Music Theory Online

The Online Journal of the

Society for Music Theory

Volume 11, Number 4, October 2005

Dmitri Tymoczko*

Voice Leadings as Generalized Key Signatures

KEYWORDS: voice leading, key signature, pitch-class interval, pitch-class space, path in pitch-class space, voice crossing, minimal voice leading, common-tone retention, interscalar path matrix, acoustic scale, Western musical notation.

ABSTRACT: A key signature is a collection of paths in pitch-class space. As such, it provides a powerful tool for thinking about voice leading, one that can be used to answer familiar objections to the very notion of a voice leading between pitch-class sets. I begin by generalizing the traditional Western system of musical notation, defining paths in pitch-class space, and showing that key signatures are collections of such paths. I then show that it is always possible to find a minimal voice leading that has no voice crossings; however, it is not always possible to avoid voice crossings while maximizing common-tone retention. I describe a general method for identifying a minimal voice leading between arbitrary chords, and show that our familiar system of diatonic key signatures implements this method. I conclude by deriving a set of key signatures for the acoustic scale. These key signatures represent minimal voice leadings from the diatonic to the acoustic collection, and are interestingly related to the familiar diatonic key signatures.

Submission received July 2005

VOICE LEADINGS AS GENERALIZED KEY SIGNATURES

A rational reconstruction of Western musical notation

DMITRI TYMOCZKO

0.1 This paper argues that key signatures provide a powerful tool for understanding voice leading. It has five sections. The first generalizes the system of Western musical notation, defining the notion of a *basic scale*—a collection of pitch classes provided with letter names. The second section shows that any key signature determines a voice leading from the basic scale to a target scale, and, conversely, that any voice leading can be expressed as a key signature. The third section describes three virtues that a voice leading or key signature may have: avoiding voice crossings, maximizing common tones, and minimizing the overall "size" of the voice leading. I show that two of these virtues are always compatible, but that all three cannot always be exemplified simultaneously. The fourth section asks how to find a minimal voice leading between arbitrary chords. I provide a general solution to this problem, and show that this solution is implemented by the familiar system of diatonic key signatures. The final section uses the same technique to develop a system of key signatures for the acoustic scale.

0.2 This investigation will require that we think rigorously about the foundations of Western musical notation. This is inherently a pedantic and somewhat painful enterprise. Readers may initially resent the author, and may wonder if he is motivated by a fetish for obscurantism or an unhealthy love of formalism for its own sake. I am not. Instead, I hope to demonstrate that some elementary features of our system of musical notation are as yet imperfectly understood. Furthermore, these features are directly related to a subject of intense current theoretical concern: the theory of voice leading. Understanding our familiar notational system, I will show, leads to new insights about the nature of pitch classes, pitch-class intervals, and voice leading.

0.3 This is chiefly because key signatures turn out to be entities of considerable intrinsic interest: collections of *paths* in pitch-class space. A path in pitch-class space resembles, but is not reducible to, a pitch-class interval. Paths are more fine-grained than intervals, allowing us to distinguish the one-semitone descending route between C and B from the eleven-semitone ascending route between them. This flexibility is quite useful when we are thinking about voice leading between pitch-class sets. Thus the humble key signature, seemingly unworthy of theoretical investigation, provides an extremely powerful instrument for investigating voice leading.

0.4 Before proceeding, I would like to acknowledge my indebtedness to the important work of Julian (Jay) Hook. Jay's 2003 paper on "[key] signature transformations" prompted my own investigation into the broader subject of voice leading. Furthermore, Jay was the first to realize that the rows of my "interscalar path matrices" can be interpreted as key signatures (Section IV).[1] The current paper brings Hook's approach together with my own, demonstrating that the theory of voice leading and the theory of key signatures are inextricably intertwined.[2]

I. Numbers, letters, clefs, and accidentals

1.1 In this section, I generalize the traditional Western system of musical notation. This process will lead me to draw distinctions that are not ordinarily drawn, largely because we take for granted special features of our familiar notational system. For example, I will distinguish scale-dependent and scale-independent measures of distance in pitch-class space, and paths in pitch-class space from pitch-class intervals. The result will be greater clarity about conventional musical notation.

1.2 I begin by providing numerical names for pitches and pitch classes. Somewhat unusually, I do so without presupposing a chromatic scale that divides pitch- and pitch-class space into discrete "steps." Instead, I develop a single, consistent set of numerical labels that can be applied to any tuning system and any chromatic universe. This system allows us to identify the diatonic scale (or more generally, a "basic scale" as defined below) prior to specifying how it is to be embedded in a larger, "chromatic" collection.[3] This

---

[1] Hook 2004.

[2] Hook also read and commented on drafts of this paper. I would also like to thank Noam Elkies, who read a draft of this paper, and who has been extremely generous in teaching me mathematics over the past year. Similarly, conversations with Cliff Callender and Ian Quinn on a variety of topics have been quite stimulating. Two of Callender's papers (2004, 2005) also sparked ideas in the present paper. Finally, my wife Elisabeth Camp provided moral support and inspiration—not to mention a very incisive set of comments on this paper.

[3] The chromatic scale becomes relevant only when defining the symbols "♯" and "♭"; see §1.14.

nicely reflects the fact that the diatonic scale was used long before it came to be interpreted as a subset of the chromatic scale.

1.3 The fundamental frequency f of a pitch can be associated with a real number p according to the equation:

$$p = 69 + 12\log_2 (f/440) \tag{1}$$

This extends the standard system of MIDI note numbers to the microtones, associating any conceivable pitch with a unique real number and any real number with a unique pitch.[4] In this continuous, linear *pitch space*, middle C corresponds to the number 60; the semitone is equal to a distance of one unit; the octave has size 12; and ascending motion in pitch corresponds to ascending motion along the real line **R**.[5]

1.4 We form *pitch-class space* by identifying, or "gluing together," all points p and $p + 12$ in pitch space. The result is the circular quotient space that mathematicians call **R**/12**Z**. We can visualize this space as shown in Figure 1.[6] Note that Figure 1 is *continuous:* although I have labeled only the familiar pitch classes of twelve-tone equal temperament, every point on the figure represents a distinct pitch class.

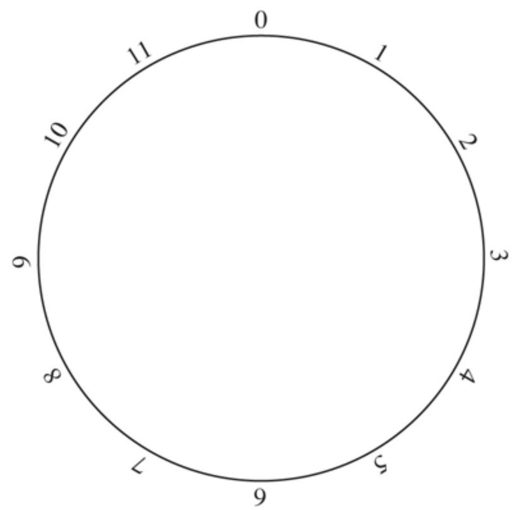
---

[4] Curiously, music theory has standard numerical names for pitch classes but not pitches: theorists typically use spelling-specific designations like "C4" and "C♯6" to refer to pitches. I use MIDI note numbers here because they are reasonably well-known, and because they are consistent with the standard numerical labels for pitch classes.

[5] According to the standard way of measuring distance in pitch space, the distance between two pitches p and q is equal to the absolute value of their difference, $|p – q|$. This is the familiar *undirected interval* between the pitches.

[6] The description of pitch-class space as a circle is not simply a metaphor and does not depend on any specific visual representation of pitch-class space. Pitch-class space is *topologically equivalent* to a circle, and thus shares with it a well-defined mathematical structure. Consequently, we can translate many true statements about pitch-class space into true statements about circles, and vice versa. For example, any method of measuring distances on any circle (that is, any mathematical metric for the circle) defines a metric for pitch classes.

Figure 1. Circular pitch-class space



1.5 We can label pitch classes using real numbers in the range $0 \leq x < 12$. These can be interpreted as clockwise arc lengths from the pitch class labeled 0.[7] Ascending motion in pitch-class space corresponds to clockwise motion on the circle; descending motion corresponds to counterclockwise motion.[8] This system generalizes the familiar integer-based system for notating pitch classes in 12-tone equal temperament, in which C = 0, C♯ = 1, and so on. In continuous pitch-class space these familiar pitch classes retain their familiar names; however, they are joined by microtones such as “C quarter tone sharp,” which is assigned the number 0.5, and “the

---

[7] Real numbers in the range $0 \leq x < 12$ form a group under addition modulo 12**Z**. This is the *quotient group* **R**/12**Z**. Two real numbers x and y are *congruent* modulo 12**Z** if there exists some integer i such that $x = y + 12i$. The quantity “$x + y$ modulo 12**Z**” is the number z, congruent to $(x + y)$ modulo 12**Z**, and lying in the range $0 \leq z < 12$. Addition modulo 12**Z** resembles integer addition modulo 12, except that non-integral values are permitted. For example, $10 + 2.5 \equiv .5$ modulo 12**Z**.

[8] It is purely a matter of convention that we associate ascending motion in pitch-class space with clockwise motion on a circle. We could just as well associate ascending motion in pitch-class space with counter-clockwise motion on an (appropriately labeled) circle. However, the distinction between ascending and descending motion in pitch-class space is not itself conventional; see note 23.

pitch class 17 cents above D," which is assigned the number 2.17.[9]

1.6 **NB.** The system of pitch-class labels described here has been defined to be *consistent with* the system of labeling pitch classes using scale degrees of the familiar 12-tone equal-tempered "chromatic scale." However, these labels do not depend on the existence of this scale, and have been defined without reference to it. The consistency of the two systems is purely a matter of notational convenience. Convenience, however, is purchased at the potential cost of confusion, and it is important to distinguish the two systems. I will use the term *semitone* to refer to a unit of length in pitch-class space that is defined without reference to any chromatic scale. A semitone is 1/12 of an octave, regardless of the chromatic scale. A "chromatic scale step" (or "chromatic step") refers to a length defined in terms of some chromatic scale (not necessarily the familiar one). Only in twelve-tone equal temperament does one chromatic scale step always equal one semitone.

1.7 We can now provide pitch classes with letter names. We choose some multiset[10] of pitch classes to serve as the *basic scale*. We order this scale by choosing some element as the first scale degree and arranging the remaining elements in "scalar order"—that is, so that the absolute sum of the intervals between successive pitch classes totals 12 or less.[11] Finally, we label the successive scale degrees of the basic scale with the letter names A, B, C, D, …

1.8 I will typically list basic scales in letter-name order. For example, I will notate the familiar white note scale as (9, 11, 0, 2,

[9] The distance between any two pitch classes *a* and *b* is usually taken to be the smallest distance between two pitches belonging to those pitch classes. This is the *quotient metric*, corresponding to the *interval class* (or *undirected pitch-class interval*) between two notes. This metric allows us to use pitch-class distances to make general statements about pitch distances. For example, "pitch class E is four semitones away from pitch class C" implies "for every pitch belonging to pitch class C there is a pitch belonging to pitch class E four semitones away from it."

[10] A multiset is a collection that may contain multiple instances of a single object. Footnote 35 motivates the use of multisets.

[11] Here we consider the intervals as real numbers in the range $0 \le x < 12$, and we add them in the normal way, rather than modulo $12\mathbf{Z}$. Given the circular model shown in Figure 1, we require that it be possible to traverse the ordering by starting at the first element, moving exclusively clockwise on the circle, and traveling no more than one circumference in the process.

4, 5, 7).[12] This indicates that letter name "A" corresponds to pitch class 9, "B" corresponds to pitch class 11, "C" corresponds to pitch class 0, and so on. If the basic scale is the pitch-class series (0, 4, 7, 0), corresponding to the C major triad with doubled root, then the letter name "A" corresponds to pitch class 0, "B" corresponds to pitch class 4, "C" corresponds to pitch class 7, and "D" corresponds to pitch class 0. Here, the letter names "A" and "D" refer to the same pitch class. Finally, if the basic scale is the equal-tempered pentatonic scale (1, 3.4, 5.8, 8.2, 10.6) then "A" corresponds to pitch class 1, "B" corresponds to pitch class 3.4, and so on. Note that the procedure described in the preceding paragraph ensures that it is always possible to rotate the basic scale so that it is in ascending numerical order. Thus (0, 0, 4, 7), (0, 4, 7, 0), (4, 7, 0, 0), and (7, 0, 0, 4) are acceptable basic scales, while (0, 4, 0, 7) is not.

1.9 We can provide pitches with letter names by appending an octave number to the letter name of the pitch class containing that pitch.[13] Thus letter name C4 corresponds to pitch number 60 ("middle C"), while letter name B3 corresponds to pitch number 59, a semitone below it.

1.10 Pitch classes in the basic scale can be identified using the familiar staff-and-clef system. A *clef* indicates that a certain staff line (or staff space) corresponds to a letter name; the next space (or line) above this corresponds to the next letter name, and so on.[14] Thus, without knowing anything about the basic scale, we can say that the pitch classes in Figure 2 are called "C," "B," "C," and "A." However, in order to translate these letter names into numerical names we must know what the basic scale is: in the conventional system, the letter names indicate pitch classes (0, 11, 0, 9). When the basic scale is (0, 4, 7, 0), then the letter names indicate the pitch classes (7, 4, 7, 0). When the basic scale is (1, 3.4, 5.8, 8.2, 10.6), then the letter names indicate the pitch classes (5.8, 3.4, 5.8, 1).

---

[12] In this paper, regular parentheses denote ordered lists, and curly braces denote unordered collections. Thus (a, b, c) is ordered, whereas {a, b, c} is not.

[13] Let a be a numerical label for a pitch as defined in §1.3. The *octave number* of a is $\lfloor a/12 \rfloor - 1$, the greatest integer less than or equal to $(a/12) - 1$. Thus "middle C" has octave number 4, as do all pitches between middle C and the next-highest C.

[14] Note that letter names "wrap around" from the end of the ordering to the beginning.

Figure 2. The letter-name series (C, B, C, A)

1.11 Clefs indicate that a staff line or staff space corresponds to a specific letter name. But they also pick out a specific *pitch* possessing this letter name. For instance, in the conventional system, the C clef indicates the staff line corresponding to "middle C," or pitch 60. The space above this line refers to the D immediately above that C, or pitch 62, and so on. We have not yet shown how to generalize the use of clefs to indicate pitches, however. Suppose, for example, our basic scale is (11, 2, 5, 8). We can use a C clef to show that some staff line corresponds to pitch class 5, the third element in the basic scale. But it is not immediately obvious whether this C clef identifies pitch 41, 53, 65, 77, or some other pitch. Once this decision has been made, though, mapping the remaining staff positions to pitches is a straightforward matter: we simply require that what is notated as an ascending step (the interval between one staff position and the next-highest staff position) correspond to an ascending interval between pitches with the appropriate letter names, spanning less than an octave.

1.12 We can rectify this problem by exploiting the system described in §1.3, writing a real number under each clef that identifies the pitch labeled it. Figure 3 demonstrates. (Alternatively, we can imagine this identification to be accomplished by unwritten convention, as in the familiar system.)[15] We will then have a fully generalized staff-and-clef system, suitable for any basic scale with three or more elements.[16]

Figure 3. Using a number to identify a clef's pitch

[15] Here and elsewhere I use the term "familiar system" to refer to the familiar (white-note) basic scale as embedded within the familiar equal-tempered chromatic collection.

[16] For basic scales with one or two elements we would need to develop an "A clef," since these systems have no letter name "C."

1.13 This generalized system ensures that there is an analogical relation between the visual appearance of the notation and the audible features of the music. In moving up by one staff position we move up by one step of the basic scale; this corresponds to a nondescending motion in pitch space.[17] If the basic scale contains no pitch-class duplications then the analogy between the visual and the aural is even tighter: moving up by one staff position corresponds to ascending by step in pitch space. This isomorphism greatly increases the legibility of the notation. (It also permits certain kinds of cross-modal puns, as in Handel's "O Thou That Tellest Good Tidings to Zion," which is said to imitate the shape of mountains in its violin lines.) The analogical relation between visual and aural will play an important role in later sections of the paper.

1.14 We can now extend the system of letter names by defining the symbols "♯" and "♭." This is rather more complicated than one might expect. The standard interpretation of these symbols, which will be adopted here, requires us to reconceive the basic scale as being embedded in (or as being *a submultiset*[18] *of*) another multiset called the *chromatic scale.*[19] The chromatic scale can be identical to the basic scale, though it is usually larger than it. For example, the familiar seven-note basic scale (9, 11, 0, 2, 4, 5, 7) is typically embedded in the larger, equal-tempered, 12-note chromatic scale {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11}.

1.15 We can order the chromatic scale as described in §1.7: choosing an arbitrary element as the first chromatic scale degree, arrange the remaining elements so that they form a nondescending

[17] I use the term "nondescending" because I am permitting basic scales to be multisets; this means that adjacent staff positions may refer to the same pitch.

[18] The mathematical definition of a "submultiset" requires that no two elements of the basic scale correspond to the same element of the chromatic scale. Thus, if the basic scale has pitch class duplications, then the chromatic scale must also have duplications.

[19] It is worth repeating that the only function of the chromatic scale in our generalized notational system is to define the symbols "♯" and "♭." Note that it is also possible to define the symbols "♯" and "♭" without regard to the chromatic scale, by stipulating that they raise or lower a pitch class by some fixed—or even contextually variable—fraction of an octave. (This may in fact be the best way to analyze the early historical meaning of these symbols, as well as their meaning to contemporary string players.) Nothing substantial rests on which definition is chosen.

series of pitch classes spanning an octave or less. Each element in the ordering is *one chromatic scale step above* the preceding element. We conceive of this ordering as circular, with the first element being one chromatic scale step above the last.

1.16 The symbols "♯" and "♭" can be appended to the right of any pitch-class letter name *X*, transforming it into a *compound letter name* that (typically) refers to a different pitch class.[20] The compound symbol "*X*♯" means "the pitch class one chromatic scale step above *X*."[21] The compound symbol "*X*♭" means "the pitch class one chromatic step below *X*." The symbols can be applied recursively. Thus "(*X*♯)♯," or "*X*♯♯," means "the pitch class two chromatic steps above *X*."

1.17 A compound letter name cannot use both the "♯" and "♭" symbols. Thus we rule out pitch-class letter names such as C♯♯♭ or F♭♯♭♭. The symbol "𝄪" is an abbreviation for "♯♯," and therefore counts as *two* accidentals. For longer strings of accidentals, we will use exponential notation: $X\sharp^{12}$ refers to the pitch class 12 chromatic steps above *X*. Non-integer exponents are also permitted. For example, the notation $X\sharp^{1.5}$ refers to the pitch class 1.5 chromatic steps above *X*.[22] This notation is useful if we want letter names for pitch classes lying outside of the chromatic scale, such as "C quarter-tone sharp" in the familiar system.

---

[20] "*X*" and "*X*♯" can refer to the same pitch class if the chromatic scale has duplicate pitch classes, or only one note.

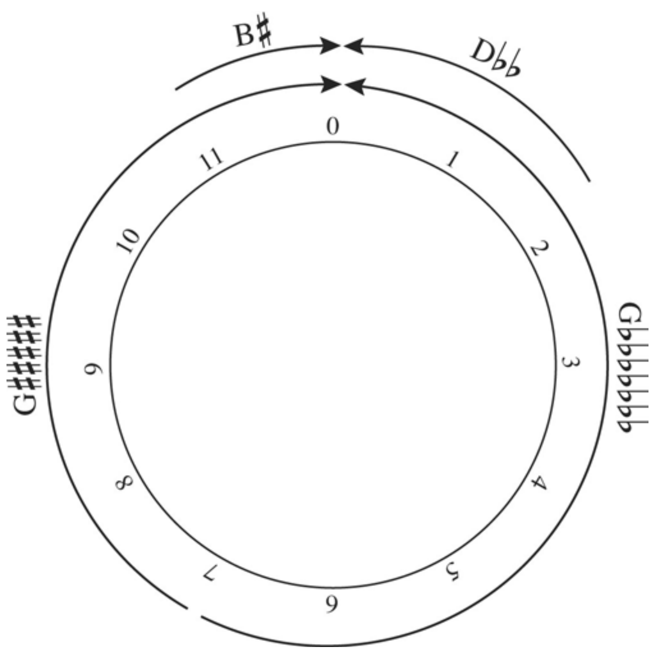
[21] The symbol "C," like the numeral "3," can be analyzed as a simple name without any internal structure. By contrast, "C♯" is a complex term like "2 + 1," whose meaning is determined by systematically combining the meanings of its component parts. The complex term "2 + 1" means "the result of adding 1 to 2," just as "4 – 1" means "the result of subtracting 1 from 4." The two complex terms *refer to* the same object, the number 3, but do so using different semantic values: the instruction "subtract 1 from 4" is different from the instruction "add 1 to 2," even though the two have the same result. In much the same way, the complex term "C♯" means "the pitch class that is an ascending chromatic step above C." This *refers to* the same object—pitch class 1—as does the complex term "D♭." However, the instruction "go up one chromatic step from C" is different from the instruction "go down one chromatic step from D."

[22] Fractional accidentals are to be interpreted in the obvious way, as fractions of a chromatic scale step. In terms of Figure 1, the fraction refers to the fraction of the length of the (smallest) ascending arc between successive steps of the chromatic scale.

1.18 Every pitch class can now be given an infinite number of letter names. For instance, in the familiar system the letter names C, B♯, D♭♭, A♯♯♯, and so forth, all refer to the pitch class 0. This contrasts with the numerical system, in which there is only one name for every pitch class.

1.19 We now come to a point that will be central to the rest of the paper. Every compound letter name can be associated with a unidirectional *path* in pitch-class space.[23] The path begins with the pitch class possessing the letter name in question; its length and direction are determined by the accidentals applied to the name. Figure 4

Figure 4. Every letter name determines a unique path in pitch-class space



[23] An ascending path in pitch-class space is the image of an ascending path in pitch space. That is, it is the path obtained by replacing every pitch in the ascending pitch-space path with the pitch class to which it belongs. The quotient metric described in footnote 9 has the highly desirable feature that every path in pitch space has the same length as its image in pitch-class space. To say that a path in pitch-class space is unidirectional is to say that it moves exclusively in the ascending or descending direction. In the intrinsic geometry of circular pitch-class space, a unidirectional path is a line segment.

demonstrates, identifying the paths corresponding to some letter names in the familiar system. B♯, for example, begins with pitch class 11 and ascends one chromatic scale step to pitch class 0, D♭♭ begins at pitch class 2 and descends two chromatic steps to pitch class 0. G♯♯♯♯♯ starts with pitch class 7 and ascends 5 chromatic scale steps to pitch class 0. G♭♭♭♭♭♭♭ starts at pitch class 7 and descends by 7 chromatic steps to pitch class 0. The two symbols G♯♯♯♯♯ and G♭♭♭♭♭♭♭ therefore determine distinct paths between the same two points in pitch-class space.

1.20 Indeed, there is an isomorphism between letter names that share a letter and unidirectional paths in pitch-class space that share a starting point. Let *L* be a letter naming a pitch class *p* in the basic scale, and let *q* be any other pitch class. Let *L** stand for the set of letter names appending accidentals to *L*. Then there exists a unique letter name in *L** corresponding to every distinct unidirectional path from *p* to *q*. Conversely, each letter name in *L** that refers to *q* corresponds to a unique unidirectional path from *p* to *q*. For example, in the familiar system, there are an infinite number of ways to refer to pitch class 0 by appending accidentals to the letter name G: G♯♯♯♯♯ (5 sharps), G♭♭♭♭♭♭♭ (7 flats), G$\sharp^{17}$ (17 sharps), G$\flat^{19}$ (19 flats), and so on. Every unidirectional path between 7 and 0 corresponds to a unique letter name beginning with the letter "G," and every letter name for pitch class 0 beginning with the letter "G" corresponds uniquely to such a path.

1.21 It is natural to appeal to paths in pitch-class space in attempting to understand the familiar system of pitch-class letter names. These paths are interesting in themselves, however, and we can refer to them without using letter names. Any real number *x* can be associated with a unidirectional path in pitch-class space, with positive numbers corresponding to ascending paths and negative numbers corresponding to descending paths. The absolute value of *x* corresponds to the length of the path as measured in (scale-independent) semitones (see §1.6). I will call this the *scale-independent* way of referring to paths in pitch-class space. In order to distinguish paths from pitch classes, I use boldface type for the former. Thus the symbol "**1**" refers to a path one semitone long that moves in the ascending pitch-class direction. "**–11**" refers to a path that descends by 11 semitones. Numbers greater than or equal to 12 or less than or equal to –12 indicate paths that move at least one complete turn around the pitch-class circle. Thus "**24**" refers to the path that moves two complete turns around the pitch-class circle in the ascending direction, returning to its starting point. Compound terms like "*a* +***b***" refer to the path ***b*** that starts at point *a*.[24] Thus "1 **+5**" refers to the path that starts at pitch class 1 and moves five ascending semitones to pitch class 6. It is clear that there is a straightforward translation between this purely numerical system and the letter-based system we have been investigating.[25]

1.22 Paths in pitch-class space generalize the familiar notion of "pitch-class interval." A traditional pitch-class interval can be reinterpreted as the *shortest ascending path* between two points in pitch-class space. Thus the shortest ascending path between pitch classes 0 and 2 is 2, the shortest ascending path between pitch classes 3 and 2 is 11, and so forth. The system described in §1.21 has names for these paths and many more besides: it allows us to specify any of the unidirectional paths connecting any two pitch classes. It therefore extends the pitch-class intervals by providing us an infinite number of *transformations* between any two pitch classes.[26]

---

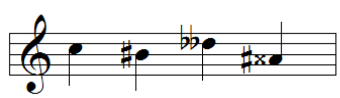
[24] Thus "*a*" refers to a pitch class, while "***+b***" refers to a path. See Lewin 1987 on the difference between labels for pitch classes and labels for pitch class intervals.

[25] There are two important differences between letter names and the numerical labels for paths described in §1.21. First, letter names allow us to refer only to paths beginning with elements of the basic scale, while numbers allow us to refer to paths with arbitrary starting points. Second, letter names measure length using chromatic scale-steps, while numbers use scale-independent semitones. Of course in the familiar system the two methods of measuring path length are equivalent.

[26] David Lewin (1987) famously conceived of intervals, not as "distances" or "relations" between fixed points in an unchanging space, but rather as active *ways of transforming* one point in the space into another. This nicely captures the motivation for thinking about paths in pitch-class space. Curiously, however, Lewin's concept of a "Generalized Interval System" specifically requires that there be *only one interval between any two musical objects.* Thus he would not consider the set of paths in pitch-class space to represent a "Generalized Interval System" for the pitch classes. I believe that Lewin was overly restrictive here; his definition of a "Generalized Interval System" does not fully embody the transformational attitude he espoused.

1.23 We can also apply the symbols "♯" and "♭" to notations for pitches in the familiar staff-and-clef system.[27] Figure 5 shows several different ways of notating pitch 72 in the familiar system. Once again, any pitch can be notated in an infinite number of ways. (This contrasts with the numerical system for labeling pitches, in which every pitch has a unique name.) Conversely, we can transform a notation for any pitch into a notation for any other pitch by adding or removing the appropriate number of accidentals. For instance, we can transform a symbol for pitch 67 into a symbol for pitch 72 by adding five sharps.[28] Note, however, that in *pitch-class* space, a notation for pitch class 7 can be transformed into a notation for pitch class 0 in an infinite number of ways: we can add five sharps, or seven flats, or seventeen sharps, or nineteen flats, and so on. By contrast, in pitch space, given a notation for pitch 67, we must add precisely five sharps to transform it into a notation for pitch 72.[29]

Figure 5. Four equivalent notations for pitch 72



[27] In addition to the staff-and-clef system, there is also "scientific pitch notation," which combines a letter name, accidentals, and a Roman numeral conveying octave information. Somewhat counterintuitively, the notation "C♭4" is typically used to mean "the pitch one chromatic step below C4," rather than "the instance of pitch class C♭ lying in octave number 4." Thus C♭4 is enharmonically equivalent to B3, not B4. A more logical system would place the octave number before the accidental, as in "C4♭." This would show that the letter name is first combined with the octave number to determine a pitch, which is subsequently altered by the application of accidentals.

[28] "Adding five sharps" here is shorthand for a process that may involve removing flats, adding sharps, or both. Thus to transform A♭♭4 (pitch 67) into A♯♯♯4 (pitch 72), we remove two flats and add three sharps, for a total change of five accidentals.

[29] This is because the group of unidirectional paths acts *simply transitively* on pitches, but not on pitch classes.

II. Voice leadings as generalized key signatures

2.1 Section I generalized the basic features of Western musical notation. In Section II, I show how to interpret key signatures within this generalized system. We will see that key signatures are collections of paths in pitch-class space, and hence can be interpreted as voice leadings between pitch-class sets. Every key signature determines a unique voice leading between pitch-class sets; and conversely, every voice leading can be expressed as a key signature. I conclude by suggesting that generalized key signatures provide a more flexible alternative to familiar conceptions of voice leading between pitch-class sets.

2.2 A *key signature* places accidentals on staff positions at the beginning of a measure or staff system. These accidentals are understood to apply by default to all staff positions corresponding to that letter name. Figure 6 illustrates. Here, one and the same notated key signature—and one pattern of letter names—is interpreted in the context of two different basic scales. Figure 6(a) is in B♭ major in the familiar system; the notation represents letter name series (B♭, E♭, B♭, E♭), indicating the pitch-class series (10, 3, 10, 3) and pitch series (70, 63, 58, 51). Figure 6(b) presents the same notated key signature, and the same series of letter names, in the context of the basic scale (9, 0, 2, 4, 7). (We imagine this scale embedded within the usual chromatic.) As always, the C clef indicates the third element of the basic scale, here pitch class 2.

Figure 6. The same notated key signature interpreted with respect to two different scales

a) the familiar system

basic scale (9, 11, 0, 2, 4, 5, 7)
letter names (B♭, E♭, B♭, E♭)
chromatic scale {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11}
pitch classes (10, 3, 10, 3)
pitches (70, 63, 58, 51)

Figure 6 (continued).

b) with a pentatonic basic scale

basic scale (9, 0, 2, 4, 7)
letter names (B♭, E♭, B♭, E♭)
chromatic scale {0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11}
pitch classes (11, 6, 11, 6)
pitches (71, 66, 59, 44)

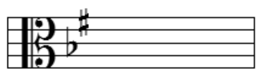
Since the basic scale has only five elements, the space below the top line in Figure 6(b) corresponds to letter name "A." Likewise, the space above the bottom line indicates letter name "E." The integer notation under the clef indicates that it refers to the pitch 62. With a little work, we can determine that the notation represents the letter-name series (B♭, E♭, B♭, E♭), corresponding to pitch classes (11, 6, 11, 6) and pitches (71, 66, 59, 44).[30]

2.3 We will require that a key signature not assign conflicting accidentals to staff positions representing the same letter name. Thus a single written key signature may be acceptable relative to some basic scales, but not others. The notation shown in Figure 7 is acceptable in the familiar system, since it indicates F♯ and B♭. It is unacceptable when the basic scale contains only four elements, since it then indicates B♯ and B♭—impermissibly assigning two different accidentals to the same letter name.

---

[30] Note that one and the same written key signature transposes the C major scale down by whole step to B♭ major, while transposing the C pentatonic scale up by whole step to D major. The exact orthographic resemblance between the two key signatures is something of a coincidence: pentatonic key signatures involving three or more accidentals no longer look exactly like their diatonic counterparts. However, there is a reason why it takes a *flat* to transpose a pentatonic scale up by fifth. This is because the pentatonic scale is the diatonic scale's complement: a voice leading associating two fifth-related pentatonic scales by way of a single descending semitone determines a complementary voice leading associating two fifth-related diatonic scales by *ascending* semitone.

Figure 7. A key signature can be acceptable relative to some scales, but not others

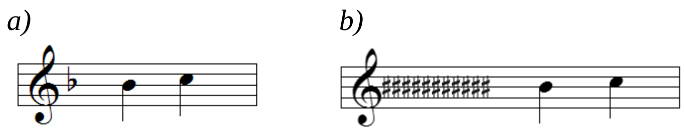


2.4 I will ignore questions about which specific staff lines to use in notating a given key signature. That is a matter of orthographic convention, tangential to the current inquiry. I will also ignore questions about pitch priority. I will say that one and the same *key signature* is used to denote two different *keys*. For example, in the familiar system, the one-sharp key signature is used to denote both E minor and G major. (It can also be used to denote F♯ locrian, A dorian, B phrygian, and so on.) I will refer to key signatures using major-key names, describing the one-sharp signature as "G major," and so on. This is merely terminological shorthand; it is not meant to privilege the major key over its relative minor key or any of the other modes.

2.5 Key signatures can be represented as ordered *n*-tuples of real numbers such that the i^{th} element of the *n*-tuple indicates the accidentals applied to the i^{th} letter name in the system. Positive numbers refer to sharps and negative numbers refer to flats. Thus, in the familiar system the 7-tuple (0, –1, 0, 0, 0, 0, 0) refers to the standard key signature for F major: the letter names A, C, D, E, F and G have no accidentals; while the letter name B has one flat, as indicated by the number –1.

2.6 We have now reached the crux of our investigation. For key-signature *n*-tuples determine *collections of paths in pitch-class space*. That is, they associate pitch classes in the basic scale with pitch classes in the "target" scale by way of specific routes through pitch-class space. For example, the standard key signature for F major, (0, –1, 0, 0, 0, 0, 0), holds six pitch classes constant (moves them by zero chromatic scale steps) and takes pitch class 11 to pitch class 10 by one descending chromatic scale step. This key signature is *not* the same as the key signature (0, 11, 0, 0, 0, 0, 0), which moves pitch class 11 to 10 by eleven ascending chromatic steps. Consequently, the two notations shown in Figure 8 are not equivalent: the first pitch shown in Fig. 8(a) is one octave lower than the first pitch shown in Fig. 8(b); the interval shown in Fig. 8(a) sounds like an ascending major second, while that shown in Fig. 8(b) sounds like a descending minor seventh.

Figure 8. Two non-equivalent key signatures



2.7 Observe that in the familiar, equal-tempered system, a key signature's target scale can be obtained by adding the i^{th} component of the key signature to the i^{th} component of the basic scale, and reducing the result modulo 12**Z**. For instance, adding the F major key signature (0, –1, 0, 0, 0, 0, 0) to the basic scale (9, 11, 0, 2, 4, 5, 7) produces the F major collection (9, 10, 0, 2, 4, 5, 7). We can write

$$(9, 11, 0, 2, 4, 5, 7) + (\mathbf{0}, \mathbf{-1}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}) \equiv (9, 10, 0, 2, 4, 5, 7) \text{ modulo } 12\mathbf{Z} \quad (2)$$

which instantiates the more general schema:

$$\text{basic scale} + \text{key signature} \equiv \text{target scale modulo } 12\mathbf{Z} \qquad (3)$$

This schema works only because in the familiar system the length of a path, as measured in terms of chromatic scale steps, happens to be the same as the length of the path as measured in scale-independent semitones (see §1.6 and §1.21). In other chromatic systems, one needs to translate the key signature *n*-tuple, which is measured in chromatic scale steps, into an *n*-tuple of scale-independent semitonal path lengths.[31]

2.8 I will now demonstrate that key signatures are extremely closely-related—indeed, virtually equivalent—to voice leadings between pitch-class sets. It follows that Western musicians are accustomed to working with such voice leadings, although they typically have not realized that this is what they are doing. This is significant, since some theorists are suspicious of the very notion

[31] For example, in an equal-tempered six-note chromatic system, one chromatic scale-step corresponds to a path that is two (scale-independent) semitones long. Thus it is necessary to multiply the values in the key signature *n*-tuple by two before adding them to the *n*-tuple representing the basic scale. For non-equal-tempered chromatic systems, the translation requires more work, since "one chromatic step" can represent various pitch-class distances.

of a "voice leading between pitch-class sets."[32] My hope is that the present discussion, by demonstrating that we need to use these voice leadings in order to explain elementary features of our system of musical notation, will help to overcome such skepticism.

2.9 To understand the relationship between key signatures and voice leadings, it is necessary to distinguish what I will call the *path-specific* and *path-neutral* conceptions of voice leading. As we will see, key signatures correspond to path-specific voice leadings. Path-neutral voice leadings are slightly more general than key signatures; they can be understood as equivalence classes of closely related path-specific voice leadings.

2.10 A *chord* is a multiset of pitch classes. A *path-specific voice leading* between two chords maps every element of the first chord to some element of the second by way of a specific unidirectional path in pitch-class space; furthermore, for every element in the second chord, at least one element in the first chord is mapped to it, again by way of some specific unidirectional path in pitch-class space. Path-dependent voice leadings can be represented using equations that precisely mirror Equations 2 and 3 above:

$$\text{source chord} + n\text{-tuple of paths} \equiv \text{target chord modulo } 12\mathbf{Z} \quad (4)$$

Here, the "*n*-tuple of paths" is an *n*-tuple of real numbers indicating the direction and magnitude of paths in pitch-class space. (Note that we refer to these paths in the scale-independent way described in §1.21.)

2.11 For example, the following equation determines a path-specific voice leading between the C diatonic and F diatonic collections

$$(9, 11, 0, 2, 4, 5, 7) + (\mathbf{0}, \mathbf{-1}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}) \equiv (9, 10, 0, 2, 4, 5, 7) \text{ modulo } 12\mathbf{Z} \quad (5)$$

This voice leading holds six pitch classes constant, while moving pitch class B down by one semitone to B♭. It therefore represents what Richard Cohn calls a *maximally smooth* voice leading.[33] Note

[32] As will become clear in §§2.18-2.20 below, I believe this skepticism derives from the fact that previous theorists have not distinguished pitch-class intervals from paths in pitch class space. Skeptical theorists have rightly felt that there is something problematic about this.

[33] See Cohn 1996. A "maximally smooth voice leading" is one in which just a single note moves, and it moves by only one semitone.

that, as in §1.21, I use boldface to designate pitch-class paths. Note also that the voice leading shown in Equation 5 is the *same* path-specific voice leading indicated by the equation

$$(0, 2, 4, 5, 7, 9, 11) + (\mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{-1}) \equiv (0, 2, 4, 5, 7, 9, 10) \text{ modulo } 12\mathbf{Z} \quad (6)$$

This is because Equations 5 and 6 both associate the same pitch classes in the source chord with the same pitch classes in the target chord, along the same paths through pitch-class space. Only the order of the respective *n*-tuples changes. This difference is purely orthographic. Chords are intrinsically *unordered* objects, and voice leadings simply specify how their elements are to be related.

2.12 Equation 7 displays another path-specific voice leading between pitch-class sets:

$$(0, 4, 7) + (\mathbf{-1}, \mathbf{0}, \mathbf{0}) \equiv (11, 4, 7) \text{ modulo } 12\mathbf{Z} \quad (7)$$

This represents what neo-Riemannian theorists might call the "L voice-leading": the third and fifth of the C major triad are held constant, while the root moves down by semitone.

2.13 Figure 8 showed that a key signature specifies not just *which* pitch classes are to be related, but also *how* they are to be related. Thus it is clear that any key signature determines *a unique path-specific voice leading* between the basic scale and the target scale. Conversely, any path-specific voice leading can be expressed using key signatures. Begin by writing the path-specific voice leading in the form

$$(a_1, a_2, \ldots, a_n) + (\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n) \equiv (b_1, b_2, \ldots, b_n) \text{ modulo } 12\mathbf{Z}$$

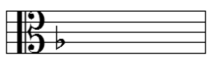
Arrange the a_i so that they form an acceptable basic scale.[34] Choose a chromatic scale containing all the a_i and b_i. Translate the path names $\boldsymbol{p}_i$, which are independent of any chromatic scale, so that they identify paths in terms of chromatic scale steps. Then apply accidentals to a staff position representing each element a_i in the

[34] Note that we can choose to call any of the a_i by the letter name "A." Consequently, there are multiple ways to arrange the a_i so that they form an acceptable basic scale. Moreover, there are multiple ways to embed this basic scale in a larger chromatic collection. Therefore, there will be more than one key signature corresponding to a single voice leading.

source chord so that the pitch class a_i moves to pitch class b_i by the appropriate path.[35]

2.14 For instance, Figure 9 provides a key signature for the neo-Riemannian "L voice leading" represented by Equation 7. The basic scale is (0, 4, 7). The letter name "A" refers to pitch class 0, letter name "B" refers to pitch class 4, and letter name "C" refers to pitch class 7. The chromatic scale is the familiar one. A single flat is applied to a staff line representing letter name A. This moves pitch class 0 down by one chromatic scale step to pitch class 11. The remaining notes in the basic scale are held constant.

Figure 9. A key signature for the "L voice leading"



basic scale (0, 4, 7)
path n-*tuple* (**–1**, **0**, **0**)
key signature (–1, 0, 0)
target scale (11, 4, 7)

2.15 Key signatures are path-specific voice leadings. Recent theory, however, has primarily been concerned with what I will call "path-neutral voice leadings." A *path-neutral voice leading* is a mapping between the pitch classes of two chords that does not specify paths by which they are to be associated. A path-neutral voice leading simply associates every element of one chord with some element of the other. We can notate path-neutral voice leadings using the notation

$$(a_1, a_2, \ldots, a_n) \to (b_1, b_2, \ldots, b_n)$$

This indicates that the *i*th pitch classes in each *n*-tuple are associated by the voice leading. Thus the notation (0, 4, 7)→(11, 4, 7) indicates that pitch classes 0, 4, and 7 are associated with pitch classes 11, 4, and 7, respectively.

2.16 The voice leading (0, 4, 7)→(11, 4, 7) is the *same* voice leading as the voice leading (0, 7, 4)→(11, 7, 4), since they both

[35] Note that we can represent every voice leading as a key signature only because we allowed multisets to serve as the basic scale. Had we not done so, we would be unable to use key signatures to represent voice leadings like (0, 4, 7, 0) + (**2**, **1**, **0**, **–1**) ≡ (2, 5, 7, 11) modulo 12**Z**.

associate the same pairs of pitch classes in the source and target chords. Once again, this is because chords are unordered objects and voice leadings merely specify which elements are related. However, the voice leading (0, 4, 7)→(4, 7, 11) is *not* the same as (0, 4, 7)→(11, 4, 7), since the two voice leadings relate different pitch classes.

2.17 It is clear that any path-specific voice leading (or, equivalently, key signature) determines exactly one path-neutral voice leading, since it associates every element of the basic scale with some element of the target scale and vice versa. Conversely, any path-neutral voice leading corresponds to an infinite class of path-specific voice leadings (or, equivalently, key signatures). This is because there are an infinite number of unidirectional paths connecting any two points in pitch class space—as we saw in §1.20.

2.18 The question now arises: when theorists talk about "voice leading," are they thinking about path-specific or path-neutral voice leadings? The question is complicated. Explicit discussions of voice leading, as for instance in Lewin (1998), Morris (1998) and Straus (2003), tend to suggest the path-neutral conception. (Lewin, for example, defines a voice leading as a set of ordered pairs of pitch classes.) However, all of these theorists accept the traditional notion of "pitch-class interval," according to which there is only one interval between any two pitch classes. This leads them, on occasion, to speak as if pitch-class intervals correspond to specific paths in pitch-class space. For this reason, much recent writing about voice leading is *de facto* concerned with path-specific voice leadings, even while using path-neutral terminology.

2.19 For example, when neo-Riemannian theorists talk about the voice leading (0, 4, 7)→(11, 4, 7), they are not simply asserting that pitch class 0 is mapped to pitch class 11. They also imagine that this mapping corresponds to a *single-semitone displacement*—that it is, as they say, a "maximally smooth" voice leading in which one note *moves by* one semitone. This notion of a pitch class "moving by" a semitone clearly suggests a path in pitch-class space. Certainly, neo-Riemannian theorists do not imagine the L voice leading to be one in which pitch class 0 moves to pitch class 11 by way of eleven ascending semitones. (After all, this is not a particularly efficient voice leading, and it involves motion by considerably more than one semitone.) Such examples suggest that theorists have often operated with the path-specific conception of voice leading, even though they have not always realized it, and even though they have not always found the most appropriate language in which to express this notion.

2.20 This observation echoes a complaint commonly made against the very notion of "voice leading between pitch-class sets": that pitch class is too crude a tool for investigating voice leading, since we cannot distinguish the myriad ways in which one pitch class can move to another. Our investigation provides a response to this complaint: as long as we are thinking about paths in pitch-class space, rather than intervals, then we can indeed distinguish the various ways in which one pitch class can move to another. And as long as we are thinking about *path-specific* voice leadings, then we can distinguish a variety of voice leadings corresponding to the same, path-neutral voice leading. What is remarkable is that we have been forced to draw these distinctions simply in order to understand elementary features of Western musical notation. For path-specific voice leadings are little more than generalized key signatures. Thus the pedantry of Section I has borne fruit in a genuinely new theoretical tool for understanding voice leading.

III. The three virtues

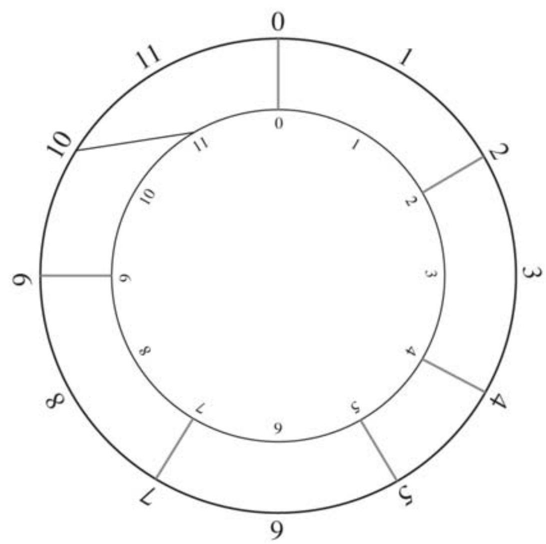
3.1 The remainder of this paper treats key signatures and voice leadings in tandem. Key signatures, rather than being the central focus of our investigation, will now serve to illustrate more general points about voice leading as such. In this section, I describe three virtues that a voice leading or key signature can have: avoidance of voice crossings, preservation of common tones, and minimization of the "size" of the voice leading. I show that, for a wide range of measures of voice-leading size, it is always possible to find a minimal voice leading that has no voice crossings. I also demonstrate that it is not always possible to preserve common tones while avoiding voice crossings.

3.2 It is highly desirable that a key signature not destroy the analogical relation between the aural and the visual. In other words, what sounds like a descending pitch interval should not be notated as an ascending interval. Thus it is reasonable to forbid key signatures such as (0, 0, 0, 0, 1, –1, 0) in the familiar system, which include both E♯ and F♭. Such key signatures make it difficult to read and conceptualize music. Figure 10 illustrates: here, what looks like an ascending step here sounds like a falling semitone.

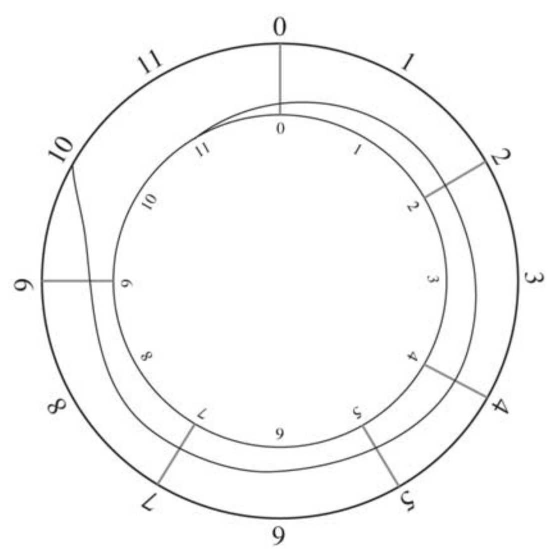
Figure 10. A very confusing key signature

Figure 11. Key signatures whose paths cross destroy the analogical relation between the aural and visual

a) the standard F major key signature



b) the alternative shown in Fig. 8(b)



3.3 A key signature preserves the analogical relation between the aural and visual only if it determines a collection of paths in pitch-class space that *do not intersect.*[36] Figure 11(a), for example, represents the familiar key signature for F major. The inner circle corresponds to the basic scale, and the outer circle corresponds to the target scale; time therefore progresses radially outward. The lines represent the key signature's paths, which do not intersect. Figure 11(b) corresponds to the key signature shown in Figure 8(b). Here, B is mapped to B♭ by eleven ascending semitones, along a path that does intersect the other paths in the key signature. As Figure 8 demonstrated, the first key signature preserves the analogical relation between the visual and the aural; the second does not. Thus we have good reason to be interested in key signatures whose associated voice leadings are crossing-free.

3.4 Second, we might prefer key signatures that use the *fewest possible number of sharps or flats* in taking the basic scale to the target scale.[37] All other things being equal, such key signatures maximize legibility. For example, it is much easier to understand the key signature of D-major, which uses just two accidentals, than the

---

[36] We will not consider two distinct paths to intersect if they meet only at an endpoint. Thus, the paths "2 **–1**" and "2 **+1**" do not cross, nor do "3 **+1**" and "5 **–1**" (the notation here is, of course, the one used in §1.21). When the basic scale contains pitch-class duplications, however, complications arise: if letter names B and C refer to the same pitch class, then {B♭, C} preserves the analogical relation between the aural and the visual, whereas {B, C♭♭} does not. This despite the fact that both pairs of letter names indicate crossing-free paths and refer to the same pair of pitch classes. (Note that the second pair has an extra flat, which reflects the size-zero "step" between B and C.) Thus avoiding crossings is necessary but not sufficient for preserving the analogical relation between the aural and the visual. In the special case where the basic scale has duplications, we must further require the key signature's paths to be appropriately distributed among letter names referring to the same pitch class. We can express this by stipulating that the key signature not be one that would have voice crossings if the "duplicated" notes were *very close together* (and still in letter-name order) rather than being exactly identical.

[37] When determining the key signature with "the smallest possible number of accidentals," we hold the basic and chromatic scales fixed; given a basic scale B and a chromatic scale C we are interested in the key signatures that use as few accidentals as possible to refer to target scale T.

enharmonically equivalent key signature of C♯♯ major—which uses 14.

3.5 When the chromatic scale is equal-tempered, a key signature minimizes the use of accidentals if and only if its associated path-specific voice leading is *as small as possible* according to the most common measure of voice-leading size. In such cases the number of sharps or flats used in a key signature is proportional to the *aggregate semitonal length* of all the paths in the voice leading. "Aggregate semitonal length" is a common measure of size that is called the L^1 or "taxicab" norm by mathematicians, and "smoothness" or "voice leading efficiency" by music theorists.[38]

3.6 Third, we might like key signatures to maximize the number of pitch classes written the same way as they are written in the basic scale. This facilitates the common practice of "pivot chord" modulation, or modulation by way of chords common to two keys. For example, the letter names {E, G, B, D} are used by both the C major and D major key signatures. By contrast, the C major and C♯♯ major key signatures contain no identically-notated "pivot notes." Of course, both scales contain the pitch classes {4, 7, 11, 2}; however, they notate them with different letter names.

3.7 A key signature maximizes "pivot notes" if and only if its associated voice leading *maximizes the number of common tones*, or zero-semitone paths. For example, the voice leading

$$(9, 11, 0, 2, 4, 5, 7) + (\mathbf{0}, \mathbf{0}, \mathbf{1}, \mathbf{0}, \mathbf{0}, \mathbf{1}, \mathbf{0}) \equiv (9, 11, 1, 2, 4, 6, 7) \text{ modulo } 12\mathbf{Z}$$

corresponding to the standard key signature for D major, has five zero-semitone paths. By contrast, the voice leading

$$(9, 11, 0, 2, 4, 5, 7) + (\mathbf{2}, \mathbf{2}, \mathbf{2}, \mathbf{2}, \mathbf{2}, \mathbf{2}, \mathbf{2}) \equiv (11, 1, 2, 4, 6, 7, 9) \text{ modulo } 12\mathbf{Z}$$

corresponding to the standard key signature for C♯♯ major, contains no zero-semitone paths. Every zero-semitone path represents a

---

[38] Roeder 1984, Lewin 1998, Cohn 1998, Straus 2003. There are other reasonable ways to measure voice-leading size as well: one might take the size of a voice leading to be the length of its largest path; or the square root of the sum of the squares of its path lengths. In the remainder of the paper, I will assume the "smoothness" metric. This is merely an expositional device. The central results and techniques of this paper are consistent with a very wide range of measures of voice-leading size.

"pivot note." The first key signature maximizes "pivot notes," while the second does not.

3.8 We now ask: which of these three virtues are mutually compatible? Can one preserve the analogical relation between the aural and visual, minimize the use of accidentals, and maximize pivot notes at the same time? Here, our inquiry into key signatures arrives at a fundamental question in the theory of voice leading. For we are asking whether it is possible simultaneously to avoid voice crossings, preserve common tones, and minimize voice-leading size. All three qualities have long been considered desiderata by theorists and pedagogues of voice leading.[39]

3.9 For an extremely wide range of methods of measuring voice-leading size, including the L^1 ("smoothness") norm, it can be shown that there will be a minimal path-dependent voice leading between any two multisets whose paths do not cross. (As the Appendix explains, this follows from the triangle inequality.) Thus, given any basic scale *B* and any chromatic scale *C*, we can always find a key signature for any target scale *T* (of the same cardinality as *B*) that minimizes accidental-use while preserving the relation between the aural and visual.[40] Since the preservation

[39] Of course, traditional rules of voice leading mandate the avoidance of voice crossings in pitch- rather than pitch-class space. However, the two are intimately connected. Crossing-free voice paths in pitch-class space are precisely those that will never have voice crossings when instantiated in pitch space. Similarly, pitch-space voice leadings between chords spanning less than an octave will be crossing-free only if they instantiate voice leadings that have no crossings in pitch-class space.

[40] When the chromatic scale is not equal-tempered, then "smoothness" is not necessarily proportional to the number of accidentals used in a key signature. However, we can ignore this complication. The result stated here—that it is always possible to preserve the analogical relation between the visual and the aural while minimizing accidental-use—holds for any chromatic scale. This is because we can transform an equal-tempered chromatic scale into any other scale of the same cardinality without changing the "size" of any key signature (as measured in accidentals), without introducing or removing voice crossings from any key signature, and without causing two key signatures that refer to the same target scale to refer to different target scales. (I simply state these facts here, without proof.) It follows that we can extend our reasoning from equal-tempered chromatic scales to all chromatic scales of the same cardinality.

of the analogical relation between the aural and visual is extremely important when notating music, I will henceforth consider only key signatures that preserve this relationship.

3.10 Now consider the third virtue: is it possible to maximize the number of "pivot notes" while preserving the analogical relation between the aural and visual? It is easy to construct examples showing that we cannot. For instance, the C major, F♯ major, and G♭ major scales all contain the pitch classes {5, 11}. However, they notate these pitch classes with different letter names: in C major they are notated {F, B}, in F♯ major they are notated {E♯, B}, and in G♭ major they are notated {F, C♭}. A little thought will show that there can be no familiar-system key signature for the collection {6, 8, 10, 11, 1, 3, 5} in which the pitches {5, 11} are assigned the letter names {F, B}, and in which the analogical relation between the aural and visual is preserved.

3.11 However, it *is* possible to minimize accidental-use while preserving common tones. Such key signatures may destroy the analogical relationship between the aural and the visual; when they do, they will involve longer paths (double sharps, triple flats, and other "higher" accidentals) than their crossing-free alternatives. For instance, in the familiar system, there is a six-accidental key signature for the diatonic collection {6, 8, 10, 11, 1, 3, 5} that notates pitch classes 5 and 11 as F and B: (1, 0, 1, 1, 2, 0, 1) or {A♯, B, C♯, D♯, E♯♯, F, G♯}. This key signature uses the same number of accidentals as the standard key signatures for F♯/G♭ major. However, it uses one double sharp, whereas the standard major key signatures use none. It also contains a notated ascending step, E♯♯–F, that sounds like a descending pitch interval.

3.12 We can be assured of finding minimal voice leadings that maximize common tones only when we use the L^1 and a few other related measures of voice-leading size.[41] There are many perfectly reasonable ways of measuring voice-leading size that do not enable us always to find such voice leadings; thus these two virtues are only weakly compatible. By contrast, avoiding voice crossings and minimizing voice-leading size are *strongly* compatible: virtually

[41] This is because such measures tend to mandate voice crossings, thereby violating the triangle inequality. This makes them somewhat less useful—though not *in principle* useless. Exactly why we should want measures of voice leading size to obey the triangle inequality is a complicated matter that deserves careful attention. Callender (2005) discusses this.

every useful measure of voice-leading size—including all of those actually employed by music theorists—allows us to find minimal voice leadings that are crossing free. This is because measures of voice-leading size typically satisfy the triangle inequality, and the triangle inequality is in turn closely connected to the avoidance of voice crossings. From the standpoint of the general theory of voice leading, then, we may have reason to consider avoidance of voice crossings more fundamental than preservation of common tones.

IV. Minimal voice-leadings and the interscalar path matrix

4.1 Section IV considers a central question in the theory of voice leading: given two chords, how does one find a minimal voice leading between them?[42] I begin by describing a solution to the general problem. I then show that the standard system of diatonic key signatures in fact implements this solution. Consequently, there is a minimal voice leading between any two diatonic collections that is equivalent to a familiar key signature.

4.2 The Appendix shows that for a wide variety of methods of measuring voice-leading size, there will always be a minimal voice leading between any two chords that is crossing-free. Thus to find a minimal voice leading between two chords we need only search the crossing-free voice leadings connecting them. There are in fact an infinite number of possibilities.[43] Fortunately, it is possible to represent these voice leadings with what I call an *interscalar path matrix.* Using such matrices, it is straightforward to find a minimal, crossing-free voice leading between arbitrary chords.

4.3 Let the "source" and "target" chords be multisets of equal cardinality. Order each multiset as described in §1.7, labeling each with "scale degree numbers" (1, 2, …, n) rather than letter names

[42] For the rest this paper, the term "voice leading" will always refer to *bijective* voice leadings. These voice leadings map every element in the source chord to *one and only one element in the target.* That is because key signatures always correspond to bijective voice leadings from the basic scale to the target scale. (Of course these scales themselves may contain duplications.) The problem of finding minimal non-bijective voice leadings between arbitrary chords is more difficult than the problem discussed here. Its solution is a matter for another paper.

[43] This is because we can add 12 to all the paths in a voice leading without creating any new crossings.

(A, B, C, …). The *interscalar path matrix* is a matrix whose rows contain the paths associated with path-specific, crossing-free voice leadings from source to target.[44] The *j*th element of row *i* is a path that takes scale degree *j* in the source chord to scale degree $j + i - 1$ in the target. (The quantity $j + i - 1$ is to be interpreted modulo the cardinality of the two chords: in a seven-note scale, scale degree 8 is the same as scale degree 1.) The resulting matrix can be used to determine every crossing-free voice leading between the chords generating it.

Table 1. A interscalar path matrix for the diatonic scale

0 steps	**0**	**0**	**0**	**0**	**0**	**0**	**0**
1 step	**2**	**1**	**2**	**2**	**1**	**2**	**2**
2 steps	**3**	**3**	**4**	**3**	**3**	**4**	**4**
3 steps	**5**	**5**	**5**	**5**	**5**	**6**	**5**
4 steps	**7**	**6**	**7**	**7**	**7**	**7**	**7**
5 steps	**8**	**8**	**9**	**9**	**8**	**9**	**9**
6 steps	**10**	**10**	**11**	**10**	**10**	**11**	**10**

4.4 Table 1 presents an interscalar path matrix relating the basic scale to itself. (Since the source and target chords are the same, we could drop the prefix "inter-," describing Table 1 as a *scalar path matrix*.) Scale degrees are numbered from pitch class 9, and correspond to the degrees of the A natural-minor scale. The first row of the matrix contains the zero-semitone paths that take scale-degree *x* in the basic scale to itself. The second row contains paths taking scale degree *x* in the basic scale to scale degree $x + 1$ in the scale. This voice leading transposes the basic scale *up by one diatonic step*, sending A to B by 2 ascending chromatic scale steps, B to C by 1 ascending chromatic step, C to D by 2 chromatic steps, and so forth. Each of the remaining rows transposes the notes of the basic scale up by some fixed number of steps. The columns of

[44] We can measure these paths either in the scale-dependent or scale-independent manner, depending on whether we are most interested in key signatures or voice leadings. In the familiar system, of course, the two are equivalent.

the matrix contain the modes of the basic scale, indicating that this scalar path matrix is what serial theorists call a rotational array.[45]

4.5 Table 1 is *complete* in the following sense: for any note of the basic scale, there is a cell of Table 1 containing a path between it and any other note of the basic scale. Thus we can specify a voice leading (i.e. row) in Table 1 simply by choosing an ordered pair of notes in the basic scale. The completeness of Table 1 implies that we can use it to generate *any* bijective crossing-free voice leading between the basic scale and any other diatonic collection. To see why this is so, consider the effect of adding a constant value to a row of Table 1. Adding 1 to all the cells in any row transposes the target chord up by semitone; subtracting 1 transposes the target chord down by semitone. Adding 12 does not transpose the target chord, but extends each path in the voice leading by one ascending turn around the pitch-class circle. Thus Table 1 allows us to generate a crossing-free voice leading that maps any pitch class x in one diatonic collection to any pitch class y in another, along any unidirectional path p connecting them in pitch-class space. There is only one such crossing-free voice leading for any (x, y, p). It follows that the rows of Table 1 generate all the bijective crossing-free voice leadings between diatonic collections.

4.6 Suppose, now, that we want to find a *minimal* voice leading between the C major collection and its transposition by t semitones. It suffices to find that row of Table 1 whose values, when added to t, are closest to 0 modulo $12\mathbf{Z}$.[46] As we have seen, adding t to each item in this row transposes the basic scale by t semitones; the result is a crossing-free voice leading between the desired collections. We then add the quantity $12c$, where c is an integer, that brings the items in the row as close as possible to 0.

[45] See Morris 1988. An interscalar path matrix differs from a rotational array in two ways. First, the source and target chords must be in scalar order, as described in §1.7. Rotational arrays, by contrast, are defined for arbitrary orderings. Second, the entries of an interscalar path matrix refer to paths in pitch-class space rather than traditional pitch-class intervals. Thus the interscalar path matrix shown in Table 3 is not a rotational array.

[46] "Closeness" implies some metric of voice-leading size. Since we are assuming the L^1 "smoothness" metric, we want to minimize the quantity $\Sigma\delta(x_i + t, 0)$, where t is the transposition in question, the x_i are the elements of the row, and δ represents "distance" as measured according to footnote 9.

4.7 For example, suppose we wish to find a minimal voice leading between the C major and G major collections. G major is $\mathbf{T}_7$ of C major, the "target scale" used to generate Table 1.[47] We therefore look for that row of Table 1 whose values, when added to 7, are closest to 0 modulo 12**Z**. Inspection shows that this is the fourth row. Adding 7 to these values gives us (12, 12, 12, 12, 12, 13, 12). Subtracting 12 yields (0, 0, 0, 0, 0, 1, 0). This corresponds to the voice leading

$$(9, 11, 0, 2, 4, 5, 7) + (0, 0, 0, 0, 0, 1, 0) \equiv (9, 11, 0, 2, 4, 6, 7) \text{ modulo } 12\mathbf{Z}$$

which is the minimal voice leading between the C major and G major collections. (Here and throughout, we continue to assume the L^1 "smoothness" norm.) Six pitches are held fixed, while F moves up by semitone to F♯.

4.8 Here's another example. Suppose we wish to find a voice leading between the C major and E♭ major collections. E♭ major is $\mathbf{T}_3$ of C major. We look for that row of Table 1 whose values, when added to 3, come closest to 0 modulo 12**Z**. This is the sixth row. Adding 3 to these values yields (11, 11, 12, 12, 11, 12, 12). Subtracting 12 from these gives us (–1, –1, 0, 0, –1, 0, 0). This corresponds to

$$(9, 11, 0, 2, 4, 5, 7) + (-1, -1, 0, 0, -1, 0, 0) \equiv (8, 10, 0, 2, 3, 6, 7) \text{ modulo } 12\mathbf{Z}$$

which is the minimal voice leading between the C major and E♭ major collections.

4.9 The rows of the interscalar path matrix can also be understood as *key signatures for the target scale*.[48] For example, the second row of Table 1, (2, 1, 2, 2, 1, 2, 2), is the key signature for B♯ major. This key signature is enharmonically equivalent to C major, referring to the same pitch classes using different letter names: in C major, pitch class 0 is assigned the letter name C, while in B♯ major it is assigned letter name B♯. Similarly, the third row of

---

[47] $\mathbf{T}_7$, of course, refers to transposition by seven semitones. Note also that the C major collection is both the source and target collection of Table 1.

[48] This fact was first noted by Hook (2004). (It follows from the fact that every path-specific voice leading corresponds to a key signature.) The description of key signatures as voice leadings, the use of interscalar path matrices to represent voice leadings, and the technique of calculating with them are my own.

Table 1, (3, 3, 4, 3, 3, 4, 4), is the key signature for A♯♯♯ major. The remaining rows are key signatures for G♯♯♯♯♯ (or G$♯^5$) major, F$♯^7$ major, E$♯^8$ major, and D$♯^{10}$ major. All of these key signatures preserve the analogical relation between the aural and visual. (Obviously they do not minimize accidentals!)

4.10 Adding 1 to each element in the key signature for *X* major transforms it into a key signature for *X*♯ major; subtracting 1 from each element turns a key signature for *X* major into a key signature for *X*♭ major. For example, the key signature for C major is (0, 0, 0, 0, 0, 0, 0). The key signature for C♯ major is (1, 1, 1, 1, 1, 1, 1). The key signature for C♭ major is (–1, –1, –1, –1, –1, –1, –1). The key signature for C♯♯ major is (2, 2, 2, 2, 2, 2, 2). The key signature for C$♯^{12}$ major—moving each pitch class by exactly one octave—is (12, 12, 12, 12, 12, 12, 12). Since Table 1 contains a key signature assigning any pitch class in the basic scale to any letter name in the range A–G, we can use it to generate *all* the diatonic key signatures preserving the analogical relation between the aural and the visual.

4.11 Suppose now we wish to find a key signature for $\mathbf{T}_{11}$ of the basic scale that minimizes accidental-use. We look for that row of Table 1 whose values, when added to 11, come closest to 0 modulo 12**Z**. This is the second row of the matrix, representing the key signature for B♯ major. Subtracting 1 from (2, 1, 2, 2, 1, 2, 2) gives us the familiar key signature for B major, (1, 0, 1, 1, 0, 1, 1).

4.12 We can continue this process for the remaining diatonic collections. Table 2 shows fourteen of the fifteen standard key signatures. (The “trivial” zero-sharp, zero-flat C major key signature is omitted.) Each key signature in Table 2 adds or subtracts a constant value from the corresponding row of Table 1. Observe that every row of Table 1 gives rise to two key signatures, one involving sharps, the other flats. Adding the trivial key signature to this collection gives us the complete set of key signatures described in music fundamentals textbooks.

Table 2. The fourteen familiar key signatures

Sharp Key Signatures **Flat Key Signatures**

C♯ maj	1	1	1	1	1	1	1	C♭ maj	-1	-1	-1	-1	-1	-1	-1
B maj	1	0	1	1	0	1	1	B♭ maj	0	-1	0	0	-1	0	0
A maj	0	0	1	0	0	1	1	A♭ maj	-1	-1	0	-1	-1	0	0
G maj	0	0	0	0	0	1	0	G♭ maj	-1	-1	-1	-1	-1	0	-1
F♯ maj	1	0	1	1	1	1	1	F maj	0	-1	0	0	0	0	0
E maj	0	0	1	1	0	1	1	E♭ maj	-1	-1	0	0	-1	0	0
D maj	0	0	1	0	0	1	0	D♭ maj	-1	-1	0	-1	-1	0	-1

4.13 With the exception of the first row—which contains "redundant" key signatures enharmonically equivalent to B major and D♭ major—all of the key signatures in Table 2 minimize accidental-use. Equivalently, the path-specific voice leadings shown in Table 2, with the exception of those in the first row, are minimal according to the L^1 ("smoothness") norm. They are complete in the following sense: at least one of the minimal voice leadings between any pair of diatonic collections is contained by Table 2. We conclude, therefore, that our familiar set of key signatures does indeed solve the problem described at the beginning of Section IV.

4.14 I will end this section by noting that the rows of Table 1 share a distinctive feature: each contains just two consecutive integers.[49] This is because the diatonic scale divides the chromatic scale into seven pieces *as evenly as possible.*[50] In general, the more evenly a chord divides the octave, the more the rows of its scalar path matrix will converge on a constant value.[51] Our investigation shows that this convergence will in turn endow the chord with special voice-leading properties: for if the values in a row of a

[49] Consequently, none of the voice leadings shown in Table 2 involve paths longer than one semitone, and none of the standard diatonic key signatures uses double sharps or double flats.

[50] Clough and Douthett 1991.

[51] Indeed, the convergence of the values in the rows of the scalar path matrix can be taken to *define* "evenness." This is precisely what Clough and Douthett (1991) does, in the special case of "maximal" evenness. Note that convergence is measured relative to some measure of voice-leading size. Clough and Douthett characterize "maximal evenness" by measuring convergence according to what mathematicians call the L^∞ metric.

chord's scalar path matrix are clustered around a single value, then there will exist a number *t* that, when added to these values, brings them all close to 0 modulo 12**Z**. As we have seen, this implies that there will be an efficient voice leading between that chord and its transposition by *t* semitones.[52] Thus the unusual evenness of the diatonic collection manifests itself in the unusually efficient voice leadings between its transpositions; the efficiency of these voice leadings in turn manifests itself in the fact that diatonic key signatures use a very small number of accidentals.[53]

V. Acoustic-scale key signatures

5.1 I conclude by showing how these methods can be used to derive key signatures for the "acoustic" (or ascending melodic minor) scale. These key signatures will constitute a complete set of minimal bijective voice leadings between diatonic and acoustic collections. As before, our basic scale is (9, 11, 0, 2, 4, 5, 7). The chromatic scale is the familiar one. The "C acoustic scale," (9, 10, 0, 2, 4, 6, 7) or A–B♭–C–D–E–F♯–G, is known to the classical tradition as the ascending form of G melodic minor.[54]

5.2 The acoustic-scale key signatures that preserve the analogical relation between the aural and visual correspond to crossing-free, bijective voice leadings between diatonic and acoustic collections. These voice leadings can be obtained from the rows of the interscalar path matrix shown in Table 3.

---

[52] See above, §§4.6-4.8. Indeed it can be shown that if *M* is a maximally even chord and if *N* is any other chord of the same cardinality, then the minimal bijective voice leading between *N* and $T_x(N)$ can be no smaller than the minimal bijective voice leading between *M* and $T_x(M)$. A proof of this statement is a matter for another paper; I state the result here merely to pique readers' curiosity.

[53] Interested readers are encouraged to explore this matter by constructing scalar path matrices for the whole tone scale, the harmonic minor scale, and the seven-note chromatic cluster.

[54] Traditionally, of course, the name "C acoustic scale" refers to the ordering C–D–E–F♯–G–A–B♭. Since we are disregarding pitch priority, we are free to rearrange the scale so that its scale-degree numbers correspond to its letter names.

Table 3. An interscalar path matrix
from the basic scale to the acoustic scale

0 steps	**0**	**–1**	**0**	**0**	**0**	**1**	**0**
1 step	**1**	**1**	**2**	**2**	**2**	**2**	**2**
2 steps	**3**	**3**	**4**	**4**	**3**	**4**	**3**
3 steps	**5**	**5**	**6**	**5**	**5**	**5**	**5**
4 steps	**7**	**7**	**7**	**7**	**6**	**7**	**7**
5 steps	**9**	**8**	**9**	**8**	**8**	**9**	**9**
6 steps	**10**	**10**	**10**	**10**	**10**	**11**	**11**

The basic scale is the familiar one; the target scale is (9, 10, 0, 2, 4, 6, 7). Scale degrees in both collections are numbered starting with pitch class 9. As before, row i contains a path-specific voice leading sending scale degree j in the basic scale to scale degree $j + i - 1$ in the C acoustic collection. For example, the first row of Table 3 corresponds to the voice leading

$$(9, 11, 0, 2, 4, 5, 7) + (\mathbf{0}, \mathbf{-1}, \mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{1}, \mathbf{0}) \equiv (9, 10, 0, 2, 4, 6, 7) \text{ modulo } 12\mathbf{Z}$$

which sends scale degree x in the basic scale to scale degree x in the acoustic collection. Similarly, the second row of Table 3 corresponds to the voice leading

$$(9, 11, 0, 2, 4, 5, 7) + (\mathbf{1}, \mathbf{1}, \mathbf{2}, \mathbf{2}, \mathbf{2}, \mathbf{2}, \mathbf{2}) \equiv (10, 0, 2, 4, 6, 7, 9) \text{ modulo } 12\mathbf{Z}$$

which sends scale degree x in the basic scale to scale degree $x + 1$ in the acoustic collection.

5.3 These voice leadings can be described as "interscalar transpositions." To understand why, recall that scalar transposition (including its most familiar form, diatonic transposition), displaces every element in a scale by some constant number of scale steps. The voice leadings shown in Table 3 are similar, displacing each element of the basic scale by a constant number of scale steps, *while changing the underlying scale in the process.*[55] Hence the name "*inter*scalar transposition."

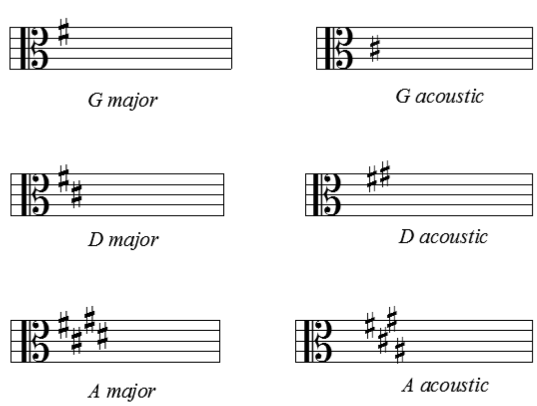
[55] This follows from the definition of the interscalar path matrix: row i of Table 3 sends scale degree j in the basic scale to scale degree $j + (i - 1)$ in the acoustic scale.

5.4 We can also interpret the rows of Table 3 as key signatures. The first row corresponds to the key signature for the C acoustic scale, containing both F♯ and B♭. The second row corresponds to to enharmonically equivalent key signature (1, 1, 2, 2, 2, 2, 2) or (A♯, B♯, C♯♯, D♯♯, E♯♯, F♯♯, G♯♯). The remaining rows provide alternate key signatures for the scale {9, 10, 0, 2, 4, 6, 7}.

5.5 The interscalar path matrix shown in Table 3 is again complete, allowing us to generate every crossing-free voice leading between the two collections. Equivalently, it allows us to generate all the acoustic-scale key signatures that preserve the analogical relationship between the aural and the visual.

5.6 Comparing Table 3 to Table 1 reveals an interesting similarity: the rows of Table 3, with the exception of the first, reorder the corresponding rows of Table 1. For example, the second row of Table 1 is (2, 1, 2, 2, 1, 2, 2), while the second row of Table 3 is (1, 1, 2, 2, 2, 2, 2). Both rows contain five twos and two ones. Analogous statements can be made about rows 3–7. This means that in the familiar system the key signatures for the acoustic scale will closely resemble those for the diatonic scale. With the exception of the first row, they will use the same number and same type of accidentals as their diatonic counterparts. Figure 12 illustrates, comparing a few representative key signatures for diatonic and acoustic scales.

Figure 12. Diatonic and acoustic-scale key signatures compared



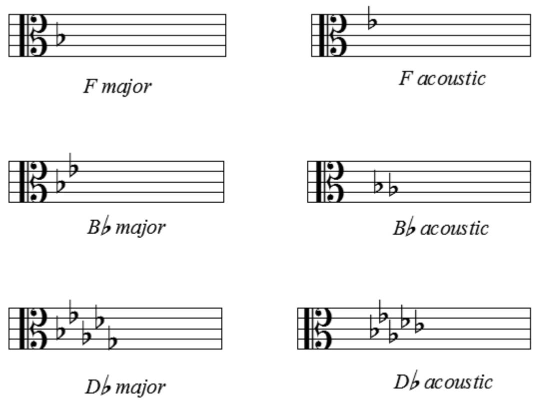


Table 4. Thirteen acoustic-scale key signatures

Sharp Key Signatures								**Flat Key Signatures**							
C ac	0	-1	0	0	0	1	0	C ac	0	-1	0	0	0	1	0
B ac	0	0	1	1	1	1	1	B♭ ac	-1	-1	0	0	0	0	0
A ac	0	0	1	1	0	1	0	A♭ ac	-1	-1	0	0	-1	0	-1
G ac	0	0	1	0	0	0	0	G♭ ac	-1	-1	0	-1	-1	-1	-1
F♯ ac	1	1	0	1	1	1	1	F ac	0	0	-1	0	0	0	0
E ac	1	0	1	0	0	1	1	E♭ ac	0	-1	0	-1	-1	0	0
D ac	0	0	0	0	0	1	1	D♭ ac	-1	-1	-1	-1	-1	0	0

5.7 Table 4 provides the full set of thirteen key signatures for the acoustic scale. In keeping with tradition, I reject those key signatures that apply two or more accidentals to a staff line—in other words, those containing double sharps, double flats, and so forth. Note that the key signature for the C acoustic scale appears on both the sharp side and the flat side, since it contains both a sharp and a flat. The thirteen key signatures shown in Table 4 preserve the analogical relation between the aural and visual. Hence, their associated path-specific voice leadings are crossing-free. The key signatures in Table 4 also all minimize accidental-use. Consequently, their associated voice leadings are minimal according to the L^1 ("smoothness") norm. Finally, the voice leadings form a complete set: Table 4 contains a minimal bijective voice leading from (9, 11, 0, 2, 4, 5, 7) to every acoustic scale.

5.8 Acoustic-scale key signatures, though rare, are not unknown. They appear in Rzewski's *Les Moutons de Panurge*, Bartok's *Mikrokosmos*, and elsewhere. The voice leadings associated with them are considerably more common, appearing throughout the literature—but particularly in the work of Debussy.[56] Just as the acoustic-scale key signatures resemble the diatonic key signatures, so too do these voice leadings resemble the minimal voice leadings between diatonic collections. Just as there is a voice leading that relates two diatonic scales by moving a single pitch up by chromatic scale-step (F→F♯, relating C major to G major), so too is there a minimal voice leading relating a diatonic scale to an acoustic scale that moves a single pitch up by chromatic scale-step (C→C♯, relating C major to G acoustic). As Table 4 shows, there exists a voice leading between diatonic and acoustic collections that lowers one note by one semitone, one that raises two notes by semitone, one that lowers two notes by semitone, and so on.

5.9 Notice again that each of the rows of Table 3 contains only two consecutive integers. This is because the acoustic scale, like the diatonic scale, divides the octave into seven nearly-even parts. The rows of an interscalar path matrix will all converge only to the extent that the matrix is constructed from *two* nearly even collections.[57]

VI. Conclusion

6.1 The preceding discussion has pursued several interrelated goals. Sections I and II provided a rigorous and systematic reconstruction of our familiar system of musical notation, thereby shedding new light on old theoretical concepts. We saw that we can name pitches and pitch classes without postulating a "chromatic scale" or quantizing pitch-class space into discrete units. We saw that familiar pitch-class letter names can be associated with paths in pitch-class space, and that these paths represent an interesting alternative to traditional pitch-class intervals. We distinguished two different conceptions of voice leading, *path-specific* and *path-neutral*. Key signatures correspond to path-specific voice leadings, since they specify not only *which* pitch classes are to be related, but also *how* they are to be related.

[56] See Tymoczko 1997, 2002, forthcoming.

[57] Interested readers can explore the matter by constructing interscalar path matrices linking the basic scale to a variety of septachords.

6.1 The rest of the paper used key signatures to illustrate aspects of the general theory of voice leading. Section III presented three virtues that voice leadings and key signatures can have: avoidance of voice crossings, preservation of common tones, and minimization of the size of the voice-leading. We saw that for a wide range of methods of measuring voice-leading size, it is always possible to avoid voice crossings while minimizing voice-leading size. However, it is not always possible to avoid voice crossings while preserving common tones.

6.3 Section IV solved an interesting and important problem: how does one find a minimal bijective voice leading between arbitrary chords? We saw that this solution is implemented by our familiar system of diatonic key signatures. Section V applied the same methodology to another case, identifying a complete set of minimal voice leadings between diatonic and acoustic scales. These voice leadings are intimately related to minimal voice leadings between diatonic collections. As we saw, this resemblance is due to the fact that both collections divide the octave into seven nearly even pieces.

6.4 Unfortunately, it is not possible here to explore further the general theory of voice leading. However, I hope to have provided some tantalizing hints about what such a theory looks like. With luck, a more thorough treatment of these matters will appear in print shortly.

Dmitri Tymoczko
310 Woolworth Center
Princeton, NJ 08544-1007
http://music.princeton.edu/~dmitri

Return to beginning of article

APPENDIX

Voice crossings and the triangle inequality

A.1 This appendix describes the relation between voice crossings and the triangle inequality. It does not provide rigorous mathematical proofs, but rather offers an informal explanation of the connection between these two concepts. A more rigorous and comprehensive treatment is a matter for another paper.

A.2 For simplicity, consider some path-specific voice leading between two pairs of pitch classes $\{a, b\}$ and $\{c, d\}$.

$$(a, b) + (\boldsymbol{x}_1, \boldsymbol{x}_2) \equiv (c, d) \text{ modulo } 12\mathbf{Z} \qquad \text{(A1)}$$

Ordered pairs of pitch classes can be represented as points on a 2-torus.[58] Equation A1 therefore determines a line segment on the 2-torus. A measure of voice-leading size measures the length of this line segment. Since we are concerned with the *smallest* voice leading (shortest line segment) between $\{a, b\}$ and $\{c, d\}$, we can assume that the path lengths $|\boldsymbol{x}_1|$ and $|\boldsymbol{x}_2|$ are less than or equal to half an octave.[59] This in turn allows us to disregard the periodic structure of the torus, representing it as a portion of the plane $\mathbf{R}^2$.

A.3 Suppose that Equation A1 has a voice crossing. The line segment joining (a, b) to (c, d) will therefore intersect the line $x = y$. Figure A1 illustrates.[60] Point (p, p) is the point of the crossing. If the measure of voice-leading size is invariant under permutation of voices, the line segment $(p, p)\rightarrow(d, c)$ will have the same length as the line segment $(p, p)\rightarrow(c, d)$. The length of the straight path $(a, b)\rightarrow(p, p)\rightarrow(c, d)$ will therefore be the same as the length of the "kinked" path $(a, b)\rightarrow(p, p)\rightarrow(d, c)$. If the measure of voice-leading size obeys the triangle inequality, then there is a direct path from (a, b) to (d, c) that is no longer than the kinked path $(a, b)\rightarrow(p, p)\rightarrow(d, c)$. Therefore we can find a voice leading

---

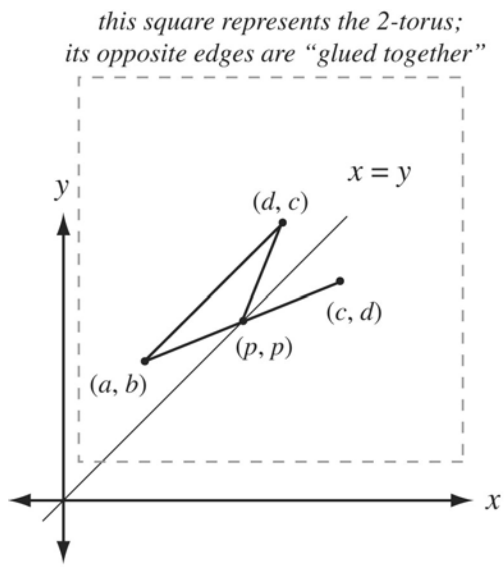
[58] An *n*-torus is a figure whose *n* dimensions are all circles. Thus a 2-torus has two circular dimensions; it can be represented in Euclidean space as the surface of a donut.

[59] This assumption is valid only when our method of voice-leading size is *nondecreasing* in each of its individual path lengths. In other words, one cannot make a voice leading smaller simply by *lengthening* one of its paths.

[60] In drawing the figure we have placed (a, b) "above" the line and (b, c) "below" it. Nothing in the argument rests on this, however.

mapping *a* to *d* and *b* to *c*, no larger than the voice leading shown in Equation A1, and not crossing the line $x = y$.

Figure A1. Pitch-class space, represented as a portion of the plane



A.4 The above argument generalizes straightforwardly to higher-dimensional voice leadings involving more than two voices: we remove voice crossings between pairs of voices by considering the two-dimensional subspaces involving these two voices alone. This procedure never introduces additional crossings into the voice leading; hence we can always reduce the number of crossings to zero without increasing the voice leading's size. Furthermore, this procedure does not change the number of "voices" in the voice leading, modeled here as the dimension of the space. Finally, note that since the argument does not utilize the circular features of pitch-class space, it can be applied to pitch space as well.

A.5 We conclude that, for many measures of voice-leading size, it is always possible to find a minimal bijective voice leading between arbitrary multisets that is crossing-free. The result holds in both pitch and pitch-class space, and requires only that our measure be nondecreasing in each of its path lengths, invariant under permutation of voices, and consistent with the triangle inequality. These conditions are so basic that it is tempting to take them to characterize every "reasonable" measure of voice-leading size.

BIBLIOGRAPHY

Block, Steven and Douthett, Jack. 1994. “Vector Products and Intervallic Weighting.” *Journal of Music Theory* 38(1): 21-41.

Callender, Clifton. 2004. “Continuous Transformations.” *Music Theory Online* 10(3).

———. 2005. “Some Thoughts on Measuring Voice-Leading Distance.” Miami: Paper presented to the 2005 Meeting of Music Theory Southeast.

Clough, John and Douthett, Jack. 1991. “Maximally Even Sets.” *Journal of Music Theory* 35(1/2): 93–173.

Cohn, Richard. 1996. “Maximally Smooth Cycles, Hexatonic Systems, and the Analysis of Late-Romantic Triadic Progressions.” *Music Analysis* 15(1):9-40.

———. 1998. “Square Dances with Cubes.” *Journal of Music Theory* 42(2): 283-296.

Hook, Julian. 2003. “Signature Transformations.” Paper presented at the national meeting of the Society for Music Theory, Madison, Wisconsin.

———. 2004. “Toward a General Theory of Key Signatures and Enharmonic Equivalence.” Paper presented at a sectional meeting of the American Mathematical Society, Evanston, Illinois.

Huron, David. 2001. “Tone and Voice: A Derivation of the Rules of Voice Leading from Perceptual Principles.” *Music Perception* 19(1):1-64.

Lewin, David. 1987. *Generalized Musical Intervals and Transformations.* New Haven: Yale University Press.

———. 1998. “Some Ideas about Voice leading between PCSets.” *Journal of Music Theory* 42(1): 15-72.

Morris, Robert. 1988. “Generalizing Rotational Arrays.” *Journal of Music Theory* 32(1):75-132.

———. 1998. “Voice leading Spaces.” *Music Theory Spectrum* 20(2): 175-208.

Roeder, John. 1984. “A Theory of Voice Leading for Atonal Music.” Ph.D. dissertation, Yale University.

———. 1987. “A Geometric Representation of Pitch-Class Series.” *Perspectives of New Music* 25(1/2): 362-409.

Straus, Joseph N. 2003. “Uniformity, Balance, and Smoothness in Atonal Voice Leading.” *Music Theory Spectrum* 25: 305–52.

Tymoczko, Dmitri. 1997. “The Consecutive-Semitone Constraint on Scalar Structure: A Link Between Impressionism and Jazz.” *Integral* 11: 135-179.

———. 2002. “Stravinsky and the Octatonic: A Reconsideration.” *Music Theory Spectrum* 24(1): 68-102.

———. Forthcoming. “Scale Networks and Debussy.” *Journal of Music Theory*.

Return to beginning of article

prepared by
Brent Yorgason, Managing Editor
Updated 26 October 2005