'_i \equiv r_i \mod 12$. The x and y coordinates form a rectangular grid, rather than the more typical square grid, since the distance, ρ_{pc}, between any $P_{y_0} * x_0$ and $P_{y_0} * (x_0 + h)$ is greater than that between any $P_{y_0} * x_0$ and $P_{y_0+h} * x_0$.[43] For example, the distance between $P_7 * 6$,

[42] This is an adaptation of transpositional combination for ordered sets in a continuous space. (Cohn 1991)

[43] Let $\text{ic}(h) = 6 - \mid 6 - h \mod 12 \mid$. Then $\rho_{\text{pc}}(P_{y_0} * x_0, P_{y_0} * (x_0 + h)) = \sqrt{\frac{9}{5}} \cdot \text{ic}(h)$ and $\rho_{\text{pc}}\left(P_{y_0} * x_0, P_{y_0+h} * x_0\right) = \sqrt{\frac{8}{5}} \cdot \text{ic}(h)$.

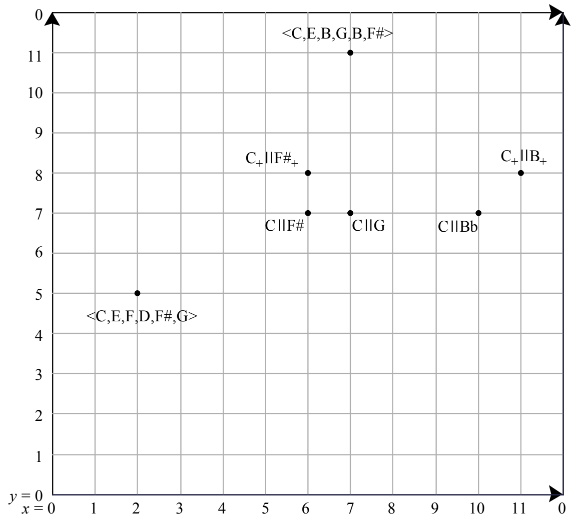


Figure 30: *Region containing T-classes of the form $/P_y * x/_T, x, y \in [0, 12)$.*

or $C \parallel F\sharp$, and $P_7 * 7$, or $C \parallel G$, is $\rho_{\text{pc}} = \sqrt{\frac{9}{5}}$, whereas the distance between $P_7 * 6$ and $P_8 * 6$, or $C_+ \parallel F\sharp_+$, is $\rho_{\text{pc}} = \sqrt{\frac{8}{5}}$. The x- and y-axes are orthogonal, however.[44]

[8.1.4] For $g(t)$ it suffices to consider the region corresponding to $x \in [0, 12), y \in [7, 8]$, shown in Figure 31. (Since Figure 31 is a cross section of the torus in Figure 30, it can be imagined as a cylinder, though one with a slight curve in the vertical direction.) Points labeled on the figure correspond to integer values of x and y. The resulting 12-tone equal tempered T-classes are identified by the set classes to which they belong. Motion along the top edge (circle), as x increases from 0, yields a cycle of four equal tempered set classes repeated three times: [0,4,8] or Aug. at $x = 0$, [0,1,4,5,8,9] or Hex. at $x = 1$, [0,2,4,6,8,10] or W.T. at $x = 2$, Hex. at $x = 3$, etc. As x varies continuously, an augmented triad gradually evolves into a hexatonic collection, which gradually evolves into a whole-tone collection, and so forth around the circle. Motion along the bottom edge (circle), as x increases from 0, yields the equal tempered set classes [0,3,7] or Maj. at $x = 0$, [0,1,3,4,7,8] or Harm.$_6$, [0,2,3,5,7,9] or V^{11}, [0,1,4,6,9] or $V_7^{\sharp 9}$, [0,1,4,5,8] or Hex.$_5$, [0,1,3,5,8] or M^9, [0,1,3,6,7,9] or Petrouch., and etc. around the circle. As x varies continuously, a major triad gradually evolves into a six-note subset of the harmonic minor, which gradually evolves into the members of a V^{11} chord, and so forth.

[8.1.5] The graph of $g(t)$ and, thus, $f(t)$, is plotted in this rectangular region. At $t = 0$, $/f(t)/_T = /P_8 * 11/_T = /\alpha/_T$. As t increases, $f(t)$ slopes gently down and to the left, wraps around the left "boundary" to the right "boundary", and converges on $/P_7 * 10/_T = /\omega/_T$, or a V^{11} chord. The distance between $f(t)$ and the equal-tempered set classes labeled in Figure 31 reveal those set classes to which the transformation is most similar at any moment. For instance, at point A, $f(\frac{5}{13})$ is closest to the whole-tone collection; at point B, $f(\frac{1}{2})$ is equally close to the augmented triad, the

[44] Let $A = /P_{y_0} * x_0/_T$ and $B = /P_{y_1} * x_0/_T$ be fixed points in Figure 30. If the x- and y-axes are orthogonal, then there is no other T-class of the form $/P_{y_1} * x_1/_T$ that is closer to A than B. In other words, if the point $B_x = /P_{y_1} * x/_T$ varies with respect to x, then $\rho(A, B_x)$ is minimized when $x = x_0$, which can be proved true using a similar approach to that of Sections 6 and 7.

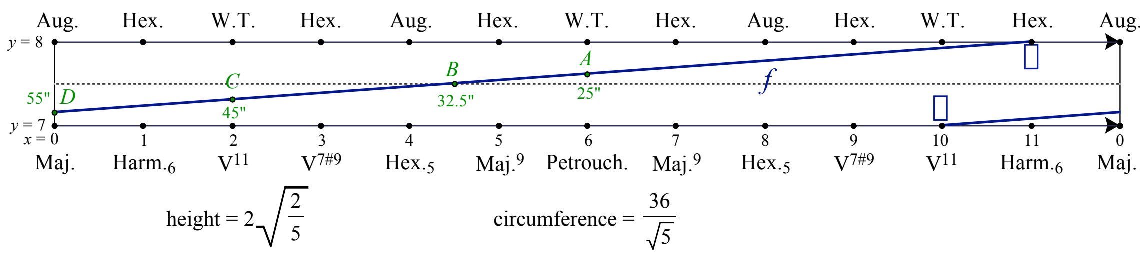


Figure 31: *Region containing T-classes of the form $/P_y * x/_T, x \in [0,12), y \in [7,8]$.*

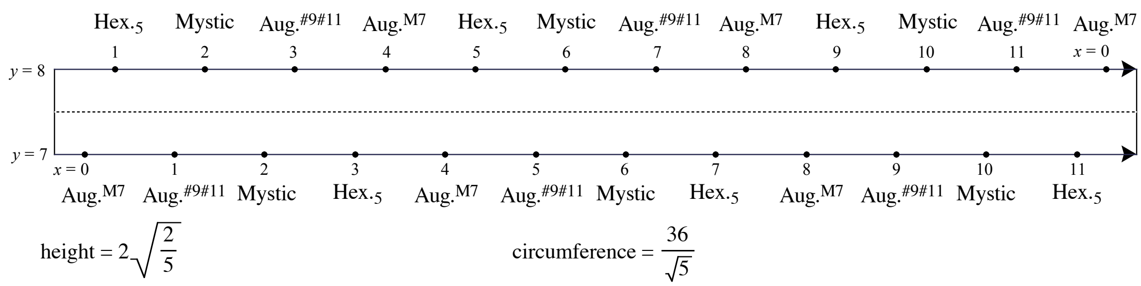


Figure 32: *Region containing T-classes of the form $/P_y \star x/_T, x \in [0,12), y \in [7,8]$.*

major ninth chord, the hexatonic collection, and its unique pentachordal subset; and at point C, $f(\frac{9}{13})$ is most similar to a V^{11} chord.

[8.1.6] Though Figure 31 is a useful representation of the space in which this interpolation travels, we can still get a better sense of this space, particularly near the center of the region, shown as a dashed line. (This dashed line corresponds to all T-classes of the form $/P_{7.5} * x/_T$ and intersects $f(t)$ at B.) Let $P_y \star x = P_y \parallel T_x(P_{15-y})$. (The third voice of the first trichord is related to the third voice of the second trichord by inversion about pc 7.5.) Then we can generate Figure 32 in a similar manner to Figure 31, plotting $/P_y \star x/_T$ with $x \in [0,12)$ and $y \in [7,8]$. (The regions in Figures 31 and 32 are exactly the same size.) As before, points labeled on the figure correspond to integer values of x and y, and the resulting 12-tone equal tempered T-classes are identified by their set classes. Motion along the bottom edge (circle), as x increases from 0, yields a cycle of four equal tempered set classes repeated three times: [0,1,4,8] or Aug.^{M7} at $x = 0$, [014579] or Aug.$^{\sharp 11}_{\sharp 9}$) at $x = 1$, [013579] or Mystic at $x = 2$, Hex.$_5$ at $x = 3$, etc. (For the bottom circle, $x = 0$ is near the far left.) Motion along the top edge (circle), as x increases from 0, yields the same equal-tempered set classes in reverse. (For the top circle, $x = 0$ is near the far right.) Allowing the triads to change from one type to the other while maintaining their position relative to one another (allowing y to vary while holding x constant) does *not* result in vertical motion as in Figure 31. Instead there is a slight slope to the left from any $/P_7 \star x_0/_T$ to $/P_8 \star x_0/_T$. (This is necessary to

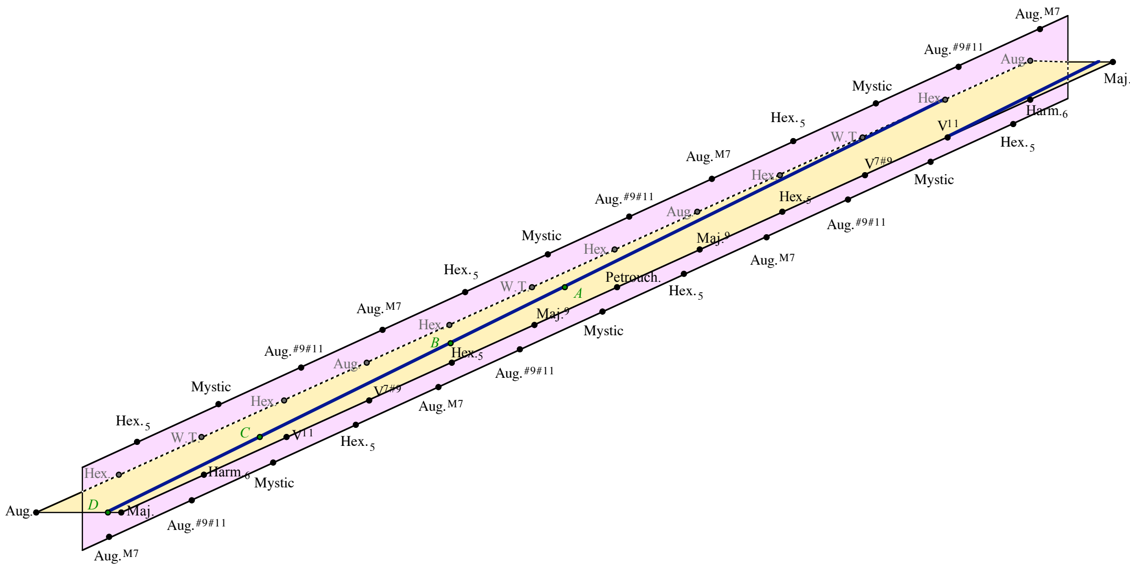


Figure 33: *Intersection of Figures 31 and 32.*

preserve distances.)[45,46]

[8.1.7] The center of the regions in Figures 31 and 32 corresponds to the T-classes resulting from all possible concatenations of $\langle 0, 4, 7.5\rangle$ with itself, since this set lies precisely halfway between $\langle 0, 4, 7\rangle$ and $\langle 0, 4, 8\rangle$. The dashed line is therefore the intersection of the two regions: $/P_{7.5} * x/_T = /P_{7.5} \star x/_T$.[47] We can get a sense of the space by allowing the two rectangles to intersect at right angles, shown in Figure 33. (The left and right "boundaries" of both rectangles are aligned so that this "boundary" meets the intersection of the two rectangles at $x = 0$. As before, the compound object should be bent so that the arrow heads and tails are coincident.) The rectangle of Figure 31 is oriented horizontally, while that of Figure 32 is oriented vertically. In addition, the path of $f(t)$ is plotted on the horizontal rectangle. The closer f is to the center of the region in Figure 31 the closer it is to the region in Figure 32.

[8.1.8] Example 5 is a gradual, though not continuous, realization of f. f passes through a total of thirteen steps in the horizontal direction around the cylinder in Figure 31 (dividing the cylinders into twelve equal sections). Since it takes 65" to complete the interpolation in Example 5, every five seconds corresponds to a single step around the cylinders. The reader is encouraged to listen to Example 5 while following along with Figure 31 and/or 33.

[8.1.9] Figure 34 provides planar cross sections of Figure 33 at the points labeled A and C, and a 3D cross section at point B. Distances on these cross sections provide a reference for how these

[45] The distance between a point on the bottom line, $/P_7 \star x_0/_T$, and the top line, $/P_7 \star x_0/_T$, is minimized at $x = x_0 + \frac{2}{3}$, where $\rho_{\mathrm{pc}} = \sqrt{\frac{8}{5}}$. For example, the point on the top circle directly above $/P_7 \star 6/_T$, or $/C \parallel F\sharp_+/_T$, is $/P_8 \star 6\frac{2}{3}/_T$, or the T-class of C_+ concatenated with the major triad a third tone above F.

[46] Allowing y to vary continuously over the interval [0,12) would yield a torus like Figure 30.

[47] In three dimensions, the intersection of two tori would be another three-dimensional object. However, we are actually dealing with two tori interacting in a four-dimensional space, in which case the intersection of the two objects is a two-dimensional circle, represented here unwrapped as a line.

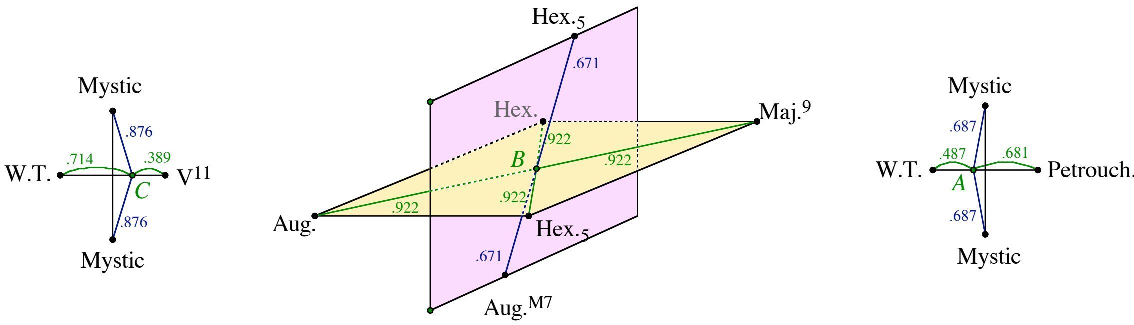


Figure 34: *Selected cross-sections of Figure 33.*

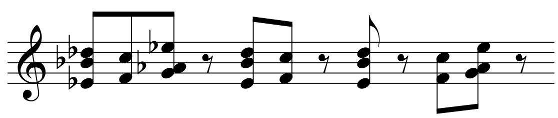


Figure 35: *Realization of a common bcset in Reich's music.*

points compare with neighboring 12-tone equal tempered set classes. A is most like the whole-tone collection, but is also similar to the Petrouchka chord and the mystic collection. Point C is most similar to V^{11}, but one can also hear traces of the mystic collection. The midpoint of f, B, lies at the intersection of the two rectangles. B is closest to Aug.^{M7} and Hex.$_5$. The next closest set classes to B are those lying on the horizontal rectangle.[48] Returning to Figure 33, at point D all six voices in f converge on a consonant triad-like sonority.

8.2 *à la* Steve Reich: continuous multiplicative operation

[8.2.1] The last example is a variation of Steve Reich's phase technique. Figure 35 is a realization of a very common beat-class set (bcset) in Reich's music, $P = \{0, 1, 2, 4, 5, 7, 9, 10\} \mod 12$. Let's consider a transformation in which every "voice" begins with beat class (bc) 0 and gradually moves to one of the bcs in the above bcset in a given time span. We can model this transformation as a continuous multiplicative operation on P. Let $M_t(P) = \{0t, 1t, 2t, 4t, 5t, 7t, 9t, 10t\}$, where $t \in [0, 12)$ and all multiplications are taken mod 12. The particular transformation described above is equivalent to $M_t(P)$ as t varies from 0 to 1. We will designate the general transformation as t varies from 0 to $12 \equiv 0$ as $h(t) = M_t(P)$.

[8.2.2] Since we can not hear a bcset changing continuously, it is necessary to sample the interpolation at evenly spaced intervals. For instance, sampling h at multiples of $t = \frac{1}{4}$ yields the following

[48] Considering T-classes not included in Figure 33, B is also a distance of $\rho_{\text{pc}} \approx 0.922$ from eight other T-classes. However, at any time, t_0, the 12-tone equal tempered T-class(es) closest to $f(t_0)$ is (are) included in Figure 33.

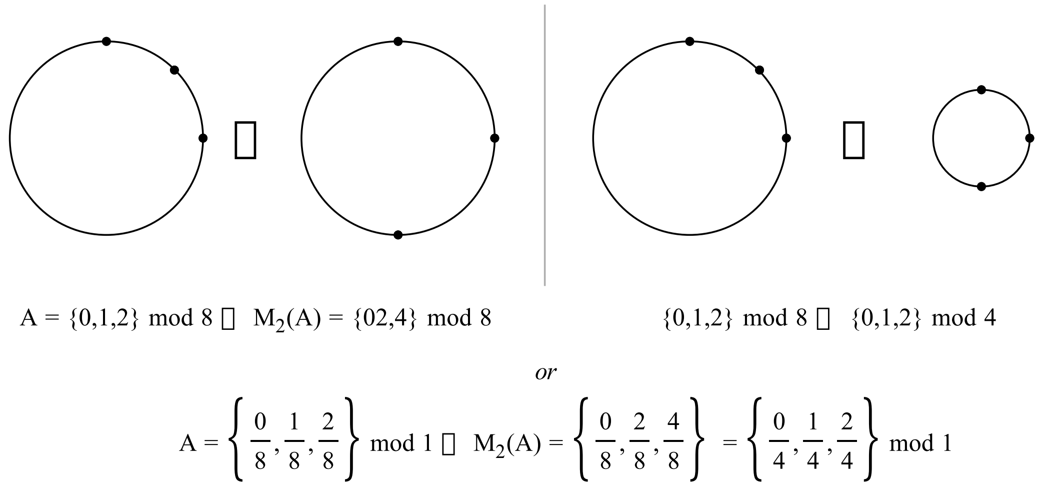


Figure 36: *A geometric perspective of M_2.*

sequence:

$$\begin{aligned}
h(0) &= \{0,0,0,0,0,0,0,0\},\\
h(\tfrac{1}{4}) &= \{0,\tfrac{1}{4},\tfrac{1}{2},1,1\tfrac{1}{4},1\tfrac{3}{4},2\tfrac{1}{4},2\tfrac{1}{2}\},\\
h(\tfrac{1}{2}) &= \{0,\tfrac{1}{2},1,2,2\tfrac{1}{2},3\tfrac{1}{2},4\tfrac{1}{2},5\},\\
&\vdots\\
h(1) &= \{0,1,2,4,5,7,9,10\},\\
&\vdots\\
h(12) &= \{0,0,0,0,0,0,0,0\}.
\end{aligned}$$

Example 6 is a realization of this process with t varying from 0 to 3, sampling h at multiples of $t = \frac{1}{20}$.

[8.2.3] As the number of sample points increases, the realization of the interpolation becomes more gradual. Taking an infinite number of samples will yield a truly continuous transformation. Though this is not possible in practice (who has the time!), we can imagine doing so in order to get a better sense of the bcsets this interpolation will pass through.

[8.2.4] There are many moments in h, say $h(\pi)$, where the resulting rhythm is difficult or impossible to intuit, but there are also many moments where h converges on simpler rhythms. For instance, since $M_2(P) = \{0,2,4,6,8,10\}$, $h(2)$ converges on a rhythm of six even divisions of the meter—e.g., quarter notes in 6/4 meter if we assume elements of P to be multiples of an eighth note. Similarly, since $M_3(P) = \{0,3,6,9\}$, $h(3)$ converges on a rhythm of four even divisions—e.g., dotted-quarter notes in 12/8 meter. At first glance $h(\frac{12}{7}) \approx \{0,1.714,3.428,5.143,6.858,8.571\}$ appears to be a very complicated rhythmic pattern. However, if we divide the meter into 7 equal divisions rather than 12, then $h(\frac{12}{7}) = \{0,1,2,3,4,5\} \mod 7$—a very audible convergence on a fairly simple meter. How can we predict the values of t that yield similar convergences?

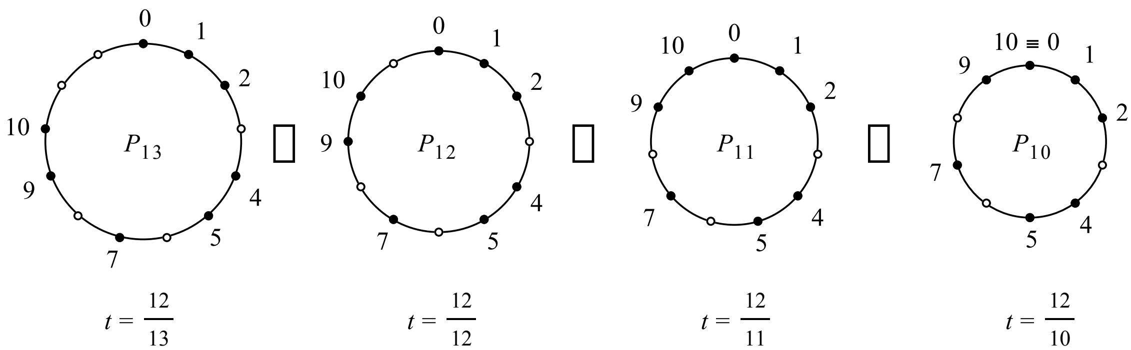


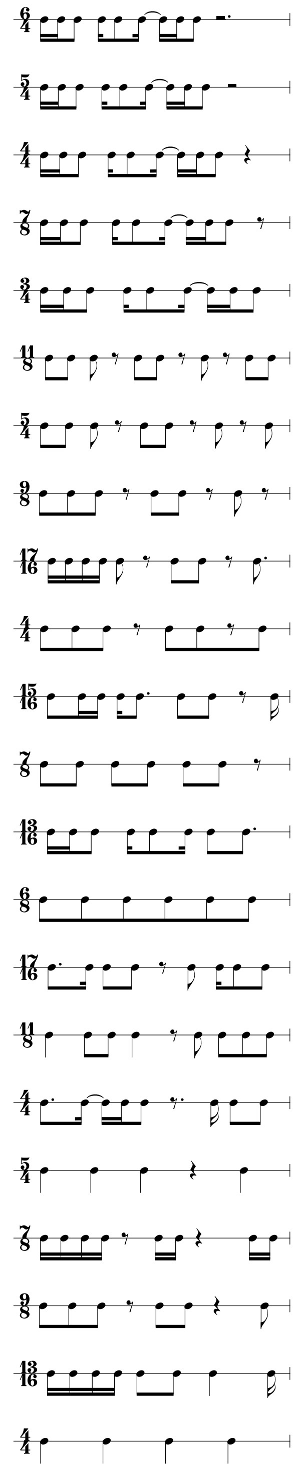
Figure 37: $h(t)$ *before and after* $t = 1$.

[8.2.5] In order to answer the question it will be helpful to take a geometric perspective of the multiplicative operation. The circle on the far left of Figure 36 has a circumference of 8. Superimposed on this circle is a chain of three beads with an arc length of 1 between each, corresponding to $A = \{0, 1, 2\} \mod 8$. The normal view of M_2 is that the chain of beads is stretched by a factor of two, while the circle remains fixed: $M_2(A) = \{0, 2, 4\} \mod 8$. The pair of circles on the right offers a different perspective. The chain remains the same length while the circle is contracted by a factor of 2 to a circumference of 4. The beads still correspond to $\{0, 1, 2\}$, but in a mod 4 environment rather than mod 8. If a single revolution around each circle corresponds to the time span of the periodicity, then it makes more sense to measure distance as the ratio of the arc length to the circumference. Accordingly, we rewrite A as $\{\frac{0}{8}, \frac{1}{8}, \frac{2}{8}\} \mod 1$ and $M_2(A) = \{\frac{0}{8}, \frac{2}{8}, \frac{4}{8}\} = \{\frac{0}{4}, \frac{1}{4}, \frac{2}{4}\} \mod 1$. For the remainder of this section we will adopt the convention of writing $Q_m = \{q_1, \ldots, q_n\}_m$ for $\{\frac{q_1}{m}, \ldots, \frac{q_n}{m}\} \mod 1$. Generalizing the situation on the right side of the figure, the operation M_x dilates a circle of circumference m to a circle of circumference $\frac{m}{x}$. Thus, $M_x(Q_m) = Q_{\frac{m}{x}}$.

[8.2.6] Now we can understand why $M_{\frac{12}{7}}(P)$ yields a subset of seven equal divisions of the rhythmic period. Since $m = 12$, $t = \frac{12}{7}$, and $\frac{m}{t} = 12 \cdot \frac{7}{12} = 7$, $M_{\frac{12}{7}}(P_{12}) = \{0, 1, 2, 4, 5, 7, 9, 10\}_7 \equiv \{0, 1, 2, 3, 4, 5\}_7 = P_7$. Any value of t that can be written as a fraction of the form $\frac{12}{k}$ will yield the bcset P_k. Figure 37 gives circle diagrams of $h(t)$, shown with black points, before and after $t = 1$. At $t = \frac{12}{13}$ the chain of beads, P, is superimposed on a circle with a circumference of 13, yielding P_{13}. The circle contracts as t increases. At $t = 1$ we have P_{12}, the rhythm in Figure 35. At $t = \frac{12}{11}$, $h(t) = P_{11}$, where the "last" bc of P, 10, is one step away from the "first" bc, 0. As the circle contracts the "last" bc continues to approach the "first" bc until the two overlap at $t = \frac{12}{10}$ yielding P_{10}, which contains only 7 bcs.

[8.2.7] Similar convergences will continue to occur at $t = \frac{12}{9}$, $t = \frac{12}{8}$, and so forth. Convergences on relatively simple rhythms occur in between these points as well. For instance, consider h at $t_{13} = \frac{24}{13}$, which occurs between $t = \frac{12}{7}$ and $t = \frac{12}{6}$. We can rewrite t_{13} as $\frac{2m}{13}$, where $m = 12$. Then $M_{\frac{2m}{13}}(P_m) = P_{m \cdot \frac{13}{2m}} = P_{\frac{13}{2}}$. But $P_{\frac{13}{2}} = M_2(P_{13})$, so $h(\frac{24}{13}) = M_2(P_{13}) = \{0, 2, 4, 8, 10, 14, 18, 20\}_{13} \equiv \{0, 1, 2, 4, 5, 7, 8, 10\}_{13}$. Similarly, $h(\frac{24}{11}) = M_2(P_{11})$. Any value for t can be written in the form $t = \frac{am}{k}$, where $a, k \in \mathbb{R}, m \in \mathbb{Z}$. Thus $h(t)$ can be written as $h(\frac{am}{k}) = M_{\frac{am}{k}}(P_m) = M_a(P_k)$. When a and k are integers and k is relatively small, as in the previous examples, we are likely to be able to intuit the resulting rhythm. However, for larger values of k, this intuition becomes less likely.

t	$h(t)$	rhythmic notation
$t_1 = \frac{1}{2}$ (15")	$\{0, 1, 2, 4, 5, 7, 9, 10\}_{24}$	
$t_2 = \frac{3}{5}$ (18")	$\{0, 1, 2, 4, 5, 7, 9, 10\}_{20}$	
$t_3 = \frac{3}{4}$ (22.5")	$\{0, 1, 2, 4, 5, 7, 9, 10\}_{16}$	
$t_4 = \frac{5}{6}$ (25")	$\{0, 1, 2, 4, 5, 7, 9, 10\}_{14}$	
$t_5 = 1$ (30")	$\{0, 1, 2, 4, 5, 7, 9, 10\}_{12}$	
$t_6 = \frac{12}{11}$ ($\approx$33")	$\{0, 1, 2, 4, 5, 7, 9, 10\}_{11}$	
$t_7 = \frac{6}{5}$ (36")	$\{0, 1, 2, 4, 5, 7, 9\}_{10}$	
$t_8 = \frac{4}{3}$ (40")	$\{0, 1, 2, 4, 5, 7\}_{9}$	
$t_9 = \frac{24}{17}$ ($\approx$42")	$\{0, 1, 2, 3, 4, 8, 10, 14\}_{17}$	
$t_{10} = \frac{3}{2}$ (45")	$\{0, 1, 2, 4, 5, 7\}_{8}$	
$t_{11} = \frac{8}{5}$ (48")	$\{0, 2, 3, 4, 5, 8, 10, 14\}_{15}$	
$t_{12} = \frac{12}{7}$ ($\approx$51")	$\{0, 1, 2, 3, 4, 5\}_{7}$	
$t_{13} = \frac{24}{13}$ ($\approx$55")	$\{0, 1, 2, 4, 5, 7, 8, 10\}_{13}$	
$t_{14} = 2$ (60")	$\{0, 1, 2, 3, 4, 5\}_{6}$	
$t_{15} = \frac{36}{17}$ ($\approx$63")	$\{0, 3, 4, 6, 10, 12, 13, 15\}_{17}$	
$t_{16} = \frac{24}{11}$ ($\approx$65")	$\{0, 2, 3, 4, 7, 8, 9, 10\}_{11}$	
$t_{17} = \frac{9}{4}$ (67.5")	$\{0, 3, 5, 6, 11, 12, 14\}_{16}$	
$t_{18} = \frac{12}{5}$ (72")	$\{0, 1, 2, 4\}_{5}$	
$t_{19} = \frac{18}{7}$ ($\approx$77")	$\{0, 1, 2, 3, 6, 7, 12, 13\}_{14}$	
$t_{20} = \frac{24}{9}$ (80")	$\{0, 1, 2, 4, 5, 8\}_{9}$	
$t_{21} = \frac{36}{13}$ ($\approx$83")	$\{0, 1, 2, 3, 4, 5, 8, 12\}_{13}$	
$t_{18} = 3$ (90")	$\{0, 1, 2, 3\}_{4}$	



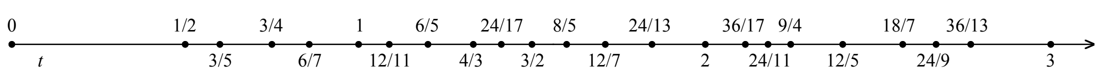
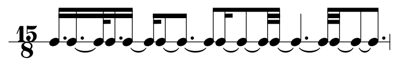


Figure 38: *Rhythmic reference points for bcset interpolation.*

For instance, $h(\frac{7}{5}) = h(\frac{7 \cdot 12}{5 \cdot 12}) = M_7(P_{60})$, or


,

is unlikely to induce an “ah-ha” response from the listener.

[8.2.8] Figure 38 lists the values of t and corresponding values of $h(t)$ where the resultant rhythm converges on relatively simple rhythms. The notation on the right column of the table is provided as an aid. Note values from one row to another are typically not equivalent. For example, the sixteenth notes in the first row are of different duration than those of the second row—it is the entire meter that remains constant in duration. Additionally, while all rhythms are oriented so that bc 0 is the downbeat, listeners may perceive some other bc as the downbeat, yielding a rotation of the notated rhythm. In Example 6, $t = 1$ occurs at 30”, $t = 2$ occurs at 60”, and so forth. The initial 30” is a gradual transformation from the opening attacks on the downbeat to the rhythm in Figure 35. After 30” the convergences listed in Table 38 become more difficult to predict. The reader is encouraged to follow along the table while listening to Example 6. (A time line is included at the bottom of Figure 38 for reference.) Example 7 is a six-minute realization of the entire process as t varies from 0 to 12 for readers who may be interested.

Return to beginning of article

References

Bloch, Stephen, and Jack Douthett. "Vector products and intervallic weighting," *Journal of Music Theory* 38 (1994): 21-41.

Callender, Clifton. "Voice-leading parsimony in the music of Alexander Scriabin," *Journal of Music Theory* 42 (1998): 219-233.

——————. "Formalized accelerando: an extension of rhythmic techniques in Nancarrow's acceleration canons," *Perspectives of New Music* 39.1 (2001): 188-210.

——————. "Simplexes and hyperprisms: A geometric model of voice-leading spaces," manuscript, 2004.

Childs, Adrian. " Moving beyond neo-Riemannian triads: Exploring a transformational model for seventh chords." *Journal of Music Theory* 42 (1998): 181-193.

Clough, John and Jack Douthett. "Maximally even sets," *Journal of Music Theory* 35 (1991): 93-173.

Cohn, Richard. "Properties and generabilities of transpositionally invariant sets," *Journal of Music Theory* 35 (1991): 1-32.

——————. "Transpositional combination of beat-class sets in Steve Reich's phase-shifting music." *Perspectives of New Music* 30.2 (1992): 146-177.

—————— "Square dances with cubes," *Journal of Music Theory* 42 (1998): 283-296.

——————. "A tetrahedral graph of tetrachordal voice-leading space," *Music Theory Online* 9.4, August 2003.

Coexeter, H.S.M. *Regular polytopes.* New York: Dover, 1973.

Douthett, Jack, and Peter Steinbach. "Parsimonious graphs: A study in parsimony, contextual transformations, and modes of limited trasposition." *Journal of Music Theory* 42 (1998): 241-263.

Gann, Kyle. *The music of Conlon Nancarrow.* New York: Cambridge University Press, 1995.

Grey, J.M. "Multidimensional perceptual scaling of musical timbre." *Journal of the Acoustical Society of America* 61 (1977): 1270-1277.

Lewin, David. *Generalized musical intervals and transformations.* New Haven: Yale University Press, 1987.

——————. "Cohn functions." *Journal of Music Theory* 40 (1996): 181-216.

——————. "Some ideas about voice-leading between pcsets," *Journal of Music Theory* 42 (1998): 15-72.

Polansky, Larry. "Morphological metrics," *Journal of New Music Research* 25.4 (1996): 289-368.

Quinn, Ian. "Fuzzy transposition of pitch-class sets." Paper delivered at the New England Conference of Music Theorists, 1996.

Roeder, John. “A theory of voice leading for atonal music.” Ph.D. dissertation, Yale University, 1984.

———. “A geometric representation of pitch-class series.” *Perspectives of New Music* 25 (1987): 362-409.

Saariaho, Kaija. “Timbre and harmony: Interpolations of timbral structures,” *Contemporary Music Review* 2 (1987): 93-133.

Shepard, Roger. “Circularity in judgments of relative pitch,” *Journal of the Acoustical Society of America* 36 (1964): 2346-2353.

Straus, Joseph. “Uniformity, balance, and smoothness in atonal voice leading.” *Music Theory Spectrum* 25 (2003): 305-352.

Taylor, Stephen. “Chopin, Pygmies, and tempo fugue: Ligeti’s *Automne a Varsovie*,” *Music Theory Online*, 3.3, 1997.

Weisstein, Eric W. “Limit.” From *MathWorld*—A Wolfram Web Resource.
http://mathworld.wolfram.com/Limit.html

Return to beginning of article

Clifton Callender
Assistant Professor of Composition
Florida State University
ccallend@mailer.fsu.edu

Return to beginning of article

Return to beginning of article

prepared by
Brent Yorgason, Managing Editor
Updated 19 September 2004