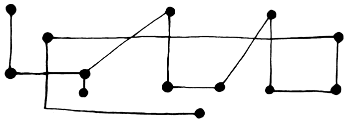
The Structure of Morphological Space



David Kant and Larry Polansky

1. Introduction

The study of contour—important in music theory, cognition, and ethnomusicology—is motivated by an interest in melodic similarity and classification. Contour theory attempts to categorize, clarify, analyze, and define basic melodic principles, as well as, more generally, morphology—musical phenomena quantifiable in some parameter(s) as *change over time*. Contour is fundamental to perception. As such, an understanding of contour relationships, such as distance functions (metrics), in contour space is essential.

A *contour* is an ordered set of directional relationships between (quantifiable) elements, prioritizing "up/down/equal" over "how much up or down." The study of contour has often consisted of categorical classification and a search for archetypes. The up/down motion of things changing in time is one way to understand morphology, or to paraphrase Henry Cowell, the "nature of melody" in terms of a restrictive yet perceptually primal feature. The study of contour relationships (particularly similarity) and categorizations is fundamentally an effort to describe contour space.

This paper presents new mathematical and computational tools to visualize and understand the *structure* of contour space. By integrating degrees of magnitude into contour representations—*ranking* (Marvin and Laprade 1987) and *n-ary contour* (Polansky and Bassein 1992)—we propose a formal unification of *contour space* with *morphological space*. Since contour is an equivalence relation, we first examine contour equivalence relations on the space of morphologies. Next, using *basis coordinate space* (or simply *basis space*), we formally describe how contour archetypes generate, and organize, all possible morphologies, creating a single, highly-structured mathematical space, suggesting a reconsideration of the usual distinctions between contour and morphology.

In this paper we offer a formal description of the *continuum of direction and magnitude*, via consideration of problems and ideas raised by contour theory regarding the structure of *combinatorial contour (CC-) space.*

2 The Study of Musical Contour

2.1 Contour Theories

Some formulations of contour are reductive, typological, and categorical (Adams 1976; Huron 1995; Seeger 1960; Kolinski 1965a, 1965), while others focus on measurable distances and relationships between contours (Polansky 1981, 1987, 1996; Marvin and Laprade 1987; Marvin 1994; Morris 1987, 1993, 2001; Johnson 2001a, 2001b; Quinn 1999). Theorists such as Morris, Marvin and Laprade, and Friedmann (1985, 1987) utilize techniques, terminology, concepts and standard transformational operations from post-tonal music theory (inversion, retrograde, embedding), whereas Polansky and others have focused on generalized mathematical tools.

2.2 Linear and Combinatorial Contour

The distinction between linear contour—adjacent directional relationships—and combinatorial contour—the network of such relationships—is fundamental. *Linear contour* (*LC*) is defined as a vector of adjacent ternary directional relationships in a *morph* (*M*),[1] a finite ordered list of values.[2,3] *Combinatorial contour* (*CC*) is defined by the half-matrix of pairwise relationships in a morph (Morris 1987, 2001; Polansky 1996, 1987, 1981; Quinn 1999; Marvin and Laprade 1987). By convention, we use –1, 0, and 1 to represent directional relationships "is less than," "is equal to," and "is greater than." The CC half-matrix is often expressed in the literature as a vector, by convention: relationships to the first element (top row), relationships to the second element (second row), etc. The morph, LC, and CC all have different lengths. For a morph of length L, the length of the LC vector L_{lc} is $L - 1$, and the length of the CC vector L_{cc} is $(L^2 - L)/2$. Example 1 shows LC and CC vectors for the morph $[3, 4, 5, 1]_m$.

Contour—both LC and CC—is an equivalence relation on all possible morphologies. Many different morphs have the same LC or CC representation, but CC further distinguishes morphs that are equivalent as LCs. CC restricts the range of magnitude variation of represented values more than LC does, and thus is a more "accurate" representation. Since CCs are equivalent to ranked morph elements, greater contour lengths L allow for greater resolution of rank and greater approximation of magnitude.[4]

Below, we describe a geometric representation which includes linear and combinatorial contours in the same space, and consider some ramifications of this unified space.

Morph	$[3, 4, 5, 1]_m$
LC	$[-1, -1, 1]_{lc}$
	[d(a, b) d(b, c) d(c, d)]
CC	$[-1, -1, 1, -1, 1, 1]_{cc}$
	[d(a, b) d(a, c) d(a, d) d(b, c) d(b, d) d(c, d)

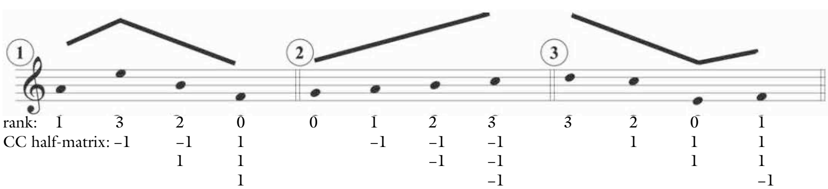
EXAMPLE 1: STANDARD VECTOR NOTATIONS FOR *MORPH* (VALUES), *TERNARY LINEAR*, AND *COMBINATORIAL CONTOUR*

2.3 DISTANCE AND SIMILARITY

Considerable attention has been paid to the measurement of *contour similarity*. Similarity functions, and more powerfully, *metrics*, are a way of understanding the structure—perceptually and mathematically—of contour space. Marvin and Laprade's *CSIM* (contour similarity measure) is based on the COM-matrix, measuring difference by counting the number of corresponding elements that are the same between two COM-matrices. Similarly, Polansky's *Ordered Combinatorial Direction* metric (OCD) (1981, 1987, 1996) allows for a number of similarity measures on morphologies of the *same* length. Morris' contour reduction algorithm (1993, 2001), Marvin and Laprade's *CEMB* (contour embedding function) and Polansky's *Unordered Combinatorial Direction* metric (UCD), as well as various ways of applying the latter's OCD metric, allow for similarity measurement between contours of *different* lengths.[5]

Example 2 shows three, simple, four-element pitch sequences (*morphs*) represented as *rankings* (bold), and CC (half-matrices, below). The LC distances (OLD)[6] and the CC distances (OCD) (both described below) are shown in Example 3.

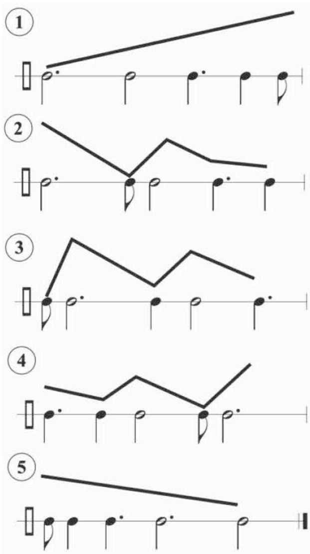
In this paper, we present some new formulations for distance in contour space, incorporating multiple features of contour, including different magnitudinal resolutions.



EXAMPLE 2: THREE SIMPLE FOUR-ELEMENT PITCH SEQUENCES. THE FIRST MELODY (RANKING: [1320]) READS: THE FIRST VALUE IS LESS THAN THE SECOND AND THIRD, AND GREATER THAN THE FOURTH; THE SECOND VALUE IS GREATER THAN THE THIRD AND FOURTH; THE THIRD VALUE IS GREATER THAN THE FOURTH. LCS ARE SHOWN AS THE DIAGONALS OF THE HALF-MATRIX

	(1, 2)	(1, 3)	(2, 3)
OLD	0.33	0.0	0.33
OCD	0.16	0.5	0.16

EXAMPLE 3: COMBINATORIAL (OCD) AND LINEAR (OLD) DISTANCES FOR ALL PAIRS OF THE THREE MORPHS



rank: [43210]
CC: [1111/111/11/1]

rank: [40321]
CC: [1111/-1-1-1/11/1]

rank: [04123]
CC: [-1-1-1-1/111/-1-1/1]

rank: [21304]
CC: [1-11-1/-11-1/1-1/-1]

rank: [01234]
CC: [-1-1-1-1/-1-1-1/-1-1/-1]

EXAMPLE 4: FIVE SIMPLE FOUR-ELEMENT DURATION SEQUENCES (MORPHS), REPRESENTED AS RANKINGS AND CCS

2.4 IMPOSSIBLE CONTOURS

Polansky and Bassein (1992) formally distinguish between possible and impossible CCs en route to determining the number of possible contours for a given length. They (and Morris [2001]) note that many CCs (Morris' COM-matrices) are impossible because of transitivity violation. Impossible contours are thus "holes" in CC-space. As CC length L_{cc} increases, the number of impossible contours grows much faster than the number of possible contours.

Impossible contours may be detected by inspection. Polansky and Bassein's mathematical proof uses a combinatorial argument to show that the number of possible CCs of length L_{cc} is "the number of ways to place L balls in h boxes with no box empty," demonstrating as well the equivalence of previously proposed CC enumerations. They also propose the question: "What is the organic construction of CC-space, avoiding either a generate-and-test approach or ranking-style enumeration?"

In this paper, we describe an endemic mathematical method for the geometric description of (possible) CC-space—including only and all possible CCs.

2.5 LENGTH (L) AND MAGNITUDE (N)

Most previous work has focused on *ternary contour* (Polansky and Bassein 1992): three distinct directional relationships (less than, equal to, greater than).[7] Polansky and Bassein also propose the idea of *n-ary contours*: those that represent any number (> 1) of intermediary magnitudinal values (approaching, eventually, the morph itself). Two properties of CC are thus *length* (L) and *magnitude* (or *resolution*: n).

Contours for morphs of $L = 3$ constitute large equivalence classes, with infinitely many "melodies" in each. As magnitude (n) increases, the set of CCs approaches a one-to-one mapping of the set of morphs. The geometric representation of CC-space described below integrates magnitude and direction into a single continuum. The first step in this approach is the observation that the set of possible CCs is a strictly proper subset of the set of LCs.

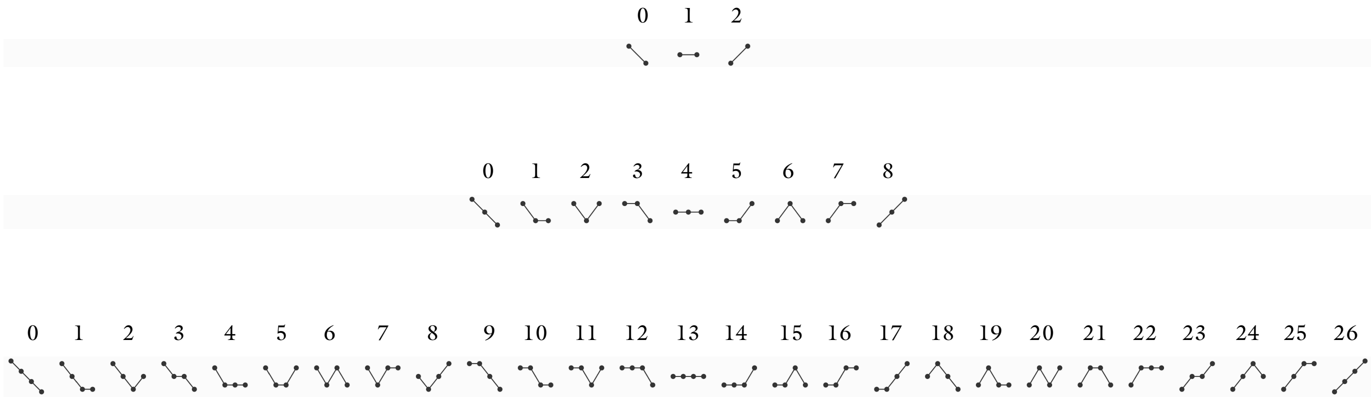
3 Linear Contour Space

3.1 REPRESENTATIONS OF LC-SPACE

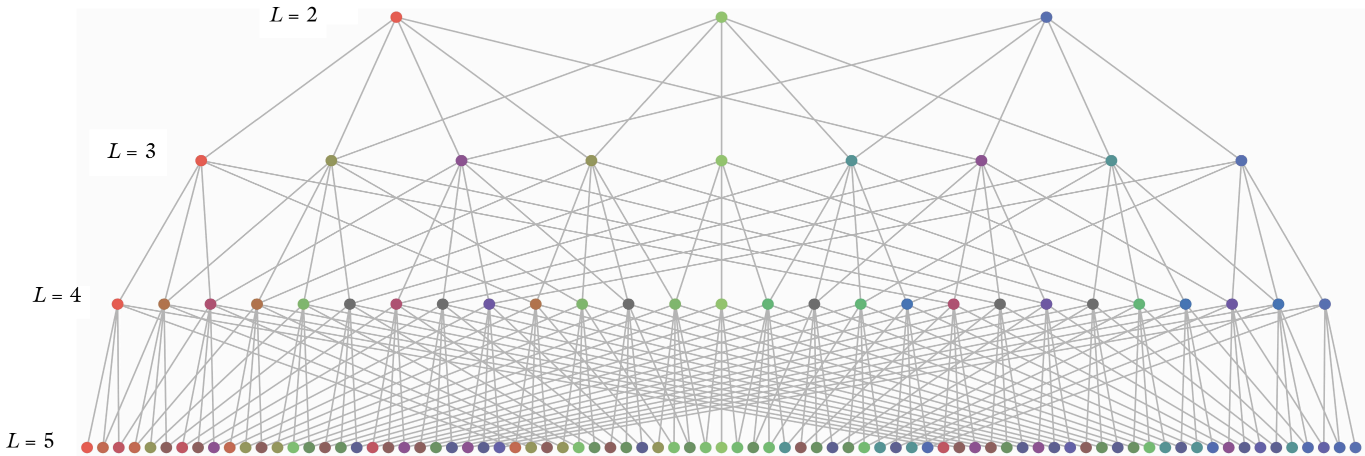
LC-space refers to the set of all linear contours along with various relations, metrics, and geometric representations. As morph length L grows, the number of LCs grows exponentially by 3^{L-1} and the length of LC vectors grows linearly by $L_{lc} = L - 1$. Both the dimensionality (contour length) and the "radius" (number of contours) of the space increase quickly.

3.2 ENUMERATION BY LINEAR INDEX

Enumeration is a common and important tool in the study of contour, usually involving *ranking* of the elements of a *morph*.[8] A simple representation of LC-space is *enumeration by linear index*, counting in a base-3, or ternary number system (Polansky and Bassein 1992). LC vectors are interpreted as base-3 numbers with digits –1, 0, 1 and converted to base-10, assigning to each LC a unique value along a number line from 0 to $(L - 1)^3$. This value is called the *linear index*. Number-lines of different lengths can be normalized and overlaid, representing LCs of any length along a single, common dimension, affording a simple unidimensional metric between them. Example 5 shows LC-space of morph lengths $L = 2, 3, 4$ enumerated by linear index.



Example 5: LC-space enumerated by linear index for $L = 2, 3, 4$;
each row corresponds to a different contour length;
LCs are enumerated along the horizontal axis by linear index;
representative morphs for each LC are drawn



EXAMPLE 6: ENUMERATION BY LINEAR INDEX.
EMBEDDING STRUCTURE OF LC-SPACE FOR L =2, 3, 4, 5.
NODES REPRESENT LCS, AND EDGES BETWEEN NODES INDICATE EMBEDDINGS,
OR WHEN A SHORTER LC APPEARS EXACTLY AS A SUBSTRING IN A LONGER LC.
LCS ARE ENUMERATED ALONG THE HORIZONTAL AXIS BY LINEAR INDEX.
COLOR REPRESENTS THE NUMBER OF "IS GREATER THAN," "IS LESS THAN," "IS EQUAL TO"
TRANSITIONS INTERPRETED AS RGB COLOR VALUES.

While linear index captures some general contour features, it is not a particularly meaningful metric because morphological relationships are distributed across intervals of powers of three—digits, representing different orders of magnitude, flip, from right-most to left-most, every $3^0, 3^1, \ldots, 3^{L-1}$ counts. As a result, LCs that represent intuitively similar contours may not be mathematically similar under linear index distance—flipping a digit to the right or to the left affects linear index by different orders of magnitude. Example 6 illustrates this by showing embeddings, or instances when a lower length LC appears as a substring in a higher length LC.[9]

3.3 GEOMETRIC INTERPRETATION OF LC-SPACE

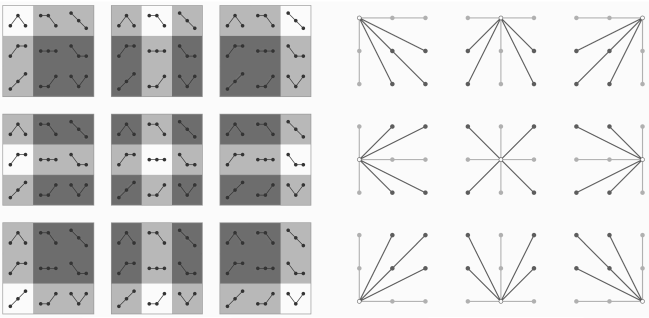
Geometric representation clarifies features of LC-space not shown by enumeration. LCs are represented as points in a hypercube of $L_{lc} = L - 1$ dimensions, bounded by 1 and –1. Axis coordinates correspond to contour values 0, –1, or 1. The nine LCs of $L = 3$, for example, comprise only two values and are easily visualized as points in two-dimensional LC-space—a square bounded by 1 and –1. All non-origin points lie either on axes or diagonals and can be categorized into two groups according to the number of on-axis, or zero elements (see Example 7a). The dimensionality of the hypercube increases with LC length, but axes remain bounded by 1 and –1. For example, each of the 27 LCs of $L = 4$ comprises three values and becomes a point on a three-dimensional cube (similarly bounded by 1 and –1). While the geometric representation is useful for low dimensions, in order to visualize the space, longer LCs require high-dimensional representations for plotting and visualization.

3.4 DISTANCE IN LC-SPACE

Distance functions—measures of contour similarity—are fundamental to understanding the structure of contour space. Under Polansky's Ordered Linear Direction (OLD) metric, which measures how closely two contour vectors align element by element and is the linear contour equivalent of both Marvin and Laprade's and Polansky's combinatorial contour metrics, CSIM and OCD, LC-space is a highly structured, symmetric metric space. We first examine the structure of linear contour, and then refigure combinatorial contour as a sparse and irregular *subset* of linear contour.[10]

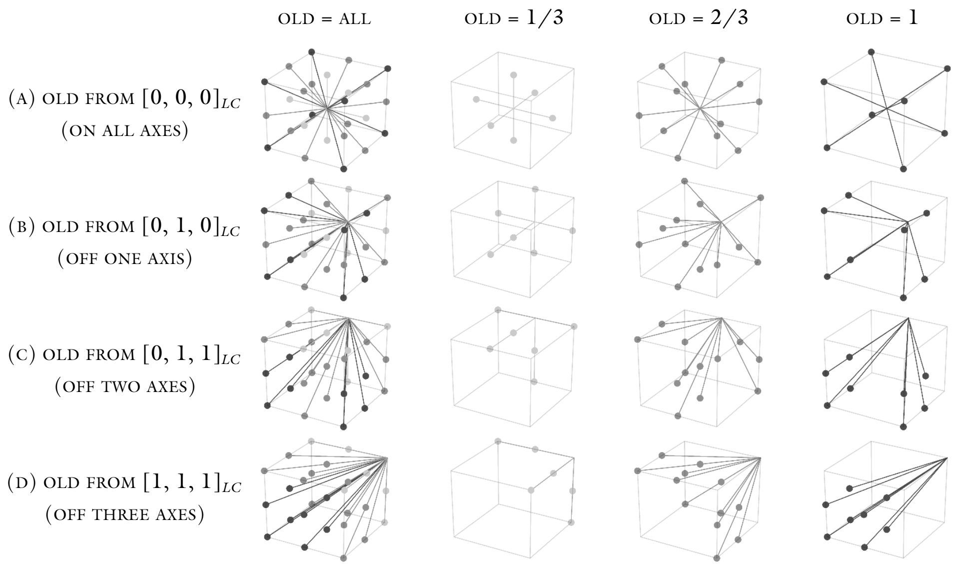
Distance in LC-space can be represented with a partition plot, showing the distances from a given LC (*source*) to every other LC in the space. The OLD between two LCs is the number of LC positions that are different, divided by the LC length ($L - 1$). Interpreted geometrically, OLD counts the number of shared axes between two points, and partitions the space into $L - 1$ subsets that share 0, 1, . . ., $L - 1$ axes with the source LC. Different source LCs partition the space differently, according to OLD distance from that particular source LC. Examples 7 and 8 show OLD partition plots for $L = 3$ and 4, visualizing how distance throughout the space changes depending on which LC is considered the source.

As the source contour changes, the partition subsets rotate around the space. The number of subsets and the number of LCs within each subset remain the same for each source. Subsets are categorized by angle, shown in the flower plots of Examples 7 and 8. When the source LC is the geometric origin $[0, . . ., 0]_{lc}$, partitioning corresponds to the number of non-zero axes in each LC. For example, for $L = 3$, the two nonzero subsets are (1) the four vertices at an OLD distance of 1 and (2) the four axial points at an OLD distance of 0.5. In $L = 4$, this expands to include one additional partition of axial points.



(A) MORPH PLOT (B) FLOWER PLOT

EXAMPLE 7: GEOMETRIC REPRESENTATION OF OLD DISTANCE BETWEEN ALL PAIRS OF CONTOURS IN LC-SPACE L = 3. MORPH PLOT (LEFT) SHOWS REPRESENTATIVE MORPHS FOR EACH LC. FLOWER PLOT (RIGHT) DRAWS LINES BETWEEN EACH PAIR, ILLUSTRATING THE ANGLE BETWEEN THEM. COLOR INDICATES OLD DISTANCE FROM THE SOURCE (WHITE) TO THE EVERY OTHER LC IN THE SPACE (LIGHT OR DARK GRAY). THERE ARE THREE OLD DISTANCE VALUES FOR L = 3 (0, 1/2, AND 1).



Example 8: Geometric representation of OLD distances between all pairs of contours in LC-space for $L = 4$. Flower plots show angles between pairs for visual clarity. Color indicates OLD distance from the source LC, located at intersection of all the lines. There are four OLD distance values (0, 1/3, 2/3, and 1). Only four plots are necessary to show distances from any of the 27 possible LCs because there are four groups, classified according to number of on-axis elements: center; off one axis; off two axes; and off all three axes. OLD distance plots are identical for each group up to rotation. All others are rotations of these four basic forms (or three because the origin is unitary)

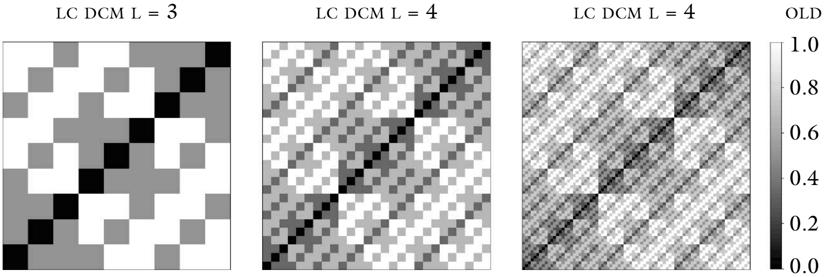
Distance from the geometric origin is called the *normal partition* and is important for two reasons. First, it organizes the space into groups of LCs whose partition plots are equivalent under rotation. As can be seen in Example 7, partition plots for the four vertices of the square are rotations of one another, and partition plots for the four points on the axes are rotations of one another. Distance throughout the entire space can be represented in terms of a smaller number ($L - 1$) of fundamental OLD distance partitions, or *components*.

Second, these components—and consequently all OLD partitions for source LCs other than the origin—are transforms of the normal partition by the transform T_v, composed of translation by the source LC vector v together with a modulus operation that wraps values around the edges of the space. Translation moves every point by the same distance and direction as the source is from the origin, and modulus wraps values from 1 to -1 and vice versa when the magnitude of translation exceeds $abs() = 1$. T_v maps each point from its location relative to the geometric origin $[0, \ldots, 0]_{lc}$ to an analogous location relative to v. This allows us to recast linear contour in terms of the structure of the normal partition together with the structure of the transform, simplifying linear contour and expressing it in terms of its underlying structure and symmetries.

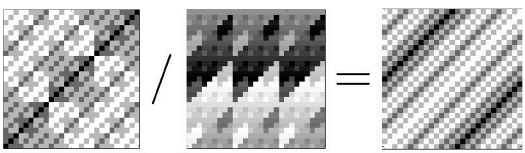
A common problem that arises in the application of contour theory to music composition is to find a contour (or all contours) a specified distance from a given contour. This requires evaluating the distance from the given contour to every other contour (as shown by a partition plot). In the case of linear contour, this is simple because every LC has the same distribution of OLD distances, regardless of its location. This becomes more difficult for combinatorial contour (Section 4).[11]

3.5 Matrix Representations for Higher L

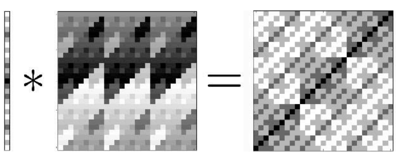
Distance partitions can also be plotted along a linear index, allowing comparisons between OLD spaces for LCs of different lengths. A *dissimilarity matrix* (*DSM*) as shown in Example 9 for $L = 3$, 4, 5 represents OLD distances between all pairwise combinations of LCs on a two-dimensional grid, with LCs enumerated along each axis by linear index. The DSM shows the OLD's self-similar, or fractal, structure, in which increasing LC length increases resolution—note that the DSM of $L = 3$ (leftmost) appears repeatedly throughout DSMs of higher L shifted in location and scaled in color intensity. This structure is the result of OLD distance together with the structure of linear index enumeration, or counting in a base-three system.



EXAMPLE 9: DISSIMILARITY MATRICES LC-SPACE $L = 3, 4, 5$. OLD DISTANCES BETWEEN ALL POSSIBLE PAIRS OF LCS WITH CONTOURS ENUMERATED ALONG EACH AXIS BY LINEAR INDEX. COLOR AT EACH POINT INDICATES OLD DISTANCE BETWEEN THE TWO LCS INDEXED ON EITHER AXIS. EACH COLUMN (OR ROW) REPRESENTS A PARTITION OF LC-SPACE WITH THE SOURCE LC OR ZERO DISTANCE IN BLACK. SINCE ALL PARTITIONS ARE TRANSFORMS OF THE NORMAL PARTITION, EACH COLUMN IS SIMPLY A PERMUTATION OF THE ELEMENTS IN THE OTHERS

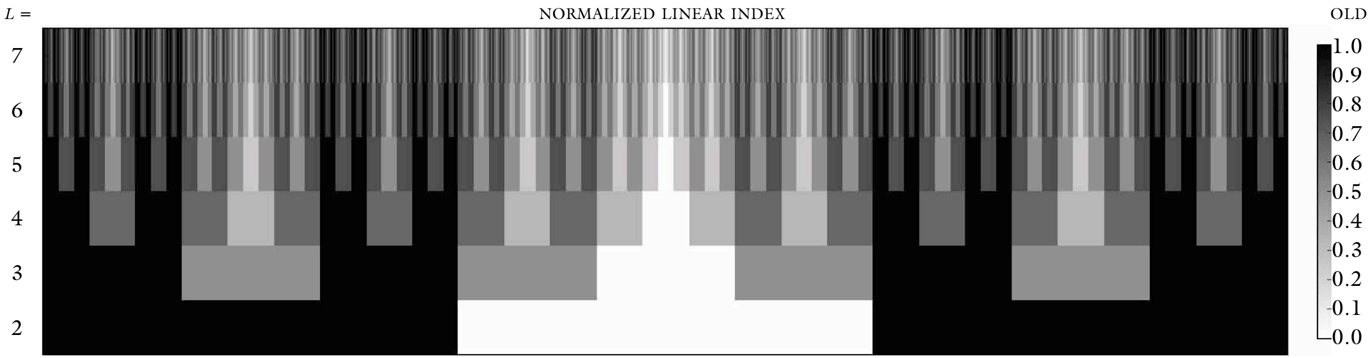


(A) UNSCRAMBLE



(B) CONVOLUTION

EXAMPLE 10: USE OF THE TRANSFORM T AS REPRESENTED BY A TRANSFORMATION MATRIX TO (A) MAP EVERY PARTITION BACK TO THE NORMAL PARTITION AND (B) REPRESENT LC-SPACE ($L = 4$) AS THE CONVOLUTION OF THE NORMAL PARTITION TAKEN OVER ALL SOURCES. T IS REPRESENTED BY A MATRIX (MIDDLE MATRIX IN EACH SUBFIGURE) IN WHICH AXES CORRESPOND TO LINEAR INDEX AND COLOR AT EACH ENTRY INDICATES THE NORMALIZED LINEAR INDEX TO WHICH ONE LC IS MAPPED UNDER T OF THE OTHER



EXAMPLE 11: NORMAL PARTITION (OLD TO THE ORIGIN [0, 0, . . .,0]LC) FOR LCS OF LENGTH *L* = 2 THROUGH *L* = 7. CONTOURS ARE ENUMERATED ALONG NORMALIZED LINEAR INDEX, ALLOWING COMPARISON BETWEEN PARTITIONS OF DIFFERENT LENGTH LCS, WHICH SHOWS THE SELF-SIMILARITY OF THE OLD AS CONTOUR LENGTH *L* INCREASES. AS THE OLD GRAIN INCREASES, SO DOES THE FREQUENCY, OR THE RESOLUTION, OF THE RECURSION. PARTITIONS OF OLD FOR SOURCE LCS OTHER THAN THE ORIGIN ARE PERMUTATIONS OF THIS STRUCTURE BY THE TRANSFORM *T*.

The transform T_v can also be represented by a *transformation matrix* whose value at each entry gives the normalized linear index to which each LC (indexed by column) is mapped under T_v (where v is the LC indexed by row). The transformation matrix unscrambles the DSM, showing that all partitions are permutations of the normal partition (Example 10a). The transformation matrix can also be used to represent LC-space as the convolution of the normal partition transformed over all possible source points (Example 10b). The normal partition, or distance to the geometric origin $[0, \ldots, 0]_{lc}$, is transformed by T_v, which maps points from the normal partition to a new origin v, and summed for all LCs. The structure of LC-space is primarily that of the normal partition, shown in Example 11 for increasing LCs of increasing lengths L. While this may seem trivial for linear contour, it becomes significant for combinatorial contour, due to "holes" in CC-space.

4 Combinatorial Contour Space

Combinatorial contour has become the most common representation used in contour theory. CC offers a greater degree of distinction—two morphs may have the same LC representation but different CC representations—and is thus a more accurate representation of a morph than LC. CC-space, consisting of all vectors that represent possible CCs, is complicated by "holes," or vacant points that represent impossible contours. CC-space is *not* a subspace of LC-space because the operations of addition and multiplication almost entirely result in points which represent impossible CCs. In this section we reconsider CCs as a *subset* of LCs, a step towards a new understanding of the structure of CC-space.

4.1 Combinatorial contour as a subset of linear contour

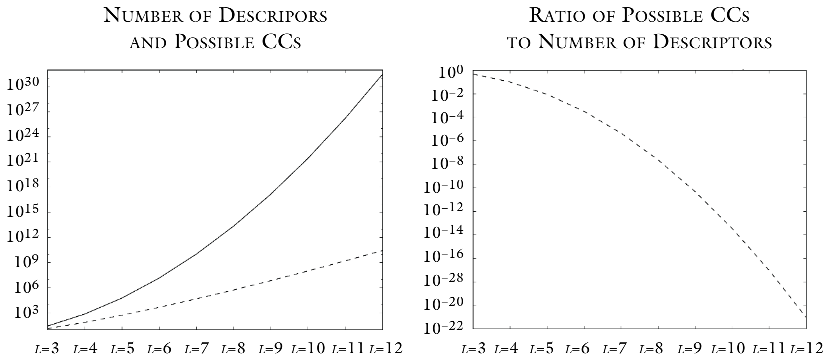
Like LCs, CCs are conventionally written as unidimensional ternary-valued vectors. Any CC vector, whether possible or impossible, also describes an LC of length L_{cc}. All CC descriptors are also LC descriptors. In other words any CC vector describes *two different equivalence classes* of two different lengths of morphs because CC and LC vectors of the same length describe morphs of different lengths. For example, the CC $[-1, -1, 1, -1, 1, 1]_{cc}$ for morph length $L = 4$ also describes an LC for morph length $L = 7$. The two identical contour descriptions are equivalence classes of morphs of different lengths. CCs are a *subset* of LCs.

LCs can assume any length (> 1), but CCs may *only* assume lengths equal to binomial coefficients of the form $\binom{L}{2}$: 3, 6, 10, 15, 21, 28, 36, ... There are *no* CCs of L_{cc} = 2, 4, 5, 7, 8, 9, 11, 12, 13, 14, ... Further, because of "impossibility" (intransitivity) (Polansky and Bassein 1992), there are relatively few CCs for a given L_{cc}—*few vectors describe possible CCs.* However, all LC vectors are possible for any L_{lc}, so holes in CC-space are occupied in LC-space by LC descriptors. For example, there are 27 ternary CC descriptors for $L = 3$, only 13 of which are possible Ccs.

As morph length L increases, the number of contour descriptors grows more quickly than the number of possible CCs, shrinking the number of possible CCs to a decreasingly small fraction of LC-space. Extending Polansky and Bassein (1992), Example 12 lists the ratio of the number of contour descriptors to the number of possible CCs as morph length increases up to $L = 12$. These values are also plotted in Example 13 on a logarithmic scale, showing how quickly the ratio decreases. Even for morphs of length $L = 12$, the ratio of possible CCs is astronomically small.

L	L_{CC}	Num Contour Vectors (LCs)	Num CCs	Ratio
2	1	3	3	1
3	3	27	13	0.481
4	6	729	75	0.103
5	10	59,049	541	9.162e–03
6	15	14,348,907	4,682	3.263e–04
7	21	1.046e+10	47,293	4.521e–06
8	28	2.288e+13	545,835	2.386e–08
9	36	1.501e+17	7,087,261	4.722e–11
10	45	2.954e+21	1.022e+08	3.461e–14
11	55	1.744e+26	1.623e+09	9.301e–18
12	66	3.090e+31	2.809e+10	9.090e–22

EXAMPLE 12: TABLE SHOWING NUMBER OF LCS AND CCS, UP TO 12, EXTENDING POLANSKY AND BASSEIN (1992)

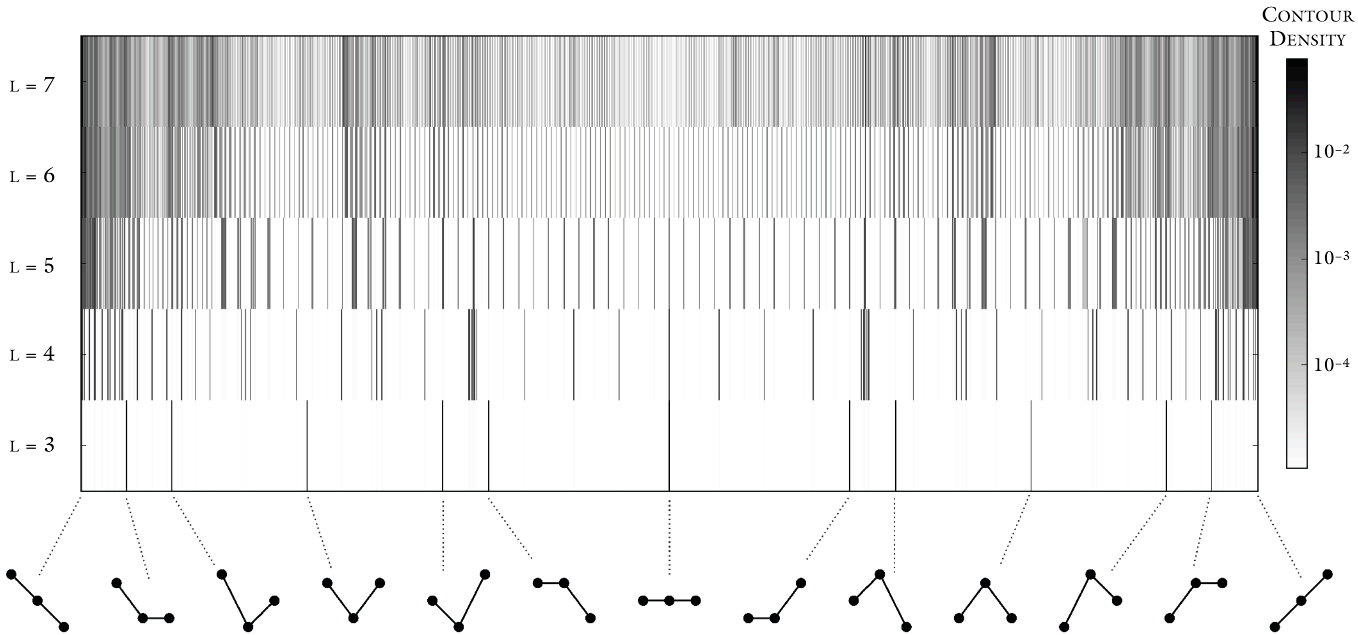


EXAMPLE 13: CURVES OF THE NUMBER OF LCS (VECTORS) AND NUMBER OF CCS

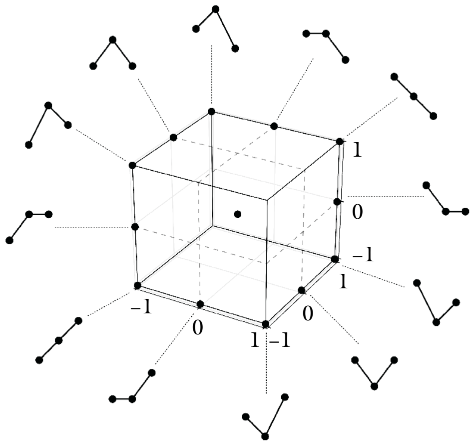
4.2 VISUALIZATIONS OF IMPOSSIBLE CCS IN CC-SPACE

The structure of CC-space—where the holes are located and why—remains an open question. Example 14 plots the locations of possible CCs within the space of CC descriptors as enumerated by linear index for increasing morph lengths *L*. The lowest row, for example, shows the locations of the 13 ternary CCs of length $L = 3$.

This structure can also be represented geometrically by extending the geometric representation of linear contour from Section 3.3 to combinatorial contour. CCs are plotted as points in a hypercube of $L_{cc} = (L^2 - L)/2$ dimensions, bounded by 1 and –1. The locations of the 13 possible CCs of morph length $L = 3$ are illustrated in Example 15, forming an easily visible symmetric subset. Non-marked points at intersections of grid lines are the impossible CCs. The space of possible CCs is thus irregular—points have neighbors in different locations and at different distances away. Under this representation, those holes would be occupied by LCs if we considered this to also be a plot of LC-space.



EXAMPLE 14: ENUMERATION ALONG THE NUMBER LINE FOR CCS $L = 3$ THROUGH 7, ILLUSTRATING HOLES IN CC-SPACE. COLOR INDICATES THE DENSITY OF CONTOURS AT EACH POINT. COUNTS ARE NORMALIZED SEPARATELY FOR EACH ROW

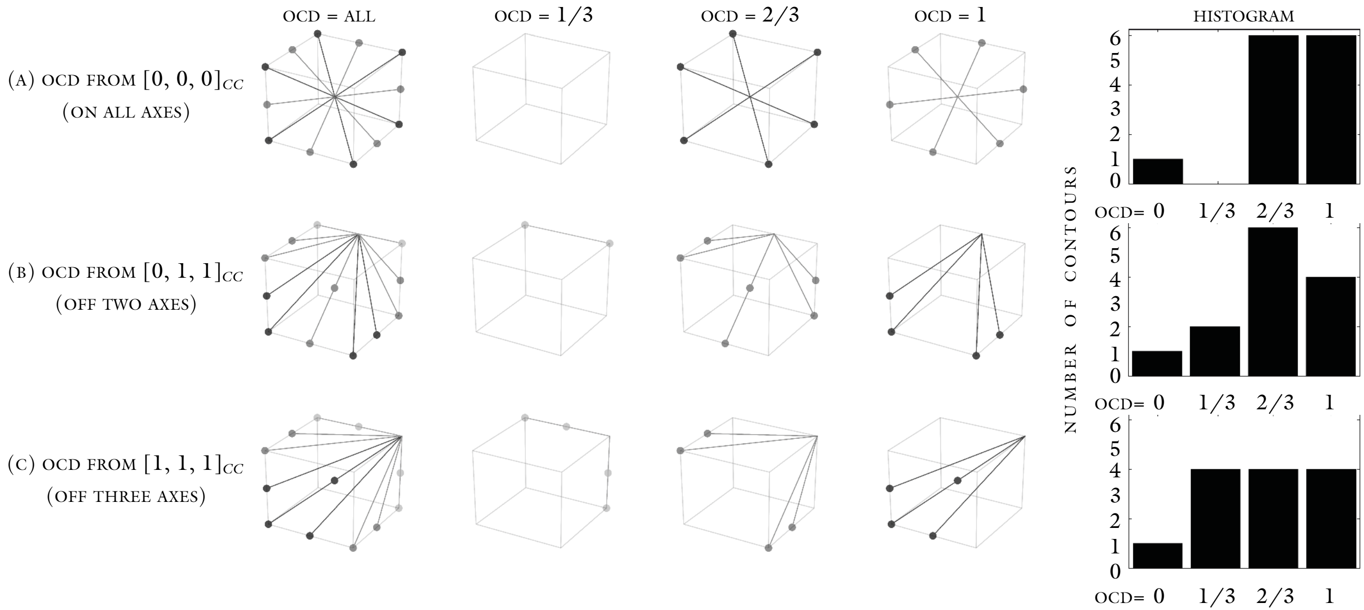


EXAMPLE 15: GEOMETRIC INTERPRETATION OF THIRTEEN TERNARY CCS $L = 3$ PLOTTED ON THE CUBE, SHOWING HOLES IN THE SPACE (IMPOSSIBLE CCS). MORPH LENGTH $L = 3$, $L_{CC} = 3$. $L_{LC} = 3$ AS WELL, BUT IN THIS CASE REPRESENTING MORPHS OF $L = 4$. THE ORIGIN $[0, 0, 0]_{CC}$ IS IN THE CENTER OF THE CUBE. POSSIBLE CCS ARE MARKED WITH DOTS. HOLES (IMPOSSIBLE CCS) ARE UNMARKED INTERSECTIONS OF GRID LINES

4.3 DISTANCE IN CC-SPACE

The OLD is trivially extended to measure distance in CC-space: the number of equal entries in the two CC vectors divided by the CC vector length (Polansky's OCD [1981, 1987, 1996] or Marvin and Laprade's CSIM [1987]). Unlike LC-space under OLD, CC-space is not highly structured and symmetric under OCD due to holes (impossible CCs).

Example 16 shows partition plots for CC-space $L = 3$, each plot representing distance from a given source CC to every other CC in the space. CC-space, however, is not as easily translated as LC-space, as shown by the histograms, which count the total number of CCs at each OCD distance. CCs do not fully populate points in LC-space, so changes in orientation—with different origins—affect distance distributions, and consequently the shape of the space. Importantly, unlike LC-space, CC-space components (and partitions) are *not* permutations of one another (nor of a single normal partition) because CC-space is not closed under the transform T_v. T_v maps some possible CCs to impossible CCs, and some impossible CCs to possible CCs. This makes navigation of CC-space, whether compositionally or analytically, complicated, because moving by a given distance from a source CC usually arrives at an impossible CC.



EXAMPLE 16: GEOMETRIC REPRESENTATION OF OCD DISTANCE IN CC-SPACE, $L = 3$. COLOR INDICATES OCD DISTANCE FROM THE SOURCE CC, LOCATED AT INTERSECTION OF ALL THE LINES. SINCE CCS DO NOT FULLY POPULATE THE SPACE, THERE ARE ONLY 3 CC COMPONENTS (PARTITIONS THAT ARE UNIQUE UP TO ROTATION) WHEREAS LC-SPACE OF THE SAME LENGTH HAS 4 COMPONENTS (EXAMPLE 8). IT IS IMPOSSIBLE TO CONSTRUCT A CC IN $L = 3$ THAT HAS ONLY ONE NON-ZERO ELEMENT, SO THERE IS NO CC COMPONENT FOR OFF ONE AXIS. HISTOGRAMS FOR EACH COMPONENT COUNT THE NUMBER OF CONTOURS AT EACH OCD DISTANCE. MOVING FROM ONE SOURCE TO ANOTHER AFFECTS THE HISTOGRAM DISTRIBUTION AND SHAPE OF THE SPACE

4.4 MATRIX REPRESENTATIONS OF CC-SPACE

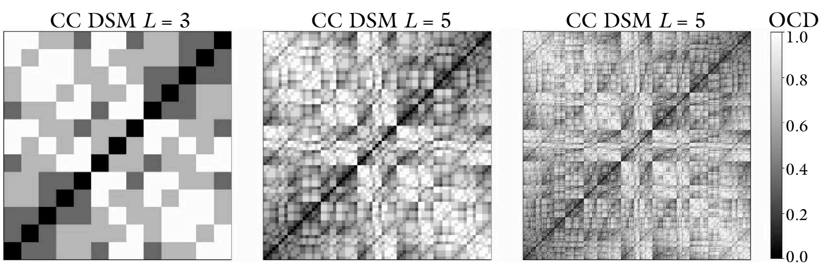
A dissimilarity matrix (DSM) can be used to visualize OCD distance for CCs of lengths greater than $L = 3$. As with linear contour, combinatorial contour DSMs represent OCD distances between all pairwise combinations of CCs on a 2-dimensional grid. Example 17 shows DSMs for CC-space $L = 3$, 4, 5, revealing that the self-similar structure observed with linear contour is less clear for combinatorial contour. The linear contour DSMs appeared as shifted and scaled versions of one another. The self-similar structure of combinatorial contour is more complex.

The reason for this structure can be illustrated by conceiving of CC-space DSM as the result of removing the impossible CCs from the LC-space DSM and squeezing the remaining possible CCs back together. A CC *impossibility mask* is a binary matrix used to represent the locations of possible CCs by masking out impossible CCs from the space of contour descriptors (LC-space), as shown in Example 18.

To summarize:

1. The set of possible CCs of length L_{cc} is a strictly proper subset of LCs whose length $L_{lc} = L_{cc}$. The CC and LC representations, however, correspond to morphs of different lengths, or Ls.

2. Distances (Hamming) between CCs are the same as those between LCs of the same lengths.

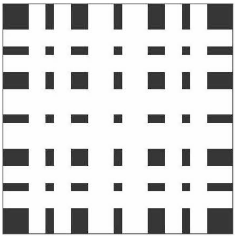


EXAMPLE 17: DISSIMILARITY MATRICES CC-SPACE L = 3, 4, 5. CONTOURS ARE ENUMERATED ALONG EACH AXIS BY COMBINATORIAL INDEX, AND COLOR AT EACH POINT INDICATES OCD DISTANCE BETWEEN THE TWO CCS INDEXED ON EITHER AXIS. IN LC-SPACE, EACH PARTITION (COLUMN) IS A PERMUTATION OF ANY OTHER. THIS IS NOT THE CASE FOR CC-SPACE

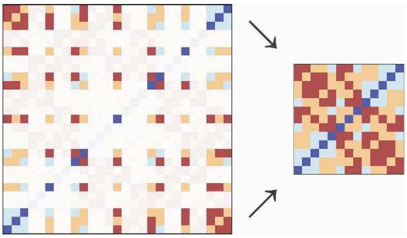
3. The structure of CC-space is inherited from LC-space but with an impossibility mask, creating holes. These holes are determined by 1) allowable length as L varies and 2) allowable transitivity (possibility) for a given length L.

4. Very few LC-spaces contain possible CCs because CCs only exist for lengths equal to binomial coefficients of the form $\binom{L}{2}$. There are many more LC-spaces than CC-spaces.

The Structure of CC-Space is the Result of a CC Impossibility Mask Applied to LC-Space of the Same Contour Vector Length



EXAMPLE 18A: THE CC IMPOSSIBILITY MASK (L = 3), DISTINGUISHING POSSIBLE (BLACK) FROM IMPOSSIBLE (WHITE) CCS



EXAMPLE 18B: CC MASK APPLIED TO THE LC-SPACE DSM (L = 4) TO PRODUCE THE CC-SPACE DSM (L = 3)BY FILTERING OUT IMPOSSIBLE CONTOURS

Normal Form	Contour	CC-Index	LC-Index	M-Index
$[2, 1, 0]_m$	$[1, 1, 1]_{cc}$	0	0	21
$[1, 1, 0]_m$	$[1, 1, 0]_{cc}$	1	1	9
$[2, 0, 1]_m$	$[1, 1, -1]_{cc}$	2	2	19
$[1, 0, 1]_m$	$[1, 0, -1]_{cc}$	3	5	10
$[1, 0, 2]_m$	$[1, -1, -1]_{cc}$	4	8	11
$[1, 1, 0]_m$	$[0, 1, 1]_{cc}$	5	9	12
$[0, 0, 0]_m$	$[0, 0, 0]_{cc}$	6	13	0
$[0, 0, 1]_m$	$[0, -1, -1]_{cc}$	7	17	1
$[1, 2, 0]_m$	$[-1, 1, 1]_{cc}$	8	18	15
$[0, 1, 0]_m$	$[-1, 0, 1]_{cc}$	9	23	3
$[0, 2, 1]_m$	$[-1, -1, 1]_{cc}$	10	24	7
$[0, 1, 1]_m$	$[-1, -1, 0]_{cc}$	11	25	4
$[0, 1, 2]_m$	$[-1, -1, -1]_{cc}$	12	26	5

EXAMPLE 19: INDEXED REPRESENTATIONS OF POSSIBLE CCS: *LC-INDEX*, *CC-INDEX*, *M-INDEX*. LC-INDEX INTERPRETS THE CC VECTOR AS A TERNARY NUMBER WITH ELEMENTS $(1, 0, -1)$; CC-INDEX INTERPRETS THE CC VECTOR AS A TERNARY NUMBER AND SORTS THEM FROM 0 TO 12; M-INDEX INTERPRETS THE MORPH NORMAL FORM VECTOR AS A TERNARY NUMBER WITH ELEMENTS $(2, 1, 0)$. FOR EXAMPLE, $[2, 1, 0]_M$ IS INTERPRETED AS $(2 \times 3^2) + (1 \times 3)$

5 The Structure of CC-Space: Magnitude and Contour

How may CC-space be visualized on its own rather than as a "swiss-cheesy" subset of linear contour, and how do conventional ideas of distance operate in that space? Because L_{cc} grows much faster than L, visualization in Euclidean coordinates quickly becomes unwieldy. For example, for morph length $L = 5$, LC length $L_{lc} = 4$ and CC length $L_{cc} = 10$. A 10-dimensional space is required to represent the 541 possible CCs by ternary descriptor vectors. More interestingly, those 541 *possible* vectors are embedded in a space of 59,049 ternary contour descriptors (58,508 or 99% of which are impossible).

5.1 ENUMERATION

Enumeration of morph elements in *normal form* (Marvin and Laprade 1987), or *ranking*, creates a simple ordering of possible CCs. Example 19 lists the thirteen possible CCs of $L = L_{cc} = 3$ in normal form notation.

Normal form rankings of ternary CCs (whose possible values range from 0 to $L - 1$) contain no empty ranks. $[1, 3, 5]_m$ and $[2, 3, 21]_m$ both become $[-1, -1, -1]_{cc}$ and $[0, 1, 2]_r$ (r denotes normal form). Rankings [0, 2, 5] and [1, 2, 3] are not in normal form—the first skips a rank, the second is not normalized. Normal form enumeration lists *only and all possible CCs* as morphs (that is, vectors which describe both direction and magnitude). All normal form rankings of length L describe possible CCs of length L_{cc}.[12]

5.2 FULL RANK AND *N*-ARY CONTOUR

An important further distinction between possible CCs is between those that *do not* compress ranking values, called *full rank*, and those that *do* compress ranking values. In the former (full rank), the magnitudes of all inter-element differences (or *deltas*) between pairs of morph normal form rankings are described *without compression*.

For example, consider the CC $[-1, -1, -1]_{cc}$ (the first is less than the second, the first is less than the third, the second is less than the third —ascending line) with morph normal form $[0, 1, 2]_m$. If CC elements are interpreted to express *magnitude* as well as direction (in this case: −1, +1, and 0), then the CC does not accurately express all the different inter-element deltas in the normal form morph.[13] The difference between the first and third normal form morph elements—rankings 0 and 2—is two ranks, not one, but a value of 2 is unavailable in the ternary representation. The CC *is not* full rank. The CC $[-1, 0, 1]_{cc}$ however, with morph normal form $[0, 1, 0]_m$, *is* full rank because all inter-element deltas accurately represent the magnitude of rank differences as well as direction (in this case, no inter-element deltas are greater than 1).

CCs that are not full rank (in ternary) can be represented as full rank if higher degrees of contour magnitude are used, in what is called *n-ary contour* (Polansky and Bassein 1992; Morris[14]). For example, $[-1, -1, -1]_{cc}$ can be represented without compressing rank values in *quinary* form, using absolute values of contour element magnitudes greater than 1: $[-1, -2, -1]_{cc}$. This quinary representation is full rank, with large enough deltas—sufficient resolution—to accurately describe inter-element magnitudinal ranking differences. In other words, the higher the possible CC values (n, or magnitude) the greater the resolution of the contour descriptor.

Formally, we define the *resolution* of a contour, denoted by n or a descriptive term (ternary, quinary, etc.), as the number of possible inter-element values used to represent the morph as a contour. In this paper, we assume all contours are symmetric—an equal number of positive and negative difference values in addition to equality—thus n is the absolute value of the largest element (either positive or negative) in the contour vector. The number of difference values is $2n + 1$. For example, *quinary* ($n = 2$) has $2 \times 2 + 1$ possible relationships (−2, −1, 0, 1, 2). In this formulation, contours have both *length* (L) and *resolution* (n).[15,16] For ternary contours (using only values −1, 0, and 1) $n = 1$. Intuitively, ternary contour is the simplest representation/reduction of magnitude, compressing all inter-element differences to a single degree of difference.[17]

Rank, a property of morphs, has been used to mean the number of unique rank values in morph normal form, or the number of unique values in a morph, which is equal to the maximum rank value in normal form, assuming ternary contour ($n = 1$). Extending rank to n-ary contour, *n-ary rank* is the maximum rank value in *n-ary normal form* (also called *n-ary ranking*), where n-ary normal form is an extension of normal form to allow skips of magnitude n.

The concept of full rank is an extension of Polansky and Bassein's (1992) impossibility from ternary contour ($n = 1$) to magnitude, or n-ary contour ($n > 1$). Possible n-ary contours must satisfy not only the transitive property of inequality but the additive property as well. Formally, a contour is *full rank* if the additive property holds for the system of distance relationships between any three morph elements.[18] Full rank is a property of contours, not of morphs. Full rank indicates whether a contour has sufficient magnitudinal delta to describe the morph normal form without compressing two different rank magnitude deltas into the same magnitudinal category.

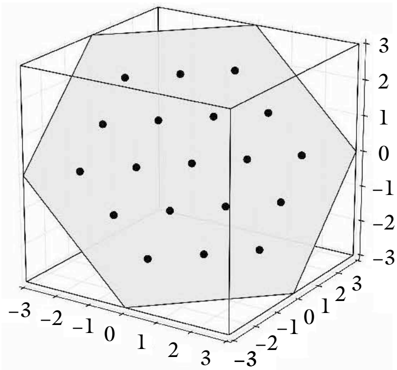
Contours may be full rank (or not) for any n. Of the thirteen ternary CCs for $L = 3$, seven are full rank, six are not (the difference of one is accounted for by the all-equal "straight line" contour, trivially of full rank). The contours that are not full rank in ternary, however, have full rank representations in higher (more resolute) n-ary contours. Example 20 lists the thirteen ternary contours as well as their full rank representations. Contours highlighted in gray are *not* full rank in ternary, requiring magnitudes of $n = 2$, or quinary contour to be full rank. Contours not highlighted are full rank in ternary ($n = 1$).

Normal Form	CC (n = 1)	CC (n = 2)	Rank
$[2, 1, 0]_m$	$[1, 1, 1]_{cc}$	$[1, 2, 1]_{cc}$	2
$[1, 1, 0]_m$	$[1, 1, 0]_{cc}$		1
$[2, 0, 1]_m$	$[1, 1, -1]_{cc}$	$[2, 1, -1]_{cc}$	2
$[1, 0, 1]_m$	$[1, 0, -1]_{cc}$		1
$[1, 0, 2]_m$	$[1, -1, -1]_{cc}$	$[1, -1, -2]_{cc}$	2
$[1, 1, 0]_m$	$[0, 1, 1]_{cc}$		1
$[0, 0, 0]_m$	$[0, 0, 0]_{cc}$		0
$[0, 0, 1]_m$	$[0, -1, -1]_{cc}$		1
$[1, 2, 0]_m$	$[-1, 1, 1]_{cc}$	$[-1, 1, 2]_{cc}$	2
$[0, 1, 0]_m$	$[-1, 0, 1]_{cc}$		1
$[0, 2, 1]_m$	$[-1, -1, 1]_{cc}$	$[-2, -1, 1]_{cc}$	2
$[0, 1, 1]_m$	$[-1, -1, 0]_{cc}$		1
$[0, 1, 2]_m$	$[-1, -1, -1]_{cc}$	$[-1, -2, -1]_{cc}$	2

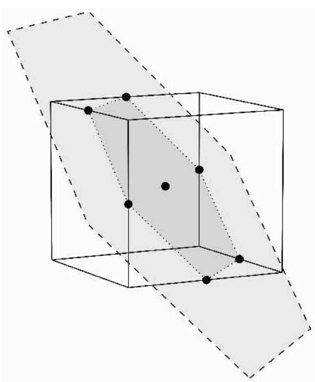
EXAMPLE 20: *FULL RANK* AND *NON-FULL RANK* TERNARY CONTOURS. NOTE THAT $N = 2$ IS THE MINIMUM NECESSARY TO REPRESENT FULL RANK TERNARY CONTOURS OF $L = 3$. N′, THE MINIMUM NEEDED FOR L-LENGTH CONTOURS OF MAGNITUDE N, IS A FUNCTION OF N AND L: $N' = (L-1)(N)$. FOR N-ARY CONTOURS OF $L = 3$, $N = 1$ (TERNARY), $N' = 2$ (QUINARY). SIMILARLY, FOR QUINARY CONTOURS, $L = 3$, $N = 2$, $N' = 4$ (NONARY), ETC.

Perceptually, n-ary contour provides increasing degrees of magnitude resolution, representing shades of perceptual relationships—for example, a *quinary* contour might represent “a lot less than,” “a little less than,” “equal to,” “a little greater than,” “a lot greater than.” Formally, magnitude (up to ranking) is implicit in the system of directional relationships of a ternary CC. Ternary CCs do not fully express the information present in the representation, and this truncation of values to *abs*() ≤ 1 is the cause of holes in ternary CC-space and the reason why CCs do not form a closed subspace of LC-space. By using full rank n-ary CC representations, CCs do form a closed subspace of LC-space. The distinction between full rank and non-full rank contours is an important step in what follows, the formulation of a “well-behaved” structure of morphological space emanating solely from contour.

GEOMETRIC ILLUSTRATION OF THE FULL RANK PLANE
REDUCED TO TERNARY CONTOUR

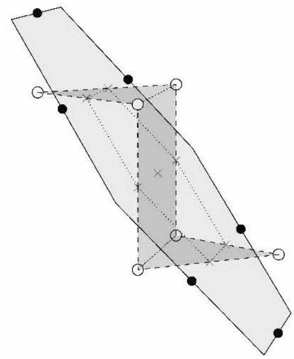


EXAMPLE 21A: FULL RANK PLANE IN *N*-ARY CONTOUR SPACE (PLOTTED UP TO $N = 3$): THE FULL RANK CCS LIE ON A PLANE ROTATED DIAGONALLY IN *N*-ARY CONTOUR SPACE

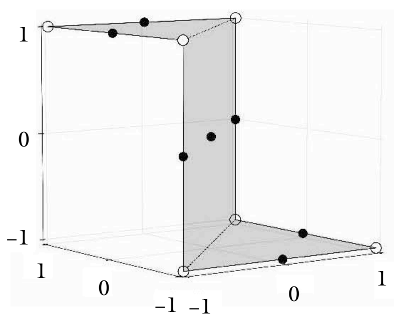


EXAMPLE 21B: INTERSECTION OF THE TERNARY CUBE AND THE FULL RANK PLANE: CCS THAT ARE FULL RANK IN TERNARY LIE AT THE INTERSECTION THE PLANE AND CUBE (PLOTTED IN BLACK)

Geometric Illustration of the Full Rank Plane Reduced to Ternary Contour (cont.)



EXAMPLE 21C: CCS THAT ARE NOT FULL RANK IN TERNARY ARE MAPPED ONTO THE TERNARY CUBE BY RANK REDUCTION. IN $L = 3$, THESE CCS (PLOTTED IN BLACK) BECOME THE AXES OF THE TERNARY CUBE (PLOTTED IN WHITE)



EXAMPLE 21D: RANK REDUCTION FROM FULL RANK TO TERNARY CONTOUR, VISUALIZED AS A FOLDING OF THE CC PLANE TO FIT INTO THE TERNARY CUBE

5.3 THE GEOMETRY OF FULL RANK

The concept of full rank is fundamental to the structure of CC-space, illustrated by plotting full rank contours in *n-ary combinatorial contour space* (CC_n-space), an extensions of CC-space to magnitudes greater than 1. n-ary CCs are represented as points in a hypercube of L_{cc} dimensions bounded by n and $-n$. For CCs where $L_{cc} = L = 3$, n-ary contour space is represented as a cube with axes extending from $-n$ to $+n$. Example 21 plots full rank CCs in n-ary CC-space $L = 3$.

Full rank contours of morph length $L = 3$ lie on a plane P that is rotated diagonally within the cube of n-ary contour space (Example 21a). This plane forms a true subspace of n-ary contour space because it is closed under addition and multiplication. The plane describes only and all possible full rank CCs. Because ternary CCs do not contain elements of *abs*() > 1, they are restricted to the unit cube bounded by ±1. Full rank ternary CCs are located at the intersection of the full rank plane and the unit cube, lying on both the unit cube and the full rank plane (Example 21b).

CCs that are not full rank in ternary form ($n = 1$) have full rank representations for $n > 1$. Although the ternary forms do not lie on the full rank plane, the corresponding full rank representations (of greater resolution) do. Full rank CCs are mapped to ternary CCs by a rank-reduction transform $R_{n\rightarrow 1}$ which maps from n-ary contour space to ternary contour space by limiting the contour elements to values of magnitude 1, or *abs*() ≤ 1 (Example 21c).

Ternary contour can be thought of as the result of a rank reduction transform $R_{n\rightarrow 1}$ applied to the full rank plane: $R_{n\rightarrow 1}(P)$ = CC_1-space. In other words, the set of full rank CCs are a kind of ur-set of contours, from which all other, lower rank n contours can be derived —in this case, the 13 possible ternary CCs. The full rank plane represents all ternary CCs, but represents some in their higher resolution full rank form, while others, already full rank, are represented directly. The 13 possible ternary CCs are in some sense the simplest forms (excluding the three LCs of length $L_{lc=2}$), lying at the center of the plane of all possible CCs.

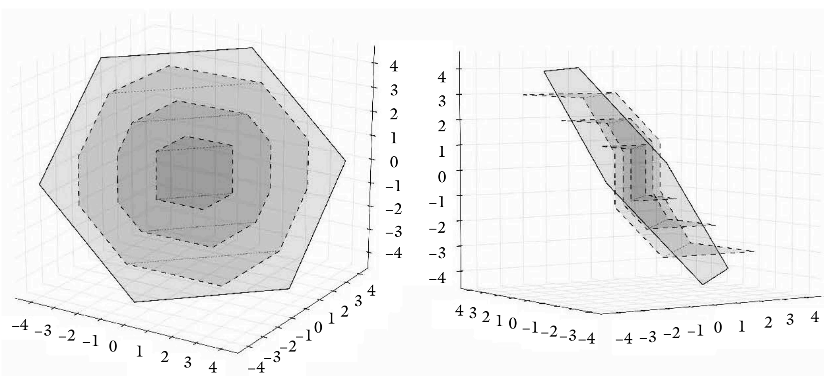
Geometrically, this transform of the full rank plane P into ternary contour space is a folding and stretching of the plane to fit into the unit cube (Example 21d). CCs of rank $n > 1$ are pushed onto to the unit cube by limiting contour elements to *abs*() ≤ 1. The resulting surface still sits diagonally in CC-space. This transform of the diagonal plane accounts for the structure and location of the possible CCs in CC-space.

The transform R, illustrated for $L = 3$ in Example 22, can be generalized to any higher value of n, creating a continuum from lower resolution n-ary contours in which ternary is the most compressed to full rank, which is not compressed at all. All CCs, for any L or n exist on this continuum. For example, as n increases, the monotonically ascending ternary CC $[-1, -1, -1]_{cc}$ eventually becomes every possible version of a monotonically ascending line, all of which are folds from higher n into ternary, or $n = 3$.

For CCs of morph length $L > 3$, full rank CCs form a lower dimensional subspace (of dimension $L - 1$) embedded within n-ary contour space (of dimension L_{cc}). For $L = 3$, the full rank subspace is a plane in 3-dimensional n-ary contour space. For morph length $L = 4$, full rank contours form a 3-dimensional subspace of a 6-dimensional n-ary contour space. As with $L = 3$, for $L > 3$ the subspace is a vector space that consists entirely and only of full rank CCs.

All possible full rank CCs span an entire lower dimensional vector subspace in the space of all CC descriptors. The set of all possible CCs of any L and any n is the same as the set of all full rank CCs for any L.

Because the subspace is a vector space, we can construct a new coordinate system for it, enabling us to work directly in the space of full rank contours and investigate the structure of the space comprising all and only possible CCs. In the next section, we construct a new space, *combinatorial basis space* (henceforth: *basis space*), making use of this change of coordinate systems from n-ary contour space to the subspace of full rank CCs.



EXAMPLE 22: RANK REDUCTION FROM FULL RANK TO SUCCESSIVE DEGREES OF N-ARY CONTOUR (PLOTTED FROM TWO DIFFERENT PERSPECTIVES)

6 Basis Space

Basis space[19] is a representation CC-space that comprises all and only full rank CCs. Through a change of coordinate systems, full rank CCs are represented by a set of basis vectors which combine arithmetically to form all possible full rank CCs. This representation separates possible from impossible CCs, describing possible CCs in minimal dimensionality and providing a visualization of contour relationships.

6.1 Change of Basis

Basis space provides a set of basis vectors $B = \{b_1, b_2, \ldots, b_{L-1}\}$ that can uniquely express any full rank CC as the weighted sum of basis vectors with scalars $a_1, a_2, \ldots, a_{L-1}$. These vectors form the axes of the basis-space coordinate system, in which CCs are represented by basis vector scalars instead of by the (n-ary) elements of CC vectors themselves. CC vectors are reconstructed by recombining basis vectors as weighted sums according to scalars:

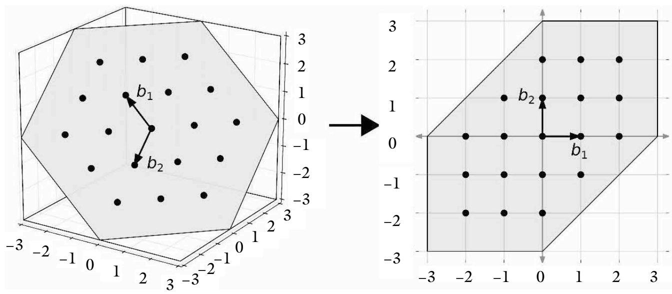
$$[a_1, a_2, \ldots, a_{L-1}]_b = a_1 \cdot b_1 + a_2 \cdot b_2 + \ldots + a_{L-1} \cdot b_{L-1}$$

Basis space significantly reduces dimensionality, requiring only $L - 1$ basis vectors. CCs are represented by $L - 1$ scalars rather than L_{cc} CC elements. More importantly: *basis space represents every full rank CC, and all contours represented in basis space are full rank CCs.* CCs that are not full rank cannot be expressed as weighted sums of basis vectors and as such have no representation in basis space. Basis space eliminates non-full rank CCs.[20]

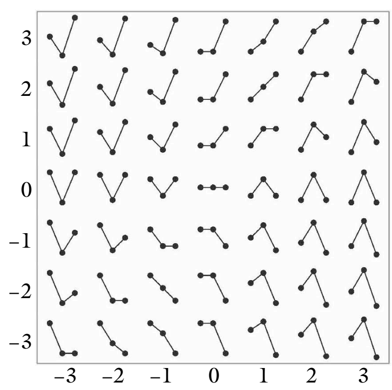
In $L = 3$, basis space is the full rank plane. For $L > 3$, basis space is the $L - 1$ dimensional subspace of full rank CCs. Interpreted geometrically, basis vectors are the *axes* of these full rank subspaces. Any full rank CC in the subspace can be found by movement along basis vector axes, and any movement along an axis will remain in the subspace. The scalars of the weighted sums correspond to the direction and magnitude of movement along basis vectors in the subspace.[21] Example 23 illustrates the change of basis from CC-space to basis space for $L = 3$. Example 24 plots basis space $L = 3$ with representative morphs drawn, showing the correspondence between basis-space coordinates and morph features. Note that larger morphs are further from the origin, and morphs gradually change shape as location rotates around the origin. (The method of construction for basis vectors is described below).

ILLUSTRATION 1: Conversion from basis space $[1, 1]_b$ to ternary *cc* $[1, 1, 0]_{cc}$:

$$\begin{aligned}[1,1]_b &= 1 \cdot b_1 + 1 \cdot b_2 \\ &= 1 \cdot [-1,0,1] + 1 \cdot [0,-1,-1] \\ &= [-1,-1,0]_{cc}\end{aligned}$$



EXAMPLE 23: CHANGE OF BASIS FROM CC-SPACE TO BASIS SPACE IN L = 3. WITHIN CC-SPACE, THE FULL RANK CCS FORM A DIAGONAL PLANE. BASIS SPACE PROVIDES TWO VECTORS B1, B2 THAT ARE CAPABLE OF EXPRESSING ANY CC OF LCC = 3 AS A WEIGHTED SUM. INTERPRETED GEOMETRICALLY, ANY FULL RANK CC CAN BE FOUND BY MOVEMENT ALONG THESE TWO VECTORS AND ANY MOVEMENT ALONG THESE TWO VECTORS WILL REMAIN ON THE PLANE (LEFT). THESE VECTORS FORM THE AXES OF A NEW COORDINATE SYSTEM, BASIS SPACE, IN WHICH FULL RANK CCS ARE REPRESENTED BY COORDINATES CORRESPONDING TO THE DIRECTION AND MAGNITUDE OF MOVEMENT ALONG BASIS VECTORS (RIGHT).



EXAMPLE 24: BASIS SPACE $L = 3$

6.2 CONSTRUCTING BASIS VECTORS

Basis vectors encode the logic of combinatorial contour—changing one value in a CC affects the system of relations between all elements. Basis vectors encode this system of transitivity by definition, since the space they represent *cannot contain intransitive CCs.* Each vector describes how each member of the system of combinatorial relations is affected when one morph value is changed. For example, raising the second element of a length $L = 3$ morph changes the delta between first and seconds elements as well as the delta between the second and third elements, but does not affect the delta between the first and third elements. This system is encoded by the basis vector $b_1 = [-1, 0, 1]$.

While basis space requires only $L - 1$ basis vectors, the basis vectors are themselves CCs of length L_{cc}. These contours are, in some sense, "primitives" or fundamental CCs of the space that combine in various ways to construct all possible full rank CCs. The basis-space axes $[1, 0]_b$ and $[0, 1]_b$ represent the CC vectors themselves: $[-1, 0, 1]_{cc}$ and $[0, -1, -1]_{cc}$.

Basis vectors may be formed from the CCs of morph-space axes [1, 0, . . ., 0], [0, 1, . . ., 0], [0, 0, . . ., 1] which correspond to these primitive morphs—morphs in which only one element is raised at a time. As such, basis space can be thought of as the space of first order combinatorial differences on morph space. While morph space also separates possible from impossible CCs (unlike CC-space), basis space transposes this structure of morph space into combinatorial contours that reveal the structure of possible CCs within the space of all CC contour descriptors (the latter not included in basis space). Example 25 shows the construction of basis vectors for any L.

The choice of basis vectors is not unique. Basis space requires $L - 1$ vectors (the minimal number required to span the entire space), which can be derived from any subset of $L - 1$ morph-space axes. Different choices of vectors reorient basis space through rotation. By convention, we use morph-space primitives that correspond to raising the second, third, fourth, . . . elements, discarding the first.

Intuitively, the dimension of basis space is one less than that of morph space because basis space represents the first order signed difference of morph space. Only the differences between morph elements are retained, not the absolute offset. As such, basis space is isomorphic to morph space under transpositional invariance, or the space of all "zeroed" morphs,[22] and is a generalization of it through the choice of basis vectors.

L	Basis Vectors $[b_1, \ldots, b_{L-1}]$
$L = 3$	$b_1 = [1, 0, -1]$
	$b_2 = [0, 1, 1]$
$L = 4$	$b_1 = [1, 0, 0, -1, -1, 0]$
	$b_3 = [0, 1, 0, 1, 0, -1]$
	$b_4 = [0, 0, 1, 0, 1, 1]$
$L = 5$	$b_1 = [1, 0, 0, 0, -1, -1, -1, 0, 0, 0]$
	$b_3 = [0, 1, 0, 0, 1, 0, 0, -1, -1, 0]$
	$b_4 = [0, 0, 1, 0, 0, 1, 0, 1, 0, -1]$
	$b_5 = [0, 0, 0, 1, 0, 0, 1, 0, 1]$
L	$b_n = [a_{i,j}$ for i in $(1, \ldots, L)$ and j in $(i, \ldots, L)]$
	where $a_{i,j} = \begin{cases} 1 & \text{if } i = n+1\,; \\ -1 & \text{if } j = n+1\,; \\ 0 & \text{otherwise}\,. \end{cases}$

EXAMPLE 25: CONSTRUCTION OF BASIS-SPACE VECTORS

6.3 TERNARY CONTOURS IN BASIS SPACE

As shown above, some ternary contours are full rank, others are not. Ternary contours that are not full rank are represented through their higher (quinary), full rank counterparts. Finding these lower rank representations, such as ternary, requires an additional rank-reduction transform to the desired resolution $R_{n\to 1}$. Converting from lower rank contours to basis space requires first integrating to a full rank contour before solving for basis coordinates.

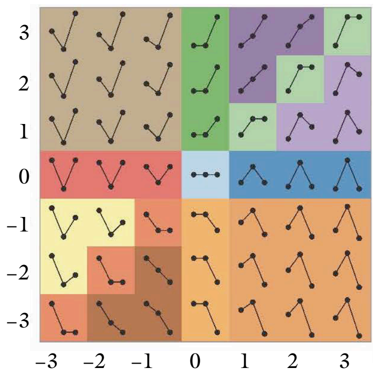
ILLUSTRATION 2: Conversion from basis space $[2, 1]_b$ to ternary CC $[-1, -1, 1]_{cc}$:

$$\begin{aligned}
[2,1]_b &= R_{n\to 1}(2\cdot b_1 + 1\cdot b_2) \\
&= R_{n\to 1}(2\cdot[-1, 0, 1] + 1\cdot[0, -1, -1]) \\
&= R_{n\to 1}([-2, -1, 1]_{cc}) \\
&= [-1, -1, 1]_{cc}
\end{aligned}$$

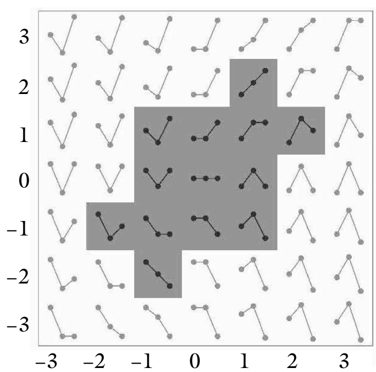
ILLUSTRATION 3: Conversion from ternary CC $[1, 1, 1]_{cc}$ to basis space $[-1, -2]_b$: First integrate $[1, 1, 1]_{cc}$ to $[1, 2, 1]_{cc}$ then solve for basis coordinates a_1, a_2:

$$\begin{aligned}
[1, 2, 1]_{cc} &= a_1\cdot b_1 + a_2\cdot b_2 \\
&= a_1\cdot[-1, 0, 1] + a_2\cdot[0, -1, -1] \\
&= [-1, -2]_b
\end{aligned}$$

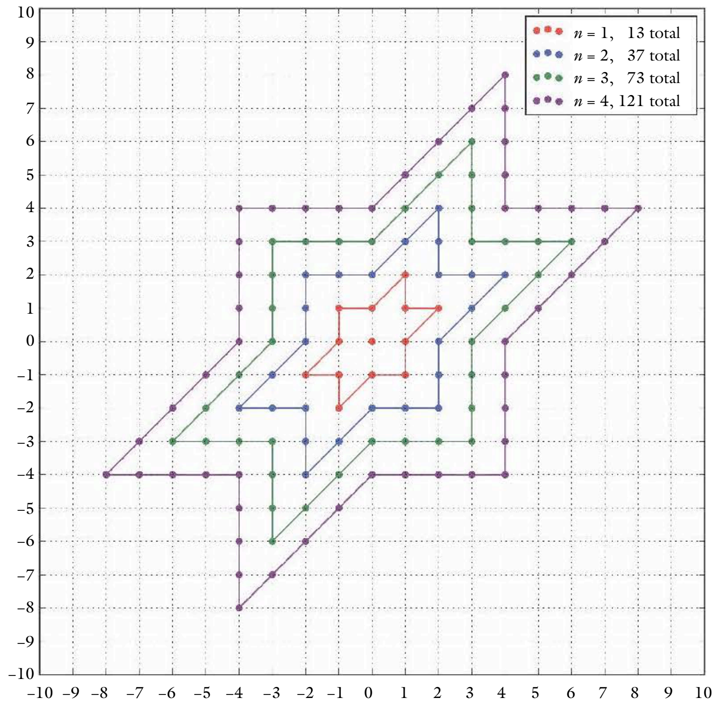
Basis Space $L = 3$



EXAMPLE 26A: TERNARY CC REGIONS IN BASIS SPACE.
CONTOURS WITHIN THE SAME COLORED REGION ARE EQUIVALENT UNDER TERNARY CONTOUR—THEY HAVE THE SAME TERNARY CC REPRESENTATION.



EXAMPLE 26B: TERNARY NORMAL FORM CCS HIGHLIGHTED IN GRAY.
THE CHART SHOWS THAT THE SIX NONFULL RANK TERNARY CCS (IN NORMAL FORM): $[2, 1, 0]_M$, $[2, 0, 1]_M$, $[1, 0, 2]_M$, AND INVERSIONS $[1, 2, 0]_M$, $[0, 2, 1]_M$, $[0, 1, 2]_M$, EXTENDING FROM THE CUBE'S VERTICES, ARE ALL ADJACENT TO SOME FULLRANK TERNARY CC. NON-FULL RANK CCS ARE CHARACTERIZED BY HAVING THREE DISTINCT NORMAL FORM VALUES. FULL-RANK CCS HAVE TWO EQUAL VALUES.
THIS DISTINCTION, BY EXTENSION, HOLDS FOR ANY N



EXAMPLE 27: STELLATED FORM IN BASIS SPACE, $L = 3$, INCREASING N. SINCE EACH N INCLUDES ALL CONTOURS OF LOWER N, COLOR INDICATES BOUNDARIES BETWEEN SUCCESSIVELY INCREASING INTEGER VALUES OF N

6.4 THE STRUCTURE OF BASIS SPACE

Basis space is unbounded, extending infinitely in all dimensions, and comprises all full rank CCs. Rank reduction defines equivalence classes on the space of all full rank CCs, subdividing it into a finite number of equivalence classes, each of which represent an infinite set of CCs. Visualizing these equivalence relations in basis space—that is, operating on full rank CC representations rather than truncated ternary representations—allows us to investigate the structure that exists before reduction.

Example 26a shows basis space divided into 13 regions, or equivalence classes, of ternary contour. Boundaries between equivalence classes are organized radially, characterized by angle about the center.

As with morphs, normal form[23] is used to represent contour equivalence classes, and basis space uses a corresponding definition of normal form—smallest integer scalars, but allowing negative values as well. The basis normal forms are the members of each equivalence class that are closest to the origin. In Example 26b, the 13 normal form ternary CCs are highlighted in gray. As scalars increase, *n*-ary contours up to any length (and any value for *n*) can be plotted in basis space (Example 27). The concept and visualization of *n*-ary contour is thus an immediate result of representation in basis space, revealing CC-space's fundamental structure as *n* increases (for some *L*). In other words, while contour equivalence class corresponds to angle, the degree of resolution—*n*-ary contour—is represented by radii lengths (conventional Euclidean distance from the origin).

6.5 STELLATED FORM AND THE NUMBER OF *N*-ARY CCS

The number of possible *n*-ary contours can be expressed by the formula:

$$\sum_{h=1}^{L} n^{h-1}\, h!\, S(L, h)$$

where L is morph length, n is contour resolution, and $S(L, h)$ is a *Stirling number of the second kind*. Inserting the n^{h-1} term generalizes, to any n, Polansky and Bassein's (1992) formula for the number of possible ternary contours for some L.[24] Intuitively, the general form of the summation shows that each $(n - 1)$-ary contour is cumulatively included in the number of contours for n (with the same L). As an example, for $n = 2$ (quinary) and $L = 3$ there are 37 contours, 13 of which are the ternary contours (with rank 1 or 2). Example 28 lists the number of n-ary contours for increasing values of L and n. (See Appendix A for an enumeration of n-ary CCs, $L = 3$).

Basis space also raises the question of cardinality: how many contours are contained in a given equivalence class? The generalized answer to this question helps clarify the significance and/or likelihood of observed contours with respect to the set of possible contours. Typological categorizations of observed melodies such as those of Huron (1995) are enriched by knowing the expected-against-observed distribution. Cardinality can be visualized as the area of the colored regions in Example 26a.

	$n = 1$	$n = 2$	$n = 3$	$n = 4$
$L = 3$	13	37	73	121
$L = 4$	75	365	1,015	2, 169
$L = 5$	541	4,501	17,641	48,601
$L = 6$	4,682	66,604	367,926	1,306,808

EXAMPLE 28: NUMBER OF *N*-ARY CONTOURS FOR INCREASING *L* AND *N*. NUMBER OF CONTOURS INCREASES EXPONENTIALLY WITH *N*, BECAUSE INCREASING *N* ADDS HIGHER ORDER TERMS TO THE SUMMATION, AND BY BINOMIAL COEFFICIENTS WITH *L*

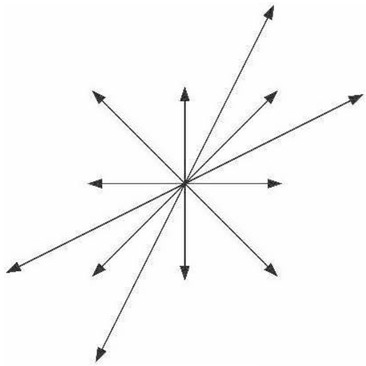
7 The Dimensions of Contour: Angle and Magnitude

Contours in basis space are organized radially, characterized by angle and magnitude from the origin (0, 0). Representing contours as vectors from the origin, the *shape* of the contour corresponds to the vector's angle (Example 29a) and the *rank* of the contour corresponds to vector's magnitude, or distance from the origin (Example 29b).[25]

Note that basis space appears squashed in quadrants where x and y have opposite signs because it is a *transposition space*. Length $L = 3$ morphs are represented in $L - 1$ dimensions by transposing, or zeroing, such that one morph element is 0, effectively representing a third orthogonal dimension. This corresponds to the stipulation that morph normal form be transposed to have zero as a minimum. For $L = 3$, basis coordinates with opposite signs combine to represent the third morph element. This form extends to and becomes more complex in higher dimensions.

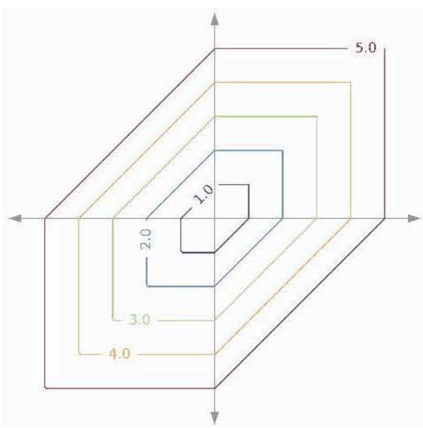
Because the dimensions of contour correspond to angle and magnitude, polar coordinates are a more direct parameterization. Example 30 shows basis space for morph length $L = 3$ plotted on a polar coordinate system. Contours are represented with coordinates (r, θ) where θ is the angle measured from the normal vector [1, 0] to basis-coordinates, and r is the contour rank or magnitude of the largest contour element. By normalizing magnitude to $1/r$ we invert and bound the space to the radius $r = 1$, placing the six rank-1 contours (for $L = 3$, the smallest possible CCs) on the outer boundary. In this representation, contours increase in size, or rank, towards the origin.

The Features of Contour Shape and Size Correspond to the Dimensions of Angle and Magnitude in Basis Space



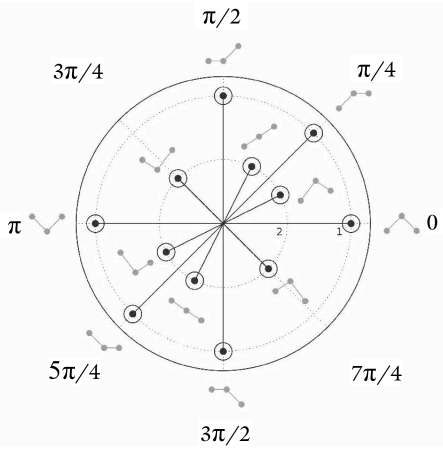
ANGLE

EXAMPLE 29A: CONTOURS AS VECTORS FROM THE ORIGIN

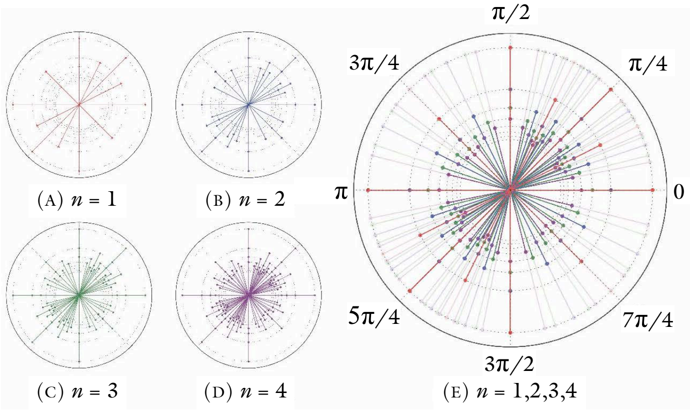


MAGNITUDE

EXAMPLE 29B: CONCENTRIC REGIONS OF INCREASING RANK, 1–5



EXAMPLE 30: BASIS SPACE $L = 3$ REPRESENTED IN POLAR COORDINATES



EXAMPLE 31: *N*-ARY CONTOUR $L = 3$ REPRESENTED IN POLAR COORDINATES. NOTE THAT THE SYMMETRY VISUALIZED EARLIER IN THE STELLATED FORM, AND ALSO IN EXAMPLE 20 AND EXAMPLE 24, IS CLEARLY SEEN IN THIS REPRESENTATION

In basis space, as n increases (for some fixed L), the space expands outward from the center. Polar coordinates map that expansion into the circle, which becomes more densely populated as n increases. Consequently, the number of unique angular differences between CCs increases, as does the resolution or the grain of the distances, because increasing n adds new possible integer values for basis-space coordinates. Converting from basis to polar coordinates, angle measures the ratio between basis coordinates, as expressed by the arctangent function. For example, if $L = 3$, $n = 1$, the only possible ratios are 1:0, 1:1, 2:1 (and their reciprocals). If $n = 2$, with possible integer rank values $\{1, \ldots, 5\}$, angular distances are generated by all possible ratios of those integers that satisfy n-ary normal form (and their reciprocals). Increasing n thus increases the number of possible fractional values (angular distances), eventually including all the rationals (and arbitrarily small distances). Example 31 plots n-ary contour morph length $L = 3$ for increasing values of n.

7.1 DISTANCE IN BASIS SPACE

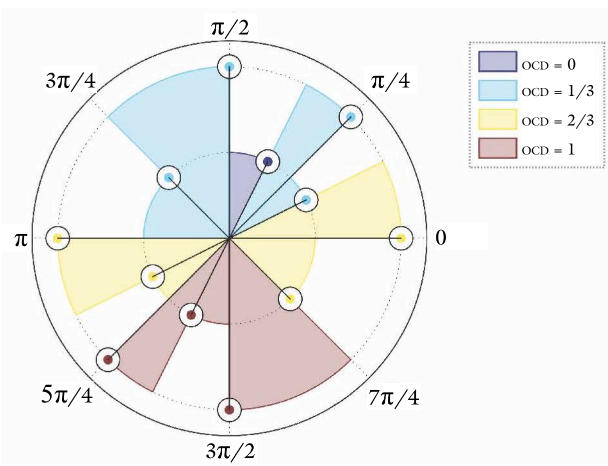
Contour theory has postulated a number of well-behaved distance functions—*metrics*—to describe the structure of contour space. Different metrics from the literature are tuned to different musical features, organizing contour space into various types of equivalence classes. Similarly, different spatial representations suggest different metrics.

Using basis coordinates, all CCs for any L, any n, have unique integer coordinates in $L - 1$ dimensions (basis vector scalars, themselves CCs). Ternary CCs occupy the central, smallest region in this space.[26] Simple Minkowski distance functions—metrics—on Euclidean spaces yield inter-CC distances between any two points in the space, providing a convenient measure which combines magnitudinal and contour differences. However, these basis-space metrics are indistinct from previously formulated distance measures which measure either magnitudinal or contour differences (or both). Additionally, standard distance functions proposed for measuring differences between CCs of different length (vectors of different dimensionality), such as Polansky's (1996) Unordered Combinatorial Direction metric (*UCD*), are, in general, less resolute and less sensitive to variations in contour shape.

CC distance in basis space may be measured solely by angle.[27] Equi-angular CCs (for some L) are equivalent under reduction of n. As n increases, so does the number of contour equivalence classes for the "ur" ternary CC.

For example, in Example 29, star vertices for ternary CCs increase in resolution (number of CCs) as *n* increases radially from the origin. Polansky's (1987, 1996) distinction between *OCD* (*Ordered Combinatorial Direction*) and *OCM* (*Ordered Combinatorial Magnitude*) correspond to angle and magnitude. OCD measures angular distance between two CCs, OCM measures magnitudinal distance. The similarity between angle and OCD raises the question: how well do the features of OCD and angular distance correspond?

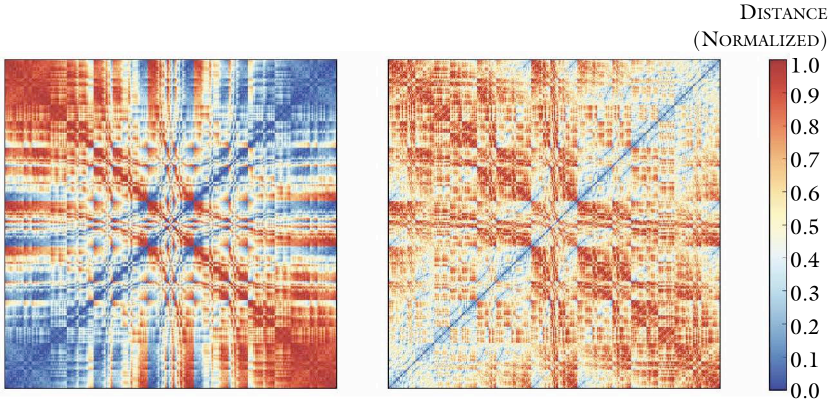
For CCs of morph length $L = 3$, angle and OCD correspond closely, as can be seen in Example 32, which plots OCD distance in polar space. In $L = 3$, OCD is a monotonically increasing function of angle over the semicircle in the positive direction from 0 to π and a monotonically decreasing function of angle in the negative direction from 0 to −π (due to inverse symmetry of CCs). In other words, if angle increases then OCD increases or stays the same; if angle decreases then OCD decreases or stays the same. The two measures always move in the same direction. Angle, however, is more resolute than OCD (which for contours $L = 3$ has only three possible non-zero distances (1/3, 2/3, 1). Equal OCD distances may often be further resolved by angle.



EXAMPLE 32: OCD DISTANCE IN POLAR SPACE. OCD IS MEASURED FROM THE SOURCE CONTOUR $[1, 2]_B$. IN $L = 3$, OCD DISTANCE AND ANGLE ARE CORRELATED—OCD INCREASES WITH ANGLE. OCD DISTANCES TO OTHER SOURCE CONTOURS ARE ROTATIONS OF THE SPACE

For higher L, angular distance between two contours (represented as vectors) can be computed from the inner product, giving the *cosine similarity*,[28] a common measure between (same length) vectors in high-dimensional spaces. Using the inner product, angle and OCD are no longer monotonically related, diverging locally. However, a comparison of DSM plots for OCD and cosine distance, which visualize distance in higher dimensional spaces by enumerating contours along linear index, shows significant correlation and identical symmetries (Example 33). While cosine similarity and OCD combine angular distances of all dimensions into one measure, angle can also be measured parametrically, or separately for each dimension, giving a vector of angular measurements, $[\theta_0, \ldots, \theta_{L-1}]$. Measured parametrically, angle and OCD correspond more closely. Within certain regions of the space, OCD and angle relate monotonically along each dimension. As angle increases or decreases along one dimension at a time, OCD generally follows.

Cosine similarity is a general, well-understood distance function—not a metric—that is often used in machine learning, optimization, and statistics. OCD, on the other hand, is a more idiosyncratic distance function—a metric—from music theory explicitly designed to measure the difference between two equal length CCs. Cosine similarity has useful analytic properties—such as being differentiable—that may be fruitful in solving musical problems in contour theory. Further, it is interesting to observe that conventional music-theoretic contour metrics (CSIM, OCD) capture some of the same inherent dimensions of contour space as cosine similarity.[29]



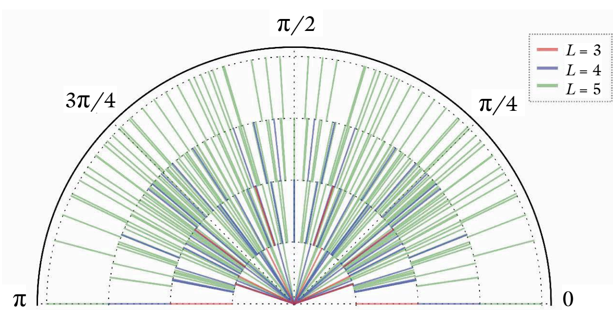
Example 33: DSM plots of angular distance (left) and OCD distance (right), $L = 5$

7.2 Representing Contours of Different Lengths

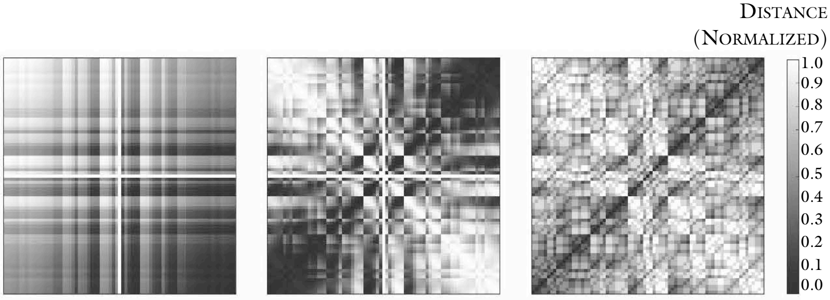
Angle in basis space can be used to organize and compare CCs of different lengths. Contours of all morph lengths L can be represented in the same space, the *unit half circle*, by measuring angular distance to an origin vector v, which takes the same general form for different contour lengths. Example 34 shows contours of morph lengths $L = 3$ through 5 in the unit half circle, with an origin vector $v = [1, 2, \ldots, L-1]_b$.

In the unit half circle, contours are organized according to their angular distance from v. Nearby contours are similar in their overall amount of deviation from v, but the deviation may be distributed across different dimensions. We know that it "wiggles," but not necessarily which contour elements do the wiggling. These contours tend to be similar in either large scale trajectory or overall amount of wiggling, but may differ in terms of which contour elements are raised or lowered. Example 35 shows DSM plots of angular distance on the unit half circle, cosine distance, and OCD for comparison.

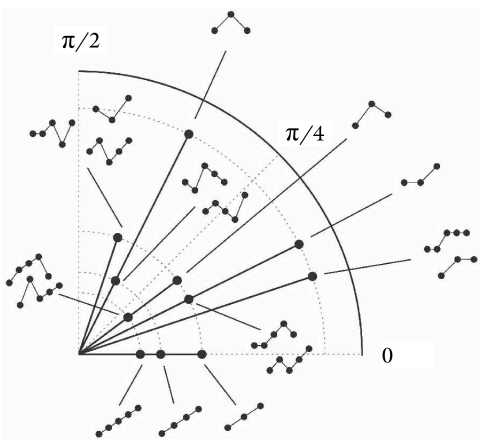
Not all contours are distinct in the unit half circle. Contours that have the same coordinates (r, θ) are said to *alias*, and contours that have the same angle θ but different magnitudes r (radius) function as fixed points across r. Example 36 shows this for a region of the unit half circle. The structure of the unit half circle is determined by the choice of basis $\beta = \{b_1, b_2, \ldots, b_{L-1}\}$ and origin vector v. Different basis representations will orient the space differently, determining which contour relationships constitute aliasing and fixed points.



Example 34: Unit half circle showing contours $L = 3, 4, 5$ (outward from center). Angle (θ) is measured to the origin vector $V = [1, 2, \ldots, L-1]_B$, which is $[1, 2]_B$ for $L = 3$, $[1, 2, 3]_B$ for $L = 4$, and $[1, 2, 3, 4]_B$ for $L = 5$



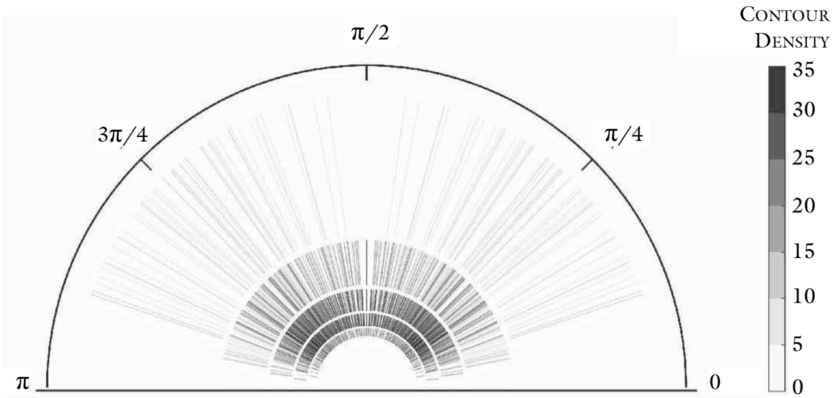
EXAMPLE 35: DSM SHOWING ANGULAR DISTANCE ON THE UNIT HALF CIRCLE (LEFT) ANGULAR DISTANCE BETWEEN CONTOURS (MIDDLE) AND OCD (RIGHT), $L = 4$



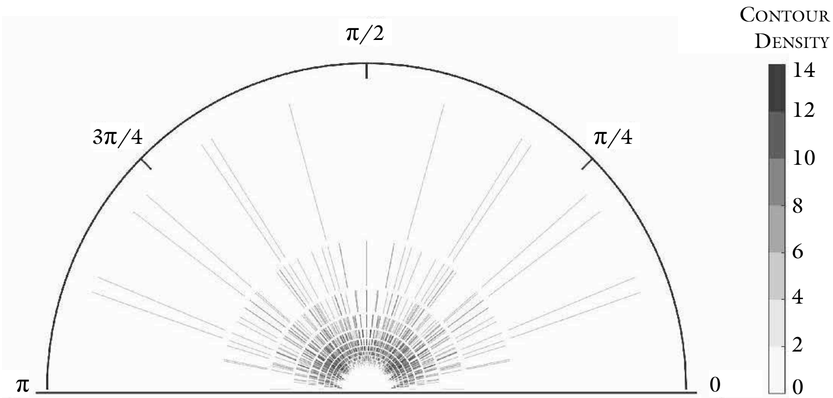
EXAMPLE 36: REGION OF UNIT HALF CIRCLE WITH SELECTED MORPHS DRAWN, ILLUSTRATING WHICH CONTOUR RELATIONSHIPS ALIAS (OCCUPY THE SAME POINT) AND WHICH CONTOUR RELATIONSHIPS FUNCTION AS FIXED POINTS (SAME ANGLE BUT DIFFERENT MAGNITUDE). BY CONVENTION, WE USE β AND $V = [1, 2, \ldots, L-1]_B$ TO MAXIMIZE UNIQUENESS, CAUSING THE FEWEST CONTOURS TO ALIAS

Increasing *L* Versus Increasing *N*

Contour Density Is Represented by Grayscale; Increasing Either *L* or *N* Fills in the Unit Half Circle, but in Different Ways



EXAMPLE 37A: INCREASING *L* = 3 THROUGH 6, *N* = 1



EXAMPLE 37B: SHOWS INCREASING *N* = 1 THROUGH 4, *L* = 4

7.3 INCREASING L VERSUS N: LENGTH AND MAGNITUDE

As n and L increase, contours populate basis space in different, but not independent ways. This can be seen in Example 37, which shows heatmap plots of contour density in the unit half circle. Increasing morph length L fills in the unit half circle outward along the radii, where as increasing resolution n fills in the unit circle towards the center in the stellated form. Compositionally or analytically, one might want to smoothly interpolate, in morphological space, between morphs of different lengths and different magnitudes, by primarily considering contour (and contour similarities). In the circular representation of contour above, fixing either n or L limits the paths that can be taken in moving between morphs, consequently defining specific sets of "reachable" points. The angular/magnitude representation of all contours (morphs), using basis coordinate vectors, allows for smooth movement between contours (and morphs), and suggests several intuitive distance relationships between all possible contours (any n, and L), and, consequently, between all morphs.

8 How Many Morphs?

In many areas of musical scholarship, such as ethnomusicology, composition, theory, and cognition, the idea of contour—non-magnitudinal descriptions of "ups and downs"—has been used to bring categorical order to the infinite set of (finite) musical shapes, or what we call *morphs*—ordered lists of measurable or known values. Contour, primary in our perception, is also an important and widely used reductive technique to formally represent large phenomenological classes of directionally-similar morphs by smaller sets of general descriptions.

All morphs can be represented as contours, with concomitant information reduction. As an extreme example, thirteen three-valued (ternary) combinatorial contours (CCs) suffice to represent any morph of length three, though in the coarsest possible reduction.

In this paper we are interested in the structure of what we call *CC-space* (combinatorial contour space), and by extension, the structure of musical morphological space. A melody can be seen as a multidimensional array of pitches, durations, and other, perhaps less prominent, parameters. Each dimension constitutes a morph, reducible in various ways (as shown in this paper) to a more general contour description.

We have focused on several problems—old and new—in the study of contour, particularly with regard to combinatorial contour. Of special interest is the problem that there are many more "impossible" contour descriptors then there are "possible" ones, making the structure of the space difficult to understand, much less visualize.

Using *combinatorial basis vectors* as axes for an multi-dimensional space, we demonstrate a clear structure for the space of possible CCs along with accompanying visualizations, suggesting various measures of distance. Practically, for example, in this space a composer or analyst might find ways of "moving smoothly" (or interpolating) between CCs of varying lengths and resolutions.

Basis space necessitates the idea of *full rank* contours—contour descriptors that are formally consistent with their equivalent rankings—extending a suggestion by Polansky and Bassein (1992): *n-ary contour*. CCs need not be limited to ternary values ("less than, greater than, or equal") but may be extended to include quinary values ("a little less than, a little greater than, and equal") and by extension, *n-ary*.

n-ary contour introduces *resolution* to contour representation. We formalize that concept as a key step in the derivation of basis vector representation of CC-space. We also generalize the Stirling number formula for the number of possible ternary CCs for a given L (Polansky and Bassein 1992) to *n*-ary CCs, showing that for every n, the number of CCs includes the number of CCs for all lesser n. All CCs represented by greater n are also members of equivalence classes of less resolute or lesser n by a "child/parent" relationship.

CCs that are "children" (more resolute) of CCs of smaller values for n are related to their less resolute "parents" by *angle* in basis space. This relationship is represented geometrically as greater n CCs (children) increasingly fill in the region defined by CCs of smaller n (parents). These related CCs increase in number radially as n increases (for a given L), subdividing the infinite set of morphs into increasingly discriminating contour representations. The smallest set of CCs—the thirteen ternary contours ($L = 3$, $n = 1$)—represents the largest equivalence classes of morphs. As n increases, the number of morphs represented by each CC approaches 1: the morph that is the same as the *n*-ary CC representation.

In this paper we have generally worked with morphs comprising non-negative integer-values, but this is an unnecessary constraint. For example, only using integer values it is simple to also explicitly and uniquely represent any rational-valued morph by contours of arbitrarily large values for n. This establishes a continuum between the set of rational-valued morphs and the set of contour representations. Further, introducing irrational values of n extends the continuum to all real-valued morphologies, including integer-, rational-, and irrational-valued morphs—a topic for further work. Contours form a one-to-one relationship with *all* morphs, including those with both integer (or rational) and real-numbered values. CC-space not only suggests several intuitive distance relationships between all possible contours (any n and L), but, by extension, between all morphs.

How many morphs are there? The set of all morphs is the same as the set of ranking descriptors of all CCs for all n, L. While there are *infinitely* many morphs, morph (and contour) vectors are, by assumption, finite in length. In the case of morphs only consisting of non-negative integer and/or rational values, the set of morphs (and, equivalently, contours) is thus the set of all finite length (L) vectors whose elemental values are of non-negative integer and/or rational magnitudes, or the *infinite union of finite sets of elements from countably infinite sets*, $\aleph_0$—resulting in a countably infinite set. However, if we allow irrational values as well, the set of morphs is the *infinite union of uncountably many finite sets*, —the cardinality of the real numbers. Alternatively, if the definition of morph allows countably infinite length, then the set of all morphs is the *infinite union of countably infinite sets* (power set) and also of cardinality .

Beginning with the most elemental (perceptual, mathematical, musical) reduction of melodic/morphological information—up, down, equal—and using the most basic representation of morphology (ternary contours), we can formally generate structure for the space of all possible morphs—melodies, rhythmic sequences, temporal forms, and so on. The structure of morphological space is the structure of combinatorial contour space.

9 Acknowledgements

This project was supported in part by a grant from the Arts Research Institute at the University of California, Santa Cruz. Thanks also to: Professor Robert Botje, Department of Mathematics, University of California Santa Cruz for invaluable advice; Beau Sievers of Dartmouth College and Andrew C. Smith of University of California Santa Cruz, for insightful feedback and conversation with us on these ideas; and *Perspectives of New Music* editors Jason Yust and Chris Stover for their questions and ideas.

Appendix A

Normal Form	CC (n = 1)	CC (n = 2)	CC (n = 3)	CC (n = 4)	Rank
$[2, 0, 0]_m$	$[1, 1, 0]_{cc}$	$[2, 2, 0]_{cc}$			2
$[3, 0, 1]_m$	$[1, 1,-1]_{cc}$	$[2, 2,-1]_{cc}$	$[3, 2,-1]_{cc}$		3
$[4, 0, 2]_m$	$[1, 1,-1]_{cc}$	$[2, 2,-2]_{cc}$	$[3, 2,-2]_{cc}$	$[4, 2,-2]_{cc}$	4
$[3, 0, 2]_m$	$[1, 1,-1]_{cc}$	$[2, 1,-2]_{cc}$	$[3, 1,-2]_{cc}$		3
$[2, 0, 2]_m$	$[1, 0,-1]_{cc}$	$[2, 0,-2]_{cc}$			2
$[2, 0, 3]_m$	$[1,-1,-1]_{cc}$	$[2,-1,-2]_{cc}$	$[2,-1,-3]_{cc}$		3
$[2, 0, 4]_m$	$[1,-1,-1]_{cc}$	$[2,-2,-2]_{cc}$	$[2,-2,-3]_{cc}$	$[2,-2,-4]_{cc}$	4
$[3, 1, 0]_m$	$[1, 1, 1]_{cc}$	$[2, 2, 1]_{cc}$	$[2, 3, 1]_{cc}$		3
$[4, 2, 0]_m$	$[1, 1, 1]_{cc}$	$[2, 2, 2]_{cc}$	$[2, 3, 2]_{cc}$	$[2, 4, 2]_{cc}$	4
$[2, 0, 1]_m$	$[1, 1,-1]_{cc}$	$[2, 1,-1]_{cc}$			2
$[3, 2, 0]_m$	$[1, 1, 1]_{cc}$	$[1, 2, 2]_{cc}$	$[1, 3, 2]_{cc}$		3
$[2, 1, 0]_m$	$[1, 1, 1]_{cc}$	$[1, 2, 1]_{cc}$			2
$[1, 0, 0]_m$	$[1, 1, 0]_{cc}$				1
$[1, 0, 1]_m$	$[1, 0,-1]_{cc}$				1
$[1, 0, 2]_m$	$[1,-1,-1]_{cc}$	$[1,-1,-2]_{cc}$			2
$[1, 0, 3]_m$	$[1,-1,-1]_{cc}$	$[1,-2,-2]_{cc}$	$[1,-2,-3]_{cc}$		3
$[2, 2, 0]_m$	$[0, 1, 1]_{cc}$	$[0, 2, 2]_{cc}$			2
$[1, 1, 0]_m$	$[0, 1, 1]_{cc}$				1
$[0, 0, 0]_m$	$[0, 0, 0]_{cc}$				0
$[0, 0, 1]_m$	$[0,-1,-1]_{cc}$				1
$[0, 0, 2]_m$	$[0,-1,-1]_{cc}$	$[0,-2,-2]_{cc}$			2
$[2, 3, 0]_m$	$[-1, 1, 1]_{cc}$	$[-1, 2, 2]_{cc}$	$[-1, 2, 3]_{cc}$		3
$[1, 2, 0]_m$	$[-1, 1, 1]_{cc}$	$[-1, 1, 2]_{cc}$			2
$[0, 1, 0]_m$	$[-1, 0, 1]_{cc}$				1
$[0, 1, 1]_m$	$[-1,-1, 0]_{cc}$				1
$[0, 1, 2]_m$	$[-1,-1,-1]_{cc}$	$[-1,-2,-1]_{cc}$			2
$[0, 1, 3]_m$	$[-1,-1,-1]_{cc}$	$[-1,-2,-2]_{cc}$	$[-1,-3,-2]_{cc}$		3
$[2, 4, 0]_m$	$[-1, 1, 1]_{cc}$	$[-2, 2, 2]_{cc}$	$[-2, 2, 3]_{cc}$	$[-2, 2, 4]_{cc}$	4
$[1, 3, 0]_m$	$[-1, 1, 1]_{cc}$	$[-2, 1, 2]_{cc}$	$[-2, 1, 3]_{cc}$		3
$[0, 2, 0]_m$	$[-1, 0, 1]_{cc}$	$[-2, 0, 2]_{cc}$			2
$[0, 2, 1]_m$	$[-1,-1, 1]_{cc}$	$[-2,-1, 1]_{cc}$			2
$[0, 2, 2]_m$	$[-1,-1, 0]_{cc}$	$[-2,-2, 0]_{cc}$			2
$[0, 2, 3]_m$	$[-1,-1,-1]_{cc}$	$[-2,-2,-1]_{cc}$	$[-2,-3,-1]_{cc}$		3
$[0, 2, 4]_m$	$[-1,-1,-1]_{cc}$	$[-2,-2,-2]_{cc}$	$[-2,-3,-2]_{cc}$	$[-2,-4,-2]_{cc}$	4
$[0, 3, 1]_m$	$[-1,-1, 1]_{cc}$	$[-2,-1, 2]_{cc}$	$[-3,-1, 2]_{cc}$		3
$[0, 3, 2]_m$	$[-1,-1, 1]_{cc}$	$[-2,-2, 1]_{cc}$	$[-3,-2, 1]_{cc}$		3
$[0, 4, 2]_m$	$[-1,-1, 1]_{cc}$	$[-2,-2, 2]_{cc}$	$[-3,-2, 2]_{cc}$	$[-4,-2, 2]_{cc}$	4

Notes

1. In this paper we use the word "morphology" and "morphological" to speak in general about the objects we refer to as "morphs."

2. In early studies of contour perception, memory and cognition (e.g., Dowling 1972, 1978; Dowling and Fujitani 1971; Cuddy and Cohen 1976; Cuddy, Cohen, and Miller 1979; Edworthy 1985), LC was usually the operative distance measure. In the music theory literature, LC corresponds to Friedmann's (1985) Contour Adjacency Series (CAS) and Morris' (1987, 2001) INT_1 function.

3. An ordered list of values could also be called a sequence or an indexed list.

4. Morris' *c-space* (normalized prime form ranking) and *COM-matrix* (representation of CC) demonstrate that for every (countable) ranking in c-space, there is a corresponding COM-matrix, or combinatorial contour, and vice versa. Marvin and Laprade's xxnormal form (Marvin and Laprade 1987; Morris 1987, 2001) and Polansky and Bassein's non-negative ternary representations are equivalent descriptions of the set of *possible* 3-valued contours (Polansky and Bassein 1992).

   Marvin and Laprade use normal form to tabulate equivalent contours (based on Morris' *c-space*), with an accompanying algorithm to reduce contours to normal form. Finally, utilizing the ranking of normal form, they enumerate a table of contour classes for contours up to length six. Morris' *contour reduction algorithm* (Morris 1987, 2001) organizes all contours, of any length, to an archetypal set of contours of a small length.

5. Other authors, notably Huron (1995), propose fixed-length typologies to represent different length contours, using statistical reductions of morphological values (an idea also used implicitly by Seeger [1960]).

6. Like CSIM and OCD, OLD (*Ordered Linear Direction*) a simple Hamming metric—a normalized count of "trivial" metrics (whose results are either 1 or 0, or "same/different"). Excluding 0, the number of values this metric assumes is $L - 1$, so the *grain* (g) of the metric is $1/(L - 1)$.

7. Also, Morris's (1987, 2001) *COM(a, b)* function.

8. See, for example, Morris's (2001, 21) definition of *c-space*, consisting of "c-pitches (cps) . . . numbered in order from low to high, beginning with 0 up to n – 1. The precise intervallic distance between the cps is indefinite, ignored, or left undefined." See also Morris (1987, 26).

9. Following Morris's *tree structure* and *contour reduction algorithm* (1987, 1993, 2001).

10. In its simplest form, Polansky's OLD metric (1981, 1987, 1996) is equivalent to those used by other theorists, such as Morris (1993) and Marvin and Laprade (1987), as well as to distance functions used in the music cognition and ethnomusicology literature and in information science (Hamming distance).

11. It is possible to extrapolate relationships between LCs using vector arithmetic. For example, the relationship LC_A is to LC_B could be extended to new LCs LC_C and LC_D by adding to LC_C the difference vector $LC_B - LC_A$. Relationships, however, are not always perceptually consistent due to the toroidal structure of the space. For instance, $[-1, -1]_{lc}$ (is less than, is less than) plus $[0, -1]_{lc}$ (is equal to, is less than) intuitively should be is $[-1, -2]_{lc}$ (is less than, is even more less than), however, because torus wraps values of *abs*() > 1 around the edge of the space to $[-1, 1]_{lc}$ (is less than, is greater than).

12. More concisely, a morph is in normal (ternary) form *iff* the minimum element is 0 and the number of unique elements = *max_element* – *min_element* + 1 (for ternary contours *max_element* = 2).

13. These relationships are implicit in and reconstructable from the combinatorial ternary direction vector.

14. Morris (personal communication, 2017) refers to these as "steps."

15. Intuitively, *L* (length) and *n* (magnitude) are analogous to the digital representation of a continuous signal, where *L* roughly corresponds to sampling rate, *n* to bit-width. Resolution is increased by either or both, eventually approaching the signal itself—or, in this case, the morph itself—by an infinitely close approximation.

16. The range of morph deltas referred to by n is not specified. Larger values of *n* only refer to greater degrees of resolution—that is, a more resolute quantization of contour. Division of a morph into *n* contour values is an implementation issue, not germane to this article. For example, a *nonary* contour (four values of "greater

than," four values of "less than," one value of equal) might have values considered as "very much less than," "not so much less than," "a little less than," "very little less than," "equal," "very little greater than," "a little greater than," "not so much greater than," "very much greater than." Alternatively, n might equally quantize (divide) a morph's total range.

17. In fact, a binary representation—[1, 0]—for "same/different" representation, sometimes called the trivial metric is even simpler, but is asymmetric and doesn't conform to the conventional definition of directional change.

18. For example, for CCs of length $L = 3$, written $[d(a, b), d(a, c), d(b, c)]$, the contour is full rank if the additive property holds for the deltas between morph elements a, b, c: $d(a, b) + d(b, c) = d(a, c)$. For CCs of length $L > 3$, written $[d(a, b), d(a, c), d(a, d), \ldots / d(b, c), d(b, d), \ldots / \ldots]$, the contour is full rank if the additive property holds for any three morph elements x, y, z: $d(x, y) + d(y,\mathrm{i}) = d(x, z)$.

19. In this paper we use the term *combinatorial basis space* to represent full rank combinatorial contour vectors—those that do not violate transitivity or additivity when interpreted as a half matrix of pairwise comparisons. This is a somewhat redundant term—any vector space has a basis—but we use "basis space" to distinguish this representation from our earlier representation of combinatorial contour space (CC-space) in L_{cc}-dimensional Cartesian coordinates.

20. While all possible CCs are represented, there will always be CCs of the same L_{cc} which lie in different rank-defined regions of basis space. However, all CCs of the same L_{cc}—even those of different rank—are contiguous.

21. Scalars and consequently, values for n, need not be integers, even though we limit ourselves here to integer representations of both contour vectors and basis-space axis values. Extension to rational and real values does not affect the formulation of the space.

22. This is the same as Callender, Quinn, and Tymoczko's (2008) *transposition space*.

23. "Prime form" (Marvin and Laprade 1987).

24. The formula for the number of possible ternary contours is:

$$\sum_{h=1}^{L} h!\, S(L, h)$$

25. Contour *rank* is related to both *size* and *resolution* depending on *normalization* (Polansky 1981). If rank is *normalized* between contours, then contour size is fixed and rank corresponds to the *resolution* of the contours. If rank is *not normalized* between contours, then contours grow in size as rank increases.

26. Morph normal form vector equivalents of CCs are likewise representable, in any magnitude, in a basis space isomorphic to CC basis space. This "normal form" space consists of all morphs with at least one zero value ("morphs invariant under transposition"). By isomorphism, CC basis space represents *all possible morphologies, in normal form*, of any length, comprising elements of arbitrary magnitudes with zero as minima. This isomorphism collapses the distinction between contour and morphology.

27. For example, Callender, Quinn, and Tymoczko (2008).

28. Cosine similarity measures the cosine of the angle between two vectors and is defined as

$$\cos(\theta) = \frac{A \cdot B}{\|A\|\,\|B\|}$$

29. A more through explication of the relationship between OCD and angular distance as well as its extensions to musical problems is a topic for further research.

References

Adams, Charles R. 1976. “Melodic Contour Typology.” *Ethnomusicology* 20/2: 179–215.

Beard, R. Daniel 2003. “Contour Modeling by Multiple Linear Regression of the Nineteen Piano Sonatas by Mozart.” PhD diss., Florida State University.

Callender, Clifton, Ian Quinn, and Dmitri Tymoczko. 2008. “Generalized Voice-Leading Spaces.” *Science* 320/5874: 346–48.

Cuddy, Lola L., and Annabel J. Cohen. 1976. “Recognition of Transposed Melodic Sequences.” *Quarterly Journal of Experimental Psychology* 28/2: 255–70.

Cuddy, Lola L., Annabel J. Cohen, and Janet Miller. 1979. “Melody Recognition: The experimental Application of Musical Rules.” Canadian Journal of Psychology 33/3: 148–57.

Dowling, W. Jay. 1972. “Recognition of Melodic Transformations: Inversion, Retrograde, and Retrograde Inversion.” *Perception and Psychophysics* 12/5: 417–21.

———. 1978. “Scale and Contour: Two Components of a Theory of Memory for Melodies.” *Psychological Reports* 85/4: 341–54.

Dowling, W. Jay, and Diane S. Fujitani. 1971. “Contour, Interval, and Pitch Recognition in Memory for Melodies.” *Journal of the Acoustical Society of America* 49/2: 524–31.

Edworthy, Judy. 1985. “Melodic Contour and Musical Structure.” In *Musical Structure and Cognition*, ed. Peter Howell, Ian Cross, and Robert West, 169–88. London: Academic Press.

Friedmann, Michael L. 1985. “A Methodology for the Discussion of Contour: Its Application to Schoenberg’s Music.” *Journal of Music Theory* 29/2: 223–48.

———. 1987. “A Response: My contour, Their Contour.” *Journal of Music Theory* 31/2: 268–74.

Hermann, Richard. 1994. “A General Measurement for Similarity Relations: A Heuristic for Constructing or Evaluating Aspects of Possible Musical Grammars.” PhD diss., Eastman School of Music.

Huron, David. 1995. "The Melodic Arch in Western Folksongs." *Computing in Musicology: An International Directory of Applications* 10: 3–23.

Johnson, Kimo. 2001a. "Morphing Combinatorial Contours." Unpublished Graduate Seminar Paper, Dartmouth College Electro-Acoustic Music Program, Hanover.

———. 2001b. "Ternary Contours." Unpublished Graduate Seminar Paper, Dartmouth College Electro-Acoustic Music Program, Hanover.

Kolinski, Mieczyslaw. 1965a. "The General Direction of Melodic Movement." *Ethnomusicology* 9/3: 240–64.

———. 1965b. "The Structure of Melodic Movement: A New Method of Analysis (Revised Version)." *Studies in Ethnomusicology* 2: 95–120.

Marvin, Elizabeth West. 1994. "Generalization of Contour Theory to Diverse Musical Spaces: Analytical Applications to the Music of Dallapiccola and Stockhausen." In *Concert Music, Rock, and Jazz Since 1945: Essays and Analytical Studies*, edited by Elizabeth West Marvin and Richard Hermann, 135–71. Rochester: University of Rochester Press.

Marvin, Elizabeth West, and Paul A. Laprade. 1987. "Relating Musical Contours: Extensions of a Theory for Contour." *Journal of Music Theory* 31/2: 225–67.

Morris, Robert. 1987. *Composition with Pitch-Classes: A Theory of Compositional Design*. New Haven: Yale University Press.

———. 1993. "New Directions in the Theory and Analysis of Musical Contour." *Music Theory Spectrum* 15/2: 205–28.

———. 2001. *Class Notes for Advanced Atonal Music Theory, Vol. I and II*. Lebanon, NH: Frog Peak Music.

Müllensiefen, Daniel, and Geraint Wiggins. 2011. "Polynomial Functions as a Representation of Melodic Phrase Contour." In *Systematic Musicology: Empirical and Theoretical Studies*, edited by Albrecht Schneider and Arne von Ruschkowski. Special issue, *Hamburger Jahrbuch für Musikwissenschaft* 28: 63–88.

Polansky, Larry. 1981. "Morphological Metrics." Lecture at Mills College Seminar in Formal Methods Series, Oakland.

———. 1987. “Morphological Metrics: An introduction to a Theory of Formal Distances.” In *Proceedings of ICMC*, 197–205. San Francisco: International Computer Music Association.

———. 1996. “Morphological Metrics.” *Journal of New Music Research* 25/4: 289–368.

Polansky, Larry, and Richard Bassein. 1992. “Possible and Impossible Melody: Some Formal Aspects of Contour.” *Journal of Music Theory* 36/2: 259–84.

Quinn, Ian. 1999. “The Combinatorial Model of Pitch Contour.” *Music Perception: An Interdisciplinary Journal* 16/4: 439–56.

Schmuckler, Mark A. 1999. “Testing Models of Melodic Contour Similarity.” *Music Perception* 16/3: 295–326.

———. 2010. “Melodic Contour Similarity Using Folk Melodies.” *Music Perception: An Interdisciplinary Journal* 28/2: 169–93.

Seeger, Charles. 1960. “On the Moods of a Music-Logic.” *Journal of the American Musicological Society* 13/1–3: 224–61.