Tuukka Ilomäki On the Similarity of Twelve-Tone Rows Studia Musica 30 Sibelius Academy

ISBN 952-5531-32-5 (paperback)
ISBN 952-5531-33-3 (pdf)
ISSN 0788-3757

Yliopistopaino
Helsinki 2008

Tuukka Ilomäki

On the Similarity of Twelve-Tone Rows

Studia Musica 30

SIBELIUS ACADEMY

Tuukka Ilomäki

On the Similarity of Twelve-Tone Rows

Studia Musica 30

Sibelius Academy

ISBN 952-5531-32-5 (paperback)
ISBN 952-5531-33-3 (pdf)
ISSN 0788-3757

Yliopistopaino, Helsinki 2008

SIBELIUS ACADEMY
DocMus Department
Studia Musica 30

Ilomäki, Tuukka
On the Similarity of Twelve-Tone Rows
274 + 26 pages

ABSTRACT

The relations of twelve-tone rows are of theoretical, analytical, and compositional interest. While relations based on the properties and transformations of rows have been widely studied, less attention has been paid to relations based on similarity. Formal similarity measures can be used to explicate ways of being similar.

This study presents an analysis and categorization of 17 similarity measures for twelve-tone rows. Nine of them are new. The categorization of the similarity measures suggests the notion of different conceptions of twelve-tone rows. Five such conceptions are identified and explicated: vector, ordered pairs, subsegments, subsets, and interval contents. Similarity measures could thus be grouped into families based on the conception that they suggest.

The similarity of twelve-tone rows allows two interpretations: comparison of the properties of the rows and the measurement of their transformational relations. The latter could be conveniently formalized using David Lewin's Generalized Interval Systems as the framework. This allows the linking of the discussion on permutations in mathematics and computer science because the measurement of the complexity of a transformation coincides with the notion of presortedness of permutations.

The study is in three parts. The first part gives an overview of the types of relations between twelve-tone rows, and presents a formalization of twelve-tone rows and row operations in terms of group theory.

The second part focuses on the properties of similarity. By way of background a review and criticism of the literature on similarity in music theory is presented. The transformational approach and the metric are promoted. It is shown that transformational similarity measures create perfectly symmetrical spaces since every row is related to the other rows by precisely the same set of transformations. Since most of the similarity measures discussed in this study are dissimilarity measures of the distance between rows, the mathematical concept of the metric is applicable; many similarity measures define a metric. One of the main findings is that any metric for twelve-tone rows that is transformationally coherent under the operations generating row classes also defines a metric for those row classes.

The third part discusses the similarity measures and the respective conceptions in detail. While the study focuses on the similarity of twelve-tone rows, the possibilities of extending the measures to the examination of other ordered pitch-class sets are also discussed. The work concludes with some examples of their analytical application.

Acknowledgments

This work was carried out in the DocMus Department at the Sibelius Academy during the period between 2002 and 2008. During the academic years 2004–2005 and 2005–2006 I studied at the Eastman School of Music under the auspices of the Fulbright program. In all these years I had the privilege to work with wonderful people who have given me a prodigious amount of help.

First of all, I wish to express my gratitude to my supervisors Prof. Marcus Castrén and Prof. Dave Headlam, both of whom have followed the project closely from its inception to its end. Their supervision and support have made the experience delightful. Prof. Robert Morris functioned as both a consulting supervisor and a reviewer of the final manuscript. His insight, suggestions, and criticism have shaped much of the work. Prof. Robert Peck, the second reviewer of the final manuscript, gave valuable criticism and inspiring suggestions. Prof. Eero Saksman read some critical portions of the manuscript and was also an inexhaustible source of encouragement. I am particularly grateful to him for his refusal to accept hand-waving as a method of proving theorems. The numerous Friday-afternoon discussion sessions with Prof. Ciro Scotto were inspiring and helped me clarify my ideas. The discussions with Dr. Tuire Kuusi were also helpful. My colleague Yki Kortesniemi gave me numerous lessons in the art of scientific writing, and provided solid advice on various issues during the various stages of this work. Joan Nordlund revised the language of the final manuscript.

I am deeply grateful to the teachers and my fellow students at the DocMus Department, the Pythagoras Graduate School, and the Eastman School of Music, who provided an inspiring and challenging working environment.

I wish to express my gratitude to the following institutions for their financial support: the Academy of Finland through the Pythagoras graduate school, the Sibelius Academy Innovation Centre, the DocMus department, the Fulbright Center, and the Leonora and Yrjö Paloheimo foundation.

Finally, I would like to thank my family and friends for copious things, most of which have absolutely nothing to do with either twelve-tone rows or similarity, and which could never be expressed in terms of group theory.

Table of Contents

Part I: Background

Definitions and conventions

The literature on post-tonal music theory is plagued by a plethora of different notations. Every book presents its own set of definitions and conventions. This study is no exception. The conventions I have used derive mostly from mathematics, in which the notational conventions are relatively well established.

Pitch classes are denoted by the integers $0, 1, \ldots, 11$. When the clarity of spacing requires it, A and B stand for the integers 10 and 11, respectively. I use *fixed-zero notation* throughout the text. Hence, 0 denotes pitch class C, 1 denotes pitch class $C\sharp$ (or $D\flat$), etc.

The practice of fixed-zero notation may need some justification. In particular, labeling of some pitch class as 0 has been criticized for being "theoretically suspect, in that it implicitly asserts that there is always one a priori tonic pitch-class" (Lewin 1977, 43).[1] My response is that labeling a pitch class 0 does not give it any special status, let alone a "tonic" status, just as we do not a priori assign such a status to pitch class C.[2] Transposition T_0, as will be demonstrated later on the basis of group theory, has a special status. However, there is no reason to jump to the conclusion that pitch class 0 also has a special status just because they both happen to contain the same symbol "0". Similarly, inversion I_0 does not have any special status over the other inversions.

An unordered set is denoted by curly braces $\{\}$ and the elements are separated by commas. For instance, pitch-class set $\{0, 1, 2\}$ contains pitch classes 0, 1, and 2. An ordered set is denoted by parentheses $()$ and the elements are separated by commas. The cardinality of a set is denoted by $\#$: for example, $\#\{0, 1, 2\} = 3$. Sometimes I define a set by setting some criterion: for example, the expression $\{n \mid n < 2\}$ denotes a set of integers smaller than 2.

I have used the T_n/T_nI-classification of set classes (Forte 1973b; Rahn 1980). I use the format 6-1[012345], and sometimes the shorter formats 6-1 and [012345], in referring to set classes,

[1]For example, Babbitt (1961b) and Perle (1991) use moveable-zero notation, which has the side effect of assigning an a priori referential status to some pitch class.

[2]In the *set* of integers $\{0, 1, 2, \ldots, 10, 11\}$ number 0 has no special status. I will discuss the notion of an unstructured set versus sets with some structure defined in more detail in section 1.2.

and Donald Martino's (1961) labeling for the six all-combinatorial[3] hexachord set classes: A = 6-1[012345], B = 6-8[023457], etc.

The M-operation denotes the multiplicative operation or the "cycle-of-fifths transformation" which maps pitch class 0 into pitch class 0, 1 into 5, 2 into 10, and so on. The twelve operations that David Lewin (1966) labels $M_1, M_2, \ldots, M_{12}$ are referred to as Lewin's M-operations.

I differentiate between an *ordered pitch-class interval* and an *unordered pitch-class interval*, referring to the latter as an *interval class*. There are twelve ordered pitch-class intervals but only seven interval classes.

The interval-class vector of a pitch-class set is written between brackets. For instance, the interval-class vector of pitch-class set $\{0, 1, 2, 3, 4, 5\}$ is [543210]. I do not include the interval class 0 in the interval-class vector.[4] I use brackets for both set classes and interval-class vectors, and the context will show which interpretation is intended.

A twelve-tone row is some ordering of the twelve distinct pitch classes.[5] Twelve-tone rows are usually referred to in uppercase letters, and the elements of a row in lowercase letters. For example, p_n denotes the nth element of a twelve-tone row P (bearing in mind that every row begins with a "zeroth" pitch class), hence $P = p_0p_1p_2p_3p_4p_5p_6p_7p_8p_9p_{10}p_{11}$.

A twelve-tone row has twelve *order positions*. An *order number* denotes the order position of a pitch class, and an *order-number row* enumerates the order positions of each pitch class (see Section 2.1.1). For example, the pitch class in the first order position of a row has order number $\mathbf{0}$ and $\boldsymbol{x_n}$ denotes the nth order number of an order-number row $\boldsymbol{X}$. I have adopted Andrew Mead's convention of writing order numbers and order-number rows in bold.

The term "twelve-tone row," or simply "row," refers to any one of the 479001600 twelve-tone rows. *Row class* denotes a set of rows that are related by a set of row operations. Unless otherwise specified, a row class denotes a set of rows that are related by transposition, inversion, retrograde, and their combinations. The term "row form" denotes a member of such a row class and is used only when the context implies the pertinent one. The row class of row P is denoted by $[P]$.

For any twelve-tone row or segment, the function INT denotes the succession of ordered pitch-class intervals between its adjacent pitch classes.[6] The ordered pitch-class intervals are enumerated between angle brackets. A segment of length n has INT of length $n - 1$. For instance, the INT of row 5409728136AB is $\langle 11 \;\; 8 \;\; 9 \;\; 10 \;\; 7 \;\; 6 \;\; 5 \;\; 2 \;\; 3 \;\; 4 \;\; 1 \rangle$. The *unordered INT contents* of a row or segment denote the unordered set of ordered pitch-class intervals between its adjacent pitch classes. Correspondingly, the *unordered interval-class contents* of a row or segment denote the unordered set of unordered pitch-class intervals between its adjacent pitch classes. For instance, the unordered INT contents of row 5409728136AB are $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$ and its unordered interval-class contents are $\{1, 1, 2, 2, 3, 3, 4, 4, 5, 5, 6\}$. The *unordered interval contents* of a row refers both

[3] See Babbitt (1955).

[4] Allen Forte (1973b) and several other authors would call this the interval vector of a pitch-class set.

[5] Milton Babbitt (1946/1992, viii) decided not to use the terms "row" or "series," but introduced the term "set." Unfortunately, since pitch-class sets are commonly abbreviated as sets, addressing twelve-tone rows as sets is somewhat confusing. In addition, in mathematics a set denotes an unordered entity.

[6] INT corresponds to Morris' INT_1 (see definition 3.10.1 in Morris (1987)).

to its unordered INT contents and to its unordered interval-class contents.

A *segment* is an ordered pitch-class succession. A *subsegment* of a row is a segment in which the pitch classes are in the same order as in the row, but the pitch classes of a subsegment do not need to be contiguous in the row. For example, both rows $5409728136\mathrm{AB}$ and $5406918237\mathrm{AB}$ contain the (non-contiguous) subsegment $540913\mathrm{AB}$.

I extend the notion of intervals to order numbers. Unless specified otherwise, I have used unordered order-number intervals. Hence, the order-number interval between order numbers $\mathbf{0}$ and $\mathbf{2}$ is 2, as is the order-number interval between order numbers $\mathbf{2}$ and $\mathbf{0}$. However, the ordered order-number interval between order numbers $\mathbf{2}$ and $\mathbf{0}$ would be -2.

I have used some standard mathematical notation. The expression $\binom{n}{k} = \frac{n!}{k!\cdot(n-k)!}$ denotes the number of ways to select k distinct elements from n elements (see Section 1.4 in Liu (1968), for example).

The expression $\lceil x \rceil$ (the "ceiling function") denotes the smallest integer that is greater than or equal to x. Similarly, the expression $\lfloor x \rfloor$ (the "floor function") denotes the largest integer that is smaller than or equal to x.

In this work the symbol $\subseteq$ denotes a *subset* and the symbol $\subset$ denotes a *proper subset*: in other words, $A \subset B$ implies $A \neq B$.

CHAPTER 1

Introduction

1.1 Rows and their relations

Typically, an analysis of a twelve-tone composition that is oriented toward the pitch-class organization involves discovering the rows of which the piece is composed. When we have identified the rows, we may proceed to analyze some of their properties. Since, in general, composers use more than one row in a twelve-tone composition, it is of interest to explore the relations between the rows of the piece in addition to analyzing the properties of the individual rows. Why these rows? Is there some special property that these two rows share? What would be the most compelling way to describe the relation between them? What can we say about the transformational relations between them? How is the row succession related to the form? What is perhaps not such an obvious question is whether the rows are similar or dissimilar. I will suggest that this is a relevant question and will provide the means with which to tackle it.

The realms of twelve-tone rows and the relations between them could be characterized as a "combinatorial explosion." We begin with only twelve distinct pitch classes. These pitch classes can be ordered in $12! = 479001600$ different ways: hence, there are 479001600 distinct twelve-tone rows. If we proceed to explore the relations between rows we have a total of $479001600 \cdot 479001600 = 229442532802560000$ distinct (ordered) pairs of rows. Further, if we proceed to examine row operations, there are $12!! = 479001600!$ possibilities, and for any row there are $479001599!$ row operations that may carry it to any other.

These are daunting numbers for a human; there is no way we can examine each pair of rows or even each row separately. In coping with this multitude, a typically human approach is to place rows and their relations into categories.[1] Consequently, a comprehensive theory of twelve-tone rows must allow for the conceptualization of rows and of their relations.

[1]The same approach is also used in other branches of science, for example, the division of species into families, suborders, orders, etc., in biology.

As far as rows are concerned, Milton Babbitt introduced the idea of a *permutational approach* (as opposed to a combinational approach), according to which a twelve-tone row is fundamentally understood as an ordering of the twelve pitch classes (Babbitt 1960, 248). Every twelve-tone row contains the very same twelve pitch classes, only differently ordered. This notion is widely accepted: the formal details are discussed in Chapter 2.

On the one hand, a row is simply an ordering of the twelve pitch classes. On the other hand, multiple aspects of rows emerge from this simple notion, which becomes clear when we try to relate different ones.[2] Rows may be related because they belong to the same row class, begin with the same pitch class, have the same or similar unordered INT contents or unordered interval-class contents, share some subsegments, and contain related imbricated pitch-class sets, for example. None of these ways of being related is either perceptually or conceptually a priori to the others, but naturally, some of them may be more relevant in an analysis of some specific composition.

The relations between rows have been an inseparable part of the twelve-tone system from the very beginning. To begin with, members of the second Viennese school created *row classes* using row operations: they were aware of some of the special properties that a particular row may have, and they created associations between rows that were similar in some respect. In particular, composers have used rows that can be related in multiple ways. Typically, a composer might employ two rows that are related by a transformation, and also associated by similarity or some shared property.

In the theory-building process I will proceed to formalize these notions of relatedness. Section 1.2 outlines some general aspects of creating a *space* that formalizes an approach to relating rows. The methodology used in this work is discussed in Section 1.3, and Section 1.4 outlines the objectives of the study.

1.2 Row spaces

1.2.1 The formalization of row spaces

From the theoretical perspective there is a need to maintain a distinction between twelve-tone rows and ways of conceptualizing their relations. I make this distinction explicit by distinguishing between the concepts of a *set* and a *space*. A set is simply a collection of items with no internal structure, and no relation is defined between its elements. By imposing some structure on the set we define a space. The structure may be the definition of a set of relations or distances, for example. The set of 479001600 distinct twelve-tone rows is an amorphous mass and defining a relation between rows constitutes imposing a structure on it.

Let me note here that, even if this work relies heavily on the work of David Lewin and, in particular, on his notion of *Generalized Musical Intervals* (Lewin 1987), the terminology I employ is somewhat different. Lewin uses the term "space" to denote an unstructured collection of elements, thus obliterating the distinction I make between a set and a space. Nevertheless, he makes the

[2] A word of caution at the very outset is appropriate. The notion that two rows are "related" does not necessarily imply the existence of a mathematical *relation*, it merely suggests an informal relatedness. I have taken great care to distinguish between these two notions. When I wish to invoke the mathematical notion of relation I am very explicit about it.

following comment about conceptualizing a musical space.

> In conceptualizing a particular musical space, it often happens that we conceptualize along with it, as one of its characteristic features, a family of directed measurements, distances, or motions of some sort. (Lewin 1987, 16)

I subscribe to this view, but suggest that it is precisely the directed measurements that turn a set into a space. For instance, ordered pitch-class intervals turn the set of pitch classes into a pitch-class space.

We can conceptualize twelve-tone rows in multiple ways, and a space formally defines how we relate them. There is no *a priori* conception of twelve-tone rows. In deciding how to relate them we bring some aspect of rows to the fore, which is made particularly clear in Part III when I examine the various similarity measures and group them together based on the conception of a row that they impose.

A space also implies a conceptualization of the rows – either explicitly or implicitly. I emphasize the fact that "space" and "conceptualization" are two distinct notions. For example, we might conceptualize a twelve-tone row as a vector (see Chapter 8), but this does not inevitably lead to some specific space – an infinite number of spaces can be created based on this concept. On the other hand, a space defined by the DISPLACEMENT similarity measure, for example, does imply (see Section 8.3) the conceptualization of twelve-tone rows as vectors. Hence, a conceptualization of the twelve-tone rows is more rudimentary than a twelve-tone row space.

I divide twelve-tone row spaces into three main categories: spaces defined by row operations, spaces defined by properties, and spaces defined by similarity. I outline the basic properties of each of these three categories in Sections 1.2.2, 1.2.3, and 1.2.4, which provides the context for a more detailed examination of the row spaces that the similarity measures create – the main topic of this work. I will take a close look at the properties of similarity in Part II, and then examine a number of similarity measures in detail in Part III.

I emphasize the fact that my categorization of row spaces is based on the formal aspects of the spaces. We could take one conception of rows and create spaces of different types based on it. For instance, let us consider the unordered INT contents of rows. On the one hand, we could define a space in which two rows are related if they have identical unordered INT contents. The result is a space defined by property, and hence it is an equivalence relation. On the other hand, we could create a function that produces a numeric value that depicts the similarity of the unordered INT contents. The result is then a space defined by similarity.

1.2.2 Row spaces by operation

The distinction between a set and a space outlined above is related to another distinction introduced in the literature, namely that between an *object* and an *operation*. Daniel Starr justifies this as follows.

> I find it both fruitful and intuitive to conceive of an *operation-object* duality, in which "operation" is a concept subject to general discussion, while "object" arises from the

> discussion of specific works. Thus, to approach various general aspects of twelve-tone or related types of music, I stress what I consider the *operations* that we apply to sets, rows, partitions, etc., rather than those *objects* themselves, or, for that matter, their classification, which is the topic most often considered. (Starr 1978, 1–2)

The gist here is that we intuit that twelve-tone rows and row operations manifest different characteristics: we can compose operations to form new operations but we do not compose rows; rows may be invariant under some operation but we do not have a corresponding notion for operations; a row operation has an inverse operation but no such concept exists for twelve-tone rows. David Lewin makes a similar point about pitch classes and pitch-class operations: "Now while the pc's do not exhibit algebraic behavior, the intervals do" (Lewin 1977, 33). In the case of pitch classes the distinction is somewhat obscured by the fact that the mathematics of both pitch classes and pitch-class operations involves integers modulo 12. Similarly, in the case of twelve-tone rows the distinction is somewhat obscured by the fact that the mathematics of both rows and row operations involves permutations. I will discuss the permutational nature of rows in Section 2.1.3, and as I will show in more detail in Section 2.2.2, row operations can be naturally and intuitively formalized as a group of permutations acting on the set of twelve-tone rows.

In addition, row operations may be used to obtain information about the objects themselves. Walter O'Connell describes the role that the transformations play when a composer tries to find invariances in his musical materials.

> The phycisist reveals these symmetries (or their absence) by subjecting his equations to various transformations, and discovering what properties each leaves unaffected. The unchanging property is said to be 'conserved', or to be 'invariant' under that transformation. Perhaps similar methods can aid the composer in his preliminary examination of the possibilities latent within his material. (O'Connell 1962, 35)

The same approach could be applied in the examination of the various properties of twelve-tone rows. For example, it would certainly be interesting to determine which row operations preserved the unordered INT contents or unordered interval-class contents either in all rows or in some selected subset, such as the set of all-interval rows.

We could turn the set of twelve-tone rows into a space by using row operations. Typically, we would require the row operations to define a group (see Section 2.2.2), and the group action to satisfy certain criteria (see Definition 2.2). The group of row operations induces a permutation group on the set of twelve-tone rows (see Section 2.2.3), and the permutation group induces an equivalence relation: a row is related to another row if and only if there is a permutation that maps the former to the latter. This ensuing equivalence relation defines the space we are seeking.

The role of row operations goes beyond creating row classes: they can be used, for example, to create and analyze row transformations, to analyze the connections between existing rows or row classes, and to create similarity measures for twelve-tone rows.

1.2.3 Row spaces by property

Twelve-tone rows can be grouped based on some property that they may have. Typically, we would consider two rows related if and only if they shared a certain property. Consequently, the resulting relation would be an equivalence relation.

For example, rows may be grouped by their unordered INT contents or unordered interval-class contents, their properties related to symmetry or invariance, or their imbricated set-class contents. A classification of twelve-tone rows in terms of the knots they define is a novel approach, which also defines an equivalence relation (Jedrzejewski 2006, 106). I will briefly discuss the issue of unordered INT contents in Section 12.3.

The spaces defined by both operations and properties are closely linked. Indeed, it is easy to devise a space defined by a property and one defined by a group of row operations that are, in fact, identical. For example, if we take the ordered succession of pitch-class intervals as a property, the ensuing equivalence relation is identical to the equivalence defined by the group of transpositions. Indeed, the examination of whether certain properties are kept invariant in an operation is one way of examining the operation.

Josef Hauer's (1925, 1926) *tropes* provide another example that would define (if they were so formalized) a row space defined by a property: two rows are of the same trope if they contain the same (unordered) hexachords.

A major category of row relations is combinatoriality, which also combines the operations and properties of rows. The origins of combinatoriality lie in Schoenberg's practice of using inversionally related row forms, as he describes in his essay "Composition with twelve tones" (Schoenberg 1975). Since then, due to the work of Donald Martino (1961), Milton Babbitt (1974), Starr and Morris (1977, 1978), and others, it is probably the most thoroughly studied aspect of twelve-tone rows.

1.2.4 Row spaces by similarity

It is possible to define a space in which the twelve-tone rows are related by similarity. For every pair of rows we could define a numeric value that denotes their degree of similarity. In this case we would not have a binary relation, but a space that was structured by the similarity values.

The most straightforward method would be to assign every pair of rows a value denoting their similarity, which is what I intend to do. I will also discuss the idea of transformational similarity, to which I will add a further layer: I will first define the transformational relations between rows and then extract the similarity values from them, thereby building a row space defined by similarity on top of one defined by transformations.

I wish to stress that the notion of "being similar" encompasses more than what is formalized as space by similarity. For example, two rows could be similar because they are related in some row space defined by a property (such as unordered INT contents).

1.3 Methodology

The methods I have used in this study are of a mathematical nature. The twentieth century saw an increase in the use of mathematical methods in music theory. In fact, way back in 1946 Milton Babbitt wrote in his dissertation as follows.

> But this monograph is, likewise, not a mathematization of the twelve-tone system, although superficially it may appear so. A true mathematization would require a formulation and presentation dictated by the fact that a twelve-tone complex is a permutation group, and would be shaped by the structure of this mathematical model. Within this framework, many of the problems and proofs would be presented in the language of number theory and combinatory analysis, and there is no question but that this approach would represent the definitive way of dealing with the matter, from a standpoint of rigor and manipulative efficiency. (Babbitt 1946/1992, ii)

Babbitt set the direction in his writings that later generations have followed. Robert Morris explicates reasons for using mathematics as follows.

> The reason for the mathematics is that it clearly and elegantly models the entities, relations, and constructions used in post-tonal music [...] I take the view that formal definitions and proven theorems provide the most secure foundation for understanding the uses and implications of our topic. In addition, those who own personal computers or write programs may find my formulations useful in designing programs to aid their analytic and compositional requirements. (Morris 1987, xii)

Mathematical methods have also been incorporated into elementary text books on music theory. John Rahn repeats the same tenet in his text book on atonal theory.

> Consistent definitions are offered in place of informally opaque concepts, and proven theorems are offered in place of previously obscure assumptions. (Rahn 1980, v)

A mathematically trained observer simply cannot help thinking in terms of the mathematical structures that musical practice suggests. In Section 2.2.2 I will give some concrete reasons why group theory is and should be used in the study of row operations. There is a hint of such inevitability in the writings of Babbitt.

> The reader with a knowledge of elementary finite group theory will recognize at once the necessary invocation of cosets and imprimitive systems. (Babbitt 1961b, 75–76)

Music theory typically makes very little reference to existing mathematical results, and thus corresponding results are reinvented by the authors. Writers of music-theory treatises have introduced new proprietary concepts and notations, and have then applied the results to musical objects. Similarity measures are a good case in point, since most of them have a correlate in mathematics or computer science, which is usually not referred to. Furthermore, the literature is replete with different definitions and terminology.

I intend in this work to avoid reinventing mathematical concepts and results and hence I will stick to mathematical terminology and standard mathematical notation.[3] There is therefore a need to translate the musical structures into the language of mathematics. This decision will probably please mathematically trained readers, but unfortunately readers with a background in music theory will be faced with yet another set of terminology and notation.

Finding a balance between issues assumed and issues explained is difficult in a work like the present one. On the one hand, a mathematically trained reader might consider the explanations unnecessary while the music theorist might find the formalism next to unintelligible. Given the need to navigate between Scylla and Charybdis I offer as a solution a series of appendices that provide a gentle introduction to some of the pertinent mathematical concepts used in this work in the hope of making it accessible to a wider audience. I have usually omitted proofs of well-known mathematical results from the body of the text, and have rather given a pointer to the relevant literature.

The prerequisites of this work are mappings, modular arithmetics, the basics of naïve set theory, and elementary combinatorics. The appendices provide brief introductions to some more advanced topics, such as permutations, relations, partially ordered sets, and graph theory. Group theory constitutes a major component of this work. Although I introduce the pertinent concepts, fluency in group theory will make the text more accessible.

This work is about twelve-tone music and the structures found in it. I introduce mathematical language and methods only to clarify the issues arising from the musical context. Consequently, I offer no new mathematical results, with the exception of introducing a new metric for permutations with some very interesting properties (see Section 9.3).

Computer applications also play a vital role in this work. Typically, we need to resort to computer applications in order to obtain the distribution of some property of rows or some relation between them. Due to the sheer numbers of rows and pairs of rows, finding the distribution by hand is usually impossible. Naturally, I provide an exact formula giving the results whenever such a formula can be found.

All of the software used in this work was developed by the author with the exception of an algorithm devised by Gara Pruesse and Frank Ruskey (1997), and an implementation of it given by Kenny Wong and Frank Ruskey, which was released under the terms of version 2 of the GNU General Public License.

1.4 On the objectives of this study

I have seven principal objectives in this work. First, I present a categorization and a framework for the formalization of twelve-tone rows and their relations which sets the stage for an analysis of similarity relations between twelve-tone rows.

[3]Unfortunately, in a few isolated cases this is not possible. For example, in order to avoid a clash between the mathematical term *transposition* and the musical term *transposition*, I have renamed the former an *exchange*. Nevertheless, I have adhered to the mathematical concept even if I have renamed it.

Secondly, I offer a comprehensive discussion of the similarity of musical objects in general, and of twelve-tone rows in particular, and this leads me to challenge some views about the nature of similarity found in music theory. In particular, I will tackle the issue of equivalence versus similarity: does the former imply the latter? In addition, I will promote the use of the metric as a tool with which to analyze the properties of similarity measures.

Thirdly, I analyze the properties of the existing measures of twelve-tone row similarity and propose several new ones. In this I have found it helpful to link the discussion to some relevant concepts in mathematics and computer science: three of the new measures (DERANGEMENT, CAYLEY DISTANCE, and ULAM'S DISTANCE) are borrowed from these disciplines. The other new measures (PITCH-CLASS DISPLACEMENT, GENERALIZED ORDER INVERSIONS, DIVISIONS, FRAGMENTATION, NESTINGS, and INTERVAL DISPLACEMENT) arose from the extending of existing measures to incorporate new dimensions, the formalization of some ideas suggested in the literature, and the failure of existing similarity measures to observe certain types of similarity.

Fourthly, I introduce and argue for the notion of different conceptions of twelve-tone rows. I identify and explicate some conceptions of rows that the various similarity measures (and other ways of relating them) discussed in the literature suggest, which then enables me to group the measures into "families."

Fifthly, I argue for a transformational approach to similarity, which goes particularly well with the idea of a permutational approach. I use David Lewin's notion of Generalized Interval Systems to define transformational relations between rows, and this enables me to connect the conceptions of twelve-tone rows and similarity measures to transformational procedures.

The literature on twelve-tone rows and their relations contains some errors and misconceptions. My approach is different from previous ones in the sense that I separate the mathematics and the music. My sixth objective is not to write a theory of row operations that is parallel to mathematical group theory, it is rather to formulate the theory in the language of mathematics, and then to translate the concepts of music theory to the corresponding mathematical concepts. This approach will help me to avoid some of the inaccuracies and misconceptions found in the existing literature: I will insist on preciseness and painstakingly avoid taking shortcuts that might blur the concepts.

Finally, I will challenge the notion of the isomorphism of pitch classes and order numbers and will offer several arguments against accepting this notion. This will provide a better understanding of the two realms of twelve-tone rows.

I have confined myself to the examination of the twelve-tone-row similarity measures. I nevertheless point out some ways in which these could be extended to pitch-class sequences other than twelve-tone rows. Such extension would require discernment between two possibilities: the permutational and the non-permutational. In all of the measures of twelve-tone-row similarity discussed in this work it is immaterial that twelve-tone rows are permutations of precisely twelve elements – they could be applied to permutations of any (finite) number of distinct elements. Consequently, the measures could be extended to cover the similarity of permutations of any number of elements as long as the entities contained the same ones and no duplications. In a non-permutational approach the pitch-class sequences to be compared may have different pitch-class contents or pitch-class duplications, or they could be of different lengths. Some of the similarity measures could also be extended to apply to such situations. I would like to stress, however, that in that case we would lose the speciality of the permutational approach and would be back to the combinational approach.

Andrew Mead aptly summarizes the motivation for exploring the properties of rows and row transformations.

> [W]e can have faith that a close scrutiny of the twelve-tone topography will enrich our sense of its syntax, our familiarity with the terrain will enable us to imagine compelling pathways through it, and we will be able to make music with the freedom, power, and conviction such knowledge affords. (Mead 1989, 227)

The same could be said about the motivation for studying the relations between twelve-tone rows. The chart of the pathways through the topography of twelve-tone rows remains incomplete, and I have only scratched the surface of the ways in which we might conceive of them and their relations. If anything, this work will suggest that the set of twelve-tone rows could be forged to exhibit multiple topographies.

CHAPTER 2

Twelve-tone rows and row operations

The focus in this chapter is on the formalization of twelve-tone rows and row operations in the light of permutations and group theory; these formalizations will prove useful in Part III in which similarity measures are discussed. For readers who are not familiar with the basic properties of permutations, the pertinent concepts are introduced in Appendix A.

2.1 Formalizing twelve-tone rows

2.1.1 Representations of rows

A twelve-tone row is a linear ordering of the twelve pitch classes. It is represented in various ways in the formalization process. This section considers three different ways of representing twelve-tone rows using numeric notation, all of which have their own uses. If necessary, it is easy to switch between the different representations.

Figure 2.1 shows the first theme of the first movement of Alban Berg's *Lyric Suite*. The first twelve pitch classes form a twelve-tone row that is labelled P. The first representation is as a *set of ordered pairs*, introduced by Milton Babbitt (1960). Row P is written below using this representation.

$$P = \{(0,5),(1,4),(2,0),(3,9),(4,7),(5,2),(6,8),(7,1),(8,3),(9,6),(10,10),(11,11)\}$$

The first number in each pair denotes the order number, and the second number denotes the pitch class.[1] For example, the order number of the first pitch class of a row is **0** and the first pitch class with the order number **0** is 5.

The above pairs are written in ascending order by order number. Extracting the succession of pitch classes from these pairs results in the second representation, the *pitch-class row*, as shown

[1]The number denoting the order number in the ordered pairs is not written in bold.

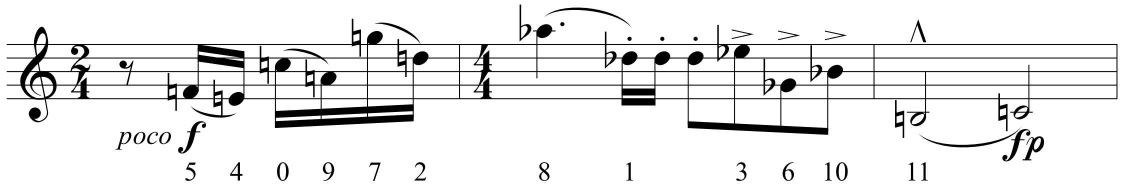


Figure 2.1: The first theme of Alban Berg's Lyric Suite.

below:

$$P = 5\ \ 4\ \ 0\ \ 9\ \ 7\ \ 2\ \ 8\ \ 1\ \ 3\ \ 6\ \ 10\ \ 11.$$

Here the order numbers are implied: the usual left-to-right ordering suggests that the pitch class at order position **0** is 5, and that at order position **1** is 4, etc.

As a set of ordered pairs is an unordered set, the order in which the pairs are enumerated is immaterial. Hence, they can be rearranged without changing the identity of the set. The same pairs as given above are written below in ascending order with respect to the pitch classes:

$$P = \{(2,0),(7,1),(5,2),(8,3),(1,4),(0,5),(9,6),(4,7),(6,8),(3,9),(10,10),(11,11)\}.$$

Extracting the succession of order numbers from these pairs results in the third representation, the *order-number row*, as shown below:

$$P = \mathbf{2\ \ 7\ \ 5\ \ 8\ \ 1\ \ 0\ \ 9\ \ 4\ \ 6\ \ 3\ \ 10\ \ 11}.$$

This representation enumerates the order positions of each pitch class from 0 to 11. For example, the order position of pitch class 0 is **2**, and of pitch class 1 is **7**. The interpretation is that pitch class 0 is the third pitch class (bearing in mind that the first one has the order number **0**).

Hence, if we need to switch between a pitch-class representation and an order-number representation of a row, the representation as a set of ordered pairs functions as a "mediator" since we can arrange the pairs by order number or pitch class to obtain a pitch-class row or an order-number row. Alternatively, as I show in Section 2.1.3, we could use the notion of inverse permutations to toggle between pitch-class representations and order-number representations of rows.

2.1.2 The duality of pitch classes and order numbers

Undoubtedly, representation as a pitch-class row is the most familiar of the three representations. It is the approach that the early serial composers used when writing arrays of row forms with a pen on paper using conventional notation.

While the notion of order numbers goes back to the writings of Babbitt in the 1960s[2], the *locus classicus* of the order-number representation is the prominent series of two articles by Andrew Mead (1988, 1989). Mead argues that pitch classes and order numbers are two dimensions of rows that

[2]In particular, see Babbitt (1960).

must be treated equally, and goes as far as claiming that there is an isomorphism between the two.

> Fundamental to the following discussion is the isomorphism between the abstract structure of unordered pitch classes and the abstract structure of unpitched order numbers inherent in Babbitt's description. (Mead 1988, 97)

Strictly speaking, the use of the word "isomorphism" is misplaced in this context. Two algebraic structures are isomorphic if and only if there exists an isomorphism from one to the other; these structures could be groups or rings, for example. Isomorphism between two structures denotes that they are structurally identical. Neither the set of pitch classes nor the set of order numbers has such a structure, however. They are sets (in the mathematical sense) with no algebraic structure, and hence, by definition, cannot be isomorphic. Instead, as I will argue in Section 2.2.5, row operations have a group structure, and a group of pitch-class operations may be isomorphic to a group of order-number operations.

In addition, we do not necessarily treat the two domains equally. For example, pitch classes 0 and 11 are adjacent, but order numbers $\mathbf{0}$ and $\mathbf{11}$ are not considered adjacent unless we consider a row to be a cyclic structure, which is not necessarily the case.[3] Consequently, even if it were possible to apply the concept of isomorphism to pitch classes and order numbers, they would not be isomorphic. I will return later to this issue with some new arguments.

Nevertheless, the notion of the two domains proves important in this study. I will show in Part III that even if we usually represent rows as pitch-class rows, in the case of similarity the order-number rows have had much more prominence in the literature. This coincides with Babbitt's permutational approach: all twelve-tone rows contain precisely the same elements and it is the ordering relations that give them their individual characteristics.

Finally, I should mention one curiosity: twelve-tone rows are usually written as pitch-class rows, but the majority of the similarity measures discussed in Part III are defined in the order-number domain since they measure the similarity between the order relations of the pitch classes in rows.

2.1.3 Applying the theory of permutations

The permutational nature of the twelve-tone system is well served by the use of the mathematical theory of permutations. As discussed in Appendix A, permutations have two related meanings: a linear ordering and a bijective mapping from a (finite) set into itself. Conceptually, the idea of a twelve-tone row corresponds to the notion of permutation as a linear ordering (rather than as a mapping). As discussed above, the twelve-tone row could be seen as a pitch-class row or as an order-number row, being a linear ordering of the pitch classes in the former sense and of the order numbers in the latter.

Since both pitch-class rows and order-number rows are linear orderings, they are best represented by the one-line notation. For example,

$$5\ 4\ 0\ 9\ 7\ 2\ 8\ 1\ 3\ 6\ 10\ 11 \tag{2.1}$$

[3] For example, when speaking about the unordered INT contents (as in the context of "all-interval rows") of a row we assume that a row has eleven intervals, which implies that it is not a cyclic structure.

stands for a pitch-class row and

$$(2.2) \qquad \mathbf{2\ 7\ 5\ 8\ 1\ 0\ 9\ 4\ 6\ 3\ 10\ 11}$$

stands for an order-number row. Both represent the same twelve-tone row.

There is a natural one-to-one correspondence between the linear orderings of n elements and the bijective mappings from n elements onto themselves. Hence, we could interpret the linear ordering (pitch-class row) in Formula 2.1 as a mapping as follows:

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 5 & 4 & 0 & 9 & 7 & 2 & 8 & 1 & 3 & 6 & 10 & 11 \end{pmatrix}.$$

Correspondingly, we could interpret the linear ordering (order-number row) in Formula 2.2 as a mapping as follows:

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 2 & 7 & 5 & 8 & 1 & 0 & 9 & 4 & 6 & 3 & 10 & 11 \end{pmatrix}.$$

These two permutations are inverse permutations: their product yields the identity permutation. This observation proves useful in computations involving both pitch-class rows and order-number rows. Similarly, the interpretation of a linear ordering as a mapping proves useful in Section 2.2.5 when I proceed to define the application of a row operation to a row: a binary operation between a row operation (permutation as a mapping) and a twelve-tone row (permutation as a linear ordering) is defined as if the latter were also a permutation as a mapping, and again it is used only as shorthand. Nevertheless, it should still be kept in mind that, strictly speaking, linear orderings and mappings belong to different domains: mappings define a group structure but linear orderings do not.

2.2 Formalizing row operations

2.2.1 Background

A significant part of twelve-tone theory involves the analysis of *row operations*. Indeed, several central concepts in the theory, such as *row class* and *invariance*, make use of the concept of row operations. In particular, obtaining new rows via row operations has been an essential part of the twelve-tone system since its inception. To provide for variety and compositional materials, composers generally use several different rows in twelve-tone compositions, and usually, although not always, these are related by some (well-defined) row operation. In order to provide variance a composer needs the means to generate a set of related rows.

The conception of row operations evolved during the 20th century. The classic row operations – transposition, inversion, retrograde, and their combinations – have their roots in the contrapuntal tradition. According to the transformational view, a row operation transforms twelve-tone rows into other twelve-tone rows in some well-defined and predictable manner and is thus described as an algorithm. These algorithms are not complicated: for example, “transpose every pitch class by n semitones” is an algorithm for transposition T_n.

Despite their apparent straightforward nature, there is no agreement on the definition of row operations. Several authors have suggested different sets of row operations – promoting some and debunking others. These sets are not gathered in a random fashion, but are intended to form a "coherent whole." One method for creating this coherence is to define operations in terms of groups. In this section I discuss why and how group theory could be used to formalize row operations and to explicate what is meant by a coherent whole. Indeed, I will argue that to make a coherent whole is to have a group structure.

I have four goals. First, the literature on row operations has been developed piecemeal over time and it is relatively fragmented. Authors have suggested various families of row operations. I will expand on this discussion and introduce a more general notion – one that includes as many as 479001600 pitch-class operations and 479001600 order-number operations.

Secondly, while group theory was an integral part of twentieth-century music theory, its applications have been relatively limited, its role being almost that of an epiphenomenon. My aim is to examine some of our common musical wisdom in terms of group theory, an excursion that will reveal some grey areas in the thinking on row operations. In particular, I will ask the following two questions. Can we always compose row operations to obtain new row operations? Are the pitch-class dimension and the order-number dimension as "isomorphic" as has been claimed in the literature?

Thirdly, the literature on row operations contains some errors and misconceptions. My approach is novel in the sense that I separate the mathematics and the music. I make no attempt to write a theory of row operations that is parallel with mathematical group theory, but rather formulate the theory in the language of mathematics and then translate the concepts of music theory into the mathematical concepts, thereby avoiding some of the inaccuracies and misconceptions found in the existing literature.

Fourthly, I make a distinction between the row operations and their "effect" on the set of twelve-tone rows – how they transform rows into other rows, and present the latter in terms of a *permutation group*. This concept has been hinted at in the literature. For example, Milton Babbitt writes as follows.

> The totality of twelve transposed sets associated with a given S constitutes a permutation group of order 12; as such it is closed, disjunct with regard to any other collection of sets T derived from a set whose intervallic succession differs from that of any member of this totality. (Babbitt 1960, 249)

However, the full details of the structure have not been thoroughly studied and the relations between twelve-tone rows, row operations and the permutation group are sometimes blurred – indeed, in the above citation the permutation group is associated with the rows themselves and not with the permutations of the set of twelve-tone rows, as should be the case. The careful separation of all the components and their roles is needed in order to produce a comprehensive theory of twelve-tone operations. It is also my conviction that such a separation reflects current thinking on row operations.

Perhaps the closest precedent for the present work is the permutational model of voice leading developed by Henry Klumpenhouwer (1991). He bases his model on the Generalized Interval System devised by David Lewin (1987). I will show in Section 2.2.3 that we cannot define row operations in terms of a Generalized Interval System, however, unless we focus on some specific subsystem of rows and row operations (such as the 48 classic row operations and some specific non-symmetric row class). The Generalized Interval System is not versatile enough to embrace a general theory of twelve-tone rows and row operations.[4]

2.2.2 The group structure of row operations

We intuit row operations to have certain properties, which match the requirements of a group structure. In the following I discuss a number of concepts from group theory and explore how they could be used to examine row operations.

Group theory is a branch of abstract algebra that concerns the properties of structures with a binary operation that satisfies certain requirements. It holds a prominent position in mathematics and has applications in virtually all exact sciences. It is also a theory that frequently features in the literature on music theory. Milton Babbitt (1946/1992, 1960, 1961b) was the first to advocate its use in the context of twelve-tone music, and David Lewin (1987), Robert Morris (1987, 2001), and Henry Klumpenhouwer (1991) provide extensive discussions. In particular, Klumpenhouwer gives a clear exposition of group theory using permutations. However, none of these authors provides quite the scope that is required here: it is my intention to translate the concepts of music theory into conventional mathematical terminology, not vice versa.

The following discussion has two aims. I will first review some fundamentals of basic group theory, and then present my argument as to *why* we should use group theory and how its concepts correspond to some of our musical intuitions.

A note on notation is due. When the context is purely that of group theory, whenever possible, I will use uppercase letters for sets and lowercase letters for elements of those sets. If the elements are sets I will use uppercase letters. When I apply these results to rows and row operations, however, I will use the (relatively) standard notation. Hence, I will use uppercase letters for rows and row operations even if they are elements of a set.

I will begin by defining a group.

DEFINITION 2.1 A *group* is a nonempty set G with the following four properties: (i) G has a binary operation. We say that G is *closed* under this operation. In other words, for every element a and b in G the result of the binary operation $a \cdot b$ is a member of G. (ii) The binary operation is associative. In other words, for every element a, b and c in G the following equation holds: $a \cdot (b \cdot c) = (a \cdot b) \cdot c$. (iii) G contains an identity element e such that the equations $a \cdot e = a = e \cdot a$ hold for all $a \in G$. (iv) Every element $a \in G$ has an inverse element a^{-1} such that $a^{-1} \cdot a = e = a \cdot a^{-1}$.

I claimed at the beginning of this section that the properties of row operations match the require-

[4]Nevertheless, I use the Generalized Interval Systems extensively in Part III of this work, in which I discuss the similarity of twelve-tone rows.

ments of a group structure. This claim is backed up by the following observations on transpositions.[5] (i) Row operations can be composed: we expect that a composition of two row operations is also a row operation. This correlates with the requirement of closure in a group (i.e., that the group is closed under the binary operation). In the case of transpositions, the composition of two transpositions is a transposition. For example $T_4T_2 = T_6$. (ii) In the composition of three row operations the order in which they are composed is insignificant. For example, $T_1(T_2T_3) = T_6 = (T_1T_2)T_3$. Incidentally, the order of composition is not to be confused with the order of the row operations themselves. The latter is known as commutativity and it means that we can reorder the row operations ad libitum. While transpositions are commutative, row operations, in general, are not. For example, it is well known that the order of transpositions and inversions cannot be changed without changing the result.[6] (iii) The existence of an identity element, a row operation that does nothing, is necessary for the inverse operations. In the case of transpositions, T_0 is the identity element.[7] (iv) For every row operation, we can undo or reverse it, returning to the point of origin. This idea corresponds to the idea that every operation has an inverse operation that undoes it. Every transposition has an inverse transposition: the inverse element of T_n is T_{12-n}. For example, $T_4T_8 = T_0$.

Group theory codifies a natural and coherent set of properties that correlate with the properties we intuit row operations to have, which is exactly why we want to use it.

2.2.3 Operations acting on rows

In Section 1.2.2 I committed myself to the idea of distinguishing between objects and operations, in this case twelve-tone rows and row operations. Now it is time to link these two concepts. I prepared the ground in the previous section by examining the group structure of row operations, and will now show that *group action* may be used to formalize the way in which row operations transform rows into other rows. I will begin by defining group action formally.

DEFINITION 2.2 If G is a group and S is a set, then the mapping $G \times S \to S \colon (g, s) \mapsto g \cdot s$ is called the (left) action of the group G on the set S if the following two criteria are satisfied. (i) If e denotes the identity element of the group G, then $e \cdot s = s$ for all $s \in S$, and (ii) $g_1 \cdot (g_2 \cdot s) = (g_1 g_2) \cdot s$ for all $g_1, g_2 \in G$ and for all $s \in S$.

In essence, in the action of group G on set S we multiply the elements of set S by an element of group G and the result is an element of set S.[8] The two criteria of the definition guarantee that the resulting binary operation has an orderly and therefore predictable structure. The first criterion states that the multiplication by the identity element of group G leaves every element in S intact, and the second that multiplying an element first with an element g_2 of G and then with another element

[5] Similar observations could be made about virtually any set of operations: transpositions were chosen as the example due to their simplicity and familiarity.

[6] For an excellent discussion about the relation of transpositions and inversions, see Lewin (1987, 46–59).

[7] When I refer to groups in general, I will use e to denote the identity element, and when referring to row operations, I will use T_0.

[8] Multiplication is used here as the generic word for a binary operation; depending on the elements of group G and set S, "compose" or "add" might be more appropriate.

g_1 of G gives the same result as multiplying by the element g_1g_2 of G (since group G is closed under composition, g_1g_2 is a member of G). Clearly, this requirement is related to the associativity of groups.

Let us illustrate the action of a group by using the group of transpositions and the set of twelve-tone rows. Let group G be the group of transpositions and let set S be the set of all twelve-tone rows. The action of group G on set S is to apply a transposition to the twelve-tone rows. The first criterion of group action requires that applying T_0 – the identity element of group G – to any row should keep the row intact, and the second that for any two transpositions T_m and T_n, applying their composition $T_{n+m} = T_nT_m$ to any row should yield the same result as applying first transposition T_m and then transposition T_n.

It is important to note what is under definition here is the multiplication of two entities of different qualities: an operation and a twelve-tone row. I have defined an *external* binary operation on the set of twelve-tone rows given the multiplication of elements that are not members of the same set.

The notation of the left group action is in accordance with left orthography:[9] we take a row P and when we apply a transposition to it we write is as T_nP. As a convention, we do not write the multiplication symbol $T_n \cdot P$.

A note on terminology is due. David Lewin (1987) distinguishes between *transformations* and *operations*. In his definition 1.3.1 a function from a set S into S is termed a transformation, and if a transformation is one-to-one and onto (that is, a bijection), it is an operation on S. The terminology here is slightly different since a group of row operations is neither from nor into the set of twelve-tone rows. The row operations induce a permutation group on the set of twelve-tone rows, however. In Lewin's terminology, the permutations in the permutation group (induced by row operations) are transformations and, since permutations are one-to-one and onto, are also operations. Nevertheless, I refer to row operations as operations or transformations practically synonymously: the former is associated more with the relations of row operations, and the latter with the "effect" that the row operations have on twelve-tone rows.

Let us now consider some possible properties of the action of group G on set S. A group action is *transitive* if for every $s_1, s_2 \in S$ there exists $g \in G$ such that $gs_1 = s_2$, in which case the cardinality of group G cannot be smaller than the cardinality of set S.[10] Further, the group action is *free* if for all $g_1, g_2 \in G$ and for all $s \in S$, $g_1s = g_2s$ implies $g_1 = g_2$. In other words, no two elements of a group map any element similarly. If a group action is free, the cardinality of group G cannot be larger than the cardinality of set S. Finally, let us say that a group action is *simply transitive* if it is both transitive and free.

In terms of twelve-tone rows and row operations, the action of a group of row operations on the set of twelve-tone rows is transitive if for any given two rows P_1 and P_2 there is a row operation F such that $FP_1 = P_2$. The action of a group of row operations on the set of twelve-tone rows

[9]Correspondingly, I could have defined right group action.

[10]It should be noted that the transitivity of a group action is a different notion from the transitivity of a relation. However, as they are defined in different realms – group action and (binary) relations – the context should always indicate in which sense the word "transitivity" is used.

is free if no two distinct row operations transform any row similarly. For example, the action of the 48 classic row operations on the set of all twelve-tone rows is neither transitive nor free. It is not transitive since, for example, none of the operations maps row $0123456789\mathrm{AB}$ (the "chromatic scale") into row $05\mathrm{A}3816\mathrm{B}4927$ (the "circle of fifths"), and it is not free since both operations T_0 and RI_{11} map row $0123456789\mathrm{AB}$ into itself. However, if we take a non-symmetric row class, the action of the 48 classic row operations is simply transitive: the action is free since all operations transform a selected row form into different row forms, and it is transitive since there exists a row operation that transforms a selected row form into any row form in the row class.

I examined the relations between row operations in Section 2.2.2. Group action opens up another perspective. The focus in this section is thus on the effect that a row operation has on the set of twelve-tone rows as a whole – how it transforms rows into other rows and how these transformations could be viewed as permutations of the set of twelve-tone rows – rather than on row operations themselves.

A row operation transforms twelve-tone rows into other twelve-tone rows, thereby producing a *permutation* of the set of twelve-tone rows. Informally, we might think that if we enumerate all twelve-tone rows in some order, a row operation "reorders" the set of twelve-tone rows. For example, if we transpose all twelve-tone rows by a semitone, we would obtain the very same twelve-tone rows but in a different order.

Thus, we could think of a group action, the application of a member of the group to a set of elements, as a permutation of that set of elements. This permutational nature is well-known and could be summarized in the following two observations. First, when an element of the group permutes the element of the set, no two distinct elements are permuted to the same element.[11] For example, if X and Y are two distinct twelve-tone rows, then T_nX and T_nY are also two distinct twelve-tone rows – in other words, $T_nX = T_nY$ if and only if $X = Y$. Secondly, for every element of the set there is exactly one element that maps into it under the action of a given element of the group. For example, if X is a row and T_n is a transposition, there is exactly one row Y such that $T_nY = X$.

The action of a group on a set induces a *permutation group*, the elements of which are permutations and the binary operation is the composition of these permutations. The four requirements of a group (Definition 2.1) are thus satisfied. First, when we compose two permutations defined by two elements of the group the result is a permutation defined by the composition of the two elements, which is condition (ii) of Definition 2.2 above. Secondly, permutations are mappings and we know from the properties of mappings that they are associative. Thirdly, the identity element of the group of operations induces a trivial permutation that keeps every element fixed, which is condition (i) of Definition 2.2 above. This permutation is the identity element of the permutation group. Finally, the permutations defined by inverse operations induce permutations that are inversions of each other. We can derive this from condition (ii) of Definition 2.2 above by setting $b = a^{-1}$. Hence

[11] This follows directly from the second criterion of group action: $gs_1 = gs_2$ implies $g^{-1}gs_1 = g^{-1}gs_2$ implies $s_1 = s_2$.

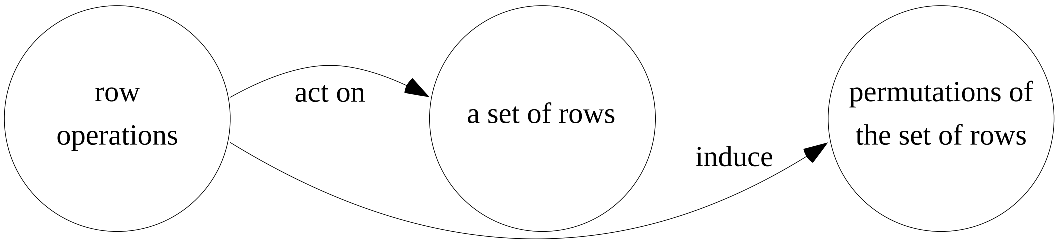


Figure 2.2: A group of operations, the set of twelve-tone rows, and the induced permutation group in pictorial form.

the following theorem is now proved.

THEOREM 2.1 A group acting on a set induces a permutation group on that set.

Again, we can illustrate the theorem by using the group of transpositions acting on the set of twelve-tone rows. The composition of the two permutations induced by the operations T_n and T_m equals the permutation induced by the operation T_{n+m}. Transposition T_0 is the identity element of every group of row operations and it induces the identity permutation. Finally, the permutation induced by transposition T_n is the inverse of the permutation induced by transposition T_{-n} (the inverse of transposition T_n).

I should stress that the group acting on a set and the resulting permutation group are two distinct groups. In fact, they are not necessarily even isomorphic, although a homomorphism from the first to the latter is easily obtained. If group G acting on a set S results in a permutation group H then the mapping $f : G \to H$, in which element $g \in G$ is mapped into the permutation it induces, is a homomorphism.

2.2.4 Twelve-tone rows, row operations, and permutations of rows

Figure 2.2 presents in a pictorial form the above concepts of the group, the set, and the induced permutation group, translating them into terms of row operations, rows, and permutations of a set of rows. Thus, a group of row operations (depicted on the left) acts on a set of twelve-tone rows (depicted in the middle). The group action induces a permutation group (depicted on the right).

It is extremely important to note the three separate entities in Figure 2.2. The first is the group of row operations, the second is the set of twelve-tone rows, and the third is the permutation group with permutations of rows as its elements. The danger lies in confusing the group of row operations with the induced permutation group – while in most cases (but not all) these two are isomorphic, they are nevertheless conceptually separate entities. We usually think of row operations as elements of the group of row operations. For example, T_2 is the row operation that maps pitch class 0 into pitch class 2, pitch class 1 into pitch class 3, etc. The permutation group that operation T_2 induces is simply an enumeration of which row is transformed into which row. For example, if we apply row operation T_2 to all twelve-tone rows, in the resulting permutation row $P = 5409728136\text{AB}$ is transformed into row $T_2P = 762\text{B}94\text{A}35801$, row $T_2P = 762\text{B}94\text{A}35801$ is transformed into row $T_4P = 9841\text{B}6057\text{A}23$, etc.

As stated earlier, the group of row operations and the resulting permutation group are usually isomorphic. However, a typical case in which they are not is when the focus is on some subset of twelve-tone rows. For example, the *Lyric Suite* row $P = 5409728136\mathrm{AB}$ is invariant under operation RT_6 and therefore there are only 24 distinct forms in its row class. If we now allow the group of 48 classic row operations to act on this set of rows (the 24 members of the row class), the induced permutation group contains only 24 distinct members. Hence, in this case the group of classic row operations and the induced permutation group are not isomorphic. If the group of 48 classic row operations acts on the set of all twelve-tone rows, it is isomorphic to the induced permutation group.

This separation into three different components is an attempt to formalize the way in which we conceptualize row operations. If it was only the formal aspect that was under consideration, it would be possible to define them purely in terms of the permutation group, and thus to define operation T_2 as the permutation of the set of all twelve-tone rows that transforms row $P = 5409728136\mathrm{AB}$ into row $T_2P = 762\mathrm{B}94\mathrm{A}35801$, row $T_2P = 762\mathrm{B}94\mathrm{A}35801$ into row $T_4P = 9841\mathrm{B}6057\mathrm{A}23$, etc. However, I am convinced that this is not the way to conceive of row operations: transposition T_2 is the row operation that maps pitch class 0 into pitch class 2, pitch class 1 into pitch class 3, etc. We would then apply this operation to twelve-tone rows. This group of transpositions is the same group that we would use to transpose pitch-class sets, for example, except that the set that the group acts on is different in these two cases. Hence, an accurate formalization of this conceptualization requires the three components outlined above.

More technically, as the row operations act on the set of 479001600 twelve-tone rows, the resulting permutation group is a subgroup of the group that is isomorphic to group $S_{479001600}$. Many of the properties of row operations can be explained in terms of that group. For instance, the row classes are the orbits of the group of row operations, the symmetries of rows correspond to the nontrivial kernel in the restriction of the group action on an orbit, and so on. The full development of the properties of group $S_{479001600}$ with respect to the row operations lies beyond the scope of this work, however.

2.2.5 Pitch-class operations and order-number operations

There is a need to clarify two categories of row operations: pitch-class operations and order-number operations. Furthermore, these two concepts play a prominent role in the development of the two Generalized Interval Systems for the transformational representation of similarity measures discussed in Section 5.4. The two categories of row operations are defined as follows.

DEFINITION 2.3 A row operation is a *pitch-class operation* if the mapping of the pitch classes is the same in every row, and an *order-number operation* if the mapping of the order numbers is the same in every row.

For example, transposition T_1 is a pitch-class operation, since pitch class n is transformed into pitch class $T_1(n)$ in every row. Similarly, retrograde is an order-number operation, since order number $\mathbf{n}$ is transformed into order number $R(\mathbf{n})$ in every row.[12] According to this definition, transposition T_0 is both a pitch-class operation and an order-number operation: it is the identity operation, which is customarily labelled transposition T_0 but which could just as well be labelled rotation r_0.

Note that the idea of the division of row operations into pitch-class operations and order-number operations is not new.[13] What is new here is that any permutation is allowed to act as a pitch-class operation or an order-number operation. Consequently, there is a total of 479001600 pitch-class operations and 479001600 order-number operations.

Not all row operations are pitch-class operations or order-number operations: examples include compositions of a pitch-class operation and an order-number operation (which is not the identity operation) and the exchange operation.[14]

Let us now explore the idea of a pitch-class operation in more detail. By definition, each pitch class is transformed identically in every row. Therefore, we can define the operation by enumerating how each one is transformed. Assume that pitch class 0 is transformed into pitch class x_0, pitch class 1 into pitch class x_1, etc. We could then represent the operation as a permutation that defines how each pitch class is to be transformed:

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ x_0 & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 & x_8 & x_9 & x_{10} & x_{11} \end{pmatrix}.$$

Each pitch class in the upper row is transformed into the corresponding pitch class of the lower row.

For example, in transposition T_2 pitch class 0 is transformed into pitch class 2, pitch class 1 is transformed into pitch class 3, etc. We could then represent transposition T_2 as a permutation that defines how each pitch class is to be transformed:

$$T_2 = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 0 & 1 \end{pmatrix}.$$

In Section 2.2.3 I applied a row operation to a row and defined it as a (left) action of a group. A given pitch-class operation is represented as a permutation, and a row as a permutation denoting a pitch-class row: the action of the row operation is then a multiplication of the row by the operation from the left. For example, if we apply transposition T_2 to the *Lyric Suite* row $5409728136\mathrm{AB}$, we

[12] $R(\mathbf{0}) = \mathbf{11}$, $R(\mathbf{1}) = \mathbf{10}$, and so on.

[13] See footnote 3 in Headlam (1985) for a brief history of the development of extended pitch-class operations and order-number operations.

[14] The exchange operation was discovered by three authors independently of each other: Walter O'Connell (1962), Larry Solomon (1973), and Michael Stanfield (1984, 1985). (See also Headlam (2006).) All three authors present geometrical symmetries as the background for the operation; in fact, Solomon describes it only in terms of geometry, whereas O'Connell and Stanfield also provide an algebraic formula. The geometric and algebraic definitions are, naturally, equivalent.

obtain[15]

$$
\begin{aligned}
T_2P &= \overbrace{\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 0 & 1 \end{pmatrix}}^{T_2} \overbrace{\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 5 & 4 & 0 & 9 & 7 & 2 & 8 & 1 & 3 & 6 & 10 & 11 \end{pmatrix}}^{P} \\
&= \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 7 & 6 & 2 & 11 & 9 & 4 & 10 & 3 & 5 & 8 & 0 & 1 \end{pmatrix}.
\end{aligned}
$$

As discussed in Section 2.1.3, twelve-tone rows are permutations in the sense of linear ordering. Here I also use the two-line notation for rows (implying permutation in the sense of a mapping).

Similarly, a given order-number operation is represented as a permutation, and a row as a permutation denoting an order-number row: the action of the row operation is a multiplication of the row by the operation from the left.

It follows naturally from these observations that the composition of two pitch-class operations is a pitch-class operation (since the composition of two permutations is a permutation). Correspondingly, the composition of two order-number operations is an order-number operation.

According to the terminology introduced in Section 2.2.3, the action of *any* group of pitch-class operations on the set of twelve-tone rows is free. This follows from the group structure of permutations and the fact that the application of a pitch-class operation to a row is defined as a group of permutations acting on permutations. Therefore, if P is a permutation representing a row and F and G are permutations representing row operations, then

$$FP = GP \Rightarrow FPP^{-1} = GPP^{-1} \Rightarrow F = G,$$

and thus no two distinct pitch-class operations transform any row similarly.

All possible pitch-class operations are permutations of the group S_{12}. The action of this group on the set of twelve-tone rows is transitive, since for any two permutations P and Q representing rows there is a pitch-class operation F such that $FP = Q$, since defining $F = QP^{-1}$ gives

$$FP = (QP^{-1})P = Q(P^{-1})P = Q.$$

Similarly, the action of *any* group of order-number operations on the set of twelve-tone rows is free, and the action of the group of all possible order-number operations S_{12} is transitive. Since the action of a group of order-number operations is both transitive and free, it is simply transitive.

It was noted in Section 2.1.2 that there is a duality in the pitch-class realm and the order-number realm: twelve-tone rows can be represented equally well as pitch-class rows or as order-number rows. It was also noted that pitch-class operations and order-number operations are identical, but the former are applied to pitch-class rows and the latter to order-number rows. Is it then possible to draw the conclusion that any operation is equally applicable to both domains? Andrew Mead claims

[15]See Appendix A for a walk-through of composing permutations.

that this is the case.

> By describing twelve-tone rows as collections of pairs of order numbers and pitch classes, he [Milton Babbitt] provided a means of conceiving of the transformations of rows as the result of identical sorts of operations performed in two dimensions. It follows logically from this description that the two dimensions, conceived abstractly, are isomorphic, and operations normally associated with one dimension can be equally applied to the other. (Mead 1989, 181)

However, consideration of the family of twelve row operations described by David Lewin and known as M-operations shows that this is not necessarily the case.[16] Namely, Lewin's M-operations are defined on the order-number domain and they should be applied to that domain only, since applying them to the pitch-class domain would force us either to drop transpositions from the catalogue of row operations, or to accept a dramatic increase in the number of operations available. Hence, the least damage is done if we drop the claim that any operation can be equally well applied to both domains.

[16]See Lewin (1966) and Mead (1989).

Part II: Similarity

CHAPTER 3

Similarity and its properties

3.1 Introduction

The ability to evaluate the similarity or resemblance between musical objects, such as pitch-class sets or twelve-tone rows, is an essential skill for a music analyst. The notion of resemblance between musical objects is intuitive in the sense that it is easier to spot the resemblance between them than to elaborate on why they resemble each other. The study of similarity measures could be seen as an effort to explicate this notion.

Creating similarity measures per se is not the aim here: I am more interested in the notions of similarity behind them. Nevertheless, there are at least two reasons for proceeding to create formal measures. First, a similarity measure as an explication is useful for solving disputes. If two people disagree on whether two entities are similar with respect to property X, then either they have different conceptions of what it means to be similar with respect to property X or one of them is mistaken. A similarity measure provides a means of clarifying the characteristics of property X, and enables us to check whether the entities are similar.

Secondly, a formal similarity method provides the tools with which to examine what it means to be similar with respect to some property. Are most entities similar in this respect, or is the similarity something extraordinary? Are there degrees of similarity? Answering these questions adequately requires us to formalize the notion of similarity and to examine the distribution of values in a similarity measure (see Section 6.3).

Two musical gestures on a musical surface that are similar in one respect may be dissimilar in another. For example, the sequences of the same pitch classes may have different contours. Nevertheless, I adopt John Rahn's (1989, 9) view, echoed by Eric Isaacson (1996, ¶3), that an effective context-sensitive measure will need to be based on a suitable context-free similarity measure. Correspondingly, similarity measures are typically established for classes of objects, such as pitch-class sets or twelve-tone rows, and not for musical gestures. These are, of course, abstractions, and the immediate perceived similarity of their realizations in the pitch space depends on the properties of the realizations – in addition to the similarity of the underlying abstract structures. Nevertheless, two pitch realizations of set classes/twelve-tone rows cannot be perceptually similar in all respects unless the underlying set classes/twelve-tone rows are similar.

The majority of music-theory literature concerning similarity is about the similarity of set classes, and references to twelve-tone rows (or pitch-class segments in general), pitch sets, and contours are much more scattered. Consequently, most of what has been written on the properties of similarity, and the aims, scope and limits of the research, is based on the similarity measures for set classes. Hence, even if the intention here is to lay the foundations for the similarity measures for twelve-tone rows (to be discussed in Part III), much of the discussion involves those for set classes.

Nelson Goodman has argued that there is no similarity per se, and that there is only similarity with respect to a given property (Goodman 1972b). I take a neutral position in this debate as I will describe similarity in terms of similarity measures. Similarity measures are technical tools, and each of them focuses on similarity with respect to some property. Furthermore, I acknowledge in the following that twelve-tone rows may be similar with respect to one property and dissimilar with respect to another; hence, a claim that two rows are similar – without specifying in what respect they are similar – would be poorly defined.

The rest of Part II concerns the properties of similarity measures. Conceptions of the nature of similarity are reflected in the criteria by which the measures are evaluated. Using this discussion as a basis, I will introduce and analyze a selection of similarity measures for twelve-tone rows in Part III.

3.2 Terminological issues

The terminology concerning similarity measures suffers from a lack of consistency in the literature on music theory. At least, the terms "similarity relation" (Rahn 1979–80; Quinn 2001), "similarity function" (Lord 1981; Demske 1995a; Buchler 1997), "similarity index" (Teitelbaum 1965; Morris 1979–80; Roeder 1987), "similarity index function" (Demske 1995a) and "similarity measure" (Castrén 1994; Buchler 1997; Kuusi 2001) have been used in overlapping and mutually incompatible ways.

The terms "relation" and "function" are borrowed from mathematics where they have firmly established meanings. The formal details of these two terms are discussed in Appendix B. Here it suffices to informally note that a (binary) relation on a set defines which of the elements of that set are related. A function is a mapping in which every element of the domain of the function is mapped into an element of its range. Ian Quinn points out that if we wish to resort to (axiomatic) set theory we could say that functions are a subset of relations (Quinn 2001, 111). However, there is a clear distinction between the everyday mathematical use of the words "relation" and "function": if something is a function mathematicians call it a function, not a relation, unless the context (such as axiomatic set theory) absolutely requires it. Therefore, addressing similarity relations, that are actually functions, as relations is as misleading as speaking about animals flying airplanes when we mean that the pilots are humans. Hence, if we accept the common (mathematical) usage of the terms "relation" and "function," then the very use of the phrase "similarity relation" is unfortunate since only a few of the similarity relations are actually relations.

For example, Allen Forte's similarity relation R_p is a (binary) relation on the set of set classes. It defines a set of pairs of set classes that are related (Forte 1973b). In contrast, those that are called similarity relations are typically not relations on a set, but are functions from the Cartesian product of the set of set classes to some range of values – typically a subset of nonnegative real numbers. For example, for every pair of set classes Robert Morris' similarity measure ASIM gives a number that describes their degree of relatedness (Morris 1979–80).

I prefer the term "similarity measure," which comprises both similarity relations and similarity functions.[1] Hence, the use of the terms "relation" and "function" would correspond to their use in mathematics. When there is a need to distinguish between similarity measures for set classes, twelve-tone rows and row classes, they are referred to as set-class similarity measures, twelve-tone-row similarity measures, and row-class similarity measures. I will also abbreviate twelve-tone-row similarity measures as row measures, and row-class similarity measures as row-class measures.

3.3 Twelve-tone rows and similarity

My aim in this work is to discuss the similarity of twelve-tone rows. The similarity of rows is different from the similarity of set classes in some important ways. The discussion is based first and foremost on the similarity of individual rows, not row classes. Set classes are equivalence classes that comprise one or several pitch-class sets. Hence, the level of abstraction is higher in set-class similarity measures than in row measures.

There is also an equivalence relation in the realm of twelve-tone rows: the row class. Consequently, we need to consider the connection between similarity and equivalence relations, and in particular, to explore whether equivalence implies similarity or vice versa.

Equivalence implies similarity in the realm of pitch-class sets: according to practically all similarity measures, equivalent pitch-class sets (members of the same set class) are maximally similar. However, this turns out not to be the case for twelve-tone rows: membership in the same row class does not guarantee similarity. In fact, in several cases two rows belonging to the same row class turn out to be maximally dissimilar. Hence, the similarity of twelve-tone rows is different from that of set classes.

It should not come as a surprise that two members of the same row class are not necessarily conceived of as similar. The perception of twelve-tone rows has been studied empirically every now and then, and a long series of psychological experiments has shown that subjects have difficulties in recognizing retrograde-related row forms, for example. In their pioneer studies Francés (1958) and Chailley (1961) referred to the difficulties the subjects had in discriminating twelve-tone rows

[1] Michael Buchler makes a similar distinction in which the similarity function is the most general category that corresponds to the similarity measure in this context. He then divides similarity functions into similarity relations that are relations in the mathematical sense, and similarity indexes that are functions in the mathematical sense (Buchler 1997, 18).

as disastrous for twelve-tone music. Their argument was that the twelve-tone system was conceptual but not perceptual (de Lannoy 1972, 13). Diana Raffman reiterates the same argument and concludes that twelve-tone music is artistically defective (Raffman 2003). Later Dowling (1972), de Lannoy (1972), and Krumhansl, Sandell and Sergeant (1987), for example, tested perception of twelve-tone rows with varying results depending on the difficulty of the task.

However, it could be argued that in a *musical* context the discrimination of row forms is not necessary. The aim in listening to serial music is not to identify row forms on-the-fly. Indeed, Schoenberg himself resented the idea of people “counting tones” in his pieces and did not even want to disclose his method at first (Schoenberg 1975, 214). Therefore the criticism of Francés and Chailley misses the mark.

One could draw the conclusion that equivalence (membership in the same row class) does not guarantee immediate perceptual similarity: equivalence is based on a convention. Naturally, similarity does not guarantee equivalence, either. Two rows may be similar even if they are not members of the same row class.

Ian Quinn argues that equivalence should be at the one end of similarity.

> We are led, therefore, to understand that the relationship between the predicates IS–SIMILAR–TO and IS–EQUIVALENT–TO is a closer one than theorists often acknowledge. (Quinn 2001, 118)

Quinn has set-class similarity measures as his context and, accordingly, he draws his evidence from them; hence it is unclear whether he intends his statement as a statement about the properties of similarity in general or only as a statement about the similarity of pitch-class sets or set classes. However, in terms of the similarity of twelve-tone rows this is clearly not the case, since practically any of the published row measures (including the new ones introduced in this work) could be used to demonstrate that equivalent rows are not necessarily similar. As the empirical experiments cited above show, equivalent rows are not necessarily perceptually similar. Therefore, at least with respect to twelve-tone rows, equivalence and similarity must be considered two distinct concepts.

Similarity measures could be used to evaluate rows between two different row classes or within a row class. In the latter case some row forms may be more similar to each other than to others. Indeed, examining how similar or dissimilar the rows within a row class are is an interesting way to analyze row classes. Furthermore, a composer might wish to utilize the cases in which two members of a row class are similar (in addition to being equivalent), or even to design his row class in such a way that it provides distinct row forms that are similar in some respect. For example, I will show in Part III that some rows in Alban Berg’s *Lyric Suite* are related both by equivalence (since they belong to the same row class) and by similarity.

Twelve-tone rows are abstractions. Hence, the perceptual similarity of the realizations of rows depends on the way they emphasize or shroud the similarity.[2] Let us take a simple example: if two rows both begin with pitch classes F and E, this shared feature is perceived more readily if

[2]For a review of the relation between set-class similarity measures and the perceptual similarity of instances of set classes, see Chapter 2 in Kuusi (2001).

in both realizations the pitch classes are presented in the same way (the same pitches, the same instrumentation, etc.).

Equivalence relations and similarity complement each other. These two concepts converge in the context of similarity measures for row classes (see Section 6.2).

3.4 Properties of similarity measures

Assumptions about the nature of similarity lurk behind the definitions of similarity measures: these assumptions are sometimes implicit and sometimes brought explicitly into the discussion. There is also a need to consider the relation between similarity measures and similarity: what does a similarity measure measure and how does it do it?

A significant sub-theme in the literature on similarity measures concerns the analysis of their desirable properties. This discussion covers both the descriptive aspect in terms of the properties they have and the prescriptive aspect in terms of the properties we would like them to have, and the way in which similarity measures differ from equivalence relations, for example.

As for the bulk of the discussion on similarity measures deals with the similarity of unordered sets of pitch classes (and, in a few cases, also of unordered sets of pitches), the discussion on their desirable properties has taken place in the context of pitch-class sets. However, Buchler notes the applicability of the criteria to similarity in other objects (Buchler 1997, 18).

I will review the existing literature on the properties of similarity measures in Section 3.6, and discuss what criteria a "satisfactory" similarity measure should fulfill. I will then analyze the proposed criteria and discuss their applicability, mutatis mutandis, to the similarity of twelve-tone rows. The discussion also gives the reasoning behind the choices made in the formulation of the similarity measures in Part III.

One of the major strands in the discussion to follow is the building of an argument for using the mathematical concept of the *metric*, which is discussed in detail in Chapter 4. While the metric is not a sufficient criterion in itself, it embodies some essential properties of what we would consider to be a well-behaving set of values describing the similarity of objects.

3.5 Similarity and dissimilarity

It has been pointed out that many set-class similarity measures in fact measure dissimilarity and not similarity (Buchler 1997, 31). In a dissimilarity measure the greater the value, the more dissimilar the objects are, while in a similarity measure the greater the value, the more similar they are.

Similarity and dissimilarity are seen here as two sides of the same coin, only the focus is different: a similarity measure evaluates the number and significance of the shared features between two musical objects, and a dissimilarity measure evaluates the number and significance of the differentiating features. Nevertheless, they both represent the same continuum: at the one end is similarity and at the other end is dissimilarity.

I use the term "similarity measure" both for measures based on similarity and for those based on dissimilarity, and "dissimilarity measure" when I wish to emphasize that it is based on dissimilarity. As explained in Chapters 4 and 5, the two main innovations in this work – the metric and the transformational approach – assume the use of dissimilarity measures.

The relatedness of similarity and dissimilarity is demonstrated by attempts to transform similarity measures into dissimilarity measures and vice versa. For example, Rahn modifies his own similarity measure ATMEMB to produce the dissimilarity measure DATM (Rahn 1989, 2–3). John Ward, in turn, modifies a number of dissimilarity measures to produce similarity measures (Ward 1992).

3.6 Previous studies

In the following I review the previous literature on the properties of and criteria for similarity measures, and introduce and defend my own approach.

It has been noted that authors are far more inclined to point out the unsatisfactory features of similarity relations than to offer an analysis of the conditions they should meet (Castrén 1994, 17). However, Eric Isaacson (1990) and Marcus Castrén (1994) ventured to create a set of requirements for a successful similarity measure. Isaacson was the first to specify explicit criteria for a well-mannered measure, of which he gives three. Castrén is more fine-tuned, giving six criteria, one of which is divided into four sub-criteria. In addition, Richard Hermann has developed a taxonomy for the classification of similarity measures, but it does not impose any requirements (Hermann 1994). Michael Buchler has commented extensively on the criteria put forward by Isaacson and Castrén, noting that they "were, quite naturally, reflections of their own ideas regarding how two pcsets might be related" (Buchler 1997, 19). Indeed, it would certainly be cumbersome to propose a similarity measure that even the author considered inferior. Ian Quinn has given an extensive criticism of the aforementioned authors' approaches to similarity.

Isaacson posits the following criteria for similarity measures.

> A function measuring the similarity of interval-class content between pc set-classes should: (1) provide a distinct value for every pair of sets; (2) be useful for sets of any size; and (3) provide a wide range of discrete values. (Isaacson 1990, 1)

The two first criteria could be considered requirements for the *domain* of the similarity measure, and the third criterion could be considered a requirement for its *range*. Castrén's first three criteria approximately match the three put forward by Isaacson. In the following I will first consider the domain (Section 3.6.1) and the range (Section 3.6.2), and then in Section 3.6.3 I will turn to Castrén's fourth criterion dealing with the internal coherence of similarity measures. Castrén's two last criteria are specifically related to the similarity of set classes, and consequently are not of interest here.

3.6.1 Domain

Castrén accepts Isaacson's first criterion; his second criterion C2) states that a similarity measure should "provide a distinct value for every pair of SCs," and here only Isaacson's word "sets" is replaced by "SCs" (Castrén 1994, 18). Buchler also accepts the first criterion, although he admits certain exceptions to this rule (Buchler 1997, 64).

In terms of the similarity of twelve-tone rows this criterion could be interpreted as a requirement that a similarity measure should provide a value for every pair of twelve-tone rows. As I will show in Part III, most row measures easily satisfy this criterion.

Castrén and Buchler also accept Isaacson's second criteria in principle. Castrén's criterion C1) is a slightly attenuated version in that it requires only that a similarity measure should "allow comparisons between SCs of different cardinalities," and does not require that it should be useful for sets of *any* size (Castrén 1994, 18). Similarly, Buchler puts forward some valid arguments against Isaacson's second criterion (Buchler 1997, 21). For example, not all (otherwise potentially useful) measures cope with set classes of cardinalities 0, 1, or 2. In this respect, Isaacson's second criterion is too strict.

It could be argued that similarity measures do not need to be valid for all set classes. Buchler elaborates on the difficulties of comparing trichords to set classes larger than hexachords (Buchler 1997, 21). He also mentions four set classes that are difficult to relate to others: 0-1[], 1-1[0], 11-1[0123456789A], and 12-1[0123456789AB].[3] Assume, for example, that we are comparing the empty set (or even the pitch-class set $\{0\}$) to two pitch-class sets of cardinality 10. What meaningful comparisons could we make? What musical purpose would it serve to obtain a value denoting the similarity of set classes 0-1[] and 10-1[0123456789], for example?

An ideal similarity measure certainly should give the degree of similarity for all pairs of set classes. Perhaps we should take a more practical perspective, however. A superb similarity measure for hexachords (and hexachords only) is certainly a better addition to the music analyst's toolbox than a dubious measure for all set classes. Similarly, I will introduce in Section 12.4 a row measure that can applied only to rows with identical unordered INT contents. Nevertheless, one could argue that a similarity measure of limited applicability, with an acknowledgment of its inner workings and awareness of its possible limitations, is more useful than a "one size fits all" type of measure.

Isaacson's second criterion – that a measure should be useful for sets of any size – reveals a major difference between the realms of pitch-class sets and twelve-tone rows. Not all similarity measures for pitch-class sets satisfy this criterion – a major reason being the different cardinalities. The set of twelve-tone rows is much more homogeneous, however, and we encounter no problem that is comparable to the cardinality problem. However, some of the similarity measures for twelve-tone rows could be extended to segments (ordered pitch-class sets of a length that is smaller than 12), and in this case we would need to consider the issue of different cardinalities.

Another significant difference between the realms of pitch-class sets and twelve-tone rows is the cardinalities of the domains: there are 4096 pitch-class sets (or 224 set classes) and 479001600 twelve-tone rows.

[3]While, in principle, I accept Buchler's point of view, it could be argued, for example, that set class 11-1[0123456789A] is considerably more similar to set classes 10-1[0123456789] and 12-1[0123456789AB] than to set class 3-11[037].

3.6.2 Range

Isaacson's third criterion is the most controversial and therefore the most interesting. Castrén replaces Isaacson's expression "wide range of discrete values" with the expression "a comprehensible scale of values" in his criterion C3), and further divides it into the following four sub-criteria (Castrén 1994, 18).

C3.1) all values are commensurable

C3.2) the end points are not just some extreme values, but can be meaningfully associated with maximal similarity and dissimilarity

C3.3) the values are integers or other easily managable numbers

C3.4) the degree of discrimination is not too coarse or unrealistically fine.

Castrén questions Isaacson's third requirement, deeming it "more of a recommendation than a condition on a par with the two previous ones" (Castrén 1994, 17). Similarly, Buchler also interprets Castrén's rules as preference rules rather than structural conditions (Buchler 1997, 25) – thus refining Isaacson's third criterion.

Isaacson and Castrén appear to offer two closely related flavors of the same basic argument. In support of Isaacson we could, for example, state that the usability of Forte's relations is diminished by their very limited scale: two set classes are either related or they are not. An analyst would certainly appreciate a better resolution. On the other hand, the expression "wide range" is somewhat vague, and therefore it is difficult to agree or disagree with it.[4]

Castrén offers more refined criteria for the range. His criterion C3.1) simply indicates that the range should be unequivocal for all pairs in the domain. He stresses his point by showing that value 5 in Rahn's measure MEMB_2 may denote maximal or minimal similarity, depending on the cardinality of the set classes compared (Castrén 1994, 19). The values that MEMB_2 gives for set classes of different cardinalities are simply not commensurable, making comparison difficult.

Castrén's second sub-criterion C3.2) refines Isaacson's third criterion by suggesting that the minimum and maximum values of a similarity measure should be meaningful. This seems only a practical matter, intended to make a similarity measure easier to use in the everyday life of an analyst. The interpretation of its values is easier if the maximum ones are "meaningful," or at least easily conceived of. However, in many cases this means that some scaling needs to be done, since not all measures naturally have their extremes falling at 0 and 100, for example. However, for the similarity measures that are dissimilarity measures, the value 0 constitutes a very natural lower limit of the scale, and well suits the conception that the number of differences between two identical objects is zero. Furthermore, one of the requirements of the metric is precisely that the distance between identical objects must be zero.

[4]Presumably Isaacson's own similarity measure IcVSIM represents what he would consider an approximately ideal resolution. Between set classes of cardinalities 3 to 9, the number of distinct values produced by IcVSIM is 124 (Isaacson 1992, 81). This is not far from Castrén's proposed scale of 101 distinct values.

The issue of maximum value is more controversial. We encounter several similarity measures for twelve-tone rows that are based on counting differences in the rows, for example different pitch classes at corresponding order positions and the number of order inversions (see Sections 8.2 and 9.2). These have different maximum values. However, scaling the values of these measures would distort their "natural meaning."

Castrén's third sub-criterion C3.3) is even more practical than the other two since it calls for easily manageable values.

The debate on the kinds of values that a similarity measure should produce has been going on for rather a long time. It arose in its first incarnation when Charles Lord was distressed by Teitelbaum's use of square roots in his *similarity index*, which resulted in irrational numbers (Lord 1981, 111). Isaacson also expressed the fear that the irrational values produced by his similarity measure $\mathrm{I}\textsc{c}\mathrm{VSIM}$ might trouble some readers (Isaacson 1996, 19). Later, with respect to Teitelbaum's similarity index, Quinn pointed out that square roots arose rather naturally due to Pythagorean distance (Quinn 2001, 142). Furthermore, it should not be forgotten that, for example, values 2 and $\sqrt{2}$ as real numbers are equally precise – even if the decimal representation of $\sqrt{2}$ might entice us into thinking otherwise. Hence, there should not be anything wrong with irrational numbers per se, it is only that their use leads to the need for rounding them off.[5] Castrén accepts square roots as meeting his criterion C3.3) (Castrén 1994, 21).

Castrén's fourth sub-criterion C3.4) calls for a reasonable degree of discrimination or, using a term Rahn borrowed from optics, resolution (Rahn 1979–80, 486). An exceptionally wide array of distinct values could give a false idea of the "true" resolution of the measure.

Castrén lists his sub-criteria "in order of decreasing importance" (Castrén 1994, 31). Hence, the fourth one should be the least significant. However, it invites us to pause for a moment to consider the nature of the values that similarity measures give us.

Let us consider Isaacson's $\mathrm{I}\textsc{c}\mathrm{VSIM}$. It is defined as the standard deviation of the interval-difference vector. As such it is infinitely precise (unless we introduce rounding errors) as there is no way that we could make it more precise. This is different from measuring the length of a stick, for example, since we can accomplish the task only with limited precision. There is no such problem with standard deviation.

Rounding the numbers does not indicate the preciseness of a similarity measure: it rather indicates our faith in what great distinctions we can make based on the measures. Hence, the issue of resolution arises not from the algorithm but from the relation of the measure to the idea of similarity.

Let us look at a concrete case. The BADNESS OF SERIAL FIT similarity measure, or BSF (see Section 9.3), gives the following values for the *Lyric Suite* row $5409728136\mathrm{AB}$ and two carefully selected rows:[6]

$$BSF(5409728136\mathrm{AB}, 3\mathrm{B}68\mathrm{A}7102594) = 1791647$$

[5] Of course, irrational numbers cannot be represented as decimal numbers; $\sqrt{2}$ is a different number than 1.41421.

[6] These two rows were selected as the ones producing the most minuscule difference possible with the *Lyric Suite* row using the BADNESS OF SERIAL FIT. In particular, given the distribution of values in BSF, the difference of these values is truly minimal.

and

$$BSF(5409728136\text{AB}, 3\text{AB}187690542) = 1791648.$$

The resolution of this particular measure is exceptionally good as is gives as many as 569573 distinct values ranging from 1 to 479001600. The difference between values 1791647 and 1791648 is minimal. It is difficult to imagine that – on the basis of these values – we could give any interpretation of why one of the rows would be used and not the other.

Castrén points out that we should not decide beforehand what the resolution of measures should be (Castrén 1994, 17) – a view I endorse. The similarity measures for twelve-tone rows turn out to exhibit a wide array of ranges. Since they are mostly "natural products" – values that denote some concrete quantity of difference such as the number of different pitch classes at corresponding order positions – they may remain unscaled. I would instead promote understanding of what the values are. Every similarity measure analyzed in Part III is accompanied with a description of the distribution of its values. A uniform scale would be of no help if we could not relate the values to the other values. The properties of the distribution are discussed in detail in Section 6.3.

Curiously, even if row measures have a considerably larger domain than set-class similarity measures, their ranges are typically smaller (Lewin's row measure BADNESS OF SERIAL FIT being a notable exception). For example, the range of Morris' ASIM is 106 distinct values, and that of Isaacson's IcVSIM is 124 distinct values. Only a few of the row measures examined in this work, such as BADNESS OF SERIAL FIT, CORRELATION COEFFICIENT, and SUBSET CONTENT DIFFERENCE (see Sections 9.3, 9.6, and 10.3), have a range of more than one hundred distinct values.

3.6.3 Internal coherence and aspects of similarity

Castrén's fourth criterion states that a similarity measure should "C4) produce a uniform value for all comparable cases" (Castrén 1994, 18). We could interpret this as a requirement of internal coherence: the results of the measure should always be the same in cases that we intuit as comparable or equivalent.

What is "a comparable case" is not self-evident, and it depends on the aspect of similarity that is in focus. Z-related set classes provide a prime example. These set classes are comparable cases in similarity measures based on interval-class contents, but are not in those based on subset-class contents. Indeed, the standard argument against the former is precisely that they do not discriminate between Z-related set classes.

In Section 3.7 I will introduce the concept of *transformational coherence*, which allows the formalization of one aspect of what could be meant by a "comparable case" in the realm of twelve-tone rows.

Castrén further elaborates on the uniformity of values.

> This criterion states that a similarity measure should not be affected by SC properties other than those it professes to measure. If it adopts as its basis a certain aspect of similarity, it should produce the same value for all SC pairs whose type of similarity is

uniform from the point of view of the chosen aspect. (Castrén 1994, 23)

Taking Z-related set classes once again as an example, if a similarity measure professes to measure the interval-class contents of set classes but produces different values for Z-related set classes, then it is affected by some other properties of set classes than those it professes to measure.[7]

We are faced here with the question of the nature of similarity. In particular, is there a single "universal" similarity, or are there several different aspects? The above citation from Castrén implies the latter. Isaacson explicitly focuses on the intervallic similarity of pitch-class sets (Isaacson 1990, 2). The other main stream approach is, of course, based on the subset contents. Orpen and Huron differentiate between "in what way two things are similar" and "to what *degree* two things are similar," which they label the qualitative and quantitative aspects of similarity (Orpen and Huron 1991, 2–3).

Quinn seems to disagree. He writes that "*all* aspects of similarity [...] are deeply and inextricably interrelated" and that "the similarity relations [...] speak with a single extensional voice, regardless of what they profess to measure" (Quinn 2001, 155).

Similarity relations are without doubt interrelated. However, the depth of that interrelatedness is open to question. With regard to pitch-class sets I would be inclined to say that the interrelatedness is not pervasive enough so as not to leave room for different aspects of similarity, or not to allow for two pitch-class sets to be similar with respect to some aspect and less similar with respect to another.

Figure 3.1 presents the Z-related set-class pairs and the corresponding values of ATMEMB and relative ATMEMB.[8] Since the set classes are Z-related, the interval-class contents are identical in each pair. Set-class similarity measures based on the interval-class vector deem the set classes of each pair maximally similar. ATMEMB, however, bases the measurement of similarity on the subset contents of set classes, and the values of ATMEMB between Z-related set classes range from a minimum of 0.649 to a maximum of 0.911.[9]

Furthermore, if we put the values into perspective and take into consideration their distribution in each cardinality, the similarity of the subset-class contents of the Z-related pairs turns even weaker. Of the set classes of cardinality 6, according to ATMEMB, the Z-related set classes 6-Z4[012456] and 6-Z37[012348] are more similar than 475 pairs of the 1225 pairs of distinct set classes of cardinality 6, they are equally similar to 31 other pairs, and less similar than 718 pairs. Therefore, the relative ATMEMB value is 0.41 and hence the Z-related set classes 6-Z4[012456] and 6-Z37[012348] have clearly rather dissimilar subset contents.[10] To take another angle, set class

[7] Quinn (2001, 155) refers to this passage as "the best exemplar of the Myth of Intension in the whole of the similarity-relation literature," which is simply a misreading. The criterion simply concerns whether we understand the structure of the algorithm producing the values.

[8] The notion of relative values is discussed in more detail in Section 6.4.

[9] Since ATMEMB measures the similarity of set classes, the maximum value 1.0 represents maximal similarity and the minimum value 0 represents maximal dissimilarity.

[10] The relative ATMEMB value is $0.41 \approx 1 - \frac{718}{1225}$ if we base the calculation of the relative similarity on the fact that there are 718 pairs of distinct set classes of cardinality 6 that are more similar than 6-Z4[012456] and 6-Z37[012348]; if we base the calculation on the fact that there are 475 pairs of distinct set classes of cardinality 6 that are less similar than 6-Z4[012456] and 6-Z37[012348], the relative ATMEMB value is $0.39 \approx \frac{475}{1225}$.

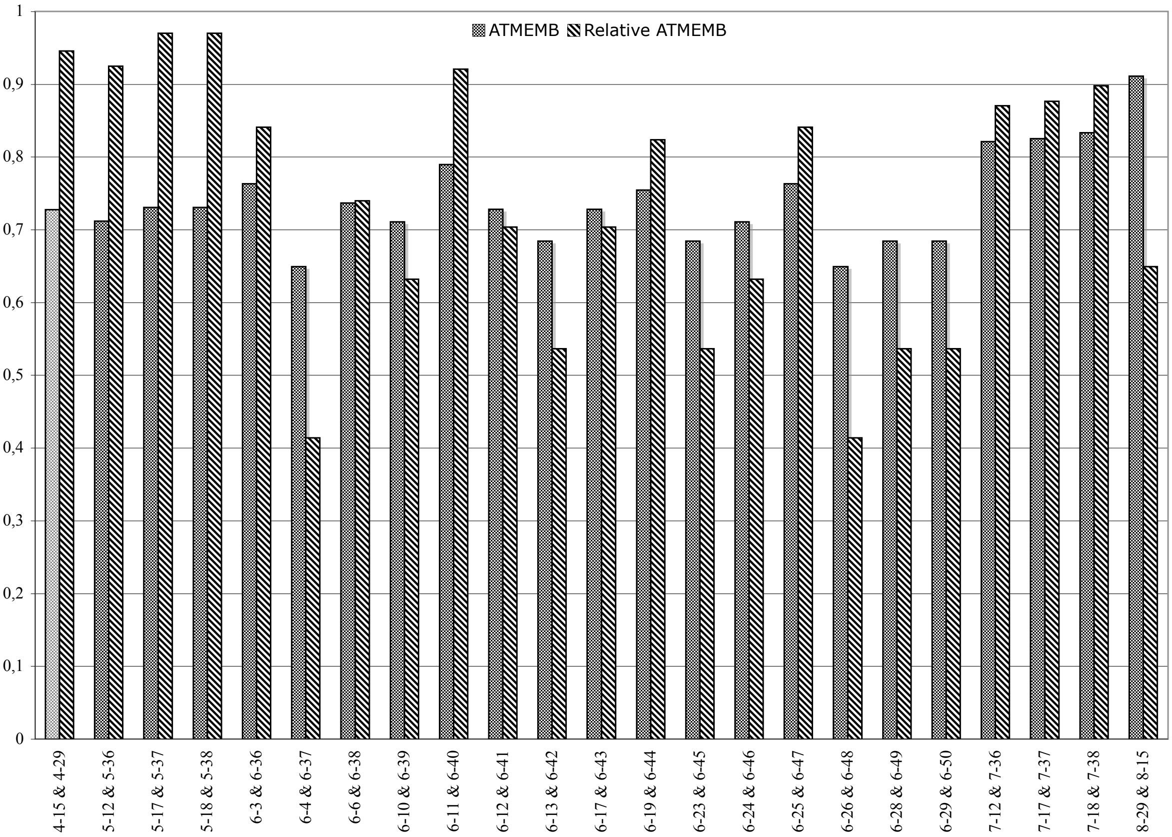


Figure 3.1: ATMEMB values and relative ATMEMB values of Z-related set classes. The gray bar represents the ATMEMB values. The striped bar represents the relative ATMEMB values: how similar the Z-related set classes are compared to all set classes of their cardinality.

6-Z10[013457] is more similar to its Z-related counterpart 6-Z39[023458] than to 18 other hexachord set classes, but less similar to 6-Z39[023458] than to 29 other hexachord set classes.

Thus one could conclude that the various similarity measures measure different properties of the set classes. If we insist on the tenet that there exists only a single type of similarity, we would have to draw the conclusion that at least one of these two measure types – the ones building on the interval-class contents and the ones building on the subset-class contents – is defective.

The existence of different aspects of similarity also depends on the domain. The case for their existence between twelve-tone rows seems to be even stronger than that for between set classes. Basically, there have been only two approaches to the similarity of pitch-class sets: subset-class contents and the interval-class vector. It comes as no surprise that the two approaches are interrelated, since subset-class contents is simply an extension of the interval-class vector: it is not confined to the embedding of set classes of cardinality 2, but concerns the embedding of set classes of all cardinalities. As the discussion on similarity measures in Part III of this work shows, approaches to the similarity of twelve-tone rows come in many more flavors.

In the context of twelve-tone rows, it seems rather evident that there are different aspects to similarity. Figure 3.2 shows two rows 5409728136AB and 63BA81724590. Arguably, in the pitch-class dimension the corresponding entries in the rows are relatively close, but in the order-

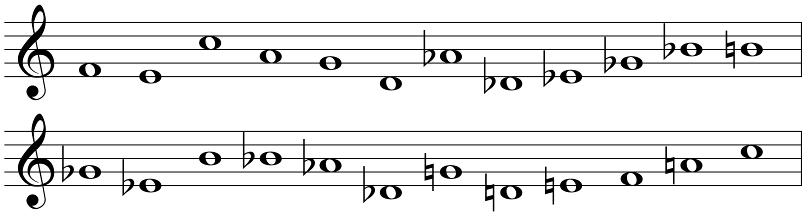


Figure 3.2: Twelve-tone rows $5409728136\mathrm{AB}$ and $63\mathrm{BA}81724590$ in which the corresponding pitch classes are close (always only a semitone difference) but the corresponding order numbers are not. Hence, the rows are similar in the pitch-class space but not in the order-number space.

number dimension they are far from each other. To take another example, rows $5409728136\mathrm{AB}$ and $\mathrm{BA}6318279045$ are related by retrograde. In both of them every pitch class has the same neighbors: pitch class 5 has only one neighbor, 4, pitch class 4 has neighbors 5 and 0, and so on. The two rows are maximally similar in this respect. On the other hand, the order of every pair of pitch classes is different in the two rows: in the former pitch class 5 precedes pitch class 4 but in the latter pitch class 4 precedes pitch class 5, and so on. The two rows are maximally dissimilar in this respect.

The question of the existence of different aspects of similarity is somewhat of a chicken-and-egg kind of problem. If we assume that there is only one kind of similarity we are led to believe that the differing values of similarity measures are due to a defect in one or more of them, while if we assume that there are several types we are led to believe that the differing values of similarity measures affirm the assumption.

Quinn's approach is to "look at the forest, not at the trees." From that perspective and in the case of the similarity of set classes, the overall topologies drawn by the various similarity measures may certainly have common shapes. However, on the detailed level we do find deviations that cannot be ignored. In the case of twelve-tone rows I acknowledge the multiple facets of similarity. It is, of course, always possible to claim that the similarity measures are faulty. However, the view of the forest also emerges from these very same faulty measures. It would seem to beg the question to selectively consider the measures faulty only when they disagree and not when they agree.

3.6.4 The role of algorithms

The "core" of most of the similarity measures is the algorithm or formula that produces the value when we give musical objects as input. There are several different set-class similarity measures only because there are several different algorithms: the properties of the algorithm are the differentiating factors between the similarity measures, especially given the fact that there are, in practice, only two approaches to the similarity of set classes: interval-class contents and subset-class contents. The differences in the similarity measures arise from the different treatment of the interval-class contents and subset-class contents.

An algorithm, per se, is not right or wrong.[11] However, we wish to obtain one that gives re-

[11] Rather, algorithms could be ranked according to their efficiency. However, those typically used in similarity measures are simple enough so that the issue of efficiency is negligible.

sults that match our intuitions. For example, the several existing similarity measures based on the interval-class vector are all (presumably) founded on the conviction that the interval-class contents of set classes is an essential differentiating factor, and hence the more similar the interval-class vectors are, the more similar the pertinent set classes should be. Without such a conviction we could not claim that the similarity measures should correspond to our intuitions.

The problem with intuition, of course, is that it lacks intersubjectivity, and that different people have different intuitions: the intuition of some researchers stresses the importance of the interval-class contents, while that of others stresses the importance of the subset-class contents. Isaacson elaborates on the role of intuition.

> Having invoked intuition, we must acknowledge that this is a sticky area, since what we call intuition is largely subjective. There will always be situations where people's intuitions differ, sometimes because of differences in musical experience, sometimes because of different choices from among multiple possible hearings. But while it seems somewhat slippery to insist, "Well, that's how *I* hear it," absent some objective measure–whatever that would be–it will have to do. (Isaacson 1996, note 8)

The idea of using algorithms is to find something more concrete than just an intuition of what is similar. However, as Castrén notes, the similarity measure has a sort of built-in paradox, since in the evaluation of its outcome we need to resort to the intuitive estimations of the experienced similarity of set classes (Castrén 1994, 12).

If we were only interested in the values that similarity measures produce we might discard the algorithms altogether and simply use our intuition to evaluate the similarity of two musical objects. Instead, a similarity measure giving values that correspond to our intuition also serves as a potential explanation of why they should be similar. For example, we could argue along the lines that two chords sound similar since the set classes they belong to have similar interval-class vectors and we hear the same interval classes (intervals between pitches that can be reduced to the same interval class) in the two chords. Furthermore, an algorithm serves as a heuristic device as it allows us to predict how similar musical objects are.

Nevertheless, the central role of algorithms has provoked criticism. Quinn laments the approach of "intuitions about algorithms."

> [...] there is no solid objective basis on which competing similarity relations can meaningfully be judged, and intuitions are the best substitute for such a basis. As to the kinds of intuitions usually mentioned, it is notable that they seem to be mostly intuitions about algorithms or about properties of the set of possible outputs as a whole. What is apparently missing here is any mention of intuitions about what similarity relations have to say about real-world pc sets. (Quinn 2001, 143)

Furthermore, in the context of justifying the irrational numbers of Teitelbaum's similarity index by the fact that they arise from the use of the Pythagorean formula (which is used to calculate spatial distances), Quinn asks for reasons why we should accept a certain mathematical model.

> [A] disturbing methodological problem remains: why should the justification for a mathematical model come from its metaphorical properties and meanings rather than from its results? (Quinn 2001, 142)

While there certainly appears to be some – to use Quinn's own expression – mathematical fetishism in the literature on similarity measures, I believe that his reproach is unfair for two distinct reasons.[12]

First, the discussion about the properties and the applicability of set-class similarity measures is, in general, furnished with real-world examples of how some similarity measures inappropriately evaluate some set classes as too similar or too dissimilar. Typically, a proposed similarity measure is defended by presenting a set of set-class pairs the similarity or dissimilarity of which seems to be well represented by the measure (and less well represented by the existing measures). For example, set classes 3-10[036] and 6-30[013679] have enjoyed wide publicity as being the test bed for the outputs of algorithms. Furthermore, as Castrén explicitly states, the final evaluation of a similarity measure happens precisely by comparing the results – how the similarity estimations correspond to the intuitive estimations of the researcher, not the internal mechanism of the algorithm.

> The final evaluation is a combination of many individual assessments: Is the scale of values from a given comparison group credible when related to that from another?; If some SC pair X is experienced to be closely similar, pair Y even more so and pair Z nothing but, do the corresponding values seem to reflect this meaningfully? (Castrén 1994, 12–13)

To use Quinn's dichotomy of intuitions, the final evaluation that Castrén describes in this passage is certainly closer to "intuitions about what similarity relations have to say about real-world pc sets" than "metaphorical properties of a mathematical model."

Secondly, we do not need to rely on our intuition in the analysis of algorithms. We could, instead, use mathematical knowledge to examine the properties of the formulas – without falling into mathematical fetishism.

Let us consider the similarity measures that are based on the interval-class vector. In general, all similarity measures follow the (informally stated) principle that set classes with major differences in their interval-class vectors are less similar than those with only minor differences. The differences between them arise precisely from the way in which they are different: two interval-class vectors may differ a little everywhere or they may be mostly the same but differ greatly at some point. The question is how we should find a balance between these two types of being different.

In order to illustrate the various approaches to balancing the differences, let us compare the similarities of set classes 6-Z6[012567], 6-Z10[013457], and 6-Z13[013467], and in particular, set class 6-Z6 to the two set classes 6-Z10 and 6-Z13. Table 3.1 presents the interval-class vectors of

[12] Quinn gives as an example Rahn's comment, "After considering with distaste exponentiation, multiplication, and addition, I chose multiplication as probably the least vicious means of combination" (Rahn 1979–80, 486). Rahn does not enunciate his reasons for considering some operations less vicious than others. However, he may have had in mind that these three means of combination each have characteristic behavior in scaling the balance between small and large differences, and the choice was made to obtain a moderate scaling of these differences.

set class	interval-class vector
6-Z10[012567]	[333321]
6-Z6[013457]	[421242]
6-Z13[013467]	[324222]

Table 3.1: The interval-class vectors of set classes 6-Z6[012567], 6-Z10[013457], and 6-Z13[013467].

set-class pair	ASIM	s.i.	IcVSIM
6-Z6, 6-Z10	0.267	3.46	1.414
6-Z6, 6-Z13	0.200	3.74	1.528

Table 3.2: The values of ASIM, the similarity index, and IcVSIM for set classes 6-Z6[012567], 6-Z10[013457], and 6-Z13[013467]. All three measures are dissimilarity measures: the value 0 represents maximal similarity, and the larger the value the more dissimilar the set classes are. When set classes of the same size (cardinalities 3 through 9) are compared, the maximal values of ASIM, the similarity index, and IcVSIM are 1, 8.48, and 3.46, respectively.

these three set classes. All entries of the interval-class vectors of set classes 6-Z6 and 6-Z10 differ by 1 or 2. Summing the absolute values of the differences results in 8. In contrast, in the interval-class vectors of set classes 6-Z6 and 6-Z13 all entries in the "even" interval classes are equal but those in the "odd" interval classes differ by 1, 2, or 3. Summing the absolute values of these differences results in 6. In cases like this, in choosing the similarity measure we encounter the problem of weighing up which type of difference is more cogent: a number of small differences that appear everywhere (such as between set classes 6-Z6 and 6-Z10) or a few larger differences that appear only here and there (such as between set classes 6-Z6 and 6-Z13).

Table 3.2 presents the values of the similarity measures ASIM, the similarity index, and IcVSIM when they are given the pairs of set classes (6-Z6, 6-Z10) and (6-Z6, 6-Z13) as inputs. If we look at the values we see that, according to ASIM, set class 6-Z6 is more similar to set class 6-Z13 than to set class 6-Z10. In contrast, according to the similarity index and IcVSIM, set class 6-Z6 is more similar to set class 6-Z10 than to set class 6-Z13.

Let us now consider how this disparity is based on the different approaches to balancing between many small differences versus a few large ones. In order to do this we need to examine the pertinent algorithms. According to ASIM, the sum of the absolute values of the differences between the corresponding entries of the interval-class vectors is divided by the sum of the counts of intervals in them. Hence, we obtain

$$\begin{aligned}\mathrm{ASIM}(\text{6-Z6, 6-Z10}) &= \frac{|4-3|+|2-3|+|1-3|+|2-3|+|4-2|+|2-1|}{15+15}\\ &= \frac{8}{30} \approx 0.267\end{aligned}$$

and

$$\begin{aligned}\mathrm{ASIM}(\text{6-Z6, 6-Z13}) &= \frac{|4-3|+|2-2|+|1-4|+|2-2|+|4-2|+|2-2|}{15+15} \\ &= \frac{6}{30} = 0.2\end{aligned}$$

as the ASIM values.

Teitelbaum's similarity index is defined as the square root of the sum of the squares of the differences between the corresponding entries of the interval-class vectors. Hence, we obtain

$$\begin{aligned}\text{s.i.(6-Z6, 6-Z10)} &= \sqrt{(4-3)^2+(2-3)^2+(1-3)^2+(2-3)^2+(4-2)^2+(2-1)^2} \\ &= \sqrt{12} \approx 3.46\end{aligned}$$

and

$$\begin{aligned}\text{s.i.(6-Z6, 6-Z13)} &= \sqrt{(4-3)^2+(2-2)^2+(1-4)^2+(2-2)^2+(4-2)^2+(2-2)^2} \\ &= \sqrt{14} \approx 3.74\end{aligned}$$

as the similarity indexes for the set-class pairs.

In the ASIM similarity measure the absolute values of the differences between the entries in the interval-class vectors are simply summed and the sum is then scaled. This type of distance (without the scaling) is known as *taxicab distance*. In contrast, in Teitelbaum's similarity index the differences between the entries in the interval-class vectors are squared, and the final value is the square root of the sum of these squares. This type of distance is known as *Euclidean distance*. The process of squaring the differences of the entries in the interval-class vectors emphasizes large differences. Therefore, as set class 6-Z6[012567] has one instance of interval class 3 and set class 6-Z13[013467] has four instances, this difference is emphasized in Teitelbaum's similarity index and, as a result, set classes 6-Z6[012567] and 6-Z13[013467] are classified as more dissimilar than set classes 6-Z6[012567] and 6-Z10[013457].

IcVSIM is based on standard deviation, which is a statistical function that describes how the values spread out around the average. We thus calculate the standard deviation of the differences in the interval-class vectors

$$\sigma = \sqrt{\frac{\sum(IdV_i - \overline{IdV})^2}{6}},$$

in which IdV_i denotes the differences between the entries in the interval-class vectors and $\overline{IdV}$ denotes the average of the differences. The existence of a few large differences turns out to be more significant here than the existence of several small differences, since IcVSIM involves squaring like Teitelbaum's similarity index.[13]

[13] As Isaacson points out, IcVSIM is a scaled variant of Teitelbaum's similarity index when the set classes have the same cardinalities (Isaacson 1990, 19).

For set classes 6-Z6 and 6-Z10, IdV(6-Z6, 6-Z10) = [-1 1 2 1 -2 -1] and $\overline{IdV} = 0$, thus

$$\begin{aligned}
&\mathrm{IcVSIM}(\text{6-Z6, 6-Z10}) \\
&= \sqrt{\frac{(-1-0)^2+(1-0)^2+(2-0)^2+(1-0)^2+(-2-0)^2+(-1-0)^2}{6}} \\
&= \sqrt{\frac{12}{6}} = \sqrt{2} \approx 1.414
\end{aligned}$$

is the IcVSIM value.

Similarly, for set classes 6-Z6 and 6-Z13, IdV(6-Z6, 6-Z13) = [-1 0 3 0 -2 0] and $\overline{IdV} = 0$, thus

$$\begin{aligned}
&\mathrm{IcVSIM}(\text{6-Z6, 6-Z13}) \\
&= \sqrt{\frac{(-1-0)^2+(0-0)^2+(3-0)^2+(0-0)^2+(-2-0)^2+(0-0)^2}{6}} \\
&= \sqrt{\frac{14}{6}} \approx 1.528
\end{aligned}$$

is the IcVSIM value.

In sum, the three similarity measures ASIM, the similarity index, and IcVSIM all represent different strategies for coping with balancing the different ways of having differences between interval-class vectors.

The different approaches to utilizing the interval vector have not yet been exhausted. For instance, the more instances of some interval class there are in a set class, the less important the additional ones are, and hence the counts should be scaled (Isaacson 1996, ¶9). Furthermore, the salience of some interval classes might lead the weighting of the different interval classes in the interval-class vectors (Isaacson 1996, ¶11). Block and Douthett (1994) discuss some uses of intervallic weighting.

Let us now return to the issue of the justification of the mathematical formulas in similarity measures. Quinn dismissed Euclidean distance by arguing that it represented the largely unquestioned transfer of a spatial metaphor into the evaluation of the similarity of set classes (Quinn 2001, 142). However, even if it has its roots in the measurement of spatial distance, it is a method used to balance the differences between the different dimensions of multidimensional entities, and its use extends far beyond the realm of spatial measurement. Goldstone notes that geometric models, even if they have been criticized, have been among the most influential approaches to the analysis of similarity in cognitive psychology (Goldstone 1999, 763).[14]

In effect, the six interval classes are treated as six dimensions in the similarity measures. The balancing of multiple dimensions has been discussed in studies on similarity in psychology. Gold-

[14]Goldstone mentions three other approaches to similarity: the contrast model, the alignment-based model, and transformational distance. The transformational approach is discussed in Chapter 5.

stone gives the following formula for distance in n-dimensional space (Goldstone 1994, 138):

$$D_{i,j} = \left[\sum_{k=1}^{n} |X_{ik} - X_{jk}|^p \right]^{(1/p)} . \tag{3.1}$$

According to Goldstone, the choice of the exponent p depends on the stimuli: those composed of dimensions that are psychologically fused together or have very small value differences are often best modeled by setting p equal to 2, which results in the Euclidean distance. Stimuli that are composed of separable dimensions are often best modeled by setting p equal to 1, which results in the taxicab distance. The dimensions of similarity measures based on the interval-class vector are the instances of interval classes in a set class. Consideration of the interval classes as dimensions that are psychologically fused together would suggest Euclidean distance.

Goldstone's formula 3.1 is known in mathematics as the distance in the L_p space. Its usefulness lies in the fact that as the factor p can be any real number greater than or equal to 1, it provides a way of adjusting the balance between a few large differences and several small ones. Petri Toiviainen (1996) and Larry Polansky (1996) used the distance in the L_p space under the moniker the Minkowski metric, and Clifton Callender (2005) undertook the task of finding the best value of p to measure voice-leading distance.

In sum, I consider mathematical fetishism to be a minor thread in the discussion on similarity measures, which is rather characterized by a careful search for measures that correspond to our perception of the similarity of set classes. Researchers have tried different ways of balancing the differences in the interval-class vectors and in the subset-class contents of set classes.

3.6.5 Transitivity

One of the most controversial issues concerning similarity relations has been transitivity. The list of participants in this debate includes Allen Forte, John Rahn, Richard Hermann, Thomas Demske, Michael Buchler, Robert Morris, and Ian Quinn.

Forte was the one who started the discussion, and his similarity relations, R_p, R_0, R_1 and R_2, were among the first similarity measures.[15] In standard mathematical terminology, described in Appendix B, Forte's four relations are binary relations: he notes that all four are non-transitive (Forte 1973b, 53).

John Rahn reaffirms the non-transitivity of similarity relations.

> Relations among sets are here sliced four ways: they can be context-free or context-dependent, and they can be equivalence relations, which are transitive, or similarity relations, which are non-transitive. (Rahn 1979–80, 483)

While Rahn appears here to discuss only genuine relations in mathematically correct parlance, in the following pages he positions Eric Regener's common-tone criteria, David Lewin's generalization

[15] Richard Teitelbaum presented his similarity index in 1965, several years before Forte's book. Forte cites Teitelbaum's article but does not refer to the similarity index when discussing his own similarity relations.

of the common-tone criteria, and the embedding criteria starting from the work of Robert Morris, under the umbrella of "relatively context-free relations." This clearly extends the term "relation" to cover methods of analyzing the similarity of pitch-class sets that are not, mathematically speaking, relations.

Hermann repeats the same tenet: "By definition, similarity relations lack transitivity" (Hermann 1995, ¶4). Here the context includes similarity functions such as REL, even if later in the article he suggests that the term similarity [relation] should apply only "for formal relations that possess reflexivity and symmetry, but lack transitivity as mathematicians would have it," and that the term "resemblance relation" should be used for similarity functions (Hermann 1995, note 2). Similarly, Buchler states that "All similarity functions are symmetric and non-transitive," and "a relation of X to Y and Y to Z does not necessitate or imply a like relation from X to Z" (Buchler 1997, 19). In Buchler's terminology, a similarity function can be either a relation or a function.

In sum, it seems that the non-transitivity of similarity relations and (inadvertently) similarity functions was well established until Ian Quinn gave the discussion on transitivity a novel turn by elevating it to the status of a myth that needed to be debunked (Quinn 2001, 122).

Since the debate on the transitivity of similarity relations has involved some confusion in the use of mathematical terminology, it would be appropriate to give the formal definition of transitivity at this point. The following definition is extracted from Definition B.3 of Appendix B.

DEFINITION 3.1 Let S be a set and R be a binary relation on S. Relation R is transitive if $(s, t) \in R$ and $(t, u) \in R$ imply $(s, u) \in R$.

Transitivity is a property that binary relations may have.[16] Therefore the question of whether similarity relations (those that are not binary relations) are transitive or not is simply nonsensical, a misuse of words.[17] Trying to figure out the answer is as meaningful as, to quote Noam Chomsky's famous example, trying to figure out whether colorless green ideas sleep furiously or not (Chomsky 1957, 15).

Some authors have used networks as a context for similarity. Morris, for example, thus describes the non-transitivity of similarity relations: "...even though A is similar to B and B is similar to C, A is not obliged to be similar to C" (Morris 1987, 103).[18] A network is, in fact, one way of representing a binary relation, and hence in this context transitivity is an applicable term. However, if we were to discuss similarity in terms of *distance* in the network, the context would no longer represent a binary relation and hence we would have to drop the concept of transitivity.

Similarly, the term reflexivity cannot be applied to functions. However, it could be interpreted as a requirement that the value that a similarity function gives to two identical objects must be one that we interpret as similar. A symmetric function is a function that gives the same value with any permutation of its variables. Hence, even if technically the symmetry of a relation is different from

[16]The word "transitivity" has other meanings in mathematics. However, in the context of (binary) relations its use is unequivocal.

[17]In a sense this is a logical step from inadvertently considering similarity functions as relations.

[18]See also section 1.2.1.2 in Morris (2001).

the symmetry of a function, the symmetry of similarity relations and the symmetry of similarity functions both express the same idea.

Furthermore, insisting on the transitivity of similarity relations creates a terminological overlap with equivalence relations. Similarity relations are reflexive (everything is similar to itself) and symmetric (if A is similar to B then believably B is also similar to A). Therefore, claiming that similarity relations are transitive results in duplicating the definition of an equivalence relation. Indeed, Orpen and Huron make transitivity the criterion separating equivalence relations and similarity relations (Orpen and Huron 1991, 2).

There are two opposing sides here: supporters of the transitivity and supporters of the non-transitivity of similarity relations. “Therefore, for the record, and for what it is worth: intransitivity *per se* does not strike me as overwhelmingly problematical” (Demske 1995b, ¶4). Quinn comments directly on Demske and writes, “...therefore, for the record, and for what it is worth: intransitivity *per se* actually does strike me as problematic” (Quinn 2001, 123).

Nevertheless, the pertinence of the term indicates that there is some property of similarity functions that the authors want to express. Indeed, as Quinn notes, they have certain properties that are evocative of transitivity. He writes, “...we have good reason to *want* similarity relations to be transitive,” even if in the following sentence he admits that similarity relations “therefore cannot possibly be transitive” (Quinn 2001, 122). While he acknowledges this disparity he – somewhat paradoxically – still insists on debunking the myth of transitivity.

One way to preserve the transitivity of similarity relations would be to consider them fuzzy relations – even if then we would be changing topic. Describing similarity in terms of fuzzy set theory is certainly a viable approach, and could be used as an alternative to the metric. However, Quinn admits that similarity relations interpreted as fuzzy relations are not transitive even in the fuzzy sense.

I will next attempt to reconstruct what I conceive of as some possible reasons for trying to twist the concept of transitivity also to apply to functions. I will then argue that we should drop the word “transitivity” altogether from the discussion (unless we are discussing genuine relations, such as Forte’s four relations), and then show that there are better ways to achieve (at least some of) these aims by using a concept borrowed from topology.

As long as we are using the concept of transitivity we have two alternatives: transitivity or non-transitivity. A relation is either transitive or it is not. The property is black and white: there are no almost transitive relations. No matter how large the relation is, a single counterexample suffices to render it non-transitive.

There are plenty of everyday examples of transitive relations: “having the same birthday” is one. If A and B have the same birthday and B and C have the same birthday, then we can deduce that A and C have the same birthday. Similarly the relation “is younger than” is transitive: if A is younger than B and B is younger than C, then we can deduce that A is younger than C.

There are also plenty of everyday examples of non-transitive relations. Quinn’s example of friendship (Quinn 2001, 127) applies here: in general if A and B are friends and B and C are

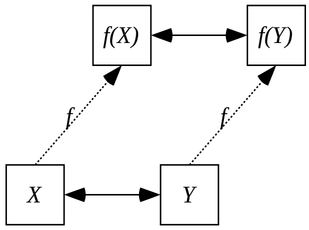


Figure 3.3: Transformational coherence in two-dimensional space.

friends, there is nothing we can say with absolute certainty about A and C – they might be friends or they might not. Even if people are loosely organized in cliques and coteries we have plenty of cases of non-transitivity: the wife's bowling friends most probably have never met the members of the husband's poetry club.

The core of the issue of transitivity is that similarity relations do not really fit in either one of the categories. If A is related to B and B is related to C, then it is *not* the case that there is nothing we can say about A and C. Since I argued above that similarity relations are not really relations, let me rephrase the issue in terms of similarity. If A is similar to B and B is similar to C, then it is not the case that there is nothing we can say about the similarity of A and C. We expect them to be at least somewhat similar.

I assume that it is exactly this property of similarity relations or similarity functions that Quinn means to describe by using the word transitivity.[19] A better approach than to misuse the word would be to turn to topology – another field of mathematics – for help. A phenomenon very similar to similarity relations (pardon the pun) has been defined unequivocally and studied thoroughly in this context. It is known as the *metric*, and is a mathematical formalization of the intuitive notion of the concept of distance. If we rephrase similarity in terms of closeness, the metric formalizes the notion: if A is close to B and B is close to C, then it is not the case that there is nothing we can say about the closeness of A and C. We expect them to be at least somewhat close. I will discuss the metric in detail in Chapter 4.

3.7 Transformational coherence

With respect to twelve-tone rows, one way of defining "a comparable case," discussed in Section 3.6.3, is to require that the mutual degree of similarity between two twelve-tone rows must remain the same in the usual row transformations, for example transposition, inversion, and retrograde. I term this requirement *transformational coherence*, and it is a necessary but not a sufficient condition for similarity measures. For example, we might require that a musical object is always maximally similar to itself; this kind of requirement cannot be expressed sensibly in terms of transformations.

Figure 3.3 illustrates idea of the transformational coherence in two-dimensional space. The dis-

[19]However, Quinn also makes a stronger claim – his "Natural Kinds Hypothesis" – that set classes divide naturally into groups of mutually similar set classes; the truth of this hypothesis cannot be logically derived from the revised interpretation of transitivity discussed here, but it is (or is not) a contingent property of set classes.

tance between the squares X and Y remains unchanged if we move both by the same transformation f.

Let us then consider two twelve-tone rows A and B. Assume that they are similar to some degree x. If we transpose both rows by T_1, we would expect the degree of similarity of the transposed rows T_1A and T_1B to be the same x as the degree of similarity of the original rows.

We could define the transformational coherence of a similarity measure formally as follows. Let us specify the transformational coherence under a group of operations: this poses no problems since a group-theoretical framework for row operations was established in Part I of this work. As discussed in Section 2.2.3, a group of operations induces an equivalence class – the row classes.

DEFINITION 3.2 Let d be a similarity measure on a set of musical objects S, and let G be a group of transformations. Similarity measure d is transformationally coherent if $d(x, y) = d(g(x), g(y))$ for all $x \in S$, $y \in S$ and $g \in G$.

The core of Definition 3.2 is that similarity relations are invariant with respect to transformations. Assume that we have two pitch-class sets or twelve-tone rows and we transpose both by n semitones. The requirement of transformational coherence states that the transformed twelve-tone rows (or pitch-class sets or other musical objects) must be precisely as similar or dissimilar as the original twelve-tone rows (or pitch-class sets or other musical objects).

I will discuss the metric for row classes in Section 6.2 and will show that transformational coherence plays a crucial role in proving the fact that we can define the similarity of row classes sensibly.

CHAPTER 4

Similarity measures and the metric

4.1 Similarity measures as distances

As discussed in Section 3.5, many similarity measures are, in fact, dissimilarity measures. A natural interpretation is that the amount of dissimilarity is the "distance" between the objects: the more similar two objects are, the closer they are and, correspondingly, the more dissimilar two objects are, the greater is their distance.

The idea of similarity as distance has been referred to several times in the literature. For example, John Rahn modifies his own ATMEMB similarity measure to a dissimilarity measure, DATM, in order to be able to discuss similarity in such terms (Rahn 1989, 2–3). Damon Scott and Eric Isaacson discuss their similarity measure *Angle* as "measuring how 'far apart' the sounds of various pitch-class sets are" (Scott and Isaacson 1998, 111). Robert Morris discusses the possibility of measuring the distance in the networks that the similarity relations define (Morris 1987, 105). Some authors even mention explicitly the concept of the metric (Roeder 1987; Orpen and Huron 1991). Eytan Agmon defines an interval/distance system in such terms even if he does not mention the term itself (Agmon 2002, 221). In addition, David Lewin utilizes Schwartz inequality in connection with his similarity measure REL even if he does not develop his theory in terms of the metric (Lewin 1979–80).[1]

When we examine a given similarity measure in terms of the metric, we are interested in two questions: "Does the measure define a metric?" and "Why does it define a metric?" The former is answered by the theorems or lemmas, and the latter by the proofs of the theorems, which makes them worth writing and reading. In general, the proofs of the theorems give more information than the theorems themselves.

[1]The triangle inequality can be derived from the Schwartz inequality: see Rudin (1987, 49), for example.

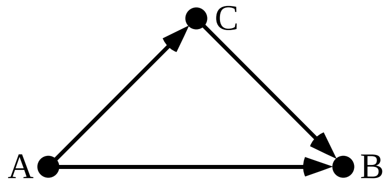


Figure 4.1: The triangle inequality.

4.2 The definition of the metric

The metric is the mathematical formalization of the concept "distance." In general, a set of distances satisfying the requirements of the metric is a "well-behaving" one, with no surprising or counter-intuitive properties. I will first discuss the formal properties of the metric, and then consider it in terms of similarity measures, arguing that its properties match the properties we intuit similarity to have.

I will begin by defining the metric formally.

DEFINITION 4.1 The *metric* on a set X is the mapping $d : X \times X \to \mathbb{R}_+ \cup \{0\}$ that satisfies the four following requirements for all $x, y, z \in X$: (i) $d(x,x) = 0$, (ii) $d(x,y) = d(y,x)$, (iii) $d(x,z) \leq d(x,y) + d(y,z)$, and (iv) $d(x,y) = 0$ implies $x = y$. If d satisfies only the first three requirements, it is called a *pseudometric*.

The distance between two points is the nonnegative (positive or zero) length of the shortest path between them. Distances define a metric if they satisfy the four requirements given in Definition 4.1. The first of these is that the distance from a point to itself is zero. The second is symmetry: the distance from A to B must equal the distance from B to A. The third requirement is known as *triangle inequality*: the distance between points A and B cannot exceed the sum of the distance from A to C and the distance from C to B. The points A, B, and C in the two-dimensional plane in Figure 4.1 illustrate triangle inequality: the distance between A and B is certainly smaller than that between A and C plus the distance between C and B. Thus, if A is near C and C is near B, then A must be relatively near B. The fourth requirement is that non-identical points must have a positive (non-zero) distance.

The requirement of nonnegative values also follows from conditions (i), (ii), and (iii), since if $d(x,y) = -k < 0$,

$$0 = d(x,x) \leq d(x,y) + d(y,x) = -2 \cdot k < 0,$$

which is a contradiction.

The metric is a formalization of closeness or nearness. Its four requirements correspond to our intuition of the relations associated with closeness. First, any object is close (or as close as possible) to itself; hence, the distance between an object and itself is zero. Secondly, closeness is a symmetric relation: if A is close to B then B is also close to A. Thirdly, triangle inequality can be given at least two interpretations. On the one hand, it implies that closeness is a "semi-transitive" relation:

if A is close to C and C is close B, then it would be counterintuitive if A and B were very far apart. On the other hand, it could be interpreted as requiring that the straight path should be the shortest: when going from point A to point B, diverting to point C is not going to make the path any shorter. Finally, the role of the fourth requirement is to make a distinction between closeness and sameness.

4.3 The metric and similarity measures

Let us now interpret the properties of similarity measures in terms of the metric. The metric is defined as a mapping $d\colon X \times X \to \mathbb{R}_+ \cup \{0\}$. We can read two requirements into this definition. First, as the domain of the mapping is the Cartesian product $X \times X$, the metric requires that a real value is defined for all pairs of objects. This coincides with the first requirement imposed by Isaacson that similarity measures should "provide a distinct value for every pair of sets" (see Section 3.6.1). However, if this requirement were to pose problems we could always restrict the domain of the mapping to the "unproblematic" subset of the domain (for example, in the case of set classes we could define the domain of a similarity function as the set classes of cardinalities 3 through 9).

Secondly, the metric requires that all values are non-negative, i.e., positive or zero. With regard to distances, a negative distance could be interpreted as a distance in the "opposite direction." However, if we consider only the distance without any defined direction, negative values would not make any sense. Similarly, if a similarity measure measured the dissimilarity of musical objects, it would be difficult to imagine what negative dissimilarity would mean. Hence, it comes as no surprise that almost none of the published similarity measures gives negative values.[2]

Let us now consider the four requirements of the metric in terms of similarity measures. The first requirement is that the distance between an object and itself is zero. This is very straightforward and intuitive. The difference between the object and itself is naturally zero: no step is required in order to move to the place where one already is, or no change is required to keep an object intact.

Zero as maximal similarity or identity provides an advantage for dissimilarity measures over similarity measures since it provides a natural end point of the scale (see Section 3.6.2). The value denoting maximal dissimilarity depends on the similarity measure and its properties, and it is not dictated by the metric.

The second requirement is that of symmetry: the distance from A to B must be the same as the distance from B to A. This expresses the idea that distance is not directional, and that it is the same both ways.

The criterion of symmetry is occasionally stated explicitly, but practically always implicitly, in the literature concerning the criteria for similarity measures. Richard Hermann makes it one of the definitive features of similarity relations: "Similarity between entities exhibits reflexivity, and symmetry but lacks transitivity" (Hermann 1994, 1), and Michael Buchler states that all similarity measures are symmetric (Buchler 1997, 19). Marcus Castrén discusses his RECREL similarity measure in terms of unordered pairs of set classes making the measure essentially symmetrical (Cas-

[2] Robert Morris' CORRELATION COEFFICIENT similarity measure (see Section 9.6) for ordered pitch-class sets being a notable exception since its range is $[-1, 1]$.

trén 1994, 102). In general, the mathematical formulas defining the values of similarity measures are formulated in such a way that symmetry results. In Section 11.3 I will discuss a similarity measure for twelve-tone rows, SCATTERING, which is the exception that proves the rule since it turns out to be non-symmetric.

The third requirement of the metric is triangle inequality which is, in a sense, the core. We could consider triangle inequality from at least two angles. First, it states that the distance is defined as the length of the shortest path. Assume that function d defines a metric on a set including the points A, B, and C. Now triangle inequality states that $d(A, B) \leq d(A, C) + d(C, B)$, which we could interpret as stating that the shortest path from A to B cannot be longer (but it may be shorter) than the path from A to B that goes via point C.

The second interpretation is a limitation of the distance between A and B when we know the distance between A and C and the distance between C and B. In other words, the straight path (ever since Euclid the straight path has been considered the shortest path) from A to B cannot be longer than the sum of the paths from A to C and from C to B (since, if it were, we could replace the longer path with the composite of the paths from A to C and from C to B). Or, from the opposite angle, if A and B are extremely dissimilar, then it is not possible to find a third element C that would be very similar to both A and B.

It should be noted that the distance from A to B may, of course, be considerably smaller than the sum of the distances from A to C and from C to B. As an extreme case, A and B may be identical points, in which case their distance is zero. Triangle inequality only gives the maximum distance from A to B when the distances from A to C and from C to B are known, but does not give any minimum distance.

It is exactly this second interpretation that expresses what I assume Quinn wanted to achieve by insisting on the transitivity of similarity relations (see Section 3.6.5).

Accorging to Quinn, "Existing relations of pc set-class similarity are quite akin to friendship" (Quinn 2001, 127). While this statement is most probably intended as a loose analogy, there is an important difference between the friendship relation and the metric that deserves comment. In the case of the friendship relation we might imagine a situation in which the two closest friends B_1 and B_2 of a person A cannot stand each other. However, the metric would not allow this kind of irregularity or discontinuity. Hence, we have to decide which of the models we are to defend. I would argue that the metric as an established theory would provide a better model, and furthermore, the majority of the similarity measures that are dissimilarity measures in fact satisfy its requirements.

Finally, the fourth requirement of the metric is that two non-identical entities should not have zero distance. The idea is that we want to have a non-zero distance between non-identical musical entities. However, this fourth requirement is not a critical one – we may allow violations, in which case we have a *pseudometric* but not a metric. For example, all similarity measures that are based on the interval-class vectors of set classes do not discriminate between Z-related set classes. However, we could turn a pseudometric on the set of set classes into a metric on the set of interval-class vectors (assuming that the similarity measure always gives two non-identical interval-class vectors a non-zero value). Of course, it would seem desirable to judge non-identical set classes to be less than maximally similar, but this is not possible when only interval-class contents are under consideration (Isaacson 1996, ¶5).

In sum, certain prominent properties of similarity can be conveniently formalized using the mathematical concept of the metric. I do not by any means claim that a similarity measure that does not define a metric cannot be compositionally or analytically useful. I merely propose that the metric provides a powerful tool with which to analyze the behavior of a similarity measure.

4.4 The metric and set-class similarity measures

I will now briefly consider similarity measures for pitch-class sets in terms of the metric in order to illustrate that many of these measures indeed do define a metric or a pseudometric. I will also provide a case in which the requirements of the metric are not satisfied, and this gives us some background for the next section on scaling.

First, it should be noted that only a dissimilarity measure can define a metric (or a pseudometric): the first requirement of the metric is that the distance between an element and itself must be zero. Therefore, similarity measures such as AK and MEMB_2 (Rahn 1979–80), REL (Lewin 1979–80) and $\cos\theta$ (Rogers 1999) cannot define a metric. Secondly, similarity measures based on the interval-class vector cannot define a metric on the set of set classes since they do not make a distinction between Z-related set classes. They can define a pseudometric, however, if the other requirements of the metric are satisfied.[3] Furthermore, some similarity measures such as IcVSIM and RECREL give the value zero to distinct set classes that are not Z-related. Thirdly, the requirement of symmetry does not pose problems since all the similarity measures for set classes are defined symmetrically. Therefore, we only need to consider triangle inequality to see if a dissimilarity measure defines a (pseudo)metric.

If a similarity measure gives the value zero to two distinct objects (and therefore it could only define a pseudometric), a violation of triangle inequality might arise when two objects with a zero distance have different distances to other objects. This is the case with Michael Buchler's SATSIM similarity measure (Buchler 1997). Figure 4.2 presents the SATSIM values for set classes 3-1[012], 3-10[036], and 4-28[0369]. The distance between set classes 3-10[036] and 4-28[0369] is zero, but they have different distances to set class 3-1[012], hence the following violation of triangle inequality:

$$\begin{aligned}\mathrm{SATSIM}([012],[036]) + \mathrm{SATSIM}([036],[0369]) &= 0.5 + 0 < 0.567\\ &= \mathrm{SATSIM}([012],[0369]).\end{aligned}$$

Consequently, SATSIM does not define a metric. Buchler acknowledges the existence of zero distances between distinct set classes, but since he does not discuss similarity measures in terms of

[3] Alternatively, they can define a metric on the set of interval-class vectors. However, this alternative is usually not promoted since – after Allen Forte's original formulation (Forte 1964) – pitch-class sets are customarily classified using the T_n/T_nI relation, not their interval-class vectors.

	3-1[012]	3-10[036]	4-28[0369]
3-1[012]	0	0.5	0.567
3-10[036]	0.5	0	0
4-28[0369]	0.567	0	0

Figure 4.2: The SATSIM values for set classes 3-1[012], 3-10[036], and 4-28[0369].

	3-8A[026]	4-25[0268]	5-13A[01248]
3-8A[026]	0	0	45
4-25[0268]	0	0	47
5-13A[01248]	45	47	0

Figure 4.3: The RECREL values for set classes 3-8A[026], 4-25[0268], and 5-13A[01248] (using the T_n classification of set classes).

the metric, he does not consider this phenomenon problematic.

Figure 4.3 presents the values of Marcus Castrén's RECREL similarity measure for set classes 3-8A[026], 4-25[0268], and 5-13A[01248].[4] RECREL gives the value 0 to set classes 3-8A[026] and 4-25[0268], but does not give identical values to these two set classes with respect to other set classes. Therefore, for example, we get the following violation of triangle inequality:

$$\begin{aligned}\mathrm{RECREL}([0268],[026]) + \mathrm{RECREL}([026],[01248]) &= 0 + 45 < 47\\ &= \mathrm{RECREL}([0268],[01248]).\end{aligned}$$

RECREL also has a few "genuine" violations: Figure 4.4 shows the RECREL values for set classes 4-2A[0124], 4-2B[0234], and 4-Z15B[0256], which give the following violation of triangle inequality:

$$\begin{aligned}\mathrm{RECREL}([0124],[0234]) + \mathrm{RECREL}([0234],[0256]) &= 4 + 40 < 46\\ &= \mathrm{RECREL}([0124],[0256]).\end{aligned}$$

Since triangle inequality is not satisfied, RECREL does not define a metric.

Isaacson defines his $\mathrm{I{\scriptsize C}VSIM}$ similarity measure as the standard deviation of the differences in the interval vectors of the set classes (Isaacson 1990, 16). As such, it gives the value zero to distinct set classes if the difference between the entries of the interval vectors is a constant.[5] However, this

[4] RECREL is defined in terms of T_n set classes.

[5] A constant difference denotes that there is no variance in the differences between the entries of the interval-class vectors. In total, there are 27 pairs of set classes, such that the difference is a non-zero constant. In addition, the difference between the entries of the interval-class vectors of Z-related set classes is, of course, always zero. For example, set classes 6-1[012345] and 7-1[0123456] have interval-class vectors $[543210]$ and $[654321]$, respectively. These two set classes could be considered relatively similar – ASIM gives them the value 0.167. On the other hand, set classes 3-4[015] and 6-43[012568] have interval-class vectors $[100110]$ and $[322332]$, respectively, and could thus be considered relatively dissimilar – ASIM gives them the value 0.667. It is quite counterintuitive that, according to $\mathrm{I{\scriptsize C}VSIM}$, set classes 3-4[015] and 6-43[012568] are maximally similar if we consider that the latter contains instances

	4-2A[0124]	4-2B[0234]	4-Z15B[0256]
4-2A[0124]	0	4	46
4-2B[0234]	4	0	40
4-Z15B[0256]	46	40	0

Figure 4.4: The RECREL values for set classes 4-2A[0124], 4-2B[0234], and 4-Z15B[0256] (using the T_n classification of set classes).

results in no violations of triangle inequality since the set classes that have the distance zero always have the same distances to the other set classes. Hence, as no violations of triangle inequality are found, IcVSIM defines a pseudometric.

Some similarity measures are based on a "norm" defined on the set classes (or rather on the interval-class vectors). A norm is a mathematical equivalent of the magnitude or size or length (in some sense) of an object. The magnitude should be understood here in an abstract sense: for example, interval-class vectors are objects in a six-dimensional space, and the entries of an interval-class vector are the magnitudes of a set class on those dimensions. Correspondingly, a similarity measure that is a norm measures the differences in magnitude of the set classes.[6]

The first three examples are norms. Teitelbaum's similarity index (1965, 88) defines a norm known as the Euclidean norm (or, more technically, as the L_2 norm). The "size" or "magnitude" of interval-class vector $x = [x_1\, x_2\, x_3\, x_4\, x_5\, x_6]$ is

$$|x| = \sqrt{x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2}$$

and consequently the distance between two interval-class vectors $x = [x_1\, x_2\, x_3\, x_4\, x_5\, x_6]$ and $y = [y_1\, y_2\, y_3\, y_4\, y_5\, y_6]$ is

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2 + (x_4 - y_4)^2 + (x_5 - y_5)^2 + (x_6 - y_6)^2},$$

which is precisely Teitelbaum's similarity index.

Morris' SIM defines a norm that is known colloquially as the "taxicab norm," or the "Manhattan

of all six interval classes whereas the former contains only instances of three.

[6]Strictly speaking, we are slightly cutting corners in labeling similarity measures as norms. A norm is generally defined in a *vector space*, one of properties of which is that it is closed under addition and scalar multiplication. In other words, we must be able to add vectors and multiply them by a scalar. It is not obvious that the interval-class vectors are members of a vector space. For example, the interval-class vectors of set classes 3-1[012] and 4-1[0123] are [210000] and [321000], respectively; adding them in the most obvious way results in [531000], but no set class has such an interval-class vector. Similarly, the multiplication of an interval-class vector by a scalar is not well defined. Consequently, we have two choices: we could accept the existence of anomalous interval-class vectors or we could associate interval-class vectors with elements of the vector space $\mathbb{R}^6$, make the necessary calculations in that space, and finally apply the results to the interval-class vectors. The latter approach seems more credible, but the details of its formalization lie outside the scope of this work.

norm[7].” The “size” of an interval-class vector $x = [x_1\, x_2\, x_3\, x_4\, x_5\, x_6]$ using the taxicab norm is

$$|x| = |x_1| + |x_2| + |x_3| + |x_4| + |x_5| + |x_6|$$

and consequently the distance between two interval-class vectors $x = [x_1\, x_2\, x_3\, x_4\, x_5\, x_6]$ and $y = [y_1\, y_2\, y_3\, y_4\, y_5\, y_6]$ is

$$d(x, y) = |x_1 - y_1| + |x_2 - y_2| + |x_3 - y_3| + |x_4 - y_4| + |x_5 - y_5| + |x_6 - y_6|,$$

which is precisely Morris’ SIM. Lord’s sf and Morris’ ASIM are scaled variants of SIM, and I will take a closer look at them in the next section.

Teitelbaum’s similarity index and Morris’ SIM belong to a family of norms known as L_p norms, the general formula for which is

$$|x| = \sqrt[p]{\sum x_i^p}$$

in which $p \geq 1$.[8] We obtain Teitelbaum’s similarity index by setting the value of p to 2, and correspondingly we obtain Morris’ SIM by setting the value of p to 1. The choice of p depends on how we wish to balance several small differences versus a few large differences. In general, the larger p is, the more weight is given to large individual differences of elements, and less to the total number of elements that are different.

Several other dissimilarity measures define a metric or a pseudometric, including Castren’s $\%\mathrm{REL}_2$ (Castrén 1994), and Rogers’ $\mathrm{IcVD1}$ and $\mathrm{IcVD2}$ (Rogers 1999). The $\%\mathrm{REL}_2$ measure defines only a pseudometric since $\%\mathrm{REL}_2$(3-5[016], 4-9[0167]) $= 0$, for example, and hence the fourth criterion of the metric ($d(x, y) = 0$ implies $x = y$) is not satisfied.

There are some dissimilarity measures that do not satisfy triangle inequality, however, which brings me to the issue of scaling – the topic of the next section.

4.5 Scaling the values

If desired, the values of similarity measures can be scaled to some suitable level. There are two basic types of scaling: linear and non-linear. Linear scaling means multiplying the values by some positive constant. For example, the equation

$$\mathrm{sf}(x, y) = \frac{1}{2} \cdot \mathrm{SIM}(x, y)$$

shows that similarity function sf is a linearly scaled version of SIM (and vice versa). Non-linear scaling, on the other hand, means modifying the values in a way that cannot be expressed as multiplying them by some constant. Taking the square root of values is an example of non-linear scaling.

Linear scaling can be used to adjust the maximum value of a similarity measure. As discussed in

[7] The name comes from the idea of measuring the distance that a taxi driver must drive in a rectangular grid of streets.

[8] Note that the formula for the L_p norms resembles the formula for the distance in L_p space (see Formula 3.1 on page 49). Indeed, the L_p norm defines a space in which distance between elements is calculated using a formula that is identical to it.

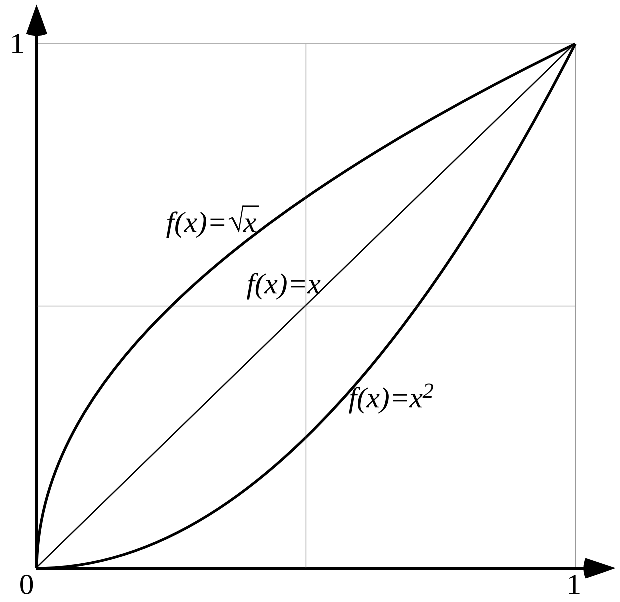


Figure 4.5: Graphs of the functions $f(x) = x^2$, $f(x) = x$, and $f(x) = \sqrt{x}$ on the domain $[0, 1]$.

Section 3.6.2, Castrén takes this approach and scales the values of similarity measures to the interval $[0, 100]$ for easy comparison.

This type of scaling does not affect triangle inequality. If a similarity measure d satisfies triangle inequality, then so does the scaled similarity measure $k \cdot d$: for any $k > 0$,

$$d(X, Y) + d(Y, Z) \geq d(X, Z) \Leftrightarrow k \cdot d(X, Y) + k \cdot d(Y, Z) \geq k \cdot d(X, Z).$$

In contrast, non-linear scaling may "make or break" triangle inequality. Let us consider a hypothetical similarity measure d and its non-linearly scaled variant e, which is defined by the formula $e(x, y) = d(x, y)^2$. Assume that d presents the following distances: $d(a, b) = 0.5$, $d(b, c) = 0.5$, and $d(a, c) = 1$. Since

$$d(a, b) + d(b, c) = 0.5 + 0.5 = 1 \geq d(a, c)$$

that triangle inequality holds. However, it does not hold in the scaled variant $e = d^2$ since

$$e(a, b) + e(b, c) = 0.5^2 + 0.5^2 = 0.25 + 0.25 = 0.5 < 1^2 = 1 = e(a, c).$$

Let us now consider squares and square roots in more detail in order to understand non-linear scaling better. Figure 4.5 shows the graphs of the functions $f(x) = x^2$, $f(x) = x$, and $f(x) = \sqrt{x}$ on the domain $[0, 1]$. The function $f(x) = x$ represents the case of no scaling. The function $f(x) = x^2$ squares the values and therefore emphasizes the differences between them. Therefore, in the above example, the similarity measure d satisfies triangle inequality but the squared values do not. The function $f(x) = \sqrt{x}$ behaves in the opposite way as it diminishes the differences between the values. There might thus be a similarity measure f that does not satisfy triangle inequality, but

d_1	[012]	[013]	[025]
[012]	0	2	4
[013]	2	0	2
[025]	4	2	0

d_2	[012]	[013]	[025]
[012]	0	2	6
[013]	2	0	2
[025]	6	2	0

d_3	[012]	[013]	[025]
[012]	0	$\sqrt{2}$	$\sqrt{6}$
[013]	$\sqrt{2}$	0	$\sqrt{2}$
[025]	$\sqrt{6}$	$\sqrt{2}$	0

Figure 4.6: The distances between set classes 3-1[012], 3-2[013], and 3-7[025] using the functions d_1, d_2, and d_3. The interval-class vectors of these set classes are [210000], [111000], and [011010], respectively. Measure d_1 is the taxicab metric on the interval-class vectors, measure d_2 is the sum of the differences of the squares of the corresponding entries in the interval-class vectors, and measure d_3 is the Euclidean metric.

the non-linearly scaled similarity measure $\sqrt{f}$ does. In fact, in the above example, the similarity measure $e = d^2$ does not satisfy triangle inequality but its square root $\sqrt{e} = \sqrt{d^2} = d$ does.

In order to illustrate the effect of squares and square roots, let us consider the distances between set classes 3-1[012], 3-2[013], and 3-7[025] using three distance functions labeled d_1, d_2 and d_3. These three distance functions "build" the Euclidean distance on the interval-class vectors bit by bit. Let us start by summing the differences between the corresponding entries, then the squares of the differences, and we finally take the square root of the sum. In particular, let us see whether triangle inequality

$$d_k([012], [013]) + d_k([013], [025]) \geq d_k([012], [025])$$

holds for each of the three distance functions d_1, d_2, and d_3. Figure 4.6 shows these distances.

The distance function d_1 is the taxicab metric, that is the sum of the absolute values of the differences in the corresponding entries in the interval-class vectors. In other words,

$$d_1(X, Y) = \sum_{i=1}^{6} |x_i - y_i|$$

where X and Y are set classes and x_i and y_i are the corresponding entries in their interval-class vectors (d_1 is also equal to Morris' SIM). It is well known that d_1 defines a metric; hence, triangle inequality holds. In the case of set classes 3-1[012], 3-2[013], and 3-7[025],

$$d_1([012], [013]) + d_1([013], [025]) = 2 + 2 \geq 4 = d_1([012], [025]).$$

Distance function d_2 is a modification of the taxicab metric in which the squares of the differences of the corresponding entries in the interval-class vectors are summed. In other words,

$$d_2(X, Y) = \sum_{i=1}^{6} (x_i - y_i)^2$$

where X and Y are set classes and x_i and y_i are the corresponding entries in their interval-class vectors. Now triangle inequality no longer holds. For example, in the case of set classes 3-1[012], 3-2[013], and 3-7[025],

$$d_2([012], [013]) + d_2([013], [025]) = 2 + 2 < 6 = d_2([012], [025]).$$

Finally, distance function d_3 is the Euclidean metric that is the square root of the sum of the squares of the differences of the corresponding entries in the interval-class vectors. In other words,

$$d_3(X, Y) = \sqrt{\sum_{i=1}^{6} (x_i - y_i)^2}$$

where X and Y are set classes and x_i and y_i are the corresponding entries in their interval-class vectors (d_1 is also equal to Teitelbaum's similarity index). Now triangle inequality holds again. For example, in the case of set classes 3-1[012], 3-2[013], and 3-7[025],

$$d_3([012], [013]) + d_3([013], [025]) = \sqrt{2} + \sqrt{2} > \sqrt{6} = d_3([012], [025]).$$

These three examples illustrate how squaring and square rooting affect the values. A distance function that violates triangle inequality might be "corrected" by taking the square root.

It should be noted that neither linear nor non-linear scaling of the values affects the "order" of similarity evaluations. If, for example, according to some similarity measure A is more similar to B than to C, then A is also more similar to B than to C according to a scaled variant of the measure.

Finally, let us consider the motivation for using scaling. First, we might use linear scaling to obtain comprehensible values as the minimum and/or maximum values. Secondly, we might use non-linear scaling to adjust the balance between several small differences versus a few large differences. Taking the squares into the similarity measure d_2 above emphasizes the large differences at the cost of the small ones. The process of finding the best scaling and the adjustment of the relative weights of small and large differences is part of the final evaluation of the results described in Section 3.6.4. Thirdly, taking the square roots into the similarity measure d_3 above results in a metric. Finally, there are cases in which scaling the values by taking logarithms gives a more comprehensible distribution of the values, as illustrated in Section 9.3 in the context of the discussion on the BSF similarity measure.

CHAPTER 5

The transformational approach to similarity

5.1 Transformational theory

Transformational music theory started to take shape in the last few decades of the 20th century. Attempting even to outline the scope of current transformational theory or its history in twentieth-century music theory would be out of place here; however, there are certain properties of this approach that are relevant when it is applied to similarity measures.[1]

Figure 5.1 reproduces David Lewin's Figure 0.1 depicting "two points s and t in a symbolic musical space" and "[t]he arrow marked i symbolizes a characteristic directed measurement, distance, or motion from s to t" (Lewin 1987, xi). The core of the transformational approach is that the focus is not on points s and t, but on the transformation that transforms s into t.

The entity-oriented and transformation-oriented approaches are not opposite, but are two angles to the same phenomenon, as evinced in the following comment by Robert Morris.

> In my view, the valorization of transformation over entity, or vice versa, is largely ideological, especially when this binary opposition is aligned with others such as noun versus verb, passive versus active, Cartesian versus Phenomenological. For, while groups of transformations act on musical entities, it is the changes in an entity's content that allow us to infer that a transformation is afoot. (Morris 2003a)

For example, the entity-oriented statement "pitch class 9 precedes pitch class 6 in row P and pitch class 6 precedes pitch class 9 in row Q" and the transformation-oriented statement "the transformation that transforms row P into row Q changes the order of pitch classes 9 and 6" state the very same thing through the idioms of the two approaches.

Correspondingly, similarity measures for twelve-tone rows are stated in entity-oriented terms in

[1]For an overview of the precursors of transformational theory see Morris (2003b).

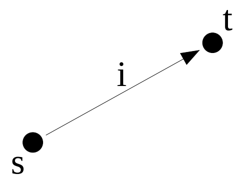


Figure 5.1: David Lewin's example of two points s and t in a symbolic musical space with the arrow marked i symbolizing a directed measurement from s to t.

the existing literature. I will show in Part III that several of the proposed measures have a transformational reading even if the original authors do not explicitly give them in transformational terms.

Milton Babbitt's work on the transformational approach to twelve-tone rows is seminal. His notion of twelve-tone rows as permutations is inherently transformational: a row constitutes a linear ordering of the pitch classes that define a set of order relations, and row operations are transformations of these order relations. Babbitt sowed the seed of this transformational idea in his dissertation (Babbitt 1946/1992), and explicitly presented it in a series of articles in the 1960s (Babbitt 1960; Babbitt 1961b).

The surge of interest in transformational theory at the end of the 20th century was largely due to the momentum created by the work of David Lewin. In order to explicate his musical insights he developed a technical vocabulary with an unprecedented power of expression, elegance and precision. I will use his concept of *Generalized Musical Intervals and Transformations* to formalize my transformational approach to the similarity of twelve-tone rows.

5.2 Transformation theory and similarity

The dichotomy of entity versus transformation gives rise to two conceptual approaches to the similarity of musical objects. The first approach, based on entities, is to define and compare their properties. The more properties they share, the more similar they are. The second approach is to devise a transformation that maps one entity into another and then their similarity is evaluated somehow based on the magnitude or complexity of the transformation. The smaller or less complex the transformation is, the more similar the two entities are, and the more complex it is, the less similar they are. Dissimilarity of entities then equals the amount of change or modification needed to transform one object into the other. Nothing needs to be changed to keep an entity as it is, and the more dissimilar the entities are the more alternations we need to make. Orpen and Huron suggest this approach: in measuring similarity we evaluate "how much 'tinkering' is required in order to reach identity" (Orpen and Huron 1991, 5).

The transformational approach is well established in psychology. In his overview of the research on similarity in cognitive psychology, Goldstone (1999) presents four approaches. He cites geometrical models as the most influential in terms of analyzing similarity. These models include traditional similarity measures as they measure the distance between entities. He also mentions featural, alignment-based, and transformational models, and it is last of these that I will adopt here.

In the realm of twelve-tone rows we could consider the geometric and transformational approaches two sides of the same phenomenon. For example, assume pitch class p is in one row at order position $\boldsymbol{n_1}$ and in another row at order position $\boldsymbol{n_2}$. If we compare the properties of the rows, we could say that the difference between the order positions of pitch class p is $|\boldsymbol{n_1} - \boldsymbol{n_2}|$. In transformational terms we would note that the transformation that maps one row into the other moves pitch class p by $|\boldsymbol{n_1} - \boldsymbol{n_2}|$ order positions. Hence, we could define a similarity measure in both terms, as I do. Naturally, it is necessary to prove that the non-transformational and transformational definitions are equivalent.

The transformational approach is applicable as long as the similarity measure can be interpreted in terms of measuring the magnitude of the transformation.

The measurement of set-class similarity is traditionally carried out according to first approach of comparing the properties (interval-class vectors or subset-class contents) of the set classes. I argue here that the transformational approach is often better suited to the evaluation of the similarity of twelve-tone rows. Furthermore, it has the advantage that it allows us to gather information about the network of distances in the whole set of twelve-tone rows. Transformations also play a crucial role in the discussion on similarity relations between row classes.

Finally, if the approach were extended to segments of pitch classes of different cardinalities, there are established methods in the transformational paradigm for comparing sequences of different lengths.

5.3 Pitch-class transformations and order-number transformations

In Section 2.1.2 I discussed the idea that pitch-class rows and order-number rows are two sides of twelve-tone rows, and I developed this idea further in Section 2.2.5 in which I examined the concepts of pitch-class operations and order-number operations. The same thinking could be applied to the similarity of twelve-tone rows: there is the similarity in the pitch-class space and the similarity in the order-number space.

Historically, the similarity of twelve-tone rows has been explored almost exclusively in terms of order relations. Indeed, according to John Ward, “Similarity among ordered sets is measured not in terms of pitch-class or interval content but in terms of ordering attributes” (Ward 1992, 77). I extend the notion of the duality of pitch-class representation and order-number representation to similarity measures. The similarity of order relations is measured by applying some formula to the order-number transformation that transforms one row into the other. Given the duality, we should be able to apply the same formula to the pitch-class transformation that transforms one row into the other. However, it turns out that we do not always have a good intuition of what the result would be when we take a measure developed for the order-number transformations and “translate” it into a measure for the pitch-class transformations. This casts serious doubt on the assumed isomorphism between the two realms.

5.4 Rows and transformations as a GIS

David Lewin’s *Generalized Interval System*, or *GIS*, provides a natural starting point for the formalization of transformational relations between twelve-tone rows (Lewin 1987). Points s and t in

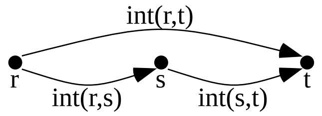


Figure 5.2: The composition of intervals in GIS. Condition (A) of GIS states that the composition of intervals from r to s and from s to t must equal the interval from r to t.

Figure 5.1 are twelve-tone rows and interval i is the transformation that transforms row s into row t. Naturally, s and t are members of the set of 479001600 distinct twelve-tone rows and transformation i is defined on the whole set, not only on the two rows. Thus transformation i (together with other transformations) turns the set of rows into a space.

Even if in the Lewinian tradition transformations and intervals are close to synonymous, row transformations are not usually represented as intervals between the rows. A natural explanation is that we only have a standard notation for transformations between rows in the same row class, such as T_4, T_4I, RT_4I, etc. However, no standard notation exists for the transformation or interval between the rows $5409728136\mathrm{AB}$ and $\mathrm{A}463592178\mathrm{B}0$, for example – the rows of Alban Berg's *Lyric Suite* and Arnold Schoenberg's *Variations* op. 31. Hence we need the notion of a *generalized* interval. An interval contains the description of how to transform one musical object (in some space) into another musical object.

A Generalized Interval System contains three elements: a space S, a set of intervals $IVLS$, and a function int that connects the two. The formal definition of GIS is given below. There is a notational difference between this and Lewin's definition: Lewin uses right orthography, but since the intervals in question are permutations I use left orthography here (see Appendix A). Hence in the composition of two functions $int(s,t)int(r,s)$ we compose the interval between r and s with the interval between s and t, not vice versa.

DEFINITION 5.1 GIS is an ordered triple $(S, IVLS, int)$ where S is a set, $IVLS$ is a group, and int is a function $int : S \times S \to IVLS$ that satisfies the following two conditions:

(A): For all $r, s, t \in S$, $int(s,t)int(r,s) = int(r,t)$

(B): For every $s \in S$ and $i \in IVLS$, there is a unique $t \in S$ such that $int(s,t) = i$.

Let us briefly review some properties of the definition. Essentially, function int connects every ordered pair of elements in space S to an interval in group $IVLS$. In other words, if we take two elements s and t in space S, function int tells us which is the (unique) interval or transformation in group $IVLS$ that transforms element s into element t. Condition (A) states that the composition of intervals from r to s and from s to t must equal the interval from r to t. This is depicted in Figure 5.2: if we decompose the interval between r and t into two intervals (the interval between r and s and the interval between s and t), the composition of the sub-intervals must equal the original interval. Condition (B) states that we should not have a surplus of intervals: given an element s in space S, for every interval i in group $IVLS$ there must be exactly one element t in space S that is at interval i from element s.

In this GIS space S is either the set of pitch-class rows or the set of order-number rows formalized as a set of permutations. The context will indicate which is the intended interpretation. Consequently, each method of evaluating the complexity of a transformation produces a pair of similarity measures: one for the pitch-class aspect and the other for the order-number aspect. It turns out that these two similarity measures coincide in two cases, but not in others.

The group of intervals $IVLS$ is the group of permutations of twelve elements, known as the symmetric group S_{12}. These are the all possible pitch-class transformations or the all possible order-number transformations. I discussed the idea of defining 479001600 pitch-class operations and, correspondingly, 479001600 order-number operations in Section 2.2.5. This bears fruit now since these operations are precisely the intervals of the GIS.

Hence, the permutations have a dual role: both twelve-tone rows and intervals are formalized as permutations. However, as discussed in Section 2.1.3, twelve-tone rows are permutations in the sense of a linear ordering, and intervals or transformations in the sense of a mapping. Consequently, space S is the *set* of twelve-tone rows formalized as a *set* of permutations with no defined internal (group) structure, while the *group* of transformations $IVLS$ is the symmetric *group* S_{12}.[2] Finally, function int is derived naturally from the rules covering the composition of permutations, as illustrated below.

Let us first define the GIS of twelve-tone rows and then prove that it satisfies conditions (A) and (B) of Definition 5.1.

DEFINITION 5.2 The GIS of twelve-tone rows is an ordered triple $(S, IVLS, int)$ where S is the set of permutations S_{12}, $IVLS$ is the group of permutations S_{12}, and int: $S \times S \to IVLS$ is defined as $int(r, s) = sr^{-1}$.

The dual role of permutations in the above definition is worthy of comment. In the expression $int(r, s) = sr^{-1}$ function int takes as its parameters two permutations r and s in the sense of linear orderings. However, its value is defined in terms of two permutations r and s in the sense of mappings. Thus, the dual nature of permutations is utilized in the interpretation of the linear orderings r and s as mappings in order to create the appropriate transformation. Nevertheless, I impose no group structure on the set of twelve-tone rows.

LEMMA 5.1 The GIS of twelve-tone rows satisfies conditions (A) and (B) of Definition 5.1.

Proof. Let r, s, and t be arbitrary twelve-tone rows. First, condition (A) of Definition 5.1 states that the composition of the intervals from r to s and from s to r must equal the interval from r to s. This is shown by the equation

$$int(s, t)int(r, s) = (ts^{-1})(sr^{-1}) = t(s^{-1}s)r^{-1} = tr^{-1} = int(r, t).$$

[2] Indeed, Oren Kolman has shown that any GIS (S, G, int) is always isomorphic to a canonical GIS (G, G, int) where its space is the same as the group of intervals but without the group structure (Kolman 2004).

Secondly, if $s \in S$ and $i \in IVLS$ are arbitrary, then for $t = is \in S$, $int(s,t) = int(s,is) = iss^{-1} = i$. Furthermore, t is unique due to the group structure of S_{12}. Thus, the GIS of twelve-tone rows satisfies conditions (A) and (B) of Definition 5.1. □

The intervals of the GIS of twelve-tone rows are transformations. Hence, with respect to the GIS of twelve-tone rows, I will use the terms "interval" and "transformation" synonymously.

Let us define the following terminology. There is a need for two Generalized Interval Systems – one for pitch-class rows and one for order-number rows – labeled the GIS of pitch-class rows and the GIS of order-number rows. For easy reference, the formal definitions are given below.

DEFINITION 5.3 The *GIS of pitch-class rows* is a GIS $(S, IVLS, int)$ where S is the set of pitch-class rows S_{12}, $IVLS$ is the group of permutations S_{12}, and int is defined as $int(P,Q) = QP^{-1}$.

The ***GIS of order-number rows*** is a GIS $(S, IVLS, int)$ where S is the set of order-number rows S_{12}, $IVLS$ is the group of permutations S_{12}, and int is defined as $int(\boldsymbol{P}, \boldsymbol{Q}) = \boldsymbol{QP}^{-1}$.

The definitions of the two Generalized Interval Systems are identical except for the interpretations of space S: in the former it is interpreted as the set of pitch-class rows and in the latter as the set of order-number rows.

Interval QP^{-1} that maps row P into row Q does not necessarily equal interval PQ^{-1}, which maps row Q into row P. However, since

$$(PQ^{-1})(QP^{-1}) = P(Q^{-1}Q)P^{-1} = id = Q(P^{-1}P)Q^{-1} = (QP^{-1})(PQ^{-1}),$$

intervals QP^{-1} and PQ^{-1} are inverse permutations. This observation will prove useful in consideration of the symmetry properties of certain similarity measures.

5.5 From a GIS to a similarity measure

The GIS of twelve-tone rows allows us to define a transformation that transforms one row into the other. I will now discuss how we can use the intervals in a GIS to define similarity measures.

Henry Klumpenhouwer discusses Lewin's conception of the interval as follows.

> Accordingly, Lewin ascribes to "interval" not only magnitude but direction. [...] The process of measuring distance and direction from one element to another quite naturally becomes the process of evaluating the means by which one element becomes another. In this sense Lewin's interval is transformational. (Klumpenhouwer 1991, 4:2)

Curiously, when we evaluate the similarity of two rows using the transformational approach we reverse the process that Klumpenhouwer describes. We *begin* with the transformation by means of which one twelve-tone row becomes another. Then we proceed to assess the magnitude of that transformation in order to evaluate the similarity of the rows.

As established in Section 5.2, the core of the transformational approach is the idea that the similarity of two rows is defined by the magnitude or complexity of the transformation or interval between them. Hence, we need to add to the GIS of twelve-tone rows a function d from the group of intervals $IVLS$ to the nonnegative real numbers; this function is the measurement of the magnitude of the transformation.

Pitch-class space provides a similar phenomenon. Transposition T_1 moves pitch class E to pitch class F, and transposition T_{11} moves pitch class F to pitch class E. The two transpositions are different; however, we could argue that they have the same "magnitudes" or "sizes." Transpositions T_1 and T_{11} have the same magnitude, but different directions. Correspondingly, we could map the twelve transpositions $T_0, T_1, \ldots, T_{11}$ into the seven interval classes; the interval classes represent their magnitudes.

The next schema formalizes this idea and it can be used to create several different similarity measures. In fact, several of the similarity measures to be discussed in Part III of this work could be considered characterizations of the transformation that transforms one row into another.

DEFINITION 5.4 (MEASURE SCHEMA) If X and Y are twelve-tone rows in a GIS and $int(X, Y) = k$, then the dissimilarity of rows X and Y is measured by the magnitude of interval k.

There are several ways of defining the "magnitude" of interval k: the transformational similarity measures discussed in Part III are different ways of evaluating the magnitude of the transformations. Since the transformations are permutations, their magnitude could be interpreted as the complexity of the permutation.

In computer science, the "presortedness" of permutations has been studied with a view to developing efficient sorting algorithms.[3] This concept of presortedness measures the extent to which a permutation differs from a referential (usually ascending) permutation, and hence it coincides with the concept of the magnitude of the interval used in this study. Therefore, we can easily translate the results on presortedness into results on the magnitudes of intervals.

As discussed in Chapter 4, the aim is for a similarity measure to define a metric. This imposes some requirements on how the magnitude of the transformation is measured. It is relatively straightforward to translate the requirements of the metric into requirements for the transformation, which is the content of the following lemma.

LEMMA 5.2 Let d be a function from group $IVLS$ of the GIS of twelve-tone rows to the nonnegative real numbers. The similarity measure on the space S of GIS, defined by d, defines a metric if the following four conditions are satisfied: (i) $d(id) = 0$, (ii) $d(x) = d(x^{-1})$, (iii) $d(x) + d(y) \geq d(yx)$, (iv) $d(x) = 0$ implies $x = id$.

Proof. First, the permutation that transforms a row into itself is the identity permutation id. Hence, if $d(id) = 0$ then the distance between a row and itself is zero.

Secondly, if $x = YX^{-1}$ is a transformation mapping row X into row Y, then $x^{-1} = (YX^{-1})^{-1} = XY^{-1}$ is a transformation mapping row Y into row X, and the requirement of symmetry in the metric translates into a requirement of equal values for inverse permutations.

[3]The word "presortedness" was coined by Mehlhorn (1979). Mannila (1985) provides a formalization of optimal sorting algorithms with respect to presortedness.

Thirdly, if x is the transformation that maps row A into row B and y is the transformation that maps row B into row C, then yx is the transformation that maps row A into row C, since

$$yx = (CB^{-1})(BA^{-1}) = C(B^{-1}B)A^{-1} = CA^{-1}.$$

Hence, the condition $d(x) + d(y) \geq d(yx)$ is triangle inequality.

Finally, the condition $d(x) = 0$ implies $x = id$ does not allow the value zero for distinct rows. Thus, the four criteria are equivalent to the four criteria of the metric. □

Thus, the transformational approach combines well with the metric, but it provides neither necessary nor sufficient conditions for it: I will give an example of a transformational similarity measure that does not define a metric (see the discussion on the SCATTERING similarity measure in Section 11.3), and a non-transformational similarity measure that defines a pseudometric (see the discussion on the INTERVALLIC DISTANCE similarity measure in Section 12.2).

I will conclude this section on a historical note. I detect a germ of the transformational approach in Morris (1987). Indeed, some of the formulas for similarity measures that Morris provides have a transformational reading since they are based on a comparison of identity transformation and a function transforming one row into another. However, at the time of writing Morris did not intend a transformational interpretation.[4]

Similarly, John Ward bases his formulas for similarity measures directly on those of Morris. Hence, we could read the formulas in transformational terms.

> In every similarity measure, two sets, P and Q, are the objects of comparison. Whether the sets are pc sets or mpc sets, they must be converted to order mappings of the type $\mathrm{OM_PP}$ and $\mathrm{OM_PQ}$. If P(F, Q) is a function that evaluates the similarity between ordered sets P and Q, it can also be expressed as F($\mathrm{OM_PP}$, $\mathrm{OM_PQ}$), which itself can be expressed as F(P0, Pn). (Ward 1992, 78)

Here, $\mathrm{OM_PQ}$ is an order-number mapping that transforms row P into row Q, and it is compared to the identity mapping, labeled here P0, that keeps every row intact. However, Ward's treatment of similarity measures does not indicate measuring the similarity of rows in terms of measuring the complexity of their transformational relation.

5.6 Left invariance

The concept of *left invariance* is useful in terms of analyzing similarity measures.[5] The following definition is adapted from Chapter II of Critchlow (1985) to left orthography.

DEFINITION 5.5 Metric d on S_n is left invariant if $d(\pi, \sigma) = d(\tau\pi, \tau\sigma)$ for all $\pi, \sigma, \tau \in S_n$.

[4]Morris in private communication.

[5]The concept of left invariance is often known as *right invariance* since right orthography is usually used. However, as I use left orthography here I define the concept as left invariance.

The idea of left invariance is to guarantee that the distances between objects do not depend on how the objects are labeled. In Definition 5.5, permutation τ is applied to permutations π and σ in order to "relabel" the entities in them. For example, if permutation τ maps n into m we interpret it as relabeling n as m in both permutations π and σ.

Let us first consider left invariance with respect to pitch-class rows. I would stress that the above definition is a metric on the *group* S_n. In the case of pitch-class rows the variables π and σ are interpreted as arbitrary pitch-class rows and variable τ is interpreted as an arbitrary pitch-class operation. In this context left invariance means that we are thinking purely in permutational terms and only the ordering relations of the twelve pitch classes matter.

Curiously, similarity measures based on the GIS of order-number rows provide left invariance for pitch-class rows. In other words, even if we do not customarily think in such terms, any pitch-class operation, such as a transposition, could be seen as relabeling the pitch classes. Hence, the application of any pitch-class operation to pitch-class rows amounts to a relabeling of the pitch classes, but the order relations between the elements of the rows are not changed. For example, if pitch class 6 is in order position **0** in one row and in order position **2** in another, we could say that the distance between these order positions is $2 - 0 = 2$. If we now transpose both rows by T_1, the distance between the order numbers of pitch class 7 in the transposed rows will be $2-0 = 2$. Hence, in this measurement it is immaterial how the pitch classes are labeled.

Let us now consider rows $p_0p_1p_2p_3p_4p_5p_6p_7p_8p_9p_{10}p_{11}$ and $p_1p_0p_2p_3p_4p_5p_6p_7p_8p_9p_{10}p_{11}$. These two rows identical except for the order of the first two pitch classes. If we transpose both rows by the same pitch-class interval, we obtain rows $q_0q_1q_2q_3q_4q_5q_6q_7q_8q_9q_{10}q_{11}$ and $q_1q_0q_2q_3q_4q_5q_6q_7q_8q_9q_{10}q_{11}$. These two rows are again identical except for the order of the first two pitch classes. With respect to the order relations, only the pitch classes were relabeled. The same would happen if any pitch-class operation was applied to rows $p_0p_1p_2p_3p_4p_5p_6p_7p_8p_9p_{10}p_{11}$ and $p_1p_0p_2p_3p_4p_5p_6p_7p_8p_9p_{10}p_{11}$.

Left invariance therefore means that we are thinking purely in combinatorial terms and do not utilize the relations of the pitch classes. In particular, none of our knowledge of the intervallic relations between pitch classes is used in a left-invariant similarity measure. For example, with respect to order relations in twelve-tone rows it is immaterial that pitch classes 4 and 5 are adjacent but pitch classes 4 and 10 are not: the focus is on the order relations and only on the order relations.

Correspondingly, similarity measures based on the GIS of pitch-class rows provide left invariance for order-number rows. Hence, the application of any order-number operation to order-number rows amounts to a re-positioning of the pitch classes in the rows, but the relations between the corresponding pitch classes are not changed. For example, the same pitch classes are paired in rows and their retrogrades (if in two rows we pair the initial pitch classes n and m, in their retrogrades we pair the same pitch classes as the last pitch classes).

5.7 Interval-preserving transformations

Lewin discusses interval-preserving transformations at some length in Section 3.4 of Lewin (1987). The definition below reproduces his Definition 3.4.6.

DEFINITION 5.6 Given a GIS (S, IVLS, int), a transformation X on S will be called *interval-preserving* if X has this property: For each s and each t, $int(X(s), X(t)) = int(s, t)$.

Let us now examine the interval-preserving transformations in the GIS of twelve-tone rows. First, note that the requirement of an interval-preserving transformation is stronger than that of left invariance discussed in Section 5.6, and of transformational coherence discussed in Section 3.7.[6] Let F be a transformation. If F is an interval-preserving transformation then the interval between any two rows P and Q and transformed rows must be the same. In other words, $\mathit{int}(P, Q) = \mathit{int}(FP, FQ)$ must hold for all rows P and Q. If a similarity measure is transformationally coherent under transformation F, it is only required that the magnitudes of the intervals between the original rows P and Q and the transformed rows FP and FQ are the same, but the intervals do not need to be the same. In other words, if similarity measure d is transformationally coherent under transformation F it is required that $d(\mathit{int}(P, Q)) = d(\mathit{int}(FP, FQ))$ but it is not required that $\mathit{int}(P, Q) = \mathit{int}(FP, FQ)$. All (transformational) similarity measures are transformationally coherent under interval-preserving transformations but the opposite does not necessarily hold.

If we measure the interval between two rows in the order-number dimension, then the interval-preserving transformations are precisely the pitch-class operations. Correspondingly, if we measure the interval between two rows in the pitch-class dimension, then the interval-preserving transformations are precisely the order-number operations.

For instance, if we compare the rows $P = 0123456789\mathrm{AB}$ and $Q = 1023456789\mathrm{AB}$ in the order-number dimension, the first two pitch classes (order positions $\mathbf{0}$ and $\mathbf{1}$) are exchanged and both rows have the same pitch classes in the remaining order positions. Now, if we apply any pitch-class operation F to these two rows, in the resulting rows FP and FQ the first two pitch classes (order positions $\mathbf{0}$ and $\mathbf{1}$) are again exchanged and both rows have the same pitch classes in the remaining order positions. However, if we retrograde rows P and Q, in the resulting rows RP and RQ the last two pitch classes (order positions $\mathbf{10}$ and $\mathbf{11}$) are exchanged and both rows have the same pitch classes in the remaining order positions. However, in the pitch-class dimension the interval between rows P and Q is the same as the interval between rows RP and RQ.

[6] Note that "interval-preserving" is a property that transformations may have in a given GIS and thus it is independent of the similarity measures. In contrast, transformational coherency is a property that a similarity measure may have under some transformations.

CHAPTER 6

Topologies of twelve-tone rows

John Rahn describes the network of similarity relations between set classes as "staggeringly complex" (Rahn 1979–80, 494). The network of similarity relations or distances between rows is certainly more complex than that of set classes, but it is possible to give a reasonable account of the behavior of the similarity measures by observing some of their general characteristics.

This chapter considers the gamut of values that the similarity measures for twelve-tone rows produce from a variety of perspectives. The symmetry of the spaces that some of the measures induce on the set of twelve-tone rows is discussed in Section 6.1, and the idea of similarity between row classes is discussed in Section 6.2. The properties of the distribution of the values is examined in section 6.3, and the chapter ends with a brief discussion on the idea of relative similarity.

6.1 Symmetries of row spaces

I would stress once more that twelve-tone rows constitute a set. They do not have any structure *a priori*, but there are numerous ways of imposing a structure on the set in order to create a space. Accordingly, I use the plural form and refer to "symmetries of row spaces." In the present case, I use the GIS of pitch-class rows and the GIS of order-number rows to impose transformational relations and, ultimately, distances on that set.

The transformational approach allows us to draw some remarkable conclusions about the similarity relations in the whole set of twelve-tone rows. A twelve-tone row is related to every other twelve-tone row by exactly one of the 479001600 pitch-class transformations and exactly one of the 479001600 order-number transformations. Consequently, every row is related to the other rows by precisely the same set of transformations. Therefore, if we define the distance between two rows based on the transformation that maps one row into the other, every row has the same network of distances to the other rows.[1] Thus, unlike the set of pitch-class sets, twelve-tone-row spaces may be

[1]Naturally, this does not mean that if row A is at distance n from row A' then all rows are at distance n from row

perfectly symmetrical, and every row has an identical neighborhood in a transformational sense.

This symmetry is depicted by the very definition of GIS (see Definition 5.1). Condition (B) states that for every row s in space S and for every interval i in $IVLS$ there is a unique row t in space S that lies at interval i from s. Since this applies to every row it follows that every row has the same set of transformations defining the rows related to it.

However, it should be noted that it would be perfectly possible to define transformational spaces that are not symmetrical. For example, in the case of the classic group of 48 row operations some rows are related to 48 rows while others are related only to 24 rows. This is clearly a space defined by transformations, but it is not symmetric since the rows have different relations to other rows. Consequently, this space cannot be formalized using a GIS.

6.2 The similarity of row classes

Typically, in the realm of unordered pitch-class sets, similarity measures are defined specifically for set classes and not for individual pitch-class sets (even if they can then be applied to them). Correspondingly, we might ask if we could define a similarity measure for row classes and not just twelve-tone rows, and if so how would the two be related.

Row classes are equivalence classes that comprise twelve-tone rows related by a group of canonical row operations, and in this work the row class is defined by the 48 classic row operations: transposition, inversion, retrograde, and their combinations. If there is some metric defined in the set of all twelve-tone rows, the distance between (or dissimilarity of) any two is defined. However, it might happen that, while some two rows might not be particularly similar, there might be others in their respective row classes that turn out to be more similar.[2] Therefore it is also useful to consider the similarity of row classes and not only of individual twelve-tone rows.

Let us consider the following example. According to the DERANGEMENT similarity measure, the more two twelve-tone rows contain the same pitch classes in the same order positions, the more similar they are (see Section 8.2). The two twelve-tone rows $A = 0123456789\mathrm{AB}$ and $B = 10\mathrm{B}23456789\mathrm{A}$ do not contain a single pitch class in the same order position and thus, according to DERANGEMENT, they are maximally dissimilar.[3] However, rows A and $T_1B = 2103456789\mathrm{AB}$ are very similar. Even if the similarity of rows A and B might be easy to spot here, in a more complex case we might obtain extra information about rows by considering the similarity of the row classes to which they belong.

Let us start with an abstract definition of the distance between any two finite sets of any objects

A'. It rather means that *every* row has the same number of rows at distance n.

[2]Set-class similarity measures are based on the properties of set classes. When in an analytical context we wish to compare two pitch-class sets we first derive the pertinent set classes and then examine their similarity. For example, pitch-class sets $\{0, 1, 2, 3, 5\}$ and $\{6, 7, 8, 10, 11\}$ do not share a single non-trivial subset. However, by transposing the latter by 6 semitones we obtain the pitch-class set $\{0, 1, 2, 4, 5\}$ and it is easy to see that $\{0, 1, 2, 3, 5\}$ and $\{0, 1, 2, 4, 5\}$ share many subsets. A set-class similarity measure based on subset-class contents detects the similarity of the subset-class contents of the set classes to which $\{0, 1, 2, 3, 5\}$ and $\{6, 7, 8, 10, 11\}$ belong; hence, the fact that they do not share a single nontrivial subset does not mean that the underlying set classes could not have shared subset-class contents.

[3]Naturally, rows A and B are very similar with respect to the ordered dyad contents (see Chapter 9) or unordered INT contents or unordered interval-class contents (see Chapter 12), for example, but here it is only a question of whether they contain the same pitch classes in the same order positions.

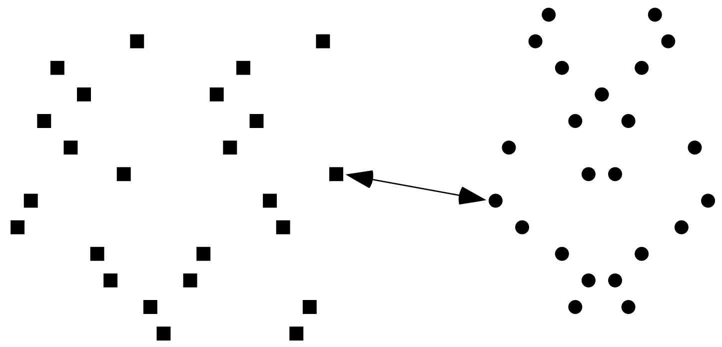


Figure 6.1: Two row classes in a symbolic musical space.

in any space. This distance is naturally defined as the distance between their closest members.[4] The formal definition is given below.

DEFINITION 6.1 Let d be a metric on set S and let X and Y be finite (non-empty) subsets of S. The distance between X and Y is the minimum of the distances between the elements of X and Y. Formally, $d(X,Y) = min\{d(x,y) \mid x \in X, y \in Y\}$.

Figure 6.1 depicts two row classes, denoted by squares and circles, in a symbolic musical space. According to Definition 6.1, the distance between the two row classes is the distance between the closest members, denoted by an arrow.

The distances between sets – if based on a metric – satisfy its first two requirements. First, the distance between a set and itself is the shortest distance between elements within that set, which is always zero. To show that $d(X,X) = 0$, let us pick any element $x \in X$. Then because d defines a metric, $d(x,x) = 0$ and, hence, $0 \leq d(X,X) \leq d(x,x) = 0$.

Secondly, the symmetry of the distances between sets ensues from the symmetry of the metric between their elements.

However, the distances between sets of elements do not necessarily constitute a metric, since the third requirement, triangle inequality, does not always hold for sets of objects. A typical example of such a situation is given in Figure 6.2, in which X, Y, and Z are sets of points in a two-dimensional plane, and the distance between sets X and Z is clearly greater than that between sets X and Y plus the distance between sets Y and Z: the sum of the lengths of the arrows between X and Y and Y and Z is clearly smaller than the length of the arrow between X and Z. Thus, if a metric is defined for (single) points in a space, triangle inequality does not necessarily hold for the distances between sets of these points.

The fourth requirement for a metric is also not necessarily satisfied. On the one hand, if the intersection of two non-identical sets is not empty (in other words they share at least one element) their distance is zero even if they are not the same set. In such a case we might only have a pseu-

[4]In the case of finite sets, closest members always exist even if they are not necessarily unique. This is not the case with infinite sets and we would need to define the distance as the infimum (the greatest lower bound) of the distances between the elements.

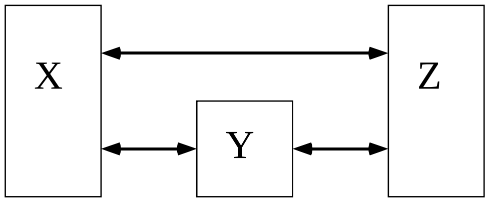


Figure 6.2: Distances between sets X, Y and Z.

dometric. Furthermore, in the case of infinite sets, two distinct sets may have a zero distance even if the distance between every pair of elements in them is strictly positive.[5] In these cases, even if triangle inequality were satisfied, we could only have a pseudometric. However, the interest here is in the similarity of row classes that are distinct sets in a finite space, and therefore the fourth criterion does not cause problems: it is satisfied for row classes if it is satisfied for rows.

It turns out, however, that the distances between row classes, in general, do define a metric provided that we start with one for rows. To see why this is the case, let us return to the idea of transformational coherence introduced in Section 3.7.

Transformational coherence was introduced in the discussion on the criteria for a similarity measure when the measure is required to give the same values in comparable cases and thus to be well-behaving. Since transposing, inverting or retrograding will always give a comparable case in the case of twelve-tone rows, this criterion of transformational coherence will provide a guarantee that a similarity measure that defines a metric will also define a metric for row classes.

It can now be proved that the distances between the equivalence classes defined by group G acting on set S_{12} satisfy triangle inequality – the crucial part of the metric.

LEMMA 6.1 Let d be a metric on set S and let G be a permutation group on set S. If the equation $d(x, y) = d(g(x), g(y))$ holds for all elements $x, y \in S$ and all permutations $g \in G$ then triangle inequality holds for the distances between the equivalence classes induced by group G.

Proof. Let X, Y and Z be equivalence classes induced by group G. Now, since the equivalence classes are finite, $d(X, Y) = d(x_1, y_1)$ for some $x_1 \in X$ and $y_1 \in Y$, and $d(Y, Z) = d(y_2, z_2)$ for some $y_2 \in Y$ and $z_2 \in Z$. Since $y_1, y_2 \in Y$ there exists a permutation g in group G such that $g(y_1) = y_2$ and $g(x_1) = x_2$. Since we required that $d(x_1, y_1) = d(g(x_1), g(y_1)) = d(x_2, y_2)$, the inequality

$$d(X, Z) \leq d(x_2, z_2) \leq d(x_2, y_2) + d(y_2, z_2) = d(x_1, y_1) + d(y_2, z_2) = d(X, Y) + d(Y, Z)$$

proves the lemma. □

This is a non-trivial result showing that if the distance between two elements in set S does not change when an operation is applied to them, triangle inequality $d(X, Z) \leq d(X, Y) + d(Y, Z)$ for

[5] For example, the sets of strictly positive and strictly negative real numbers are distinct, but nevertheless, the distance between these two sets is zero.

equivalence classes holds.

The lemma gives a sufficient condition under which triangle inequality holds for equivalence classes induced by a permutation group. Furthermore, if we translate the lemma into the language of twelve-tone rows and row operations, the equivalence classes are row classes and the condition $d(x, y) = d(g(x), g(y))$ applied in the lemma is that of transformational coherence – the one proposed earlier as the guarantee for the coherence of the distances. Thus, since equivalence classes are, by definition, distinct, we obtain a metric for finite equivalence classes that satisfies the condition of the lemma.[6]

The lemma is the guarantee that our quest for a row-class similarity measure has a sound basis. Let us begin with a condition of transformational coherence – a condition expressing a natural requirement that the degree of similarity of two twelve-tone rows should not change when the same row operation is applied to both. Then we deduce that if the distance between row classes $[X]$ and $[Y]$ is d_1 (in other words, there are rows X_1 and Y_1 that belong to row classes $[X]$ and $[Y]$, respectively, at the distance d_1) and the distance between row classes $[Y]$ and $[Z]$ is d_2 (in other words, there are rows Y_2 and Z_2 that belong to row classes $[Y]$ and $[Z]$, respectively, at the distance d_2), then the distance between row classes $[X]$ and $[Z]$ is at most $d_1 + d_2$ (in other words, there are rows X_3 and Z_3 that belong to row classes $[X]$ and $[Z]$, respectively, at most at the distance $d_1 + d_2$).

Figure 6.3 illustrates the proof of Lemma 6.1 by breaking it into small steps. Three row classes $[X]$, $[Y]$, and $[Z]$ are represented as sets of triangles, circles, and squares, respectively. In step (i) we only have the rows of the three row classes in a symbolic space. We need to show that the distance between row classes $[X]$ and $[Z]$ is not larger than the sum of the distances between row classes $[X]$ and $[Y]$ and between row classes $[Y]$ and $[Z]$.

Step (ii) involves a search for the closest elements between row classes $[X]$ and $[Y]$ (rows X_1 and Y_1) and row classes $[Y]$ and $[Z]$ (rows Y_2 and Z_2) since the distance between row classes is defined as the distance between their closest members. We cannot use rows X_1, Y_1 (or Y_2) and Z_2 to show triangle inequality: the distance between rows X_1 and Z_2 is greater than the sum of the distance between rows X_1 and Y_1 and between rows Y_2 and Z_2.

We proceed in step (iii) to find a row operation g that transforms row Y_1 into row Y_2: we know that such a row operation exists since Y_1 and Y_2 are members of the same row class.

In step (iv) we apply the same operation g to row X_1 and obtain row X_2, and in step (v) the condition of the lemma: the distance between rows $X_2 = g(X_1)$ and $Y_2 = g(Y_1)$ is the same as the distance between the original rows X_1 and Y_1, which is the same as the distance between row classes $[X]$ and $[Y]$.

Finally, in step (vi) we have found three rows, X_2, Y_2, and Z_2, such that $d(X_2, Y_2)$ is the distance between row classes $[X]$ and $[Y]$, $d(Y_2, Z_2)$ is the distance between row classes $[Y]$ and $[Z]$. We can now apply triangle inequality for rows – having started with the assumption that d defines a metric

[6]It is possible to extend the notion of transformational coherence to the musical context in which the equivalence classes are infinite. Note, however, that the distance between two disjoint but infinite equivalence classes might be 0; in such a case the condition of Lemma 6.1 guarantees only a pseudometric for the equivalence classes (see also footnote 4).

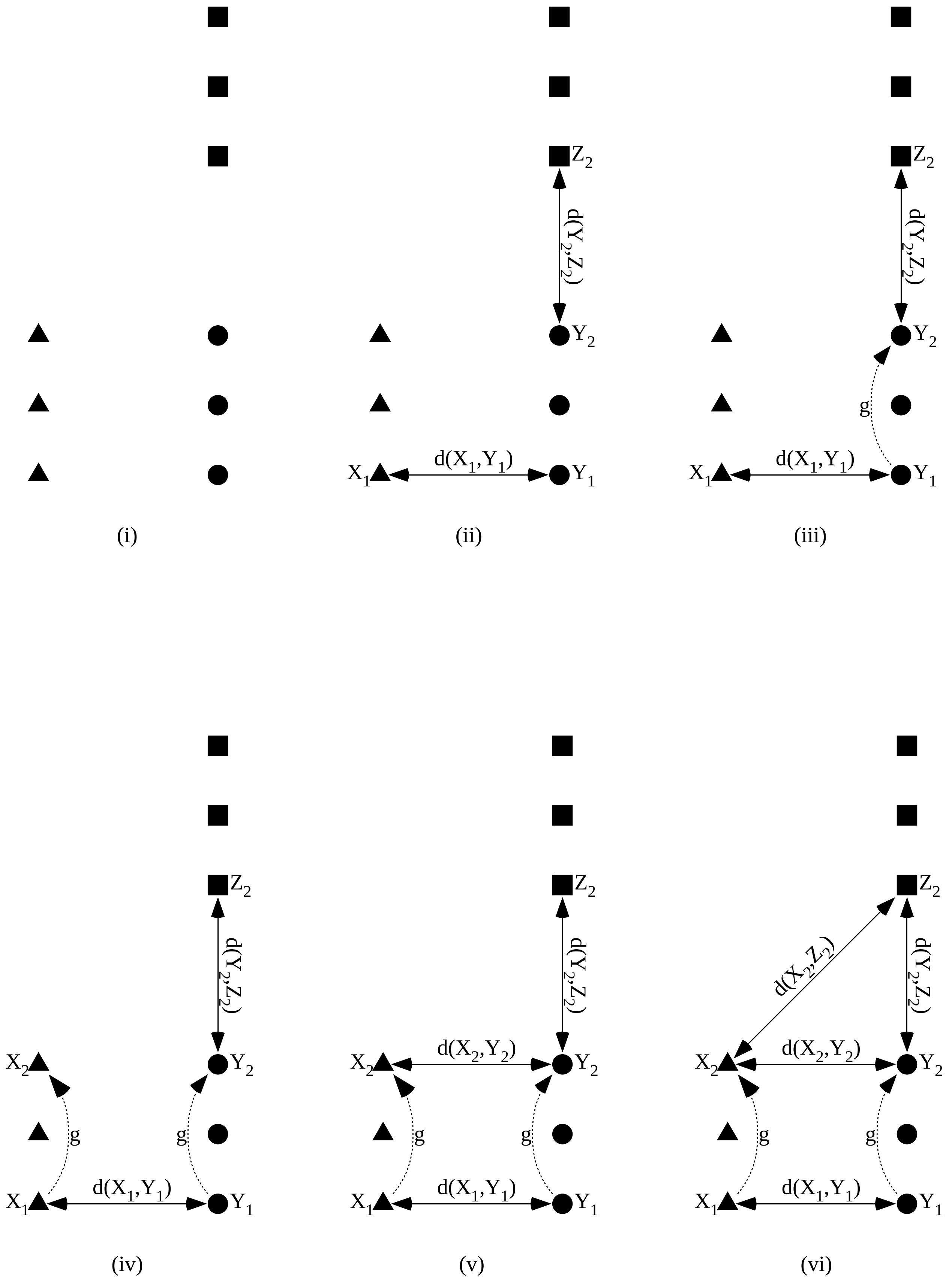


Figure 6.3: The proof of Lemma 6.1 illustrated.

for rows – thus

$$d(X_2, Z_2) \leq d(X_2, Y_2) + d(Y_2, Z_2) = d([X], [Y]) + d([Y], [Z]).$$

Furthermore, we know that the distance between row classes $[X]$ and $[Z]$ is at most the distance between rows X_2 and Z_2 since the distance between two row classes is the distance between the *closest* members and therefore cannot be larger than that between *some* members. Hence, we obtain the inequality

$$d([X], [Z]) \leq d(X_2, Z_2) \leq d(X_2, Y_2) + d(Y_2, Z_2) = d([X], [Y]) + d([Y], [Z]),$$

which shows that triangle inequality holds for row classes.

The above discussion could be summarized in a corollary that states a sufficient condition for a metric for row classes given that the starting point was a metric for twelve-tone rows.

COROLLARY 6.1 If d defines a transformationally coherent metric on a set of twelve-tone rows, then d also defines a metric for row classes.

Proof. This follows from the above discussion. □

Perhaps the most important consequence of the lemma is that it reveals a certain regularity in the distances between the members of two equivalence classes. If the condition $d(x, y) = d(g(x), g(y))$ holds for all $x, y \in S$ and $g \in G$, then every member of the equivalence class has the same set of distances from the members of another equivalence class. Thus, there is no need to examine all possible pairs of elements in two equivalence classes, and it is sufficient to examine the distances between a single member of one equivalence class and all the members of the other. If we translate this regularity into the language of twelve-tone rows and row classes, it means that any row form that belongs to one row class has precisely the same set of relations to the members of another row class as any other row in its row class.

The regularity translates into symmetry between row classes. Figure 6.4 shows two organizations of row classes in a two-dimensional space. The spatial organization (a) is copied from Figure 6.1. It turns out, however, that this type of organization is impossible. The condition of transformational coherency requires that the equation $d(X, Y) = d(g(X), g(Y))$ holds for all rows X and Y and for all transformations g. We could also interpret this requirement in such a way that if there is a row Y at the distance $d(X, Y)$ from row X, then there must also be a row at the distance $d(X, Y)$ from row $g(X)$. Clearly, this is not the case here. The rows are organized in a symmetric fashion in the spatial organization (b) in Figure 6.4. Therefore, the spatial organization of the row classes in Figure 6.3 is also misleading (even if it fulfilled its task in illustrating Lemma 6.1) since it is not symmetrical.

While the space in which the networks of distances between twelve-tones rows reside is surely significantly more complex than the two-dimensional plane, the figure illustrates an important facet

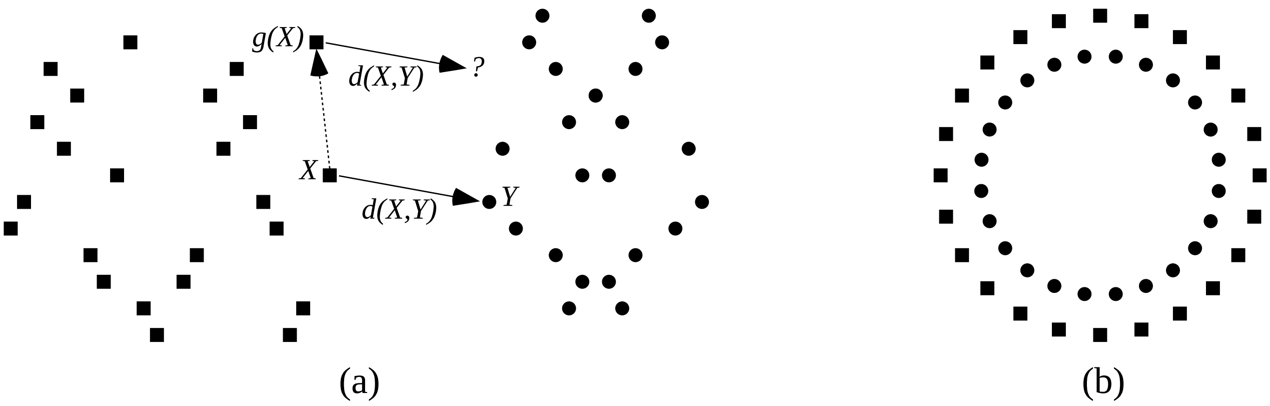


Figure 6.4: A non-symmetric and a symmetric organization of equivalence classes.

of the relations between the twelve-tone rows: the row operations "organize" the rows in a symmetric fashion with regard to any transformationally coherent method of measuring the distances between them.

It was established above in Section 6.1 that the set of twelve-tone rows is symmetrical in the sense that every row has the same set of distances to the other rows. The set of distances between the members of two row classes provides another type of symmetry.

Finally, it should be noted that the notions of transformational coherence and the similarity of row classes do not require that the underlying similarity measure defines a metric. The SCATTERING similarity measure, which is transformationally coherent but does not define a metric, is discussed in Section 11.3.

6.3 Distribution

As will be seen in Part III of this work, row measures feature very different distributions of values: there are substantial differences in the ranges of the measures, and the distributions within the ranges also have diverse properties: for example, some distributions are close to Gaussian distribution while others are exceedingly skewed.

Section 3.6.2 advocated Castrén's approach according to which we should not decide beforehand what the resolution of measures should be (Castrén 1994, 17). A similar argument could be used about the distribution. Since similarity measures give values with some very concrete meaning, there is no reason to expect some particular type of distribution to "validate" the measure. The distribution of a measure simply describes how common its type of similarity is among the rows. For example, the discussion on the DIVISIONS similarity measure in Section 10.6 will show that twelve-tone rows, on average, do not share many contiguous subsegments. Nevertheless, an analysis of the distribution provides us with an understanding of the values. For example, if a value lies in the midway of the scale, on one similarity measure it might mean that the objects are significantly more similar than on average, and on another it might mean that they are significantly less similar. Furthermore, in the case of the BADNESS OF SERIAL FIT similarity measure, a value halfway along the scale means that the rows are next to maximum dissimilarity. However, if the distribution of a similarity measure is extremely skewed, the measure is not particularly useful, because almost everything is either very dissimilar or very similar.

Standard statistical terms and methods, such as average and mean, can be used to describe a distribution, and a graph also gives a quick illustrative overview. In general, I will not modify or scale the values. However, in a few cases (such as taking the squared values of the EUCLIDEAN DISPLACEMENT similarity measure in Section 8.3 and the logarithmic values of the BADNESS OF SERIAL FIT similarity measure in Section 9.3) I will make an exception since the scaled values provide valuable information about unscaled similarity measures.

For some row measures it is possible to compile an exact formula for the distribution, while for others it is necessary to resort to a brute-force approach and use computer programs. The computation of the distribution of a single row involves comparing it with all the 479001600 twelve-tone rows.[7] While a brute-force algorithm gives an exact distribution, a computer program is less elegant than a mathematical formula. Furthermore, a mathematical formula provides information on why the distribution has the properties it has.

In transformational similarity measures for twelve-tone rows, the distribution of the values between a selected row and all the other rows is the same for all twelve-tone rows (see Section 6.1). Consequently, it suffices to compare one row (it does not matter which) to all other rows. Effectively, we are computing the distribution for the 479001600 pitch-class transformations or order-number transformations.

By way of contrast in the case of the INTERVALLIC DISTANCE similarity measure (see Section 12.2), for example, the distribution of the values varies with the rows. In order to obtain the precise distribution we would need to perform $\binom{12!}{2} \approx 10^{17}$ comparisons, which is beyond the capabilities of present computers.[8] It is therefore necessary to resort to sampling: for the non-transformational similarity measures I will compare a number of randomly generated row pairs rows, which gives a sufficiently good estimate of the distribution of the values.

In all the row-class measures examined in this work, the distribution of values between a selected row class and all others varies depending on the one selected. There are two reasons for this situation. First, if the distribution of the values varies between one row and all the others, this variance is also reflected in the distribution of the values between row classes. Secondly, the distribution of the values between row classes may vary because of their different sizes and structures: symmetric row classes contain only 24 distinct rows instead of the usual 48.

The distribution of values between row classes will be illustrated using three curves: maximum, average, and minimum. For each value, the maximum curve indicates the maximum number of row classes possible at that distance, the average curve indicates the average number, and the minimum

[7] With a relatively modern personal computer the distribution of a row measure can, in most cases, be computed reasonably quickly: the time required ranges, depending on the computational complexity of the similarity measure, from a few minutes to a few days.

[8] For example, computing the values of the INTERVALLIC DISTANCE similarity measure between a given row and all 479001600 rows using a relatively straightforward implementation and the Java programming language takes approximately four-and-a-half minutes on a PowerBook G4 running at 1.33GHz. The computation of the complete distribution would take more than 2000 years using this implementation and this equipment.

curve indicates the minimum number. Hence, none of these curves represents a distribution of the values of an existing row class: the curve depicting the distribution of any existing row class is between the minimum and the maximum.

The classic row operations divide the 479001600 twelve-tone rows into 9985920 row classes (Reiner 1985; Read 1997; Fripertinger 1992; Hunter and von Hippel 2003). Consequently, a computation of the complete distribution of the values of a row-class similarity measure would involve $9985920 \cdot 9985920 = 99718598246400$ comparisons of row classes, which is beyond the reach of modern computers. Therefore we need again to resort to sampling. Accordingly, the maximum, average, and minimum curves are obtained by comparing a number of randomly generated row classes to all 9985920 row classes.[9] Hence, the maximum and minimum values do not represent the absolute maximum and minimum values but the maximum and minimum values found in the set of randomly generated row classes. Naturally, the absolute maximum is equal to or larger than the maximum obtained by sampling, and the absolute minimum is equal to or smaller than the minimum obtained by sampling.

The main information these three curves give is the amount of variance. In some similarity measures the minimum and maximum curves are close to each other, in other cases they are not. If the curves are not close to each other, some row classes have more row classes that are close to them than others. If a similarity measure shows that there are row classes that have particularly few row classes close to them, it suggests that these row classes have some special properties that make them different from most of the others. For example, I will show in Section 8.2 that symmetry causes row classes to have fewer row classes close to them.

6.4 Relative similarity

A basic concern with similarity measures is the interpretation of the values. For example, we might compare rows 5409728136AB and $946\text{A}02\text{B}13857$, and the BADNESS OF SERIAL FIT similarity measure gives us the value 24288. Given this information, what could we then say about the similarity of the two rows? If we consider the minimum (1) and maximum (479001600) values of the BADNESS OF SERIAL FIT, the two rows seem to be more similar than dissimilar. However, the value and the scale alone do not provide us with enough information to decide about their similarity: the values are relative and have meaning only when proportioned to the distribution of the BADNESS OF SERIAL FIT values.

One strategy for avoiding the relativity of the values of similarity measures is to use relative values. A row measure indicates the distance between rows X and Y. A relative value for rows rows X and Y indicates how many rows are closer to row X than row Y is (or at least as close). Returning to the similarity of rows 5409728136AB and $946\text{A}02\text{B}13857$, using relative values would tell us that, according to BADNESS OF SERIAL FIT, there are 239477375 rows that are more similar

[9] In this study the number of randomly generated row classes that are compared to all 9985920 row classes varied between 2000 and 2500. This number seemed sufficient since the aim was to gain understanding of the general characteristics of the measures. Furthermore, as for some row measures the comparison of 2000 row classes to all 9985920 row classes took several months, it would not have been practical to take a significantly larger sample.

Figure 6.5: Two approaches to similarity: actual distance on the left and relative distance on the right.

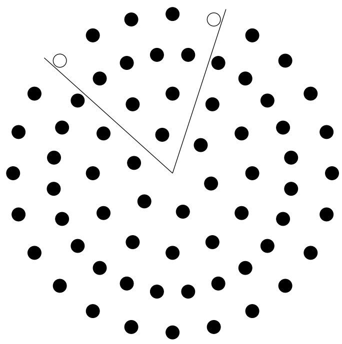


Figure 6.6: Distance as a sector.

to $5409728136AB$ than to $946A02B13857$, 43254 rows that are equally similar and 239480971 row that are less similar. Value 24288 is, in fact, the median of the distribution; hence – according to BADNESS OF SERIAL FIT – the two rows are neither similar nor dissimilar.

Figure 6.5 depicts the idea of distance between objects and the idea of relative distance on a two-dimensional plane. The picture on the left portrays the measurement of the similarity of the two white circles as a traditional distance, while the one on the right portrays the measurement of the similarity by showing the set of objects that fits inside a circle that is large enough to cover the two white circles.

To take a mundane example, if two people live 10 kilometers apart in a sparsely populated area they might be neighbors, but in a large city they certainly would not be. Hence, mere knowledge of the distance does not give the full picture. On the other hand, if we only know that two persons are neighbors we are again lacking some information.

Percentiles[10] provide a relative measurement of similarity, and Isaacson (1996) for example, uses them throughout his study.[11] However, a percentile without knowledge of the distribution of the measure has only limited value since a value belonging to percentile x might indicate close similarity with respect to one measure, but dissimilarity in another. For example, in the DERANGEMENT similarity measure the 9th percentile would give a similarity value of 10 (from a maximum of 12), denoting strong dissimilarity, but in the BADNESS OF SERIAL FIT similarity measure it would give a

[10]Here a value belongs to the nth percentile if in the total set of values there are $\frac{n}{100} \cdot 479001600$ values that are smaller than or equal to it.

[11]See also Section 8.2 in Kuusi (2001).

similarity value of 2800 (from a maximum of 479001600), denoting modest similarity. Furthermore, in DERANGEMENT the single value 10 covers 18 percentiles from the 9th to the 26th, while in the BADNESS OF SERIAL FIT the 9th percentile covers 271 values from 2800 to 3070. Hence, a similarity value must be accompanied by its distribution, otherwise we cannot fully appreciate its meaning.

I will discuss the BADNESS OF SERIAL FIT in Section 9.3, and show that it represents yet another approach to similarity: it is closer to the idea of relative similarity but is not quite the same. We might think of it as a sector as opposed to a circle, as depicted in Figure 6.6. The circle and the sector are, of course, only metaphors. The difference between them is that a circle that is large enough to contain the two rows in which we are interested may also contain rows with properties that neither of the two rows possesses, while the sector contains only rows with properties that at least one of the rows has. I will illustrate the difference between these two approaches in Section 13.2.4 with a discussion of a concrete case involving two similarity measures, ORDER INVERSIONS and BADNESS OF SERIAL FIT: the former represents the traditional distance (and we could apply the relative idea of a "circle" to it) and the latter represents the sector approach.

The overtly symmetric layout of Figure 6.6 is intentional: as discussed in Section 6.1, the rows are located symmetrically in a space in which the distances are defined by a transformational similarity measure.

Part III: Similarity measures

CHAPTER 7

An introduction to twelve-tone-row similarity measures

In Part II I discussed the properties of similarity in general, and some characteristics of similarity measures. The focus in Part III is on the more specific topic of twelve-tone-row similarity measures.

I will first summarize the existing literature in Section 7.1, and then outline my own approach in Section 7.2.

7.1 Previous research

7.1.1 Twelve-tone-row similarity measures

As noted in Section 3.4, the majority of the discussion in the music-theory literature on similarity and similarity measures deals with the similarity of unordered sets of pitches or pitch classes. Writings on the similarity of twelve-tone rows are few and far between. This section gives an overview of the existing literature: I will take a closer look at the proposed similarity measures themselves in subsequent chapters.

While Milton Babbitt did not explicitly discuss the similarity of twelve-tone rows, he was arguably the originator of the research in this area. He developed the idea of comparing twelve-tone rows by the ordered pairs they contain. Building on this notion, John Rothgeb (1967) then presented the first formal definition of a row measure called ORDER INVERSIONS (see Section 9.2).

David Lewin (1976) introduced the next one, BADNESS OF SERIAL FIT (see Section 9.3), again based on Babbitt's notion of ordered pairs. This measure is discussed extensively in Starr (1984). About ten years after Lewin two authors presented new similarity measures: John Roeder (1987) introduced INTERVALLIC DISTANCE, which is based on the pitch-class intervals between the adjacent pitch classes of rows (see Section 12.2), and Robert Morris (1987) introduced several new ones: DISPLACEMENT, CORRELATION COEFFICIENT, and SCATTERING (see Sections 8.3, 9.6, and 11.3). Later, John Ward (1992) introduced two new measures in his dissertation: SUBSEGMENT CONTENT DIFFERENCE and ORDER-INTERVAL INVARIANT N-TUPLES (see Sections 8.3

and 10.5).

Several authors have analyzed the properties of existing similarity measures and have improved their formalizations. Charles Lord (1978) devoted part of his dissertation to the analysis of ORDER INVERSIONS. John Ward (1992) analyzes the properties of all the similarity measures that existed at the time and attempted to present them all in a uniform format: first he reformulates them as similarity measures rather than dissimilarity measures (see Section 3.5), and secondly he scales them to the interval $[0, 1]$ whenever possible. He also presents tables with the distribution of the values of the similarity measures for segments of lengths between 2 and 6, but provides no further analysis of the properties of the distributions.

7.1.2 The permutational approach

All the similarity measures described above except John Roeder's INTERVALLIC DISTANCE are permutational. They do not rely on the fact that twelve-tone rows are permutations of pitch classes: only the order relations are taken into account. In fact, the measures could equally well be used to analyze permutations of twelve apples. Indeed, Robert Morris takes this position explicitly.

> Since the subject has little to do with the properties of the objects permuted—in this case, pcs—it is of general significance in many other aspects of the theory of compositional design with or without reference to pitch organization. (Morris 1987, 117)

Most of the similarity measures discussed in the present study are based on the permutational approach, with the exception of those that are based on the intervals between corresponding pitch classes (such as PITCH-CLASS DISPLACEMENT) or on the pitch-class intervals between adjacent pitch classes (such as Roeder's INTERVALLIC DISTANCE). Even the measures based on the (unordered) subset contents of rows are, in fact, permutational.

The bias toward similarity measures based on the permutational approach provoked enquiry into the idea of pitch-class/order-number isomorphism (O'Connell 1962; Stanfield 1984; Solomon 1973; Mead 1988; Mead 1989). While existing similarity measures clearly suggest the dominance of the ordering relation, I will examine whether at least some of them could be extended to the pitch-class domain as well.

7.1.3 Related similarity measures

I have used the existing research on the similarity of other ordered entities in my research on row measures. The most closely related topics are the similarities of segments and cycles. Practically all row measures can be applied to shorter segments of distinct pitch classes. In most cases, however, the segments must contain precisely the same pitch classes. Indeed, Morris (1987) and Ward (1992) specify their similarity measures for segments of any length as long as they are permutations of the same (distinct) elements. The difficulties that pitch-class duplications cause in extending similarity measures based on the permutational approach have not been fully addressed in the literature.

There are more restrictions on the application of similarity measures to cycles than to segments. Since cycles are cyclic entities, the notion of precedence is not defined, and measures such as OR-

DER INVERSIONS and BADNESS OF SERIAL FIT therefore cannot be applied. Ward (1992) presents and analyzes cyclic variants of the applicable similarity measures.

We can also learn from the research on the similarity of other ordered entities, such as melodies and contours. Orpen and Huron (1991) studied the similarity of melodies. They use a distance known as *edit distance*, which is connected to the DERANGEMENT similarity measure (see Section 8.2). It is significant that their approach works with melodies of different lengths. It also incorporates the pitch domain in addition to the order domain.

Larry Polansky (1996) introduced morphological metrics – a measurement of distances for *morphs* by which he means simply an ordered set.[1] The metrics or measurements of distances he uses are the "usual suspects" – methods that are well known in music theory and mathematics.

Elizabeth West Marvin and Paul Laprade (1987), following Friedman (1985) and Morris (1987), discuss the idea of measuring the similarity of contours. Furthermore, they generalize the contour similarity measures for contour classes, which is a direct predecessor of the idea of extending row measures to cover the similarity of row classes. Morris provides a method for analyzing the salient features of contours (Morris 1993).

7.2 The approach in this study

This study adopts an approach that offers certain benefits over the previous research on the similarity of twelve-tone rows. I have already discussed some of them. First, in Chapter 4 I showed how we could use the concept of a metric to analyze the properties of similarity measures. Secondly, I introduced the notion of row-class similarity measures in Section 6.2, and thirdly, in Section 6.3 I suggested analyzing the distributions of the similarity measures in order to provide a better understanding of the values they yield.[2]

I present as a novel idea the separation of the conception of a twelve-tone row from the measurement of similarity based on it; the details are discussed in Section 7.2.1. I then combine the conceptions of rows and similarity measures with transformational procedures in Section 7.2.2. Finally, in Section 7.2.3 I present the six twelve-tone rows I will be using as examples throughout the discussion, and discuss some technical details concerning the formalization that I will use.

7.2.1 Two stages of defining a similarity measure

One of the innovations introduced in this work is the division of the evaluation of row measures into two stages. The first stage involves explicating the conception of rows on the basis of which the distance is measured. The row measures discussed in the literature suggest different conceptions, but these conceptions and their relations have not been studied. I thus set it as one of my tasks to explicate the conceptions behind row measures.

The notion of conceptions also allows us to group similarity measures into "families." I will

[1] As Polansky explains, "Morphs are ordered shapes, such as melodies, duration series, harmonic orderings, spectra, or statistical measures of formal segregation, like the succession of mean pitches of sections of a piece" (Polansky 1996, 291).

[2] Even if the distributions of set-class similarity measures have been studied, analyses of the distributions of row measures are almost nonexistent.

$$\text{conception} \longleftrightarrow \text{similarity measure}$$

Figure 7.1: The relatedness of the conceptions of twelve-tone rows and the similarity measures.

discuss the similarity measures in Chapters 8 to 12, each of which is based on a different conception of twelve-tone rows: the vector, the set of ordered pairs, the set of subsegments, the set of subsets, and a sequence of pitch-class intervals. The second stage involves the detailed definition of how the distance is to be measured. The subsections in each chapter describe different ways of measuring the distance according to the particular conception under consideration.

A row does not dictate how it is to be conceived – a composer or an analyst is a spectator who interprets it by means of some conception, each of which brings some aspect to the fore. The conception also has significant implications for the transformational relations of rows, as discussed in Section 7.2.2.

The degree of similarity between two twelve-tone rows depends on which conception is chosen. Indeed, I will show that two rows may be maximally dissimilar according to one similarity measure but very similar according to another. Row measures clearly do not speak with a single extensional voice (compare Section 3.6.3). However, I would expect more correlation between those based on the same conception than between those based on different ones. I consider the different conceptions and the differing results of similarity measures a strength in their different analytical and compositional approaches rather than a weakness.

Figure 7.1 relates conceptions of twelve-tone rows and similarity measures in a simple pictorial form. The arrow has two heads, which symbolizes that the conceptions and the measurement of similarity are interdependent: a conception will suggest a similarity measure (the arrow from left to right) and, correspondingly, a similarity measure will suggest a conception (the arrow from the right to left).

7.2.2 Transformational and non-transformational readings

All row measures discussed in this work could be interpreted as measurements of the differences between the properties of the rows, and most similarity measures could be interpreted as measuring the magnitude of the transformation that transforms one row into the other. I will consider both approaches.

The transformational approach is applicable as long as the similarity measure can be interpreted in terms of measuring the magnitude or the complexity of the transformation. Hence, it is a question of measuring the dissimilarity of rows rather than their similarity: dissimilarity could be understood as the distance between rows (see Section 3.5).

Extensionally, the transformational and the non-transformational readings of a similarity measure coincide: both take two twelve-tone rows as input and the value does not depend on the reading. It is in the intension that they differ: the interpretations of what the measures actually measure are different. In the non-transformational reading it is that the values describe the differences between

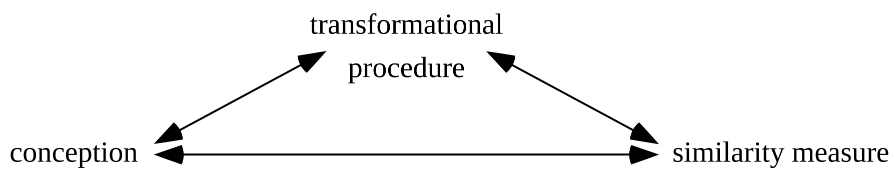


Figure 7.2: The relatedness of the conceptions of twelve-tone rows, similarity measures, and transformational procedures.

the properties of the rows, whereas in the transformational reading they describe the complexity of the transformation that transforms one row into another. Since the two readings coincide extensionally, they are merely two sides of the same coin.

Combining the transformational approach with the conception of a row gives some idea of how the transformation might be carried out. As noted in Section 5.5, sorting algorithms are closely related to the measurement of distances of permutations. Hence, given some conception of twelve-tone rows, the sorting algorithms provide us with strategies for transforming one twelve-tone row into another in practice, step by step.

Figure 7.2 extends Figure 7.1 and relates the conceptions of twelve-tone rows, similarity measures, and transformational procedures in a simple pictorial form. All three arrows have two heads, which symbolizes that the conceptions, the measurement of similarity, and the transformational procedures are interdependent: I will show that a conception will suggest a similarity measure and a transformational procedure (the arrows to the similarity measure and the transformational procedure), a similarity measure will suggest a conception and a transformational procedure (the arrows to the conception and the transformational procedure), and finally, a transformational procedure will suggest a conception and a similarity measure (the arrows to the conception and the similarity measure). The notion of conceiving a twelve-tone row based on how we would transform it into another row has interesting philosophical implications, but such a development lies beyond the scope of this study.

I will examine three transformational procedures in detail, and show that the three corresponding measures directly denote the number of steps in the procedures that are needed to transform one row into another. In these cases the relation between the similarity measure and the transformation is even more concrete than in the mere measurement of the complexity of a transformation.

7.2.3 Rows used as examples

Throughout the discussion I will illustrate the row measures by means of four rows from Alban Berg's *Lyric Suite*, a row from Arnold Schoenberg's *Variations for Orchestra* op. 31, and a row from Anton Webern's *Cantata I* op. 29. These six rows are depicted in Figure 7.3. The labels P, Q, S, T, V, and W are used to refer to them. (R is skipped in order to avoid confusion with the retrograde operation R, V is intended as a mnemonic for "Variations," and W is intended as a mnemonic for "Webern.")

The first three rows P, Q, and S of *Lyric Suite* are closely related and they allow us to illustrate

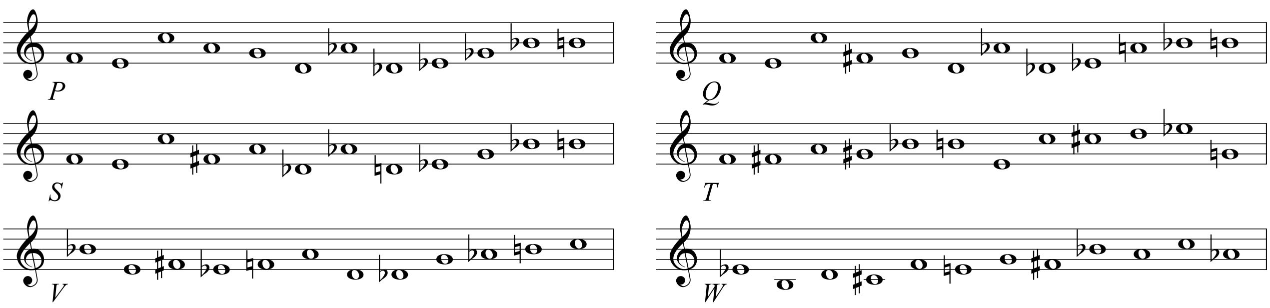


Figure 7.3: Four rows from Alban Berg's *Lyric Suite*, labeled P, Q, S, and T (the letter R is skipped over to avoid confusion with the retrograde operation), a row from Arnold Schoenberg's *Variations for Orchestra* op. 31, labeled V, and a row from Anton Webern's *Cantata I* op. 29, labeled W.

the behavior of similarity measures when the differences between rows are small and relatively easily recognizable. The fourth row T is – both intuitively and according to most similarity measures – less similar to the first three. Row V of Arnold Schoenberg's *Variations for Orchestra* and row W of Anton Webern's *Cantata I* are not related to the four *Lyric Suite* rows, and hence they provide examples of rows that that are less similar to them.

Two of the rows are symmetric: the first row P is invariant under RT_6 and the last row W is invariant under RI_{11}.[3] I will use the row classes of these two symmetric rows to demonstrate the different distributions of the values of symmetric row classes.

Throughout the discussion I will use the two Generalized Interval Systems defined in Definition 5.3 to illustrate the measurement of the similarity of rows. In particular, the pitch-class intervals and order-number intervals between rows $P = 5409728136\mathrm{AB}$ and $Q = 5406728139\mathrm{AB}$ and between rows $P = 5409728136\mathrm{AB}$ and $S = 5406918237\mathrm{AB}$ are refered to frequently. Let us therefore calculate here all the four pertinent intervals.

First, in the GIS of pitch-class rows, the pitch-class transformation that transforms pitch-class row P into pitch-class row Q can be obtained from the identity

$$QP^{-1}(P) = Q(P^{-1}P) = Q.$$

This gives transformation QP^{-1} under the rules covering the composition of permutations[4]

$$\begin{aligned} QP^{-1} &= \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 5 & 4 & 0 & 6 & 7 & 2 & 8 & 1 & 3 & 9 & 10 & 11 \end{pmatrix} \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 2 & 7 & 5 & 8 & 1 & 0 & 9 & 4 & 6 & 3 & 10 & 11 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 0 & 1 & 2 & 3 & 4 & 5 & 9 & 7 & 8 & 6 & 10 & 11 \end{pmatrix} \end{aligned}$$

which is written using the one-line notation as $QP^{-1} = 0123459786\mathrm{AB}$.

[3] I use the expression "symmetric row" to denote a row that is invariant either under RT_6 or under RI_{2k+1}. This enables me to refer to both types of invariance with a single word.

[4] See Appendix A for a walk-through of how to compose permutations.

Secondly, in the GIS of order-number rows, we first need to write rows P and Q as order-number rows as follows: $\boldsymbol{P} = \mathbf{2758109463AB}$ and $\boldsymbol{Q} = \mathbf{2758103469AB}$. The order-number transformation that maps order-number row $\boldsymbol{P}$ into order-number row $\boldsymbol{Q}$ can be obtained from the identity

$$\boldsymbol{Q}\boldsymbol{P}^{-1}(\boldsymbol{P}) = \boldsymbol{Q}(\boldsymbol{P}^{-1}\boldsymbol{P}) = \boldsymbol{Q}.$$

This gives transformation $\boldsymbol{Q}\boldsymbol{P}^{-1}$ under the rules covering the composition of permutations

$$\begin{aligned}\boldsymbol{Q}\boldsymbol{P}^{-1} &= \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 2 & 7 & 5 & 8 & 1 & 0 & 3 & 4 & 6 & 9 & 10 & 11 \end{pmatrix}\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 5 & 4 & 0 & 9 & 7 & 2 & 8 & 1 & 3 & 6 & 10 & 11 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 0 & 1 & 2 & 9 & 4 & 5 & 6 & 7 & 8 & 3 & 10 & 11 \end{pmatrix}\end{aligned}$$

which is written using the one-line notation as $\boldsymbol{Q}\boldsymbol{P}^{-1} = \mathbf{0129456783AB}$.

Let us then move to transformations that transform row P into row S. As above, in the GIS of pitch-class rows we obtain the pitch-class transformation that transforms pitch-class row P into pitch-class row S from the identity

$$SP^{-1}(P) = S(P^{-1}P) = S.$$

This gives transformation SP^{-1} under the rules covering the composition of permutations

$$\begin{aligned}SP^{-1} &= \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 5 & 4 & 0 & 6 & 9 & 1 & 8 & 2 & 3 & 7 & 10 & 11 \end{pmatrix}\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 2 & 7 & 5 & 8 & 1 & 0 & 9 & 4 & 6 & 3 & 10 & 11 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 0 & 2 & 1 & 3 & 4 & 5 & 7 & 9 & 8 & 6 & 10 & 11 \end{pmatrix}\end{aligned}$$

which is written using the one-line notation as $SP^{-1} = 0213457986\mathrm{AB}$.

Secondly, in the GIS of order-number rows, we first need to write rows P and S as order-number rows as follows: $\boldsymbol{P} = \mathbf{2758109463AB}$ and $\boldsymbol{S} = \mathbf{2578103964AB}$. This gives the order-number transformation that transforms order-number row $\boldsymbol{P}$ into order-number row $\boldsymbol{S}$ from the identity

$$\boldsymbol{S}\boldsymbol{P}^{-1}(\boldsymbol{P}) = \boldsymbol{S}(\boldsymbol{P}^{-1}\boldsymbol{P}) = \boldsymbol{S}.$$

As before, this now gives transformation $\boldsymbol{S}\boldsymbol{P}^{-1}$ under the rules covering the composition of permutations

$$\begin{aligned}\boldsymbol{S}\boldsymbol{P}^{-1} &= \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 2 & 5 & 7 & 8 & 1 & 0 & 3 & 9 & 6 & 4 & 10 & 11 \end{pmatrix}\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 5 & 4 & 0 & 9 & 7 & 2 & 8 & 1 & 3 & 6 & 10 & 11 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 0 & 1 & 2 & 4 & 9 & 7 & 6 & 5 & 8 & 3 & 10 & 11 \end{pmatrix}\end{aligned}$$

which is written using the one-line notation as $\boldsymbol{SP}^{-1} = \mathbf{0124976583AB}$.

The selection of these rows is based on convenience: I use them to illustrate the computation of the similarity measures, and in Section 14.3 I will also draw some analytical observations about them. I have also included the two symmetric rows in order to examine the differences between the symmetric and non-symmetric row classes and their distance relations.

I admit that this selection of rows is relatively limited with respect to all the different types of rows, such as multiple-order-function rows (Batstone 1972a; Batstone 1972b; Morris 1977), self-deriving rows (Kowalski 1987), super-saturated rows (Morris 1983–84), for example. Nevertheless, I argue that a more varied supply is not necessary since my approach is mainly permutational. As I stated in Section 7.1.2, I could analyze the permutations of any twelve elements. Consequently, with respect to the permutational similarity measures, these additional properties are immaterial. I will discuss the measurement of similarity based on the unordered INT contents and unordered interval-class contents of rows in Chapter 12, and in that context I will consider the different properties of rows in some more detail.

CHAPTER 8

Similarity measures based on the vector approach

The focus in this chapter is on the measurement of similarity based on the conception of a twelve-tone row as a vector.[1] The conception is described in Section 8.1, and in the subsequent Sections 8.2, 8.3, and 8.4 I discuss in more detail three similarity measures based on it. The concluding Section 8.5 considers some transformational procedures this approach suggests.

8.1 The twelve-tone row as a vector

The enumeration of which pitch classes appear in which order positions is a natural conception of a twelve-tone row. There are twelve "slots," labeled with the twelve order numbers, in which the twelve pitch classes are placed. For example, in row $P = 5409728136\mathrm{AB}$, of the *Lyric Suite*, pitch class 5 is in the first order position, pitch class 4 is in the second order position, etc. In more technical terms, we could describe this conception of a twelve-tone row as a vector in a twelve-dimensional vector space. The twelve dimensions are the twelve order positions, and what defines the twelve-tone row are the pitch classes that the vector holds on each of them.

The sole focus in this vector approach is on the absolute positions of the pitch classes – their mutual relations are immaterial. For example, both the intervals between adjacent pitch classes and the mutual order of pitch classes are ignored: these relations could obviously be derived but they are not the focus here.

Correspondingly – given the isomorphism of pitch classes and order numbers (see Section 2.1.2) – the order-number rows could also be conceptualized as vectors in a vector space. The twelve dimensions represent the twelve pitch classes and the entries in the vector are the order positions of each one.

It is straightforward thus to create similarity measures since the distances in the familiar two-

[1] In order to emphasize the conception of a twelve-tone row as a vector, I will occasionally write rows using the format $(p_0, p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_9, p_{10}, p_{11})$.

dimensional and three-dimensional spaces are customarily measured using this approach: the Euclidean distance is defined on a vector space. It is also easy to imagine that two similar twelve-tone rows are close elements in a twelve-dimensional space. Consequently, similarity measures based on the vector approach are naturally conceived of as dissimilarity measures.

It should be noted that the conception described here also allows for constructs that are *not* twelve-tone rows. For example, the pitch-class vectors could take any sequence of twelve pitch classes: we do not necessarily need to have a sequence of twelve distinct pitch classes. Similarly, the order-number vectors could allow duplicate order numbers, which would imply partial order between the twelve pitch classes: for instance, duplicate order number zero would imply that two pitch classes are both the first – their mutual order is not defined. However, some constructs involving duplicate order numbers may be less intuitive.[2] In this case the length of the vector is 12, but naturally any other length is possible.

The measurement of distance using this conception of twelve-tone rows is connected to the mathematical concept of L_p norms (this was used as an example in Section 3.6.4). Typically, we derive – by some formula – the distance between two vectors from the distances of their entries in each individual dimension. Hence, we need first to define a method for measuring the distance in each of these dimensions. If the entries in the vector are pitch classes, the distance is naturally measured using the interval classes (rather than the pitch-class intervals) interpreted as integers, and if the entries are order numbers, the distance is naturally similarly measured using the order-number intervals interpreted as integers. The total distance between two rows is then calculated based on these individual distances by using some algorithm: as discussed in Section 3.6.4, the role of the algorithm is to find a balance between several small differences versus a few large ones.

In transformational terms, examining the similarity of twelve-tone rows using this conception means examining how we should transpose (with pitch classes as the vector entries) or move (with order numbers as the vector entries) each pitch class in order to transform one row into another. Hence, the row measures discussed in this chapter are based on the premise that the more the pitch classes need to be transposed or moved in order to transform one row into another, the more dissimilar the two rows are.

Let us now consider the connection to the L_p norms. We could first think of twelve-tone rows as elements in a norm space. However, we encounter a conceptual problem, namely, the norm of a vector is its distance from a zero point. What would be the zero point in the space of pitch-class sequences of a given length? It simply does not make sense to fix one; we must resist the temptation to declare $(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$ as the zero point since that would give pitch class 0 an unwanted special status.[3] Technically it would be perfectly possible to define a norm space on the set of twelve-tone rows, but not conceptually.[4]

[2] For example, what would be the interpretation of vector $(\mathbf{6}, \mathbf{6}, \mathbf{6}, \mathbf{6}, \mathbf{6}, \mathbf{6}, \mathbf{6}, \mathbf{6}, \mathbf{6}, \mathbf{6}, \mathbf{6}, \mathbf{6})$?

[3] My unwillingness to select a zero point for the set of pitch-class vectors is analogous to David Lewin's unwillingness to select a zero point for pitch classes (Lewin 1977).

[4] Curiously, if we were to select vector $(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$ – or any vector of form $(n, n, n, n, n, n, n, n, n, n, n, n)$ for that matter – as the zero point, every twelve-tone row would have exactly the same norm.

Secondly, we could apply the concept of a norm space to the transformations, and in this case it turns out to be very natural. Its elements are either the pitch-class transformations or the order-number transformations. The trivial transformation

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \end{pmatrix},$$

which leaves every row intact (the identity element of the group of transformations), is naturally defined as the zero point. The norm of each transformation is its distance from the trivial transformation.

I introduced the concept of L_p norms in the discussion on the role of algorithms in Section 3.6.4. The three similarity measures discussed in this chapter could be considered L_p norms on the set of transformations. The distance between two rows is defined by the norm of the transformation that maps one row into the other.

The following three row measures are examined in detail next: DERANGEMENT, DISPLACEMENT and PITCH-CLASS DISPLACEMENT. DISPLACEMENT was introduced by Robert Morris (1987), and the other two are new ones introduced by the author. All three could be considered L_p norms on the set of transformations. The distance between two rows is defined by the norm of the transformation that maps one row into the other. DISPLACEMENT is the germ of the other two: we obtain DERANGEMENT by observing that DISPLACEMENT is an L_p norm in which the value of p is 1. Consequently, we can create new similarity measures by using other values for p. In addition, we obtain the new similarity measure PITCH-CLASS DISPLACEMENT by applying the idea of DISPLACEMENT to the pitch-class dimension instead of the order-number dimension. DISPLACEMENT and PITCH-CLASS DISPLACEMENT are both L_1 norms with one caveat: in the former the starting point is the order numbers, and in the latter it is the pitch classes. Since the pitch-class space is modular, we need to modify the usual L_1 norm formula to take into account the modularity. In other words, the formula for measuring the distance in each of the twelve dimensions is slightly different in the pitch-class space and the order-number space, but the algorithm by which the final value is obtained is the same in both DISPLACEMENT and PITCH-CLASS DISPLACEMENT.

8.2 DERANGEMENT

The DERANGEMENT row measure, or D, was originally invented by the author mainly to demonstrate certain properties of similarity measures (Ilomäki 2003).[5] It is probably the most straightforward similarity measure for twelve-tone rows introduced so far as it simply indicates the number of order positions in which two rows contain different pitch classes. It is based on the very simple assumption that two rows with a large number of the same pitch classes in the same order positions should appear relatively similar. In transformational terms, DERANGEMENT expresses how many order positions of pitch classes need to be altered in order to transform one row into another.

[5]This measure was originally termed "Different Pitch Classes" (Ilomäki 2003), which is more descriptive than "derangement." Nevertheless, it is relabeled DERANGEMENT here in order to emphasize its connection to a related mathematical concept that I examine in more detail below.

Let us consider rows $P = 5409728136AB$ and $Q = 5406728139AB$ in Figure 7.3. These two rows have the same pitch classes in ten order positions: **0**, **1**, **2**, **4**, **5**, **6**, **7**, **8**, **10**, and **11** (pitch classes 5, 4, 0, 7, 2, 8, 1, 3, 10, and 11). The pitch classes at order positions **3** and **9** are different (pitch classes 9 and 6): in transformational terms, we need to change the positions of two pitch classes in order to transform row P into row Q. Therefore, the DERANGEMENT value for these two rows is 2.

Having prepared the ground with these preliminary considerations I am now ready to present the formal definition of DERANGEMENT.

DEFINITION 8.1 The value of the DERANGEMENT similarity measure for twelve-tone rows X and Y is given by the formula

$$D(X, Y) = \#\{n \mid x_n \neq y_n\}$$

in which x_n is the nth pitch class of row X and y_n is the nth pitch class of row Y.

DERANGEMENT is thus the number of such order positions n in which pitch class x_n is different from pitch class y_n.

It is straightforward to prove that DERANGEMENT defines a metric. First, the values are positive real values. Secondly, it satisfies the four requirements of the metric. (i) The value of $D(X, X)$ is 0 for all rows X, and two identical rows differ at zero order positions. (ii) DERANGEMENT is clearly symmetric and hence $D(X, Y) = D(Y, X)$. (iii) Triangle inequality $D(X, Y) + D(Y, Z) \geq D(X, Z)$ holds because for every index i if $x_i \neq z_i$ then at least one of the inequalities $x_i \neq y_i$ or $y_i \neq z_i$ holds: in transformational terms, if we change n pitch classes in row X in order to transform it into row Y and then change m pitch classes in row Y in order to transform it into row Z, we need to change at most $n + m$ pitch classes in row X in order to transform it into row Z. (iv) Finally, if $D(X, Y) = 0$, then rows X and Y do not differ at any order position and therefore they must be identical. Since all four requirements are satisfied, DERANGEMENT defines a metric.

Obviously, the minimum DERANGEMENT value is zero for identical rows (denoting maximum similarity) and the maximum value is 12 (denoting maximum dissimilarity). Value 1, however, is never attained since it is impossible for two twelve-tone rows to have a different pitch class at only a single order position. Hence, the resolution of DERANGEMENT is rather poor since its range consists only of twelve distinct values: from 0 to 12 excluding 1.

What is worth noting is that the values for rows belonging to the same row class tend to be high. For example, for all twelve-tone rows X, $D(X, T_k X) = 12$ for all non-trivial transpositions T_k, $D(X, I_{2k} X) = 10$ for all even inversions I_{2k}, and $D(X, I_{2k+1} X) = 12$ for all odd inversions I_{2k+1}. The values of $D(X, RT_k X)$ and $D(X, RI_k X)$ range from 0 to 12. If row X is symmetric, then either $D(X, RT_6 X) = 0$ or $D(X, RI_{2k+1} X) = 0$ for some k. Furthermore, if a row is "almost symmetric" we will obtain a low value in either $D(X, RT_k X)$ or $D(X, RI_k X)$ for some k.[6] For

[6]A row may be invariant under operation RT_6 (retrograde composed with transposition T_6) or RI_{2k+1} (retrograde composed with an odd inversion), and may also be almost invariant under operation RT_k (retrograde composed with any transposition T_k) or RI_k (retrograde composed with an odd inversion or an even inversion).

example, row $012345789\text{AB}6$ is almost invariant under operation RI_0 since rows $012345789\text{AB}6$ and $612345789\text{AB}0$ are almost identical and the DERANGEMENT value for these two rows is 2. I will introduce a more formal approach to the issue of "near symmetry" below.

While the idea of evaluating the similarity of twelve-tone rows on the DERANGEMENT measure is a new one, the concept has been in use in mathematics and computer science for quite some time: the comparison of *strings* (sequences of characters) is an active research topic in computer science. Technically, if we consider twelve-tone rows as strings (in which the alphabet consists of the twelve pitch classes), then DERANGEMENT is equivalent to the well-known metric *Hamming distance*: it is the number of positions in two strings at which the two strings have different characters.[7] The idea of applying Hamming distance to musical objects is not new, either: Orpen and Huron used the *Damerau-Levenshtein metric* or *edit distance* – a generalization of Hamming distance – in their comparison of the similarity of melodic phrases (Orpen and Huron 1991).

The transformational approach to DERANGEMENT

Let us now consider DERANGEMENT from the transformational perspective. To that end there is a need to restate it in GIS terms. The transformational approach enables us to utilize some well-known mathematical results in analyses of its properties, and to derive a formula for its distribution.

In mathematics a *derangement* denotes a permutation (in the sense of a mapping) in which none of the elements is mapped into itself (Graham, Knuth, and Patashnik 1994, 194). For example, none of the twelve elements of the permutation

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 11 & 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 & 0 \end{pmatrix}$$

is mapped into itself and therefore it is a derangement. If a permutation has some *fixed points*, ie., some elements are mapped into themselves, it is called a *partial derangement*. For example, in the permutation

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 0 & 1 & 2 & 3 & 4 & 5 & 11 & 10 & 9 & 8 & 7 & 6 \end{pmatrix}$$

the elements from 0 to 5 are mapped into themselves whereas those from 6 to 11 are not: it is therefore a partial derangement. Hence, all permutations of twelve elements except the identity permutation

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \end{pmatrix}$$

are either derangements or partial derangements.

Naturally, the elements that are kept fixed in a permutation and those that are not are two sides of the same coin. The notions of derangement and partial derangement focus on the elements that are not kept fixed. The notion of the *stabilizer* in the theory of permutation groups, however, is defined

[7]Hamming distance is named after its inventor Richard Hamming (1950).

in terms of the elements that are kept fixed.[8] In the group of permutations of the set S the stabilizer of $s \in S$ is the set of permutations that keep s fixed. It is straightforward to prove that the stabilizer of s is a subgroup of all permutations of S. It follows now that the subgroup that keeps elements s_1 and s_2 fixed is the intersection of the stabilizer of s_1 and the stabilizer of s_2, which is also a group. It immediately follows that if S is finite, then the cardinality of any stabilizer divides the cardinality of the whole permutation group. In the present case of permutations of twelve elements, the cardinality of any stabilizer divides the cardinality of $S_{12} = 12! = 479001600$.

DERANGEMENT as a similarity measure slightly differs from the mathematical term derangement.[9] The mathematical variant denotes a permutation in which *none* of the elements is mapped into itself, while the DERANGEMENT similarity measure – as defined below – denotes the number of elements that a transformation does not map into itself.

DEFINITION 8.2 In the GIS of pitch-class rows, the value of the DERANGEMENT similarity measure for twelve-tone rows X and Y is

$$D(X, Y) = \#\{n \mid g_n \neq n\}$$

in which g_n is the nth element of transformation $int(X, Y) = YX^{-1}$.

DERANGEMENT thus defined is the number of pitch classes n that transformation YX^{-1} (the transformation that transforms row X into row Y) does not keep fixed (if transformation YX^{-1} keeps pitch class n fixed then $g_n = n$).

Let us next show that this definition of DERANGEMENT is equivalent to the non-transformational Definition 8.1. We need to show that the number of elements in transformation $int(X, Y)$ that are not mapped into themselves is the same as the number of order positions in which rows X and Y have a different pitch class. Transformation $int(X, Y)$ is a description of how each of the pitch classes needs to be changed in row X in order to transform it into row Y: hence, a pitch class is in the same order position in both rows X and Y if and only if transformation $int(X, Y)$ keeps it fixed. Therefore, the number of pitch classes that are not fixed in transformation $int(X, Y)$ is the number of different pitch classes in rows X and Y.

For example, it was shown in Section 7.2.3 that the pitch-class transformation that transforms row $P = 5409728136\text{AB}$ into row $Q = 5406728139\text{AB}$ is

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 0 & 1 & 2 & 3 & 4 & 5 & 9 & 7 & 8 & 6 & 10 & 11 \end{pmatrix}.$$

As two of the pitch classes in this transformation are not mapped into themselves, the DERANGEMENT value for rows P and Q is 2.

[8] See, for example, Cameron (1999, 4–5).

[9] In the following, I use normal font for the mathematical concept "derangement" and small caps for the similarity measure DERANGEMENT in order to distinguish between these two concepts.

Given the duality of twelve-tone rows, we can also approach them from the order-number perspective. Hence, DERANGEMENT could also be defined in terms of the GIS of order-number rows. This similarity measure would measure in how many order positions two rows have different pitch classes. The values of this measure would be identical to DERANGEMENT, since the flip side of the pitch-class perspective (the pitch classes that two rows have in the same order positions) is the order-number perspective (the order positions containing the same pitch classes): the number of pitch classes that do not appear in the same order positions in two rows equals the number of order positions with different pitch classes in two rows. In fact, it would only be necessary to change the expression "GIS of pitch-class rows" into "GIS of order-number rows," and the expression "transformation $int(X, Y)$" into "transformation $int(\boldsymbol{X}, \boldsymbol{Y})$" in Definition 8.2, to define this new measure.

As discussed in Section 6.1, since DERANGEMENT measures the magnitude of the transformation, every row has precisely the same network of distances to the other rows.

DERANGEMENT for row classes

In order to show that DERANGEMENT for row classes is well defined, we must show that it is transformationally coherent (see Corollary 6.1 in Section 6.2). DERANGEMENT is indeed transformationally coherent since the equation $D(X, Y) = D(FX, FY)$ clearly holds for all row operations: transpositions, inversion, retrograde, and their combinations. For transposition T_k, $x_i \neq y_i$ in rows X and Y if and only if $T_k(x_i) \neq T_k(y_i)$ in the transposed rows T_kX and T_kY; for inversion I_k, $x_i \neq y_i$ in rows X and Y if and only if $I_k(x_i) \neq I_k(y_i)$ in the inversions I_kX and I_kY; finally for retrograde R, $x_i \neq y_i$ in rows X and Y if and only if $x_{11-i} \neq y_{11-i}$ in the retrograded forms. Thus, DERANGEMENT is a transformationally coherent metric and therefore, according to Corollary 6.1, it defines a metric for row classes.

It can be shown that DERANGEMENT is, in fact, transformationally coherent under all pitch-class operations and all order-number operations. First, it is transformationally coherent under all pitch-class operations since, with respect to DERANGEMENT, they constitute only a relabeling of the pitch classes. If two rows have the same pitch class in a corresponding order position, they have the same pitch class in that order position after the relabeling. Secondly, it is transformationally coherent under all order-number operations since, with respect to DERANGEMENT, they constitute only a relabeling of the order positions. If a pitch class is in the same order position in two rows, then it is in the same order position in those rows after the relabeling.

Furthermore, DERANGEMENT is transformationally coherent even under the exchange operation. It is straightforward to prove this claim if we consider a row as a set of ordered pairs in which the first number denotes the order number and the second number denotes the pitch class (see Section 2.1.1). If the DERANGEMENT value for two rows is n, it means that n of the twelve ordered pairs are different and $12 - n$ of them are identical. Now, two rows X and Y share an ordered pair (a, b) if and only if rows EX and EY share the ordered pair (b, a). Hence, rows X and Y share as many ordered pairs as rows EX and EY, and consequently the DERANGEMENT value for rows X and Y is equal to its value for rows EX and EY. Therefore, DERANGEMENT is transformationally coherent under the exchange operation.

In order to calculate the DERANGEMENT value for two row classes we need to find the representatives of the row classes with the smallest value. For example, the DERANGEMENT value is 10 for rows $P = 5409728136\text{AB}$ and $V = \text{A}463592178\text{B}0$ in Figure 7.3, but for the corresponding row classes $[P]$ and $[V]$ it is 9, since it is the lowest value between any members of row classes $[P]$ and $[V]$: for example, its value for rows $P = 5409728136\text{AB}$ and $T_6V = 4\text{A}09\text{B}3871256$ is 9.

DERANGEMENT and the degree of symmetry

It was noted above that row $012345789\text{AB}6$ is almost invariant under operation RI_0, making it almost symmetric. We could take a more formal approach to the notion of "near symmetry" of rows. Namely, the invariance of row X under an operation F means that the DERANGEMENT value is zero for rows X and FX. Therefore, it is natural to define the *degree of symmetry* (or DoS) of a twelve-tone row under an operation as the number of pitch classes that are deranged in it. This is the content of the following definition.[10]

DEFINITION 8.3 The degree of symmetry of twelve-tone row X under operation F is

$$DoS(X, F) = D(X, F(X)).$$

For example, the DERANGEMENT value for row $012345789\text{AB}6$ and its retrograde inversion $RI_0(012345789\text{AB}6) = 612345789\text{AB}0$ is 2; hence, $DoS(012345789\text{AB}6, RI_0) = 2$.

If the value of $DoS(X, F)$ is zero it means that row X is invariant under operation F. The larger the value of $DoS(X, F)$, the less invariant row X is under operation F.

Invariance is usually conceived of as a binary property that a row either has or has not. I extend the concept invariance here to apply to a property that has degrees. The value 0 denotes maximal invariance and the value 12 denotes minimal invariance.

Let us now define the degree of symmetry of a row without tying it to any particular operation.

DEFINITION 8.4 The degree of symmetry of twelve-tone row X is given by the formula

$$DoS(X) = \min\{DoS(X, F)\}$$

in which F is any of the 48 standard row operations *except* the trivial transposition T_0.

Transposition T_0 has to be excluded from the definition, otherwise the degree of symmetry of any row would be 0, thereby making it useless (since $DoS(X, T_0) = 0$ for all rows X).[11] Naturally, we could extend the definition of degree of symmetry to include other row operations, such as rotations or the M-operation.

[10]The notion of the degree of symmetry of rows expressed here is different from that of the degree of symmetry of pitch-class sets; see Definition 3.6.1 in Morris (1987).

[11]We could, in fact, also define the degree of symmetry only in terms of retrograded row forms. For all rows, $DoS(X, T_k) = 12$ if $k \neq 0$, $DoS(X, I_{2k}) = 10$, and $DoS(X, I_{2k+1}) = 12$. It so happens that for every row X there is at least one operation RI_k such that $DoS(X, RI_k) \leq 10$. Namely, if the first pitch class of row X is n and the last pitch class is m, there is an inversion I_t that exchanges pitch classes n and m and, consequently, operation RI_t keeps the first and last pitch classes of row X fixed. Hence, including transpositions and inversions in the definition is redundant.

DoS	rows	row classes
0	322560	13440
2	276480	5760
4	12441600	259200
6	78336000	1632000
7	2211840	46080
8	324587520	6762240
9	26542080	552960
10	34283520	714240

Table 8.1: The distribution of the degree of symmetry of rows and row classes.

Note that every row in a row class has the same degree of symmetry. Consequently, the degree of symmetry of a row class could be defined as the degree of symmetry of any of its constituent rows.

Table 8.1 enumerates the distribution of the degree of symmetry of rows and row classes. In particular, the 322560 symmetric rows, belonging to 13440 distinct row classes, have a zero degree of symmetry.

The distribution of DERANGEMENT

We can use a well-known mathematical formula to derive the number of derangements of permutations and, consequently, we obtain a formula for the distribution of DERANGEMENT (Graham, Knuth, and Patashnik 1994, 194). The number of derangements of n objects is known as the *subfactorial* of n, notated as $!n$, and it is given by the formula

$$!n = n! \cdot \sum_{k=0}^{n} \frac{(-1)^k}{k!}. \tag{8.1}$$

For example, the subfactorial of 4 is

$$\begin{aligned} !4 &= 4! \cdot \sum_{k=0}^{4} \frac{(-1)^k}{k!} \\ &= 4! \cdot \left(\frac{(-1)^0}{0!} + \frac{(-1)^1}{1!} + \frac{(-1)^2}{2!} + \frac{(-1)^3}{3!} + \frac{(-1)^4}{4!} \right) \\ &= 24 \cdot \left(1 - 1 + \frac{1}{2} - \frac{1}{6} + \frac{1}{24} \right) = 9. \end{aligned}$$

Using Formula 8.1 we can easily derive a formula for the distribution of DERANGEMENT values. Now, if n of 12 elements *are not* mapped into themselves then there are $12 - n$ objects that *are*. For each n the formula

$$\binom{12}{12-n} = \frac{12!}{n! \cdot (12-n)!}$$

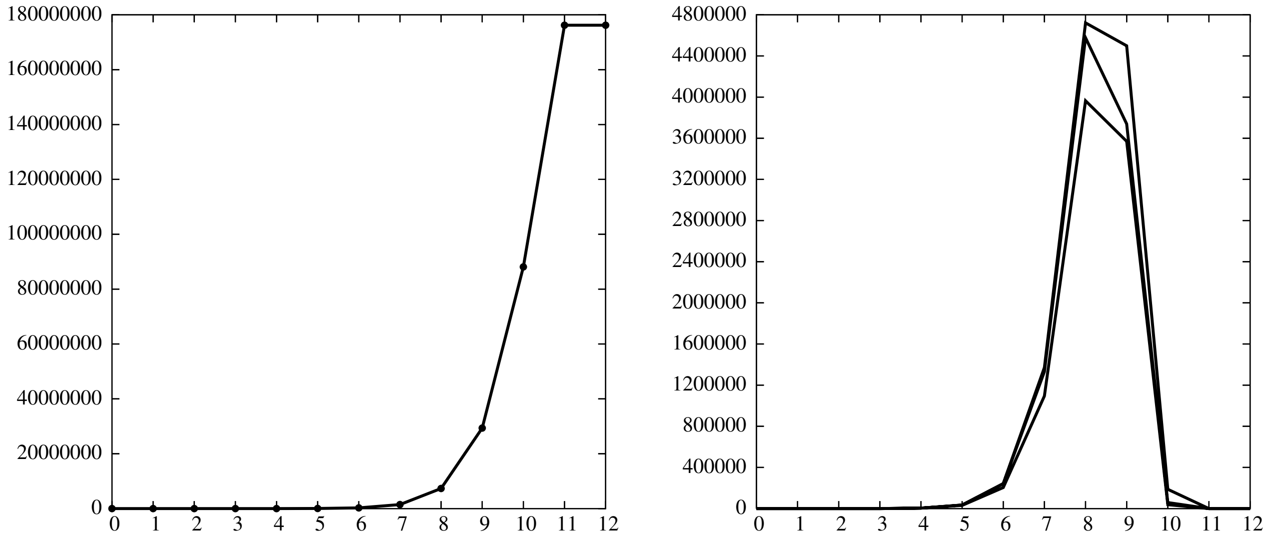


Figure 8.1: The distribution of DERANGEMENT on the left and of DERANGEMENT for row classes on the right. The former is based on a formula and hence it is precise; the latter is obtained by comparing 2500 random row classes to all other row classes. See Section 6.3 for a discussion on the three curves in this figure.

gives the number of ways of selecting $12 - n$ elements from 12 elements. These represent the fixed points – the pitch classes that are in the same order positions in two rows (or the pitch classes or order numbers that the transformation that transforms one row into the other keeps fixed). Combining this formula with Formula 8.1 gives the number of ways in which the remaining elements can be deranged. Thus, the number of twelve-tone rows that differ at n different order positions from a given row is given by the formula

$$\binom{12}{12-n} \cdot !n = \frac{12!}{n! \cdot (12-n)!} \cdot \left(n! \cdot \sum_{k=0}^{n} \frac{(-1)^k}{k!} \right) \cdot = \frac{12!}{(12-n)!} \cdot \sum_{k=0}^{n} \frac{(-1)^k}{k!}.$$

Figure 8.1 summarizes the distribution of DERANGEMENT values and the distribution of the values for row classes. The former is very skewed. It increases exponentially from the low values towards the high with two exceptions: first, there are no transformations with only one derangement since it would be impossible to keep 11 pitch classes at their positions and to change the position of only one, and secondly, the numbers of transformations inducing 11 or 12 derangements – 176214840 and 176214841 – differ only by one. This distribution could be interpreted to mean that there are only few rows that are attainable from a given row by exchanging the positions of only a few pitch classes, whereas most of the rows have only a few (if any) pitch classes in the same positions as in the given row.

Judging from the distribution, we might say that the trustworthiness of DERANGEMENT is primarily with similarity, not with dissimilarity. In other words, two rows with a very small DERANGEMENT value might be expected to be similar in many respects, but rows with a very large value might nevertheless be very similar in some other respects. For example, the value for rows

	5409728136AB	3B215476A908	A463592178B0
0	1	1	1
1	0	0	0
2	36	36	66
3	220	220	440
4	2265	2265	4455
5	17424	17424	34806
6	122278	122293	241705
7	714822	714828	1346821
8	3076230	3076995	4608486
9	5380188	5379486	3693420
10	672456	672372	55720
11	0	0	0
12	0	0	0

Table 8.2: The distribution of DERANGEMENT values for the row classes of rows $P = 5409728136\text{AB}$, $W = 3\text{B}215476\text{A}908$, and $V = \text{A}463592178\text{B}0$ (each row class is compared to the 9985920 row classes).

$P = 5409728136\text{AB}$ and $Q = 5406728139\text{AB}$ is 2, and we might expect these rows to be similar. On the other hand, the value for rows 0123456789AB and $123456789\text{AB}0$ is 12 and the rows are maximally dissimilar with respect to DERANGEMENT. However, it is rather obvious that these two rows are very similar in many respects, and this similarity can be demonstrated using several of the similarity measures that are to be discussed in subsequent sections.

The moral of these observations is that a similarity measure measures only what it measures. DERANGEMENT measures only the number of pitch classes that are in the same order positions in two rows: it does not "claim" to discern the similarity of rows 0123456789AB and $123456789\text{AB}0$, for example, and we cannot blame it for not doing so.

It is trivial to prove that no two row classes can have a DERANGEMENT value of 12 since we can fix a pitch class x in order position **0** and in any row class there are at least two distinct rows with the pitch class x in this order position. Unfortunately, for the time being it is not possible to present a pair of row classes with a DERANGEMENT value of 11 or to provide proof that no two such row classes exist.

The distribution of the DERANGEMENT values between row classes varies. Table 8.2 enumerates those for the row classes of rows $P = 5409728136\text{AB}$, $W = 3\text{B}215476\text{A}908$, and $V = \text{A}463592178\text{B}0$. Rows P and W are symmetric – row P is invariant under operation RT_6 and row W is invariant under operation RI_{11} – and row V is non-symmetric. The symmetric row classes $[P]$ and $[W]$ clearly have fewer row classes at small DERANGEMENT values than the non-symmetric row class $[V]$. Let us consider why this is so.

On the informal level, it comes as no surprise that the row classes of the symmetric rows P and W have fewer very close neighbors than that of the non-symmetric row V, for instance. We could explain this phenomenon by making the simple observation that, since symmetric row classes consist of only 24 distinct twelve-tone rows there are, in a sense, fewer possibilities to transform them into rows of other row classes: I will show below that the symmetry of the row limits the number of rows belonging to different row classes that can be obtained by exchanging only a few pitch classes.

There is a more rigorous explanation. Let us first consider a case in which the DERANGEMENT value between two row classes is 2. Let us pick an arbitrary non-symmetric row A with a degree of symmetry larger than 2. There are 66 rows belonging to as many row classes at distance 2 from row A. If a row is at distance 2 from row A, it means that it can be obtained from row A by changing the positions of two pitch classes. Now, there are $\binom{12}{2} = 66$ ways of selecting the two pitch classes. Every one of these selections results in a row belonging to a different row class from the other selections. (If the degree of symmetry of row A were 2, one selection would result in a row that was in the same row class as row A – hence the need to specify that the degree of symmetry is larger than 2.)

Let us then pick an arbitrary symmetric row A. Since row A is symmetric, it is invariant under some (non-identity) row operation F. Again, there are 66 ways of exchanging the positions of two pitch classes. Now, due to the symmetry, some selections will result in rows belonging to the same row class. Of the 66 ways of selecting the positions of the pitch classes to be exchanged, 6 are symmetric (the pairs $\{\boldsymbol{i}, R(\boldsymbol{i})\}$ for $0 \leq \boldsymbol{i} \leq 5$) and 60 are non-symmetric. Each of the symmetric selections produces rows belonging to different row classes. However, the 60 non-symmetric selections produce rows belonging to only 30 different row classes: exchanging the pitch classes at non-symmetric order positions $\{\boldsymbol{i}, \boldsymbol{j}\}$ gives exactly the same row as first exchanging the pitch classes at non-symmetric order positions $\{R(\boldsymbol{i}), R(\boldsymbol{j})\}$ and then applying row operation F. Hence, there are only $6 + 30 = 36$ row classes at distance 2 from a symmetric row class.

A similar argument could be used to show that there are $440 = 2 \cdot \binom{12}{3}$ row classes at distance 3 from a non-symmetric row class (under the condition discussed below), and only 220 from a symmetric row class. The condition mentioned above is that if a non-symmetric row can be transformed into a symmetric row by exchanging two pitch classes, the formula no longer gives the correct result.[12] For example, row $0123456789\mathrm{BA}$ can be turned into a symmetric row by exchanging either the first two or the last two pitch classes. As a result, the row class of row $0123456789\mathrm{BA}$ has only 438 row classes at distance 3.

Finally, as the mathematical formula for calculating the number of derangements in action shows, for rows that cannot be turned into a symmetric row by changing the positions of at most four pitch classes, there are $4455 = 9 \cdot \binom{12}{4} = !4 \cdot \binom{12}{4}$ row classes at distance 4. In this formula, $\binom{12}{4}$ denotes the number of ways of selecting four pitch classes to be deranged in a row and $!4$ denotes the number of ways in which the selected four pitch classes can be deranged.

[12] In fact, if a non-symmetric row can be transformed into a symmetric row by exchanging two pitch classes, then there are at least two ways of doing the exchange and these two ways are symmetric in the following sense. Assume that we can turn a non-symmetric row into a symmetric one by exchanging two pitch classes p_i and p_j. First, the positions of these two pitch classes are non-symmetric, that is, $\boldsymbol{i} \neq R(\boldsymbol{i})$, since if exchanging two symmetrically positioned pitch classes results in a symmetric row, then the original row must be symmetric, too. Secondly, if exchanging two pitch classes at order positions $\boldsymbol{i}$ and $\boldsymbol{j}$ results in a symmetric row, then so does exchanging two pitch classes at order positions $R(\boldsymbol{i})$ and $R(\boldsymbol{j})$. For example, exchanging pitch classes at order positions $\mathbf{0}$ and $\mathbf{1}$ in row $0123456789\mathrm{BA}$ results in a symmetric row and, consequently, so does exchanging pitch classes at order positions $R(\mathbf{0}) = \mathbf{11}$ and $R(\mathbf{1}) = \mathbf{10}$.

My hypothesis is that all row classes that are invariant under retrograde have the same DERANGEMENT distribution, as do all row classes that are invariant under retrograde inversion. The proof would involve an extension of the above examination of the conditions under which an exchange of pitch classes would yield rows belonging to the same row class. The hypothesis is supported by the fact that the pitch classes in all rows that are invariant under operation RT_6 have structurally identical sets of relations, in other words pitch classes at order positions $\boldsymbol{i}$ and $R(\boldsymbol{i})$ are a tritone apart. Correspondingly, the pitch classes in all rows that are invariant under operation RI_{2k+1} have structurally identical sets of relations, in other words pitch classes at order positions $\boldsymbol{i}$ and $R(\boldsymbol{i})$ are exchanged by the same inversion I_{2k+1}. However, row classes with rows that are symmetric under retrograde do not have the same distribution as those with rows that are symmetric under retrograde inversion, since the row classes of rows $5409728136\mathrm{AB}$ (invariant under operation RT_6) and $3\mathrm{B}215476\mathrm{A}908$ (invariant under operation RI_{11}) in Table 8.2 do not have the same distribution.

In sum, row classes with symmetric rows have the least number of row classes at a small distance. There is also a connection between the DERANGEMENT distribution for row classes and their degree of symmetry.

Variants of DERANGEMENT

There are a variety of ways in which me might modify the DERANGEMENT similarity measure. As discussed above, Orpen and Huron (1991) used the *Damerau-Levenshtein metric* – a generalization for calculating the number of derangements – in their analysis of the similarity of melodic fragments. They scaled the values using the function $e^{-\frac{d}{l}}$ in which d is the original distance and l is the length of the melodic fragment. This scaling, in fact, turns a dissimilarity measure into a similarity measure: identical fragments are given the value 1 and the dissimilarity values grow towards zero. Since I have argued for stating similarity measures in terms of the metric, this scaling does not seem to be a viable avenue of development.

It might be more fruitful to consider a weighted version of DERANGEMENT. The motivation for such a variant lies in the empirical tests in which it has been shown that two different melodies are more easily recognized as different if the deviance occurs either at the beginning or at the end (Pedersen 1975). Hence, we might hypothesize that the beginning of the row (and perhaps also the end of it) would carry more weight in the perceptual evaluation of the similarity of two rows, and therefore we might scale DERANGEMENT accordingly. However, in doing so we would lose the concrete and easily grasped information that gleaned from its non-scaled values: the number of order positions in which two rows have different pitch classes.

Extending DERANGEMENT

DERANGEMENT can easily be extended to segments of any length and segments with duplication and omission of pitch classes. For two segments of the same length, it gives the number of order positions in which the two segments do not have the same pitch class. With this definition, it is immaterial whether or not all pitch classes are present in the segments and whether or not there are pitch-class duplications. For instance, the DERANGEMENT value for segments 012 and 007 would be 2 since they contain different pitch classes in two order positions.

If we wish to compare two segments of different lengths we need to define a method for coping with these differing lengths. One solution is the one used in Orpen and Huron's adaptation of the *Damerau-Levenshtein metric* in which "a missing pitch class" is handled in the same way as "a different pitch class" (Orpen and Huron 1991). For example, we can handle the different lengths of segments 01234 and 0134 by interpreting both as segments of length 5, in which case the latter is interpreted as 01_34. Naturally, there are different ways of interpreting the latter segment as such. Of all the alternatives I have chosen the one that creates the smallest distance between the two segments. Threrefore, when I compare segments 01234 and 0134 I interpret the latter as 01_34 and not, for example, as $\_0134$.

8.3 DISPLACEMENT

The DISPLACEMENT similarity measure was introduced by Robert Morris (1987). He describes it as measuring "how the pcs in a segment have strayed from their original position" (Morris 1987, 119). John Ward also devoted some ten pages to it in his dissertation (Ward 1992). In the following, I will broaden our understanding of this measure by analyzing its distribution and other properties, stating it in transformational terms, discussing the metric, and extending it to row classes. I will also introduce some variants.

The assumption behind this similarity measure is that in two rows that are similar, the pitch classes should be approximately in the same order positions. Alternatively, to take a wider angle, two similar rows may have minor differences in the local order of the pitch classes but the displacements of pitch classes are local, in other words they have not strayed far from their original positions. I adopt Morris' convention of labeling the similarity measure DIS.

The following definition of DISPLACEMENT is equivalent to the one put forward by Morris, but slightly differently formulated. The rationale here is to utilize the dual nature of twelve-tone rows: DISPLACEMENT can be defined in a very straightforward manner using order-number rows. The formal definition is as follows.

DEFINITION 8.5 The value of the DISPLACEMENT similarity measure for twelve-tone rows X and Y is given by the formula

$$DIS(X,Y) = \sum_{n=0}^{11} |oint(\boldsymbol{x_n}, \boldsymbol{y_n})|$$

in which $\boldsymbol{x_n}$ is the nth order number of order-number row $\boldsymbol{X}$, $\boldsymbol{y_n}$ is the nth order number of order-number row $\boldsymbol{Y}$, $oint$ is the order-number interval between $\boldsymbol{x_n}$, and $\boldsymbol{y_n}$ and $|oint(\boldsymbol{x_n}, \boldsymbol{y_n})|$ is the order-number interval interpreted as a nonnegative integer.

Admittedly, Definition 8.5 is designed to be accurate rather than illustrative. In more colloquial terms, DISPLACEMENT is defined as the sum of the differences in the corresponding order numbers of order-number rows. Since an order-number row enumerates the positions of each pitch class in a row, the difference in the corresponding order numbers describes how far the pitch classes have strayed. Therefore, with the introduction of a minor conceptual inaccuracy, the formula for DISPLACEMENT could be rewritten as

$$DIS(X,Y) = \sum_{n=0}^{11} |\boldsymbol{x_n} - \boldsymbol{y_n}|, \tag{8.2}$$

which sums the differences between the corresponding order numbers of two rows. The inaccuracy here is that the order numbers do not define a group structure (or any other algebraic structure), and therefore the expression $\boldsymbol{x_n} - \boldsymbol{y_n}$ is not well defined.[13] This problem was resolved in the original Definition 8.5 by first referring to the order-number interval and then interpreting this interval as a nonnegative integer – the sum of integers is naturally well defined. However, for the sake of convenience I will use the slightly inaccurate format of Formula 8.2.

In rows $P = 5409728136\text{AB}$ and $Q = 5406728139\text{AB}$ in Figure 7.3, for example, ten pitch classes are in the same order positions but two pitch classes in row Q have strayed from the order positions they had in row P: pitch class 9 is at order position $\mathbf{3}$ in row P and has strayed six positions to the “right” in row Q, in which it is at order position $\mathbf{9}$. Similarly, pitch class 6 is at order position $\mathbf{9}$ in row P and has strayed six positions to the “left” in row Q, in which it is at order position $\mathbf{3}$. If we write the rows as order-number rows $\mathbf{P} = \mathbf{2758109463AB}$ and $\mathbf{Q} = \mathbf{2758103469AB}$ and apply Formula 8.2 for DIS we obtain the DISPLACEMENT value

$$\begin{aligned} DIS(P,Q) =& |2-2| + |7-7| + |5-5| + |8-8| + |1-1| + |0-0| + \\ &+ |9-3| + |4-4| + |6-6| + |3-9| + |10-10| + |11-11| = 12. \end{aligned}$$

It is straightforward to show that DISPLACEMENT defines a metric. First, the values are positive real values. The four requirements of the metric are also satisfied. (i) The value of $DIS(X,X)$ is 0 for all rows X, and no pitch class has strayed anywhere between two identical rows. (ii) DISPLACEMENT is clearly symmetric and $DIS(X,Y) = DIS(Y,X)$. (iii) Triangle inequality $DIS(X,Y)+DIS(Y,Z) \geq DIS(X,Z)$ holds because for all order numbers $|\mathbf{x_n}-\mathbf{y_n}|+|\mathbf{y_n}-\mathbf{z_n}| \geq |\mathbf{x_n}-\mathbf{z_n}|$ (this is simply triangle inequality for the one-dimensional space), and therefore it also holds for the sum

$$\sum_{n=0}^{11} |\mathbf{x_n} - \mathbf{y_n}| + \sum_{n=0}^{11} |\mathbf{y_n} - \mathbf{z_n}| \geq \sum_{n=0}^{11} |\mathbf{x_n} - \mathbf{z_n}|$$

of the differences of the order numbers. (iv) Finally, if $DIS(X,Y) = 0$ then no pitch class has strayed anywhere, and consequently rows X and Y must be identical. Since all four requirements of the metric are satisfied, DISPLACEMENT defines a metric.

The minimum DISPLACEMENT value is 0, and the maximum value is $\frac{1}{2} \cdot 12^2 = 72$. In total there are 518400 ways in which the maximum can be achieved. Two different examples follow: for any

[13]The expression $\boldsymbol{x_n} - \boldsymbol{y_n}$ implies both the existence of a binary operation between the order numbers and the existence of inverse elements of them.

row X and its retrograde RX the DISPLACEMENT value is

$$DIS(X, RX) = 11 + 9 + 7 + 5 + 3 + 1 + 1 + 3 + 5 + 7 + 9 + 11 = 72,$$

and for any row X and its rotation $r_6 X$ it is

$$DIS(X, r_6 X) = 6 + 6 + 6 + 6 + 6 + 6 + 6 + 6 + 6 + 6 + 6 + 6 = 72.$$

In these two examples the patterns of how the pitch classes have strayed are quite opposite: in row X and its retrograde RX the amount of straying ranges from 11 to 1, and in row X and its rotation $r_6 X$ every pitch class has strayed the same number of order positions.

Since for any row X and RX, $DIS(X, RX) = 72$, two rows belonging to the same row class are maximally dissimilar according to the DISPLACEMENT measure. Hence, as discussed in Section 3.3, equivalence does not guarantee similarity.

A peculiar feature of DISPLACEMENT is that all values are even numbers, a claim I will prove below when I consider DISPLACEMENT in transformational terms.[14] Therefore, while the range of DISPLACEMENT is from 0 to 72, it contains only 37 distinct values.

DISPLACEMENT could be considered a fine-tuned variant of DERANGEMENT. While with the latter we only count the number of pitch classes that have strayed away from their original positions, DISPLACEMENT takes into account how far they have strayed, thus obtaining a better resolution – there are 37 distinct values compared to the 12 in DERANGEMENT. For example, in rows $0123456789AB$ and $1032547698BA$ every pitch class in one row has been displaced by one order position with respect to the other row and, consequently, the DISPLACEMENT value for these two rows is $12 \cdot 1 = 12$. On the other hand, in rows $0123456789AB$ and $6789AB012345$ every pitch class in one row has been displaced by six order positions with respect to the other row and, consequently, the DISPLACEMENT value for these two rows is $12 \cdot 6 = 72$. The DERANGEMENT similarity measure is only able to indicate that in both cases all twelve pitch classes in one row have been displaced with respect to the other row.

The transformational approach to DISPLACEMENT

Let us now consider DISPLACEMENT from the transformational perspective. It has a natural interpretation as the measurement of the magnitude of the transformation that transforms one row into another. Indeed, even if Morris does not discuss it in transformational terms, the approach is implicit in his original definition, in which the distance is calculated as the magnitude of the transformation by measuring how far the order numbers have strayed from their natural positions in the transformation that transforms one row into another (Morris 1987, 119).

Let us consider again rows P and Q in Figure 7.3. It was calculated in Section 7.2.3 that the order-number transformation $\boldsymbol{QP}^{-1}$ that transforms order-number row $\boldsymbol{P}$ into order-number row $\boldsymbol{Q}$

[14] This observation makes it easy to spot an error in John Ward's table of maximum DISPLACEMENT values in segments of cardinalities 0 through 12 (Ward 1992, 104–105). He gives 3 – the only odd maximum value in the table – as the maximum value in segments of cardinality 3; the correct value is 4 – an even value.

is

$$\boldsymbol{QP}^{-1} = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 0 & 1 & 2 & 9 & 4 & 5 & 6 & 7 & 8 & 3 & 10 & 11 \end{pmatrix}.$$

The order numbers of the lower row denote how those of the upper row are transformed. If the values are interpreted as integers, then the sum of the absolute values of the differences between the corresponding entries of the upper and lower rows is

$$0+0+0+6+0+0+0+0+0+6+0+0 = 12,$$

which is the magnitude of this transformation. Hence, DISPLACEMENT measures how much the order-number transformation that transforms one row into the other differs from the identity order-number transformation $\mathbf{0123456789AB}$.

In the following DISPLACEMENT is defined in GIS terms. Since it measures how the order positions are changed, the GIS is that of order-number rows (see Definition 5.3).

DEFINITION 8.6 In the GIS of order-number rows, the value of the DISPLACEMENT similarity measure for rows X and Y is

$$DIS(X,Y) = \sum_{n=0}^{11} |g_n - n|,$$

in which g_n is the nth element of transformation $int(\boldsymbol{X},\boldsymbol{Y}) = \boldsymbol{YX}^{-1}$ interpreted as an integer.

I will now show that this transformational definition of DISPLACEMENT is equivalent to the non-transformational definition (see Definition 8.5). The equivalence is due to the fact that both definitions end with the sum of precisely the same differences, even if not necessarily in the same order. For example, if pitch class n is at order position $\boldsymbol{x_n}$ in row X and at order position $\boldsymbol{y_n}$ in row Y, then in transformation $int(\boldsymbol{X},\boldsymbol{Y})$ order number $\boldsymbol{x_n}$ is transformed into order number $\boldsymbol{y_n}$, and hence pitch class n contributes $|\boldsymbol{x_n} - \boldsymbol{y_n}|$ to the total sum that is the DISPLACEMENT value. The non-transformational version defines DISPLACEMENT as the sum of all differences $|\boldsymbol{x_n} - \boldsymbol{y_n}|$. Thus, the two definitions are equivalent.

As discussed in Section 6.1, since DISPLACEMENT measures the magnitude of the transformation, the result is that every row has precisely the same network of relations to the other rows.

The transformational approach provides us with a simple way of showing that the DISPLACEMENT values must be even. This is the content of the following lemma.

LEMMA 8.1 The value of $DIS(X,Y)$ is even for all rows X and Y.

Proof. Let us consider two arbitrary twelve-tone rows X and Y and the order-number transformation $int(\boldsymbol{X},\boldsymbol{Y}) = \boldsymbol{YX}^{-1}$ in the GIS of order-number rows. Since $int(\boldsymbol{X},\boldsymbol{Y})$ is a permutation it can be writen as a product of disjoint cycles (see Section A.3 in Appendix A)

$$int(\boldsymbol{X},\boldsymbol{Y}) = (p_1 \;\; p_2 \;\; \dots \;\; p_k)(q_1 \;\; q_2 \;\; \dots \;\; q_m)\dots(r_1 \;\; r_2 \;\; \dots \;\; r_n).$$

The order-number interval between the two adjacent entries p_i and p_{i+1} in a cycle is the order-number interval between p_i and $\boldsymbol{YX}^{-1}(p_i)$. Hence, the DISPLACEMENT value is the sum of the order-number intervals between the adjacent entries in the cycles (taking into account the fact that the cycles "wrap around"). We only need to show that the sum of the order-number intervals between the adjacent entries in one cycle is even: it then follows directly that adding the even values of each cycle of a permutation results in an even value.

Let us therefore examine the cycle $(p_1 \ \ p_2 \ \ \ldots \ \ p_k)$. If the length of the cycle is one, it denotes a fixed point that contributes zero to the total sum since the order-number interval between an order number and itself is, naturally, zero. We could therefore assume that the cycle is of a length that is greater than one. Now, the sum of the order-number intervals in a cycle could be odd only if there is an odd number of odd order-number intervals between the adjacent order numbers of the cycle. However, the assumption of an odd number of odd order-number intervals in a cycle leads to a contradiction since the sum of the order-number intervals must be a multiple of twelve: if we begin from the first order position in a cycle and then move by $oint(p_1, p_2)$ order positions, $oint(p_2, p_3)$ order positions, continue similarly and finally move by $oint(p_k, p_1)$ order positions, we return to our point of origin. Hence, a cycle can contain only an even number of odd order-number intervals (zero is an even number). Thus, the sum of the order-number intervals in each cycle is even, and therefore the DISPLACEMENT value must also be even. □

DISPLACEMENT for row classes

In order to show that DISPLACEMENT for row classes is well defined, we must show that it is transformationally coherent (see Corollary 6.1 in Section 6.2). DISPLACEMENT is a transformationally coherent similarity measure since $DIS(X, Y) = DIS(FX, FY)$ holds for all row operations: transpositions, inversion, retrograde, and their combinations. For transposition T_n, pitch class p has strayed k positions in row Y from its position in row X if and only if pitch class $T_n(p)$ has strayed k positions in row T_nY from its position in row T_nX. Similarly for inversion I_n, pitch class p has strayed k positions in row Y from its position in row X if and only if pitch class $I_n(p)$ has strayed k positions in row I_nY from its position in row I_nX. For the retrograde, pitch class p has strayed k positions to the left (correspondingly right) in row Y from its position in row X if and only if pitch class p has strayed k positions to the right (correspondingly left) in row RY from its position in row RX. Thus, DISPLACEMENT is a transformationally coherent metric and therefore, according to Corollary 6.1, it defines a metric for the row classes.

If we consider the extended set of row operations, DISPLACEMENT is transformationally coherent under some of them but not all. In fact, it is transformationally coherent under all pitch-class operations, such as the M-operation, since, with respect to DISPLACEMENT, they constitute only a relabeling of the pitch classes. In general, it is not transformationally coherent under order-number operations – retrograde is the only exception. For example, it is not transformationally coherent under the (non-trivial) rotations. Let us consider row $P = 5409728136\mathrm{AB}$ in Figure 7.3, its modification $P' = 4509728136\mathrm{AB}$ in which the positions of the two first pitch classes of row P have been exchanged, and the rotations of both rows $r_{11}P = 409728136\mathrm{AB}5$ and $r_{11}P' = 509728136\mathrm{AB}4$. A comparison of the DISPLACEMENT values of the original rows and the two rotated rows gives

$$DIS(P, P') = 2 \neq 22 = DIS(r_{11}P, r_{11}P'),$$

and therefore DISPLACEMENT is not transformationally coherent under rotation. The transformational coherence is broken since the rotations "wrap around." Pitch classes 4 and 5 are adjacent in rows P and P' (even if they are in a different order), but non-adjacent in rows $r_{11}P$ and $r_{11}P'$. If rows were treated as circular entities, DISPLACEMENT would be transformationally coherent under rotations. (I will return to this issue at the end of Section 8.4.)

It can be shown that DISPLACEMENT is not transformationally coherent under any other order-number operation than retrograde (and the trivial order-number operation r_0) by considering the distances between order numbers. Retrograde is the only nontrivial order-number operation that retains the order-number intervals, and DISPLACEMENT is the sum of the differences between the corresponding order numbers of rows. Therefore, DISPLACEMENT is not transformationally coherent under an order-number operation that changes the order-number interval between any two order numbers. If, for example, order-number operation $\boldsymbol{F}$ changes the order-number interval between order numbers $\boldsymbol{p_0}$ and $\boldsymbol{p_1}$, then we can construct the following two order-number rows:

$$\boldsymbol{X} = \boldsymbol{p_0 p_1 p_2 p_3 p_4 p_5 p_6 p_7 p_8 p_9 p_{10} p_{11}} \quad \text{and} \quad \boldsymbol{Y} = \boldsymbol{p_1 p_0 p_2 p_3 p_4 p_5 p_6 p_7 p_8 p_9 p_{10} p_{11}}.$$

The DISPLACEMENT value for rows $\boldsymbol{X}$ and $\boldsymbol{Y}$ is not the same as its value for rows $\boldsymbol{FX}$ and $\boldsymbol{FY}$. Hence, DISPLACEMENT is not transformationally coherent under $\boldsymbol{F}$.

DISPLACEMENT is not transformationally coherent under the exchange operation either. Let rows $P = 5409728136\text{AB}$ and $r_{11}P = 409728136\text{AB}5$ be as above. The exchange transformations of these two rows are $EP = 2758109463\text{AB}$ and $Er_{11}P = 16470\text{B}83529\text{A}$. Even a cursory glance at rows EP and $Er_{11}P$ reveals that they have very little in common. A comparison of the DISPLACEMENT values for rows P and $r_{11}P$ and for their transformations using the exchange operation gives

$$DIS(P, r_{11}P) = 22 \neq 50 = DIS(EP, Er_{11}P),$$

which confirms the observed dissimilarity.

The distribution of DISPLACEMENT

DISPLACEMENT is equivalent to a well-known metric on permutations that is known in statistics as *Spearman's footrule* (Spearman 1906). Hence, the existing research on Spearman's footrule is a useful basis for the following analysis of the properties of DISPLACEMENT.

Unfortunately, there is no formula that would give the distribution of DISPLACEMENT, which was therefore obtained using a computer program. Nevertheless, Diaconis and Graham provide some properties of Spearman's footrule that are applicable here, for example the mean of the distribution is $\frac{1}{3} \cdot 12^2 = 48$ and the maximum value is $\frac{1}{2} \cdot 12^2 = 72$ (Diaconis and Graham 1977).

Figure 8.2 shows the distribution of DISPLACEMENT values. The shape of the distribution

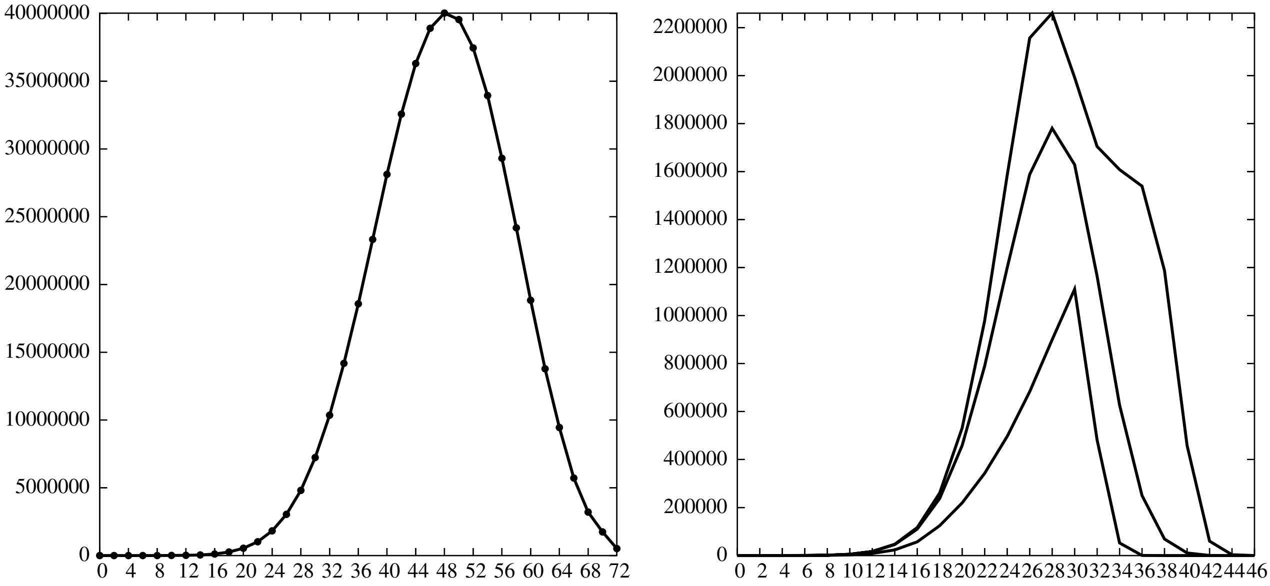


Figure 8.2: The distribution of DISPLACEMENT on the left and the distribution of DISPLACEMENT for row classes on the right. The former is obtained by computing the distances defined by all 479001600 distinct transformations and hence it is precise; the latter is obtained by comparing 2000 random row classes to all other row classes. See Section 6.3 for a discussion of the three curves in the latter.

resembles a bell curve, even if the large values dominate. Hence, given a row, there are very few that are very similar to it, relatively few that are very dissimilar to it, and the majority are neither particularly similar nor dissimilar to it. However, the distribution is not symmetric: the mean 48 is above the middle of the scale 36. Therefore, according to DISPLACEMENT the rows are, on the average, more dissimilar than similar.

As with the DERANGEMENT similarity measure, the distribution of the DISPLACEMENT values between row classes varies. The row classes of the symmetric rows (such as rows P and W in Figure 7.3) have fewer very close neighbors than those of the non-symmetric rows (such as row V in Figure 7.3). With an argument similar to that used in Section 8.2 it is possible to show that the symmetry of the row limits the number of rows belonging to different row classes that can be obtained by only small displacements of a few pitch classes.

There are relatively large differences between row classes with respect to the numbers of close and distant ones. In two of those in the sample of 2000, the distances from that of row $37B42A608519$ range from 0 to 46, and from that of row $072BA4186539$ from 0 only to 36. The most distinctive property of row $37B42A608519$ is its unordered INT contents (seven of the interval classes between its adjacent pitch classes are interval class 4). My hypothesis is thus that the unordered INT contents of the rows in that row class is related to the multiplicity of row classes with a high distance from it: since rows with a biased unordered INT contents are more rare than rows with an even unordered INT contents (I will discuss the unordered INT contents and unordered interval-class contents of rows in more detail in Chapter 12), those that can be obtained from row $37B42A608519$ by only small displacements of a few pitch classes are bound to have a biased unordered INT contents as well and, consequently, the number of close row classes cannot be as high as in row classes with an even unordered INT contents.

Variants of DISPLACEMENT

As discussed in Section 4.5, the "raw" values of a similarity measure can be scaled using various techniques. DISPLACEMENT could be considered a taxicab metric (or the L_1 norm) on the set of order-number transformations. This observation suggests that there is a family of distance functions. For example, the measurement of the distance between permutations using the L_2 norm or *Euclidean distance* is the basis of a widely used measure of correlation known as *Spearman's Rank Correlation*. In the following I apply the L_2 norm in order to define the EUCLIDEAN DISPLACEMENT similarity measure, or $EDIS$.

DEFINITION 8.7 The value of the EUCLIDEAN DISPLACEMENT similarity measure for twelve-tone rows X and Y is given by the formula

$$EDIS(X,Y) = \sqrt{\sum_{0}^{11} (\boldsymbol{x_n} - \boldsymbol{y_n})^2}$$

in which $\boldsymbol{x_n}$ is the nth order number of order-number row $\boldsymbol{X}$ interpreted as an integer and $\boldsymbol{y_n}$ is the nth order number of order-number row $\boldsymbol{Y}$ similarly interpreted.

For example, let us calculate the EUCLIDEAN DISPLACEMENT value for rows P and Q in Figure 7.3. Let us first write rows P and Q as order-number rows $\boldsymbol{P} = \mathbf{2758109463AB}$ and $\boldsymbol{Q} = \mathbf{2758103469AB}$, and then apply the formula for EUCLIDEAN DISPLACEMENT: we thus obtain the value

$$\begin{aligned} EDIS(P,Q) &= ((2-2)^2 + (7-7)^2 + (5-5)^2 + (8-8)^2 + (1-1)^2 + (0-0)^2 + \\ &+ (9-3)^2 + (4-4)^2 + (6-6)^2 + (3-9)^2 + (10-10)^2 + (11-11)^2)^{\frac{1}{2}} \\ &= \sqrt{72} \approx 8.49. \end{aligned}$$

EUCLIDEAN DISPLACEMENT features a total of 287 distinct values ranging from the minimum value 0 to the maximum value $\sqrt{572} \approx 23.9$. It therefore has a significantly better resolution than DISPLACEMENT, which features only 37 distinct values.

Let us illustrate the balancing of several small differences with a few large differences (see Section 3.6.4) by comparing DISPLACEMENT with EUCLIDEAN DISPLACEMENT. The former places more emphasis on the total number of changes, and the latter on the sizes of the changes.

Let us consider row $P = 5409728136\text{AB}$ and its two variants: six pairs of adjacent pitch classes (5 and 4, 0 and 9, etc.) have been exchanged in row $P' = 4590271863\text{BA}$, whereas only (non-adjacent) pitch classes 5 and 8 have been exchanged in row $P'' = 8409725136\text{AB}$. Hence, in comparison with row P, several small changes were introduced in row P', and only one large one in row P''. Now let us calculate the following values for DISPLACEMENT and EUCLIDEAN

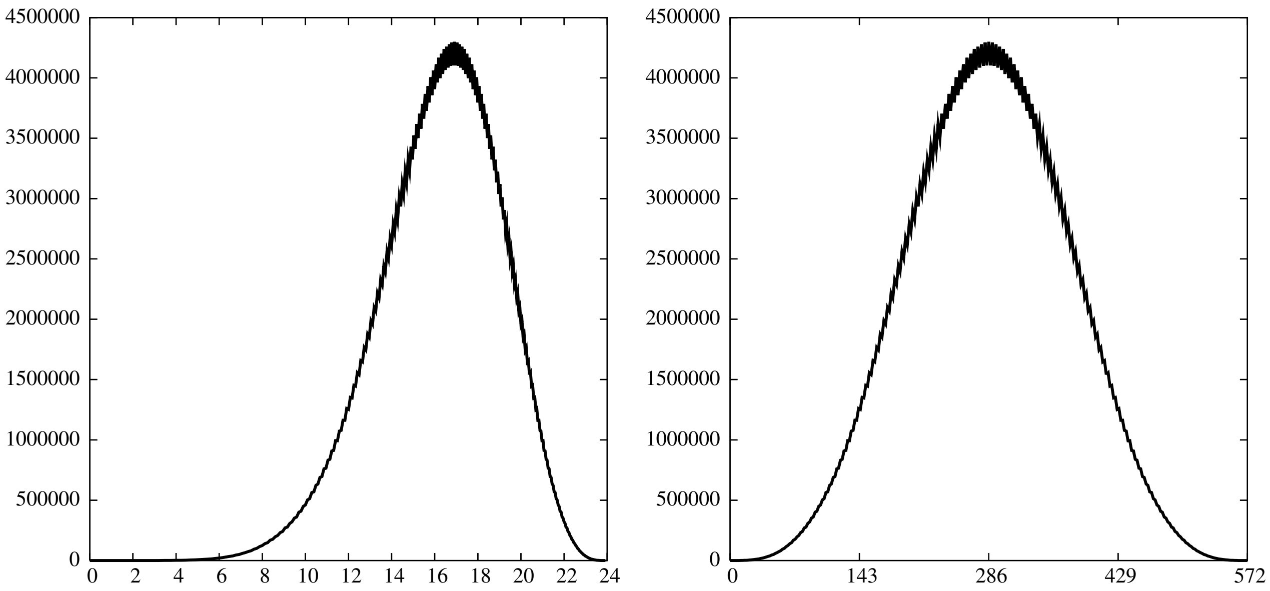


Figure 8.3: The distribution of EUCLIDEAN DISPLACEMENT on the left and the distribution of SQUARED EUCLIDEAN DISPLACEMENT on the right. Both distributions are obtained by a computing the distances defined by all 479001600 distinct transformations, and hence they are precise.

DISPLACEMENT:

$$DIS(P, P') = 12 \qquad EDIS(P, P') = \sqrt{12} \approx 3.46$$
$$DIS(P, P'') = 12 \qquad EDIS(P, P'') = \sqrt{72} \approx 8.49$$

DISPLACEMENT returns the same value for both row pairs, but EUCLIDEAN DISPLACEMENT returns a significantly higher value for the pair with only one large change. This demonstrates how EUCLIDEAN DISPLACEMENT places more weight on the sizes of the individual displacements than DISPLACEMENT: a single large displacement of a pitch class can count more than six small displacements.

EUCLIDEAN DISPLACEMENT features a less obvious type of symmetry. For the sake of illustratiion, let us define a variant of EUCLIDEAN DISPLACEMENT termed SQUARED EUCLIDEAN DISPLACEMENT: the values of SQUARED EUCLIDEAN DISPLACEMENT are squares of the values of EUCLIDEAN DISPLACEMENT. Figure 8.3 shows the distributions of both. The distribution of the values of both illustrate the effect of scaling: that of SQUARED EUCLIDEAN DISPLACEMENT is perfectly symmetrical. However, SQUARED EUCLIDEAN DISPLACEMENT does not define a metric, as EUCLIDEAN DISPLACEMENT does.

Finally, I will mention yet another member of the family of L_p norms: the L_∞ norm. This is formally defined as

$$L_\infty(X, Y) = \lim_{p \to \infty} L_p(X, Y).$$

Even if the definition of the L_∞ norm as a limit may appear complicated, it is possible to characterize L_∞ in a very straightforward manner: for each pitch class we trace how much it has been displaced and then select the maximum of these displacements as the value of L_∞. For example, between rows $P = 5409728136\text{AB}$ and $P' = 4590271863\text{BA}$ every pitch class is displaced by one order position and therefore $L_\infty(P, P') = 1$. Correspondingly, between rows P and $P'' = 8409725136\text{AB}$ pitch classes 5 and 8 both move six positions and therefore $L_\infty(P, P'') = 6$. The DISPLACEMENT value for both pairs is 12.

In a sense, L_∞ is the opposite of DERANGEMENT. It only traces the maximum displacement of any single pitch class but the total number of pitch classes that are displaced is immaterial. On the other hand, DERANGEMENT takes into account only the total number of pitch classes that have been displaced and it does not matter how much they have done so. The other L_p norms lie in between these two extremes.

Extending DISPLACEMENT

DISPLACEMENT can be easily extended to segments of any length as long as they contain the same pitch classes. If the pitch classes are duplicated we need to find the transformation that transforms one segment into the other with the least number of displacements. I will discuss the details of such an approach in Section 12.4.

If we wish to compare two segments with different pitch-class contents, we need to define what the displacement of a pitch class would mean if the other segment does not contain that pitch class. One solution could be to define a "penalty" for such a situation: if a pitch class had to be displaced altogether, the extent of the displacement could be, for instance, the length of the segment. The same solution could be used for segments of different lengths. For example, if we wish to transform segment 007 into segment 00 we need to displace pitch class 7 altogether. The penalty for such a displacement could be the length of the original segment, which in the case of segment 007 is 3.

8.4 PITCH-CLASS DISPLACEMENT

The DISPLACEMENT similarity measure is based on tracking the changes in the order positions of pitch classes. Due to the dual nature of twelve-tone rows, DISPLACEMENT suggests another similarity measure – one based on the displacement of pitch classes on the pitch-class dimension, labelled PITCH-CLASS DISPLACEMENT, or $PCDIS$ for short. While DISPLACEMENT measures how far pitch classes have strayed in the order-number space, PITCH-CLASS DISPLACEMENT measures how far they have strayed in the pitch-class space.

While DISPLACEMENT and PITCH-CLASS DISPLACEMENT have precisely the same starting point, a major difference arises from the fact that, unlike order numbers, pitch classes constitute a modular space: they wrap around but order positions do not.[15] Due to the modularity, PITCH-CLASS DISPLACEMENT measures the distance between pitch classes in terms of interval classes, not in terms of pitch-class intervals. Hence, the distance between pitch classes 1 and 11 is 2, while

[15] Or, to be precise, the pitch-class intervals induce a modular space on the set of pitch classes, whereas the space that the order-number intervals induce on the set of order numbers is not modular.

that between order numbers **1** and **11** is 10.

The formal definition of PITCH-CLASS DISPLACEMENT is given below.

DEFINITION 8.8 If X and Y are twelve-tone rows, then the PITCH-CLASS DISPLACEMENT value is given by the formula

$$PCDIS(X,Y) = \sum_{n=0}^{11} |ic(x_n, y_n)|$$

in which ic is the interval class between pitch classes x_n and y_n and $|ic(x_n, y_n)|$ is the interval class interpreted as an integer.

Again, Definition 8.8 is designed to be accurate rather than illustrative. In more colloquial terms, PITCH-CLASS DISPLACEMENT is defined as the sum of the interval classes of the corresponding pitch classes of pitch-class rows. Since a pitch-class row enumerates the pitch classes at each order position, the interval class of the corresponding pitch classes describes how far they have strayed. Therefore, with the introduction of a minor conceptual inaccuracy the formula for PITCH-CLASS DISPLACEMENT could be written as

$$PCDIS(X,Y) = \sum_{n=0}^{11} ic(x_n, y_n), \tag{8.3}$$

which sums the interval classes between the corresponding pitch classes of two rows. The inaccuracy here is that the interval classes do not necessarily define a group structure (while the pitch-class intervals certainly do), and even if they did, it would contain only the seven interval classes as its elements. (What would be the sum of interval classes 3 and 4?) This problem was resolved in Definition 8.8 by interpreting the interval classes as nonnegative integers – the sum of integers is naturally well defined. However, for the sake of convenience I will be using the slightly inaccurate format of Formula 8.3 as shorthand.

Let us consider rows $P = 5409728136\text{AB}$ and $Q = 5406728139\text{AB}$ in Figure 7.3. The pitch class at order position **3** in row P is 9, and the corresponding pitch class in row Q is 6; the interval class between these pitch classes is 3. Similarly, the pitch class at order position **9** in row P is 6 and the corresponding pitch class in row Q is 6; the interval class between these pitch classes is 3. The rows have the same pitch class at all other order positions. If we apply the $PCDIS$ formula the PITCH-CLASS DISPLACEMENT value for rows P and Q is

$$\begin{aligned} PCDIS(P,Q) &= ic(5,5) + ic(4,4) + ic(0,0) + ic(9,6) + ic(7,7) + ic(2,2) + \\ &\quad + ic(8,8) + ic(1,1) + ic(3,3) + ic(6,9) + ic(10,10) + ic(11,11) \\ &= 6. \end{aligned}$$

This value is therefore smaller than the DISPLACEMENT value of 12 for the same rows. Hence, we could say that the two pitch classes that are exchanged are further from each other in the order-number dimension than in the pitch-class dimension.

As with DERANGEMENT and DISPLACEMENT, it is straightforward to show that PITCH-CLASS DISPLACEMENT defines a metric. First, the values are positive real values, and secondly, the four requirements of the metric are satisfied. (i) The value for $PCDIS(X, X)$ is 0 for all rows since the corresponding pitch classes are the same. (ii) PITCH-CLASS DISPLACEMENT is clearly symmetric and $PCDIS(X, Y) = PCDIS(Y, X)$ since the interval class between two pitch classes does not depend on their order. (iii) Triangle inequality $PCDIS(X, Y) + PCDIS(Y, Z) \geq PCDIS(X, Z)$ holds because for all pitch classes $ic(x_n, y_n) + ic(y_n, z_n) \geq ic(x_n, z_n)$ (this is simply triangle inequality for the modular one-dimensional space), and it therefore also holds for the sums

$$\sum_{n=0}^{11} ic(x_n, y_n) + \sum_{n=0}^{11} ic(y_n, z_n) \geq \sum_{n=0}^{11} ic(x_n, z_n)$$

of the interval classes. (iv) Finally, if $PCDIS(X, Y) = 0$, then the sum must contain only interval class 0 and rows X and Y must therefore be identical. Since all four requirements of the metric are satisfied, PITCH-CLASS DISPLACEMENT defines a metric.

The minimum value of PITCH-CLASS DISPLACEMENT is 0, denoting maximal similarity. Since, due to the modularity of the pitch-class space, no pitch class can be transposed further away than a tritone, the maximum PITCH-CLASS DISPLACEMENT value is $12 \cdot 6 = 72$, denoting maximal dissimilarity. Hence, the maximally dissimilar rows are related by transposition T_6. Since for any row A and T_6A, $PCDIS(A, T_6A) = 72$, two rows belonging to the same row class are maximally dissimilar according to PITCH-CLASS DISPLACEMENT.

Since PITCH-CLASS DISPLACEMENT measures similarity based on the GIS of pitch-class rows, its value for rows related by transposition or inversion does not depend on the rows, but only on the operation. Hence, for any row X, $PCDIS(X, T_1X) = 12$. Correspondingly, for any row X and any inversion I_k, $PCDIS(X, I_kX) = 36$. On the other hand, the PITCH-CLASS DISPLACEMENT value for rows related by retrograde depends on the rows: the possible values range from 12 to 72.

All PITCH-CLASS DISPLACEMENT values (like DISPLACEMENT values) are even numbers. We can prove this claim by using a similar argument as in Lemma 8.1 in Section 8.3. Therefore, since the PITCH-CLASS DISPLACEMENT range spans from 0 to 72, it contains only 37 distinct values (like that of DISPLACEMENT).

PITCH-CLASS DISPLACEMENT could be considered a fine-tuned variant of the DERANGEMENT. While DERANGEMENT only counts the number of pitch classes that have been transposed by a non-zero interval, PITCH-CLASS DISPLACEMENT takes into account how much they have been transposed, thus yielding a better resolution – there are 37 distinct PITCH-CLASS DISPLACEMENT values, compared to only 12 distinct DERANGEMENT values.

The transformational approach to PITCH-CLASS DISPLACEMENT

Let us now consider PITCH-CLASS DISPLACEMENT from the transformational perspective. It has a natural interpretation as the measurement of the magnitude of the transformation that transforms one row into another.

Let us consider again pitch-class rows $P = 5409728136\mathrm{AB}$ and $Q = 5406728139\mathrm{AB}$. The pitch-class transformation that transforms row P into row Q is $QP^{-1} = 0123459786\mathrm{AB}$ (see Section 7.2.3). The sum of the differences (in a modular space) between the origin and destination of pitch classes gives us the PITCH-CLASS DISPLACEMENT value for this transformation:

$$\begin{aligned} PCDIS(P,Q) &= ic(0,0) + ic(1,1) + ic(2,2) + ic(3,3) + ic(4,4) + ic(5,5) + \\ &+ ic(6,9) + ic(7,7) + ic(8,8) + ic(9,6) + ic(10,10) + ic(11,11) \\ &= 6. \end{aligned}$$

Hence, PITCH-CLASS DISPLACEMENT measures the extent to which the tranformation mapping one row into the other differs from the identity transformation $0123456789\mathrm{AB}$.

Let us define PITCH-CLASS DISPLACEMENT in the following in GIS terms. Since it measures how the pitch classes are being transposed, I have chosen the GIS of pitch-class rows (see Definition 5.3).

DEFINITION 8.9 In the GIS of pitch-class rows, the value of the PITCH-CLASS DISPLACEMENT similarity measure for twelve-tone rows X and Y is

$$PCDIS(X,Y) = \sum_{n=0}^{11} |ic(g_n, n)|$$

in which g_n is the nth element of transformation $int(X,Y) = YX^{-1}$, $ic(g_n, n)$ is the interval class between pitch classes g_n and n, and $|ic(g_n, n)|$ is the interval class interpreted as an integer.

Let us now show that this transformational definition of PITCH-CLASS DISPLACEMENT is equivalent to the non-transformational (see Definition 8.8). The equivalence is due to the fact that both sum up precisely the same differences, even if not necessarily in the same order. For example, if row X has pitch class x_n at order position $\boldsymbol{n}$ and row Y has pitch class y_n at order position $\boldsymbol{n}$, then in transformation $int(X,Y)$ pitch class x_n is transformed into pitch class y_n, hence order position $\boldsymbol{n}$ contributes the modular difference $ic(x_n, y_n)$ to the total sum that is the value of PITCH-CLASS DISPLACEMENT. However, according to the non-transformational definition, PITCH-CLASS DISPLACEMENT is the sum of all modular differences $ic(x_n, y_n)$. Thus, the two definitions are equivalent.

As discussed in section 6.1, since PITCH-CLASS DISPLACEMENT measures the magnitude of the transformation, every row has precisely the same network of distances to the other rows.

PITCH-CLASS DISPLACEMENT for row classes

As with the previous similarity measures, in order to show that PITCH-CLASS DISPLACEMENT for row classes is well defined, we must show that it is transformationally coherent (see Corollary 6.1 in Section 6.2). The equation $PCDIS(X,Y) = PCDIS(FX, FY)$ clearly holds for all row operations: transpositions, inversion, retrograde, and their combinations. For transposition T_k, $ic(x, y) = ic(T_k(x), T_k(y))$ for all pitch classes x and y. Similarly for inversion I_k, $ic(x, y) = ic(I_k(x), I_k(y))$ for all pitch classes x and y. For the retrograde the very same interval classes are summed as in the original forms but in retrograded order. Thus, PITCH-CLASS DISPLACEMENT is a transformationally coherent metric and therefore, according to Corollary 6.1, it defines a metric for the row classes.

If we consider the extended set of row operations, PITCH-CLASS DISPLACEMENT is transformationally coherent under some of them but not all. In fact, it is transformationally coherent under all order-number operations such as the rotations since, with respect to PITCH-CLASS DISPLACEMENT, they constitute only a reordering of the pitch classes in the rows, and the same pairs of pitch classes in two rows are compared both before and after the reordering. Of the pitch-class operations, PITCH-CLASS DISPLACEMENT is transformationally coherent under only those that keep the interval classes invariant – and that amounts precisely to the transpositions and inversion.[16] This is analogous to the observation in Section 8.3 that the DISPLACEMENT similarity measure is transformationally coherent only under the order-number operations that keep the order-number intervals invariant.

The issue of transformational coherence brings up another difference between DISPLACEMENT and PITCH-CLASS DISPLACEMENT. In the case of the latter, it is transformationally coherent under all operations on “the other dimension,” in other words under all order-number operations, and under transpositions and inversion on the pitch-class dimension. In the case of DISPLACEMENT, again it is transformationally coherent under all operations on “the other dimension,” in other words under all pitch-class operations; in addition, it is transformationally coherent under retrograde but not under rotations (that would correspond to the transpositions). This disparity is, of course, due to the fact that the twelve order numbers do not form a modular space, as the twelve pitch classes do. In other words, there are 24 pitch-class operations that keep the interval classes between adjacent pitch classes invariant, but there are only two order-number operations that keep the (unordered) order-number intervals invariant.

PITCH-CLASS DISPLACEMENT is not transformationally coherent under the exchange operation either. Let us consider rows $P = 5409728136\mathrm{AB}$ and $Q = 5406728139\mathrm{AB}$ in Figure 7.3. Using the exchange transformation we obtain rows $EP = 2758109463\mathrm{AB}$ and $EQ = 2758103469\mathrm{AB}$. A comparison of the PITCH-CLASS DISPLACEMENT values for rows P and Q and for rows EP and EQ gives

$$PCDIS(P, Q) = 6 \neq 12 = PCDIS(EP, EQ),$$

which confirms that PITCH-CLASS DISPLACEMENT is not transformationally coherent under the exchange operation.

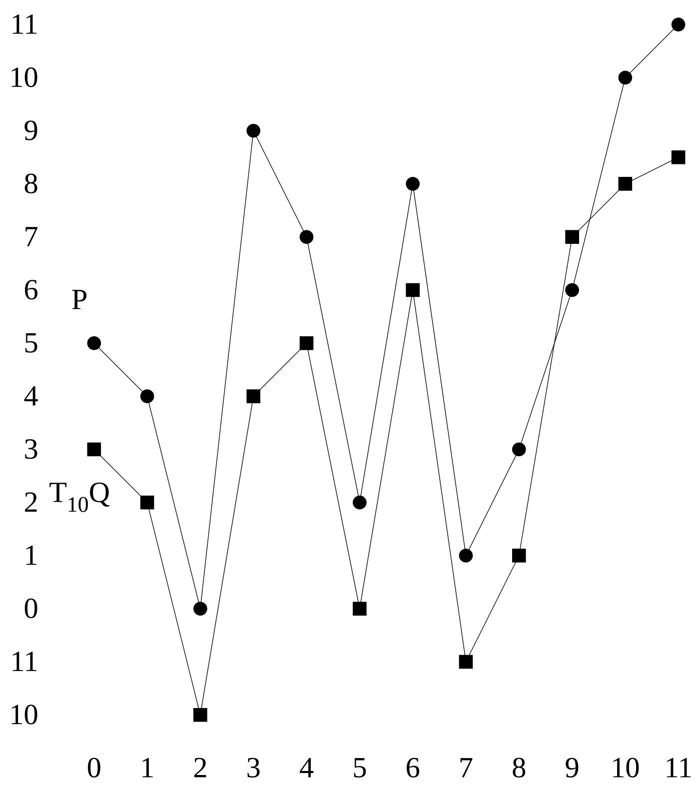


Figure 8.4: Rows $P = 5409728136\text{AB}$ and $T_{10}Q = 32\text{A}4506\text{B}1789$ as curves. The PITCH-CLASS DISPLACEMENT value is the sum of the distances between the curves at the corresponding order positions.

PITCH-CLASS DISPLACEMENT and rows as curves

Let us now take a closer look at what it is that PITCH-CLASS DISPLACEMENT actually measures. I will use the visual metaphor of a "curve" as an aid.

Figure 8.4 shows rows $P = 5409728136\text{AB}$ and $T_{10}Q = 32\text{A}4506\text{B}1789$ from Figure 7.3 as curves on a two-dimensional plane in order to visually illustrate PITCH-CLASS DISPLACEMENT. The horizontal dimension denotes the order positions (the pitch classes are ordered from left to right) and the vertical dimension denotes the pitch-class space. Since the pitch-class space "wraps around," a few pitch classes appear in two positions; this was done in order make the visualization clear by drawing the two curves as close to each other as possible.[17]

PITCH-CLASS DISPLACEMENT measures how close the corresponding pitch classes of two rows are to each other in the pitch-class space. Hence, we could argue that we are, in some sense, measuring the similarity of the curves that the two rows draw on the modular pitch-class space.

There is one caveat to this interpretation of PITCH-CLASS DISPLACEMENT, however. To continue with the spatial metaphor, PITCH-CLASS DISPLACEMENT does not actually measure the similarity of the curves: in informal terms, it rather measures how well the curves fit together based

[16]While inversion does not keep pitch-class intervals invariant does keep interval classes invariant. For an extensive discussion on interval-preserving transformations, see Chapter 3 of Lewin (1987).

[17]In order to obtain the visually most appropriate representation the paper should be wrapped around a cylinder.

on how much space there is between them.[18] These measurements are related but they do not necessarily always coincide. For example, the curves of the two transpositionally related rows $C = 0123456789\mathrm{AB}$ and $T_6C = 6789\mathrm{AB}012345$ are both lines and therefore identical as curves. However, according to PITCH-CLASS DISPLACEMENT, rows C and T_6C are maximally dissimilar. We would certainly expect the transposition level not to have an effect on the similarity of the *curves* of two rows.

We can circumvent this problem by defining a row class consisting of the twelve transpositionally related rows and comparing their similarity. The mechanism is the same as in the measurement of the similarity of "ordinary" row classes consisting of 48 rows. In this case the transposition level of the rows to be compared is immaterial: the result is the "best fit" of the two rows and could be interpreted as a measurement of the similarity of their curves. The rows could be depicted as two cylinders with dots marking the pitch classes. One cylinder is then inserted into the other and rotated until the best fit is found.

If we take the class of transpositionally related rows as a new basic unit, we need only four of these classes to examine the similarity of row classes: the classes of the prime form, inversion, retrograde, and retrograde inversion. This corresponds with the idea of the similarity of contour classes introduced by Marvin and Laprade (1987), since they also suggest that the contour class consists of the prime form, inversion, retrograde, and retrograde inversion. Hence, the class of transpositionally related rows and their concept of contour play similar roles.

The distribution of PITCH-CLASS DISPLACEMENT

Figure 8.5 shows the distribution of the values for PITCH-CLASS DISPLACEMENT and PITCH-CLASS DISPLACEMENT for row classes. The shape of the former resembles the bell curve: given a row there are very few rows that are very similar or very dissimilar to it, and the majority are neither particularly similar nor dissimilar to it. The distribution of PITCH-CLASS DISPLACEMENT is perfectly symmetrical (unlike that of DISPLACEMENT), and the middle value 36 is the highest in percentage terms. Due to the symmetry, the median and the average of values are both 36.

The maximum PITCH-CLASS DISPLACEMENT value for row classes found in the sample is 36, which is, curiously, precisely in the middle of the scale of PITCH-CLASS DISPLACEMENT for twelve-tone rows. The distribution of these values for row classes seems to have less variance than the distribution of the DISPLACEMENT values for row classes (compare Figures 8.2 and 8.5): the maximum, average, and minimum curves are clearly closer to each other in the PITCH-CLASS DISPLACEMENT Figure 8.5 than in the DISPLACEMENT Figure 8.2.

Extending PITCH-CLASS DISPLACEMENT

PITCH-CLASS DISPLACEMENT can easily be extended to segments of any length. Furthermore, there may be pitch-class duplications and the two segments do not even need to contain the same pitch classes. We cannot apply the permutational approach if there are duplications or different pitch-class contents, but the property-based Definition 8.8 of PITCH-CLASS DISPLACEMENT works

[18]This is not quite the same as the area between two curves (which could be the basis for yet another measure).

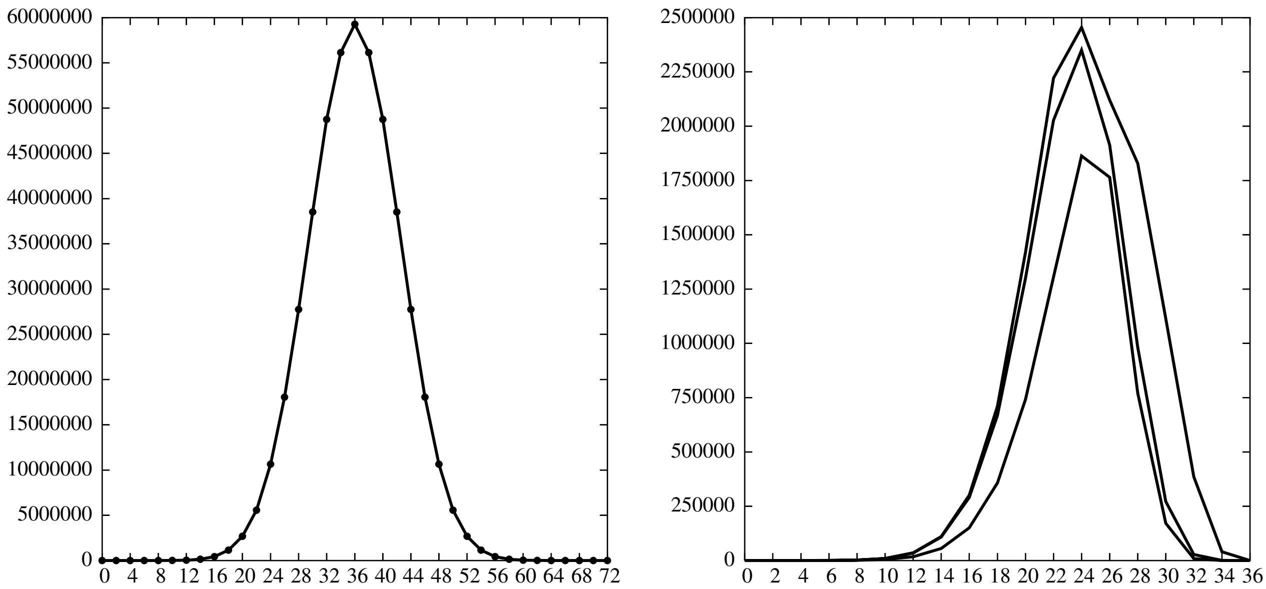


Figure 8.5: The distribution of PITCH-CLASS DISPLACEMENT (on the left) and of PITCH-CLASS DISPLACEMENT for row classes (on the right). The former is obtained by computing the distances defined by all 479001600 distinct pitch-class transformations and hence it is precise; the latter is obtained by comparing 2000 random row classes to all other row classes. See Section 6.3 for a discussion on the three curves in the latter.

well.

If we wish to compare two segments of different lengths, we could define a "penalty" for such a situation along the lines outlined in Section 8.3 in the context of extending the DISPLACEMENT similarity measure.

8.5 Transformations in the vector approach

The basic premise in designing a transformational path between two rows is that in the process the rows become gradually more and more dissimilar to the original row and more and more similar to the destination row. While for some compositional purposes it would be perfectly possible to design paths that do not take the shortest route, I will concentrate on finding paths that do not take unnecessary steps aside.

A natural approach would be to require that all pitch classes that are already in their correct order positions are not displaced in the transformation. For example, a total of seven pitch classes (5, 4, 0, 8, 3, 10, and 11) are in the same order positions in rows $P = 5409728136\text{AB}$ and $S = 5406918237\text{AB}$. I will introduce a transformational procedure for exchanging pitch classes in Section 9.4.2. Even if such a procedure is derived from the conception of rows as a set of ordered pairs, it could also be applied here, since that transformational procedure always keeps intact the pitch classes that are already in their correct order positions. Consequently, at each step more and more pitch classes will be in the same order positions as in the destination row.

CHAPTER 9

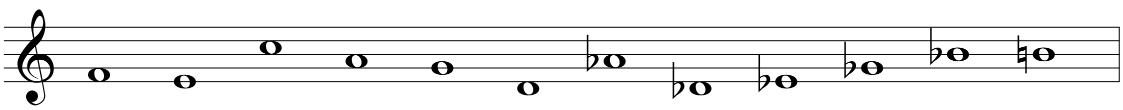
Similarity measures based on ordered pairs

This chapter concerns the measurement of similarity based on the conception of a twelve-tone row as a set of ordered pairs or dyads. The first Section 9.1 describes the conception, and the subsequent sections 9.2 and 9.3 discuss in more detail two similarity measures that are based on it. Section 9.4 introduces some transformational procedures that the notion of ordered pairs suggests, and these give rise to a new similarity measure that is discussed in Section 9.5. Section 9.6 introduces yet another measure; its relation to the notion of ordered pairs is somewhat vague, but I will show that it has strong correlation with some of the measures that are explicitly based on them. Chapter 13, which includes an extensive discussion on ordered pairs, deals with the application of graph theory to the analysis of similarity relations.

9.1 The twelve-tone row as a set of ordered pairs

In order to lay the ground for the discussion to follow on similarity measures based on the notion of ordered pairs, I will begin by considering some of their basic properties. A twelve-tone row defines an ordering on the twelve pitch classes, which in turn defines the mutual order of every pair of pitch classes. This observation is so trivial that its implications are easily overlooked. Milton Babbitt introduced the notion of the twelve-tone row as a *protocol* that defines the order in which the pitch classes appear in it (Babbitt 1962). In particular, the protocol could be described as a set of ordered pairs of type (x, y); David Lewin labels these *protocol pairs* (Lewin 1976, 252). The interpretation of such a pair is that "x must appear before y."

In a twelve-tone row, the first pitch class precedes the other eleven, the second one precedes ten, and so on. Hence, every twelve-tone row contains $11 + 10 + \ldots + 2 + 1 = 66$ ordered pairs, which is half of the total of $12 \cdot 11 = 132$ possible ordered pairs that contain two distinct pitch classes. Obviously, if i and j are two distinct pitch classes, then in a given twelve-tone row either pitch class i precedes pitch class j or pitch class j precedes pitch class i.



$(5,4)$	$(5,0)$	$(5,9)$	$(5,7)$	$(5,2)$	$(5,8)$	$(5,1)$	$(5,3)$	$(5,6)$	$(5,10)$	$(5,11)$
	$(4,0)$	$(4,9)$	$(4,7)$	$(4,2)$	$(4,8)$	$(4,1)$	$(4,3)$	$(4,6)$	$(4,10)$	$(4,11)$
		$(0,9)$	$(0,7)$	$(0,2)$	$(0,8)$	$(0,1)$	$(0,3)$	$(0,6)$	$(0,10)$	$(0,11)$
			$(9,7)$	$(9,2)$	$(9,8)$	$(9,1)$	$(9,3)$	$(9,6)$	$(9,10)$	$(9,11)$
				$(7,2)$	$(7,8)$	$(7,1)$	$(7,3)$	$(7,6)$	$(7,10)$	$(7,11)$
					$(2,8)$	$(2,1)$	$(2,3)$	$(2,6)$	$(2,10)$	$(2,11)$
						$(8,1)$	$(8,3)$	$(8,6)$	$(8,10)$	$(8,11)$
							$(1,3)$	$(1,6)$	$(1,10)$	$(1,11)$
								$(3,6)$	$(3,10)$	$(3,11)$
									$(6,10)$	$(6,11)$
										$(10,11)$

Figure 9.1: Ordered dyads of row $P = 5409728136\mathrm{AB}$ in Figure 7.3.

The following notation facilitates the discussion. If $0 \leq i \leq 11$, $0 \leq j \leq 11$ and $i \neq j$, then (i, j) denotes an ordered pair of pitch classes such that pitch class i occurs before pitch class j in a row. Therefore, either (i, j) or (j, i) is in a given row.

Figure 9.1 depicts row $P = 5409728136\mathrm{AB}$ from Figure 7.3 and its ordered dyads. The first pitch class 5 precedes the eleven other pitch classes. Since pitch class 5 precedes pitch class 4, the ordered pair $(5, 4)$ occurs in row P and the ordered pair $(4, 5)$ does not.

The protocol is a potentially useful compositional tool. As it presents the shared ordered pairs between two rows it defines the invariances of ordered pitch-class pairs that could be presented on the musical surface composed of these rows.[1]

The focus here is solely on the mutual order of the pitch-class pairs – their absolute positions are immaterial. The intervals between adjacent pitch classes also are ignored.

The notion of ordered pairs gives rise to two very different methods of evaluating the similarity of twelve-tone rows. The first method is to compare the ordered pairs in rows. Babbitt considered ordered pairs very important (Babbitt 1960; Babbitt 1961a; Babbitt 1965; Babbitt 1946/1992). The evaluation of the similarity of rows based on them is implicit his writings, but the formal definition of the ORDER INVERSIONS similarity measure was given by John Rothgeb (1967). The measure was developed further and some of its properties were examined in Morris (1987), while Ward (1992) provides some further analysis. The second method is called BADNESS OF SERIAL FIT. It was developed by David Lewin (1976), and Starr (1984) provides an extensive discussion of it.

[1] Variation V of Arnold Schoenberg's *Variations* op. 31 provides an exceptional example of the invariance of ordered dyads. The idea of this variation is that the twelve pitch classes are organized as six ordered pairs in six voices. For example, the first row form in measure 178 is divided into ordered dyads $\{98, 10, 23, 45, 67, \mathrm{AB}\}$. Since these ordered dyads are invariant in row forms $I_8V = \mathrm{A4253B671098}$ and $RT_1V = \mathrm{109823A6475B}$, the row form cannot be deciphered only on the grounds of dyads. Indeed, for this reason Tiina Koivisto (1996) and John Covach (2000) interpret some of the row forms differently.

Two other similarity measures are also based on the notion of ordered pairs. CAYLEY DISTANCE has not been used in music-theory literature before, but it is a well-known measure for permutations in mathematical contexts. CORRELATION COEFFICIENT was introduced by Robert Morris (1987), and it is far from obvious that it is actually based on the notion of ordered pairs. Indeed, a good deal of the following discussion involves attesting that it can be thus related.

ORDER INVERSIONS are examined in the following Section 9.2, and BADNESS OF SERIAL FIT in Section 9.3. Section 9.4 considers the transformational procedures that the notion of ordered pairs suggests. CAYLEY DISTANCE and CORRELATION COEFFICIENT are described in Sections 9.5 and 9.6. A graph-theoretical approach to both ORDER INVERSIONS and BADNESS OF SERIAL FIT is presented in Chapter 13.

Finally, I would point out that, due to the cyclicity of the pitch-class space, this conception cannot be extended to the realm of pitch classes. It would be most natural to claim that one order position precedes another. However, no such attribute is naturally found in the pitch classes – we certainly cannot claim that one pitch class would be higher or lower than another. Consequently, the notion of isomorphism between pitch classes and order numbers faces a challenge: how could the pitch-class space and the order-number space be isomorphic, in other words structurally identical, if we can meaningfully define a precedence relation in one of them but not in the other?

9.2 ORDER INVERSIONS

ORDER INVERSIONS or OI is based on the concept of an *order inversion*.[2] An order inversion between rows X and Y is a pair of pitch classes i and j such that pitch class i precedes pitch class j in row X but pitch class j precedes pitch class i in row Y. Naturally, for each ordered pair (i, j) that is in row X but not in row Y there is a corresponding ordered pair (j, i) that is in row Y but not in row X.

This similarity measure is based on the premise that two similar rows must have a large number of shared ordered pitch-class pairs. Hence, the degree of similarity of two rows can be examined by calculating the number of order inversions between them. The formal definition of ORDER INVERSIONS is given below.

DEFINITION 9.1 If X and Y are twelve-tone rows, then the ORDER INVERSIONS value is given by the formula

$$OI(X, Y) = \#\{(i, j) | (i, j) \in X \text{ and } (i, j) \notin Y\}$$

where i and j are distinct pitch classes.

Let us consider the order inversions between rows P and Q in Figure 7.3. The positions of pitch classes 6 and 9 are exchanged in the two rows. Pitch classes 7, 2, 8, 1, and 3 occur after pitch class 9 in row P, but before pitch class 9 in row Q; hence pairs $(9, 7)$, $(9, 2)$, $(9, 8)$, $(9, 1)$, and $(9, 3)$ of row P are inverted in row Q. Similarly, the same five pitch classes 7, 2, 8, 1, and 3 occur before pitch class 6 in row P, but after pitch class 6 in row Q; hence the pairs $(7, 6)$, $(2, 6)$, $(8, 6)$, $(1, 6)$,

[2]Interestingly, the idea of order inversions in permutations goes as far back as 1750, see Knuth (1998, 11).

and $(3,6)$ of row P are inverted in row Q. Finally, since the mutual order of pitch classes 9 and 6 is changed, the pair $(9,6)$ of row P is inverted in row Q. Therefore, the ORDER INVERSIONS value for rows P and Q is 11. These order inversions are illustrated in Figure 9.2.

Tracking the various order inversions between rows can be tricky and error prone. A foolproof way to calculate the ORDER INVERSIONS value is to enumerate the ordered pairs in both rows in a manner shown in Figure 9.1, then to calculate the number of *shared* pairs, and finally to subtract that number from 66. The set of shared pairs is the intersection of the pairs of two rows. Since the total number of 66 ordered pairs in a row comprises the shared and non-shared pairs between two rows, we can express ORDER INVERSIONS in terms of the intersection of the sets of the ordered pairs of two rows as

$$OI(X,Y) = \#\{(i,j) | (i,j) \in X \text{ and } (i,j) \notin Y\} \\ = 66 - \#(\{(i,j) \ \in X\} \cap \{(i,j) \in Y\}) \tag{9.1}$$

in which i and j are distinct pitch classes.

The range of ORDER INVERSIONS contains 67 distinct values. The minimum value is 0 denoting maximum similarity, and the maximum value is $\binom{12}{2} = 66$ denoting maximum dissimilarity. The maximally dissimilar rows are related by retrograde. Since for any row A and RA, $OI(A, RA) = 66$, two rows belonging to the same row class are maximally dissimilar according to ORDER INVERSIONS.

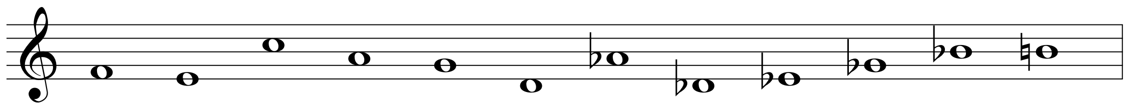
It is straightforward to show that ORDER INVERSIONS defines a metric. First, the values are positive real values, and secondly, the four requirements of the metric are satisfied. (i) Trivially, the value of $OI(X,X)$ is 0 for all rows: there are no order inversions between two identical rows. (ii) ORDER INVERSIONS is clearly symmetric and $OI(X,Y) = OI(Y,X)$. This is most easily seen in Formula 9.1 for ORDER INVERSIONS in which the symmetry is due to the symmetry of operator $\cap$ (for any sets A and B, the intersection $A \cap B$ equals the intersection $B \cap A$). If the ordered pairs (a_1,b_1), (a_2,b_2), ..., (a_n,b_n) of row X are inverted in row Y then the ordered pairs (b_1,a_1), (b_2,a_2), ..., (b_n,a_n) of row Y are inverted in row X. (iii) Triangle inequality $OI(X,Z) \leq OI(X,Y) + OI(Y,Z)$ holds since if $OI(X,Y) = m$ and $OI(Y,Z) = n$, then rows X and Y have $66 - m$ common pairs, and of these at least $66 - m - n$ are common to row Z. Thus rows X and Z have at maximum $m + n$ different pairs. (iv) Finally, if $OI(X,Y) = 0$, then rows X and Y do not have any pairs of pitch classes in a different order so they must be identical. Since all four requirements are satisfied, ORDER INVERSIONS defines a metric.

The transformational approach to ORDER INVERSIONS

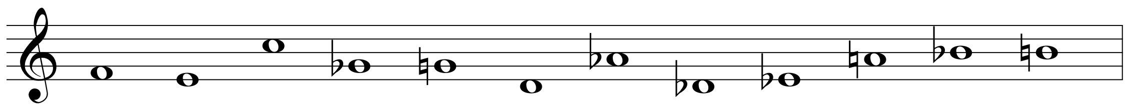
Let us now consider ORDER INVERSIONS from the transformational perspective. To that end it is necessary to restate the concept in GIS terms. This facilitates the use of some well-known mathematical results in the analysis of its properties and in the calculation of its distribution.

An order inversion is related to the mathematical concept of *permutation inversion*.[3] The formal

[3]The term permutation *inversion* should not be confused with the *inverse* of a permutation, which means the inverse



(5, 4)	(5, 0)	(5, 9)	(5, 7)	(5, 2)	(5, 8)	(5, 1)	(5, 3)	(5, 6)	(5, 10)	(5, 11)
	(4, 0)	(4, 9)	(4, 7)	(4, 2)	(4, 8)	(4, 1)	(4, 3)	(4, 6)	(4, 10)	(4, 11)
		(0, 9)	(0, 7)	(0, 2)	(0, 8)	(0, 1)	(0, 3)	(0, 6)	(0, 10)	(0, 11)
			(9, 7)	(9, 2)	(9, 8)	(9, 1)	(9, 3)	(9, 6)	(9, 10)	(9, 11)
				(7, 2)	(7, 8)	(7, 1)	(7, 3)	(7, 6)	(7, 10)	(7, 11)
					(2, 8)	(2, 1)	(2, 3)	(2, 6)	(2, 10)	(2, 11)
						(8, 1)	(8, 3)	(8, 6)	(8, 10)	(8, 11)
							(1, 3)	(1, 6)	(1, 10)	(1, 11)
								(3, 6)	(3, 10)	(3, 11)
									(6, 10)	(6, 11)
										(10, 11)



(5, 4)	(5, 0)	(5, 6)	(5, 7)	(5, 2)	(5, 8)	(5, 1)	(5, 3)	(5, 9)	(5, 10)	(5, 11)
	(4, 0)	(4, 6)	(4, 7)	(4, 2)	(4, 8)	(4, 1)	(4, 3)	(4, 9)	(4, 10)	(4, 11)
		(0, 6)	(0, 7)	(0, 2)	(0, 8)	(0, 1)	(0, 3)	(0, 9)	(0, 10)	(0, 11)
			(6, 7)	(6, 2)	(6, 8)	(6, 1)	(6, 3)	(6, 9)	(6, 10)	(6, 11)
				(7, 2)	(7, 8)	(7, 1)	(7, 3)	(7, 9)	(7, 10)	(7, 11)
					(2, 8)	(2, 1)	(2, 3)	(2, 9)	(2, 10)	(2, 11)
						(8, 1)	(8, 3)	(8, 9)	(8, 10)	(8, 11)
							(1, 3)	(1, 9)	(1, 10)	(1, 11)
								(3, 9)	(3, 10)	(3, 11)
									(9, 10)	(9, 11)
										(10, 11)

Figure 9.2: Order inversions between rows $P = 5409728136\mathrm{AB}$ (top) and $Q = 5406728139\mathrm{AB}$ (bottom) in Figure 7.3. The ordered pairs that the two rows do not share are inside the boxes.

definition of permutation inversion is given below.

DEFINITION 9.2 Let $P = p_0\, p_1 \ldots p_{n-1}$ be a permutation. If $i < j$ and $p_i > p_j$ then pair (p_i, p_j) is called an inversion of permutation P.

In colloquial terms, a permutation inversion is a pair of elements that do not appear in their "natural" order in a permutation. For example, 4 and 5 are considered to be in their natural order if 4 precedes 5. Hence, if 5 is before 4 then pair $(5, 4)$ is a permutation inversion.

This concept is not very useful in the analysis of twelve-tone rows in that a row in which pitch class 4 precedes pitch class 5 is in no way more natural than a row in which pitch class 5 precedes pitch class 4. However, it is useful for analyzing the transformation between two rows. The following definition defines ORDER INVERSIONS in terms of the GIS of order-number rows (see Definition 5.3) and permutation inversions.

DEFINITION 9.3 In the GIS of order-number rows, the ORDER INVERSIONS value for twelve-tone rows X and Y is

$$OI(X, Y) = \#\{(p_i, p_j) \mid i < j \text{ and } p_i > p_j\}$$

in which (p_i, p_j) is a pair in order-number transformation $int(\boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{Y}\boldsymbol{X}^{-1}$.

ORDER INVERSIONS is thus defined simply as the number of permutation inversions in order-number transformation $int(\boldsymbol{X}, \boldsymbol{Y})$.

Let us consider rows $P = 5409728136\text{AB}$ and $Q = 5406728139\text{AB}$ in Figure 7.3. We begin by writing these two rows as order-number rows $\boldsymbol{P} = \mathbf{2758109463AB}$ and $\boldsymbol{Q} = \mathbf{2758103469AB}$. The order-number transformation that maps row $\boldsymbol{P}$ into row $\boldsymbol{Q}$ is $\boldsymbol{QP}^{-1} = \mathbf{0129456783AB}$ (see Section 7.2.3). We then need to find all pairs of elements that do not appear in their natural order in the transformation. For example, 9 precedes 3, therefore 3 and 9 do not appear in their natural order and hence pair $(9, 3)$ constitutes a permutation inversion. Just as expected, there is a total of 11 permutation inversions in transformation $\boldsymbol{QP}^{-1}$: $(9, 3)$, $(9, 4)$, $(9, 5)$, $(9, 6)$, $(9, 7)$, $(9, 8)$, $(4, 3)$, $(5, 3)$, $(6, 3)$, $(7, 3)$, and $(8, 3)$.

We can show that the transformational definition of ORDER INVERSIONS is equivalent to the non-transformational definition (see Definition 9.1) by considering the order-number transformation $\boldsymbol{YX}^{-1}$ that maps the order numbers of row X into the order numbers of row Y. If (j, i) is a permutation inversion in $\boldsymbol{YX}^{-1}$ (and, hence, $i < j$), then in row Y there are pitch classes at order positions $\boldsymbol{i}$ and $\boldsymbol{j}$ that are in the opposite order in row X. Correspondingly, if $i < j$ and in row X there are pitch classes at order positions $\boldsymbol{i}$ and $\boldsymbol{j}$ that are in the opposite order in row Y, then in transformation $\boldsymbol{YX}^{-1}$ order number $\boldsymbol{i}$ must be mapped into a greater order number than $\boldsymbol{j}$ and hence, $(\boldsymbol{YX}^{-1}(\boldsymbol{j}), \boldsymbol{YX}^{-1}(\boldsymbol{i}))$ is a permutation inversion in $\boldsymbol{YX}^{-1}$.

More formally, we can prove the equivalence of the non-transformational definition 9.1 and the transformational definition 9.3 by utilizing the fact that the pitch classes in two rows can be relabeled without altering the ORDER INVERSIONS value: the labels of pitch classes are immaterial

mapping of a permutation.

with respect to the order inversions. For example, if pitch classes 0 and 1 are in a different order in two rows, we could relabel pitch class 0 as α and pitch class 1 as β and then we would have α and β in a different order instead of 0 and 1. As discussed in Section 5.6, it would be possible to relabel the pitch classes in the two rows X and Y by multiplying both rows by any permutation τ from the left. In particular, we could choose permutation τ to be Y^{-1} and thus obtain the relabeled rows $Y^{-1}X$ and $Y^{-1}Y = id$ (the identity permutation 0123456789AB). Since all pitch classes are in their natural order in the identity permutation, an order inversion between the original rows X and Y translates directly to a permutation inversion in permutation $Y^{-1}X$. Now recall that a pitch-class row as a permutation is the inverse of the corresponding order-number row as a permutation: we could thus write equation $Y^{-1}X = \boldsymbol{YX}^{-1}$ for the permutations. Note further that $\boldsymbol{YX}^{-1}$ is the order-number transformation that transforms row X into row Y. Thus, an order inversion of rows X and Y translates directly to a permutation inversion in transformation $\boldsymbol{YX}^{-1}$, and therefore the non-transformational definition 9.1 and the transformational definition 9.3 are equivalent.

Milton Babbitt mentions several times the fact that if we take two complementary transformations, for example T_n and T_{12-n}, then for any row X there are equally many order inversions between the rows X and T_nX and between the rows X and $T_{12-n}X$ (Babbitt 1960; Babbitt 1961a; Babbitt 1965). The transformational approach allows us to prove this easily. Since the ORDER INVERSIONS value between two rows is the number of permutation inversions in the transformation that maps one row into the other, we can prove the claim by showing that the number of permutation inversions is always the same in a permutation and its inverse.[4]

LEMMA 9.1 If σ is a finite permutation and τ is its inverse permutation, then the number of inversions in σ and τ is the same.

Proof. The proof entails showing that the number of pairs that are in their natural order is the same in both permutations.

Let us examine the permutations of the integers $1, 2, \ldots, n$. Let f be a function from the integers to their position in the permutation σ and let g be a function from the integers to their position in the permutation τ. Let the symbol $<_\sigma$ denote a relation of elements that are in their natural order in σ and let the symbol $<_\tau$ denote a relation of elements that are in their natural order in τ (and $<$ without an index denotes just the ordinary "smaller than" relation). Hence,

$$x <_\sigma y \Leftrightarrow x < y \text{ and } f(x) < f(y).$$

Similarly,

$$x <_\tau y \Leftrightarrow x < y \text{ and } g(x) < g(y).$$

Now, because σ and τ are inverse permutations, f and g are inverse functions. Hence, $x = g(f(x))$

[4]See Rothgeb (1967) for a different proof showing that the similarity of row X to row Y, measured by order inversions, is the same as that of row Y to row X, similarly measured.

and

$$\begin{aligned} x <_\sigma y &\Leftrightarrow x < y \text{ and } f(x) < f(y) \\ &\Leftrightarrow g(f(x)) < g(f(y)) \text{ and } f(x) < f(y) \\ &\Leftrightarrow f(x) <_\tau f(y). \end{aligned}$$

Thus, relation $x <_\sigma y$ holds if and only if relation $f(x) <_\tau f(y)$ holds. Because f is an bijection, it is evident that $<_\sigma$ and $<_\tau$ are relations of the same cardinality. Thus, the number of pairs in ascending order is the same in both permutations and, consequently, the number of inversions is also the same. □

ORDER INVERSIONS for row classes

Let us define ORDER INVERSIONS for row classes in the same way as previous similarity measures. In order to show that it is well defined, it must be shown to be transformationally coherent (see Corollary 6.1 in Section 6.2). ORDER INVERSIONS is transformationally coherent since the equation $OI(X, Y) = OI(FX, FY)$ clearly holds for all row operations: transpositions, inversion, retrograde, and their combinations. For transposition T_k, for every order-inversion pair (x, y) in rows X and Y there is a corresponding order-inversion pair $(T_k(x), T_k(y))$ in the transposed rows T_kX and T_kY. Similarly, for inversion I_k, for every order-inversion pair (x, y) in rows X and Y there is a corresponding order-inversion pair $(I_k(x), I_k(y))$ in the inverted rows I_kX and I_kY. Finally, the order inversions in the retrograded forms involve the same pairs of pitch classes as in the original forms, since if pitch class x precedes pitch class y in row X but not in row Y, then pitch class y precedes pitch class x in row RX but not in row RY. Thus, ORDER INVERSIONS is a transformationally coherent metric and therefore, according to Corollary 6.1, it defines a metric for row classes.

As far as the extended set of row operations is concerned, ORDER INVERSIONS is transformationally coherent under some of them but not all. In fact, it is transformationally coherent under all pitch-class operations, such as the M-operation, since with respect to ordered pairs they constitute only a relabeling of the pitch classes. In general, it is not transformationally coherent under the order-number operations – retrograde being the only exception. For example, it is not transformationally coherent under the (non-trivial) rotations. Let us consider row $P = 5409728136\text{AB}$ in Figure 7.3, and the modification $P' = 4509728136\text{AB}$, in which the positions of the first two pitch classes have been exchanged, and the rotations of both rows $r_{11}P = 409728136\text{AB}5$ and $r_{11}P' = 509728136\text{AB}4$. A comparison of the values for the two original rows and two rotated rows gives

$$OI(P, P') = 1 \neq 21 = OI(r_{11}P, r_{11}P'),$$

and therefore ORDER INVERSIONS is not transformationally coherent under the rotation. Transformational coherence is broken since the rotations "wrap around." Pitch classes 4 and 5 are adjacent in P and P' but non-adjacent in $r_{11}P$ and $r_{11}P'$. If the rows were treated as circular entities, ORDER

INVERSIONS would be transformationally coherent under the rotations, but then the order of the pitch classes would not be defined.

A similar argument to that used in Section 8.3 will show that ORDER INVERSIONS is not transformationally coherent under any other order-number operation than retrograde. If, for example, order-number operation $\boldsymbol{F}$ changes the unordered order-number interval between order numbers $\boldsymbol{p_0}$ and $\boldsymbol{p_1}$, it will change the number of pitch classes between the pitch classes at those order positions. Consequently, we can construct the following order-number rows:

$$\boldsymbol{X} = \boldsymbol{p_0p_1p_2p_3p_4p_5p_6p_7p_8p_9p_{10}p_{11}} \quad \text{and} \quad \boldsymbol{Y} = \boldsymbol{p_1p_0p_2p_3p_4p_5p_6p_7p_8p_9p_{10}p_{11}}.$$

The ORDER INVERSIONS value for rows $\boldsymbol{X}$ and $\boldsymbol{Y}$ and for rows $\boldsymbol{FX}$ and $\boldsymbol{FY}$ depend only on the unordered order-number interval between order numbers $\boldsymbol{p_0}$ and $\boldsymbol{p_1}$. Since order-number operation $\boldsymbol{F}$ changes that unordered order-number interval, the ORDER INVERSIONS value for rows $\boldsymbol{X}$ and $\boldsymbol{Y}$ is not the same as it is for rows $\boldsymbol{FX}$ and $\boldsymbol{FY}$. Hence, ORDER INVERSIONS is not transformationally coherent under $\boldsymbol{F}$.

ORDER INVERSIONS is not transformationally coherent under the exchange operation either. Let rows $P = 5409728136\text{AB}$ and $r_{11}P = 409728136\text{AB}5$ be as above. The exchange transformations of these two rows are $EP = 2758109463\text{AB}$ and $Er_{11}P = 16470\text{B}83529\text{A}$. A comparison of the ORDER INVERSIONS values for rows P, and $r_{11}P$ and their transformations using the exchange operation gives

$$OI(P, r_{11}P) = 11 \neq 33 = DIS(EP, Er_{11}P),$$

which confirms the observed dissimilarity.

The distribution of ORDER INVERSIONS

We can use a well-known mathematical formula to derive the distribution of ORDER INVERSIONS. Theorem 2.3 in Bóna (2004) states that the generating polynomial for the number of permutation inversions in a permutation is

$$(9.2) \qquad \sum_{p \in S_n} x^{i(p)} = (1+x)(1+x+x^2)(1+x+x^2+x^3)\cdots(1+x+x^2+x^3+\ldots+x^{n-1})$$

where $i(p)$ denotes the number of permutation inversions. The interpretation of the polynomial is that the coefficient of the monomial x^k denotes the number of permutations having k inversions.

By using Formula 9.2 we can derive (using a computer program since the polynomial expansion contains 12! terms altogether) the distribution of ORDER INVERSIONS. This is depicted in Figure 9.3.

Since the sum of permutation inversions in retrograde-related permutations (of twelve elements) is always 66, the distribution is symmetric. In other words, if there are n permutation inversions in permutation $p_0p_1p_2p_3p_4p_5p_6p_7p_8p_9p_{10}p_{11}$, then there are $66 - n$ permutation inversions in permuta-

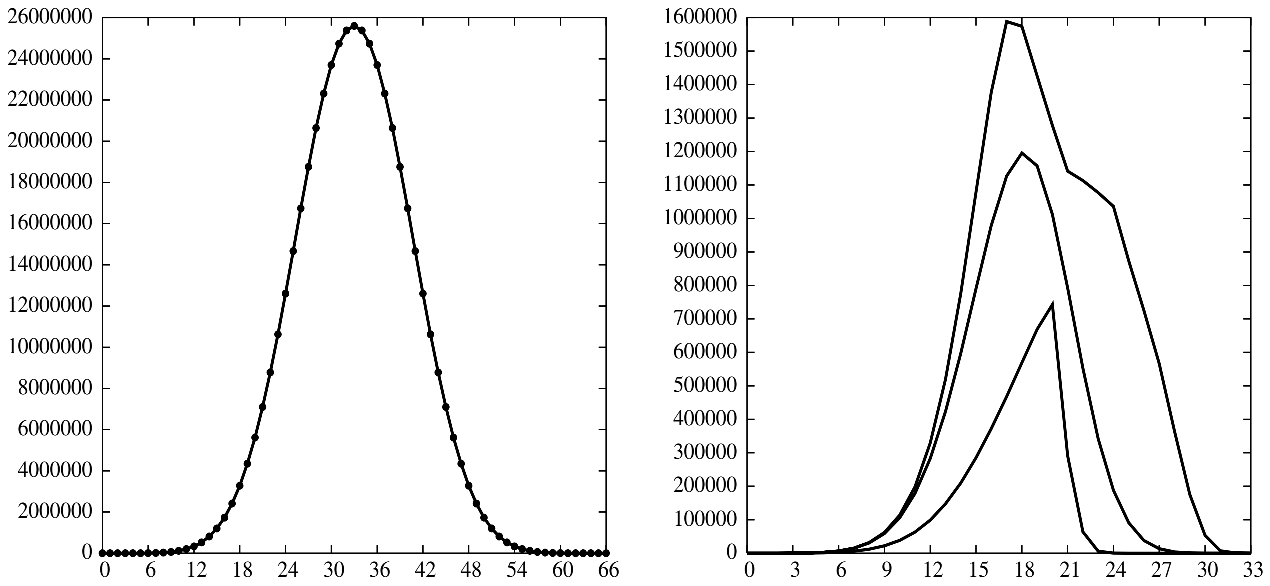


Figure 9.3: The distribution of ORDER INVERSIONS on the left and the distribution of ORDER INVERSIONS for row classes on the right. The former was obtained using Formula 9.2 and therefore it is precise; the latter was obtained by comparing 2000 random row classes to all other row classes. See Section 6.3 for a discussion on the three curves in the latter.

tion $p_{11}p_{10}p_9p_8p_7p_6p_5p_4p_3p_2p_1p_0$. Hence,

$$OI(X,Y) = 66 - OI(X,RY). \tag{9.3}$$

Donald Knuth derives a formula for the mean $\frac{n(n-1)}{4}$ of the distribution for permutations of length n (Knuth 1998, 16). Because of the symmetry of the distribution the mean can also be obtained by dividing the maximum value $\binom{n}{2}$ by 2. In this case the length is 12; hence, the mean is $\frac{12\cdot 11}{4} = 33$, which is exactly the middle of the scale.

In terms of enumerative combinatorics, an explicit formula would be far preferable to a generating polynomial. However, Formula 9.2 for the distribution of ORDER INVERSIONS is considerably more complex than Formula 8.1 for the distribution of DERANGEMENT. An explicit recursive formula can be derived in cases in which a permutation of n elements has at most n order inversions, but for the computation of the number of permutations in S_{12} with more than 12 inversions we must resort to the generating polynomial.[5]

The distribution of ORDER INVERSIONS resembles the bell curve. Hence, given a row, there are very few rows that are very similar to it or very dissimilar to it, and the majority are neither similar nor dissimilar.

For the distribution of the values between row classes we must resort to sampling. The maximum value poses an interesting problem. The absolute maximum value is 33 since, if the distance of rows X and Y is greater than 33 then the distance of rows X and RY is less than 33. Theoretically, there

[5] See Formula 2.5 in Bóna (2004).

row class	classes at 32	symmetry	hexachords	intervals
5409728136AB	264	R-symmetric	C-hexachord	11
061728394A5B	480	R-symmetric	D-hexachord	2
01235BA46789	21	RI-symmetric	6-2	4
032597648BA1	5	RI-symmetric	6-33	4
0486A2B73519	3993	semi-symmetric	F-hexachord	5
023A15B94786	9	non-symmmetric	B-hexachord	7
01235B4A6789	15	non-symmmetric	6-2	5
03691B2485A7	1	non-symmmetric	6-Z45/6-Z23	6
06493B1752A8	4	non-symmetric	6-Z29/6-Z50	6

Table 9.1: Some row classes, the number of row classes at a distance of 32 and some characteristics, symmetry properties, hexachords, and the number of distinct intervals between the adjacent pitch classes in the row.

might be two row classes at a distance of 33 but so far no such pair has been found.[6]

There are several pairs of row classes at a distance of 32 from each other. Table 9.1 shows a selection of row classes with other row classes at a distance of 32. The first column denotes the row, the second column is the number of row classes that are at a distance of 32 from the row class of the first column, and the last three columns depict the symmetry properties of the row, the hexachord contents, and the number of distinct intervals between the adjacent pitch classes, respectively. The variety of different types of rows present in Table 9.1 shows that we cannot pinpoint any property of the rows (other than the order of pitch classes) that would cause them to have row classes at a distance of 32. The table includes symmetric, semi-symmetric[7], and non-symmetric row classes, rows with two identical hexachord set classes, rows with all-combinatorial set classes, and rows with two Z-related hexachord set classes, all-interval rows and rows with only two distinct intervals between the adjacent pitch classes.

As with the previous similarity measures, the distribution of the ORDER INVERSIONS values between row classes varies. The row classes of the symmetric rows (such as rows P and W) have fewer very close neighbors than the row class of the non-symmetric rows (such as row V). Again, one might explain this phenomenon simply by observing that the symmetric rows have, in a sense, fewer possibilities of being transformed into rows of other row classes. A similar argument to that used in Section 8.2 suggests that the symmetry of the row limits the number of rows belonging to different row classes that can be obtained by only small displacements of a few pitch classes. However, the very limited data in Table 9.1 also suggests that symmetric rows have more rows at a distance of 32 than non-symmetric rows (which is in accordance with the observation that the row classes of symmetric rows have fewer very close neighbors than those of non-symmetric rows).

[6]An exhaustive search would involve $\binom{9985920}{2} = 49859294130240$ comparisons of row classes – a task beyond available computing facilities.

[7]A semi-symmetric row means here a row in which the second hexachord is a transposed retrograde of the first one.

operation	min	max
T_0	0	0
T_1, T_5, T_7, T_{11}	11	61
T_2, T_{10}	10	62
T_3, T_9	9	63
T_4, T_8	8	44
T_6	6	66
I_{2k}	5	65
I_{2k+1}	6	66

operation	min	max
T_0R	66	66
$T_1R, T_5R, T_7R, T_{11}R$	5	55
$T_2R, T_{10}R$	4	56
T_3R, T_9R	3	57
T_4R, T_8R	22	58
T_6R	0	60
$I_{2k}R$	1	61
$I_{2k+1}R$	0	60

Table 9.2: The minimum and maximum values of ORDER INVERSIONS in rows related by the 48 canonical row operations.

ORDER INVERSIONS and row operations

The values of ORDER INVERSIONS and twelve-tone operations have been an issue of interest since the early days of twelve-tone theory. Back in 1946 Babbitt discussed the condition under which pitch class a precedes pitch class b in a given row and a given row operation (Babbitt 1946/1992). Rothgeb tracked the minimum and maximum numbers of order inversions that row operations can induce (Rothgeb 1967): these numbers are reproduced in Table 9.2.

Two observations can be made about Table 9.2. First, the retrograde-related operations have complementary values in the sense that the minimum ORDER INVERSIONS value in an operation plus the maximum value in its retrograded operation always equals 66. Secondly, there is a significant variance between the minimum and maximum numbers. For example, some rows have no order inversions under RT_6 while others have as many as 60. This implies that some rows offer more possibilities for invariant pairs of pitch classes than others.

With respect to presortedness, it has been remarked as a negative property of order inversions that permutations "of the type

$$(n{+}1 \;\; n{+}2 \;\; n{+}3 \;\; \ldots 2n \;\; 1 \;\; 2 \;\; 3 \;\; \ldots n)$$

have a quadratic number of inversions, even though such sequences are intuitively almost in order and are also easy to sort using merging" (Mannila 1985, 319). For example, intuitively rows $P = 5409728136\mathrm{AB}$ and $r_6P = 8136\mathrm{AB}540972$ do not seem particularly dissimilar since the two hexachords are identical (and only the order is different). However, the ORDER INVERSIONS value for these two rows is 36, which is slightly higher than the median value 33.

9.3 BADNESS OF SERIAL FIT

The BADNESS OF SERIAL FIT similarity measure, or BSF, builds on the same property of rows as ORDER INVERSIONS: the ordered pairs of pitch classes. The idea is not to measure the differences in two rows, but rather to pick out their similar features and count the number of rows that share them. The more similar two rows are, the more common properties they have, and the

three-tone rows	protocol	rows that satisfy the protocol	BSF
012 021	01 02	012, 021	2
021 210	21	021, 201, 210	3
012 210	$\emptyset$	012, 021, 102, 120, 201, 210	6

Figure 9.4: The protocols formed by the combinations of three "three-tone rows" 012, 021, and 210. The first column shows a pair of rows, the second column shows their shared protocol, the third column shows the rows that satisfy the protocol, and the fourth column shows the BADNESS OF SERIAL FIT value for the pairs of three-tone rows.

more distinctive this combination of properties is and therefore the fewer rows there are with these properties.

BADNESS OF SERIAL FIT turns out to have very fine resolution – in fact, it is several orders of magnitude finer than any other similarity measure discussed in this work. However, this comes at the cost of significantly increased complexity in the calculation: while it is perfectly clear what we are measuring, the complexity of the actual measurement renders the relation between the rows compared and the value returned somewhat like a black box.

BADNESS OF SERIAL FIT, like ORDER INVERSIONS, is based on protocol pairs. While the latter counts the differences between two rows, the idea of the former is first to generate the shared protocol (the set of common ordered pairs) of two rows and then to count the number of rows that satisfy it. The expression "satisfying a protocol" means that if a protocol contains the set of ordered pairs $\{(a_1, b_1)$, (a_2, b_2), ..., $(a_n, b_n)\}$, then if a row satisfies that protocol its set of 66 ordered pairs must include those occurring in the protocol. The size of the protocol that rows X and Y define is $66 - OI(X, Y)$.

Let us consider rows $P = 5409728136\text{AB}$ and $Q = 5406728139\text{AB}$ in Figure 7.3. As discussed in Section 9.2, the ORDER INVERSIONS value for rows P and Q is 11; as 11 pairs are inverted, they share a total of $66 - 11 = 55$ ordered pitch-class pairs. In order to calculate the BADNESS OF SERIAL FIT value for these two rows, let us first define these 55 common pitch-class pairs as a protocol and then count the total number of twelve-tone rows that satisfy it, i.e., the total number of twelve-tone rows that have these 55 ordered pitch-class pairs. The number is 42, and thus according to BADNESS OF SERIAL FIT these two rows are very close relatives since only a tiny portion of the 479001600 twelve-tone rows satisfy the protocol they define.[8]

Let us now examine the BADNESS OF SERIAL FIT values. A row defines a protocol in itself, and it is the only row that satisfies that protocol (since no other row contains precisely the same set of ordered pairs). Thus $BSF(X, X) = 1$ for any twelve-tone row X. As is the case with ORDER INVERSIONS, retrograde-related rows always have the largest value: when a row is retrograded the order of every pair of dyads is changed and therefore retrograde-related rows do not share a single ordered dyad and the protocol defined by them is empty. Since any row satisfies an empty protocol, $BSF(X, RX) = 479001600$ for any twelve-tone row X. Hence, the BADNESS OF SERIAL FIT

[8]This case is relatively easy to calculate by hand, but a computer application is required to calculate the value of BADNESS OF SERIAL FIT for two arbitrary rows.

values range from 1 denoting maximal similarity to 479001600 denoting maximal dissimilarity.

In general, the more pairs the protocol has, the more refined it is, and the fewer rows satisfy it; this relationship is very complex, however, and the size of the protocol is a poor indicator of the BADNESS OF SERIAL FIT value (I will return to this in Chapter 13 with some formal tools to illustrate the computational complexity of BADNESS OF SERIAL FIT). For example, row $0123456789\mathrm{AB}$ defines a protocol of size 42 with both rows $0456789\mathrm{AB}123$ and $215430\mathrm{BA}9876$. However, the BADNESS OF SERIAL FIT value for rows $0123456789\mathrm{AB}$ and $0456789\mathrm{AB}123$ is 165, and for rows $0123456789\mathrm{AB}$ and $215430\mathrm{BA}9876$ it is 51840. Hence, we cannot predict its value from the size of the protocol.

David Lewin noted that the common protocol of two rows defines a partial order: the order of some but not necessarily all pairs of pitch classes is defined in it. For example, in the partial order defined by rows P and Q in Figure 7.3, the order of pitch classes 6 and 9 is not defined (since in row P pitch class 9 is before pitch class 6, and in row Q pitch class 6 is before pitch class 9). I will defer the discussion on BADNESS OF SERIAL FIT in terms of partial orders until Chapter 13.

It is easy to prove that BADNESS OF SERIAL FIT does not define a metric. First of all, the first criterion (see Definition 4.1) is not satisfied since zero is not a possible value given its range. Furthermore, triangle inequality (the third requirement) does not hold either. For a simple counterexample showing why triangle inequality does not hold, let us consider the three-tone rows using pitch classes $\{0, 1, 2\}$ and the protocols they define. Figure 9.4 lists three pairs of total orders of the set with their protocols, and the number of three-tone rows satisfying the protocol. The inequality

$$BSF(012, 021) + BSF(021, 210) = 3 + 2 = 5 < 6 = BSF(012, 210)$$

shows that triangle inequality does not hold for these three three-tone rows. It does not hold for BADNESS OF SERIAL FIT for twelve-tone rows, either. A Monte Carlo-type of survey of how often triangle inequality holds has given the estimation that it does not hold in about 28.6% of random triplets (Ilomäki and Kortesniemi 2004). In other words, if we choose three random twelve-tone rows, there is a probability of 28.6% that the distances between these rows do not satisfy triangle inequality. The second requirement of the metric (symmetry) is satisfied since BADNESS OF SERIAL FIT is defined symmetrically as the number of rows satisfying the protocol that is common to the two rows. The fourth requirement ($d(x, y) = 0$ implies $x = y$) is satisfied trivially since zero is an unattainable value.

An unique feature of BADNESS OF SERIAL FIT is that it is not limited to the examination of two rows. We might as well consider the protocol defined by three or more twelve-tone rows. In Chapter 13 I will show how the notion of partial orders makes this definition very natural.

The transformational approach to BADNESS OF SERIAL FIT

At first sight, BADNESS OF SERIAL FIT does not look like a transformational measure. However, there is a simple argument showing it has a natural interpretation as the measurement of the magnitude of the transformation that transforms one row into another.

Assume that we are measuring BADNESS OF SERIAL FIT in two rows X and Y. Hence, the first step is to define the shared protocol of the rows. Let us label this $X \cap Y$ (the technical aspects and validity of this notation will be explained in Chapter 13 in which BADNESS OF SERIAL FIT is discussed in terms of partial orders). Let us now relabel the pitch classes in such a way that row Y becomes row $0123456789\mathrm{AB}$ (see Section 5.6). Therefore, the pitch class at nth position in row Y is labeled n in both rows. The new rows will now be $Y^{-1}X$ and $Y^{-1}Y = 0123456789\mathrm{AB}$, and the BADNESS OF SERIAL FIT value for the original two rows X and Y is identical to that for the two rows $Y^{-1}X$ and $Y^{-1}Y$. Now, since order-number operations and twelve-tone rows can both be reinterpreted as permutations, we can reinterpret row $Y^{-1}X$ as the order-number operation $\boldsymbol{Y}\boldsymbol{X}^{-1}$ that transforms order-number row $\boldsymbol{X}$ into order-number row $\boldsymbol{Y}$ (since pitch-class row X interpreted as a permutation is identical to the order-number row $\boldsymbol{X}^{-1}$ interpreted as a permutation, applying order-number operation $Y^{-1}X = \boldsymbol{Y}\boldsymbol{X}^{-1}$ to order-number row $\boldsymbol{X}$ results in $\boldsymbol{Y}\boldsymbol{X}^{-1}\boldsymbol{X}$, that is the order-number row $\boldsymbol{Y}$). Therefore BADNESS OF SERIAL FIT is a measure of the complexity of the order-number operation that maps one row into the other.

This helps in determining the distribution of BADNESS OF SERIAL FIT: having first produced the distribution of the values for some row we can apply the same distribution to another row by simply relabeling the pitch classes.[9]

I will defer further analysis of the transformational interpretation of BADNESS OF SERIAL FIT to Chapter 13 in which I have more technical language at my disposal.

Logarithmic values of BADNESS OF SERIAL FIT

David Lewin suggests using logarithmic values for BADNESS OF SERIAL FIT.

> For various technical reasons, I suspect that the logarithms of these numbers would provide an even better measure, both intuitively and in light of what seem to me to be some interesting information-theoretic implications. But at the present time, I am nowhere near working out this matter to my own satisfaction. (Lewin 1976, 256)

Let us define the new similarity measure LOGARITHMIC BADNESS OF SERIAL FIT, or LOGBSF, simply as logarithmic values of BADNESS OF SERIAL FIT. Thus,

$$\mathit{LOGBSF}(X, Y) = \log(\mathit{BSF}(X, Y)).$$

For these purposes the base of the logarithm could be any real number greater than 1, but below I provide some arguments for choosing 2.

I would not like to second-guess the rationale for using the logarithmic values of BADNESS OF SERIAL FIT that Lewin had in mind. However, at least two reasons can be found: the issues of the metric and distribution.

[9] In fact, this strategy of relabeling the elements is used by Kenny Wong and Frank Ruskey in the implementation of an algorithm devised by Gara Pruesse and Frank Ruskey (1997) to calculate the number of linear extensions of a partial order – which is the mathematical equivalent of BADNESS OF SERIAL FIT (see Chapter 13).

As discussed above, BADNESS OF SERIAL FIT does not define a metric for two reasons: the value of two identical rows is not zero and triangle inequality does not hold. Using logarithmic values would solve both of these problems. First, the value of a row measured against itself is 1. Independently of what we choose as the base of the logarithm we obtain $\log 1 = 0$ and therefore requirement (i) of the metric is satisfied. Secondly, the following theorem by Alexander Sidorenko (1992, 339–340) can be used to prove that triangle inequality holds for the logarithmic values.[10]

THEOREM 9.1 (SIDORENKO) If the incomparability graph of a partial order P can be covered by the incomparability graphs of partial orders $P_1, P_2, \ldots, P_k$, then

$$e(P) \leq e(P_1)e(P_2)\cdots e(P_k).$$

Proof. See Sidorenko (1992). □

Here $e(P)$ denotes the number of linear extensions of partial order P.

In order to apply Sidorenko's theorem to the current setting, let us take $k = 2$ and simply note that the incomparability graph of $X \cap Z$ is covered by the incomparability graphs of $X \cap Y$ and $Y \cap Z$, thus giving the following corollary:

COROLLARY 9.1 If X, Y and Z are three linear orders on the same set, then the inequality

$$e(X \cap Z) \leq e(X \cap Y)e(Y \cap Z)$$

holds.

Let us now examine the triangle inequality for LOGARITHMIC BADNESS OF SERIAL FIT in more detail. We obtain the following inequality from Corollary 9.1:

$$BSF(X, Y) \cdot BSF(Y, Z) \geq BSF(X, Z).$$

Since the logarithm is a monotonously ascending function and the BADNESS OF SERIAL FIT values are positive, we can take logarithms on both sides of the inequality, and thereby obtain the following inequality:

$$\log(BSF(X, Y) \cdot BSF(Y, Z)) \geq \log(BSF(X, Z)).$$

By applying the rules of logarithms we then obtain the following inequality:

$$\log(BSF(X, Y)) + \log(BSF(Y, Z)) \geq \log(BSF(X, Z)).$$

[10] I am indebted to Graham Brightwell for directing me to this article.

However, since the LOGARITHMIC BADNESS OF SERIAL FIT values are simply logarithms of the BADNESS OF SERIAL FIT values we can write the above inequality as follows:

$$LOGBSF(X,Y) + LOGBSF(Y,Z) \geq LOGBSF(X,Z).$$

Therefore triangle inequality holds for LOGARITHMIC BADNESS OF SERIAL FIT.

These inequalities concerning LOGARITHMIC BADNESS OF SERIAL FIT also give us a better understanding of the BADNESS OF SERIAL FIT values. Namely, triangle inequality holds for BADNESS OF SERIAL FIT if the binary operation is not an addition but a multiplication. It also gives us an estimation of how its values behave.

Let us now return to the example of three-note rows in Figure 9.4: if we use addition as the binary operation triangle inequality fails since $2 + 3 < 6$, but if we use multiplication it holds since $2 \cdot 3 \geq 6$.

The second reason for using logarithmic values concerns their distribution. Of course, the values are simply scaled values of the "ordinary" BADNESS OF SERIAL FIT: scaling them using logarithms does not, in a sense, give us any new information. However, we get a better perspective by using the logarithmic values. As shown in Figure 9.5, the distribution of values in BADNESS OF SERIAL FIT is extremely skewed, while the distribution of the logarithmic values creates a centered curve that resembles the bell curve.

The logarithmic values also level off the BADNESS OF SERIAL FIT scale. It turns out that the set of possible values at the upper end of the scale is very sparse. Let us consider the upper end of the scale and use row $P = 5409728136\text{AB}$ in Figure 7.3 as an example. Row P and its retrograde RP are maximally dissimilar, therefore

$$BSF(P, RP) = BSF(5409728136\text{AB}, \text{BA}6318279045) = 479001600.$$

Let us now switch the positions of the adjacent pitch classes 4 and 5 in row RP and label the resulting row BA6318279054 as RP'. For the original row P and the retrograde of the modified row RP',

$$BSF(P, RP') = BSF(5409728136\text{AB}, \text{BA}6318279054) = 239500800.$$

The enormous difference between the values $BSF(P, RP) = 479001600$ and $BSF(P, RP') = 239500800$ would imply that rows RP and RP' are very different – at least with respect to row P. However, this is not the case since the only difference between them is the order of the adjacent pitch classes 4 and 5. Using the logarithmic values scales this difference. Using logarithm to base 2 produces the values $LOGBSF(P, RP) \approx 28.84$ and $LOGBSF(P, RP') \approx 27.84$. The difference between them is precisely 1. Incidentally, $LOGBSF(RP, RP') = 1$, and it shows why 2 was selected as the base: the smallest possible change that can be made to a row is to flip the order of two adjacent pitch classes. The original row and the row with the flipped dyad together then define a protocol that they satisfy but no other row does. Hence, the BADNESS OF SERIAL FIT value for these two rows is 2. Using logarithm to base 2 the logarithm of 2 is 1 and, hence, the LOGARITHMIC BADNESS OF SERIAL FIT value for these two rows is 1.

Incidentally, these three rows provide yet another illustration of triangle inequality, since using multiplication instead of addition as the binary operation produces

$$BSF(P, RP') \cdot BSF(RP', RP) = 239500800 \cdot 2 = 479001600 = BSF(P, RP),$$

and therefore triangle inequality holds for the logarithmic values.

I argued at the end of Section 3.6.2 that scaling may distort the natural meaning of a similarity measure. While this is also true in the case of BADNESS OF SERIAL FIT, the metric property and the bell-shaped distribution of the logarithms of the BADNESS OF SERIAL FIT values speak for making an exception in this case.

BADNESS OF SERIAL FIT for row classes

Like the previous row measures, BADNESS OF SERIAL FIT is transformationally coherent, consequently BADNESS OF SERIAL FIT for row classes is well defined. The equation $BSF(X, Y) = BSF(FX, FY)$ clearly holds for all row operations: transpositions, inversion, retrograde and their combinations. As with the previous measures defined in the order-number realm (DISPLACEMENT and ORDER INVERSIONS), BADNESS OF SERIAL FIT is transformationally coherent under all pitch-class operations, such as the M-operation, since with respect to the protocol defined by the rows they constitute only a relabeling of the pitch classes. Finally, retrograde-related rows have the same shared pairs of pitch classes as the original forms, but reversed: if pitch class x precedes pitch class y in both rows X and Y, then pitch class y precedes pitch class x in both rows RX and RY. Therefore $BSF(X, Y) = BSF(RX, RY)$. Thus, BADNESS OF SERIAL FIT is transformationally coherent but as it does not define a metric for rows it does not define a metric for row classes either.

In general, BADNESS OF SERIAL FIT is not transformationally coherent under order-number operations – retrograde is the only exception. For example, it is not transformationally coherent with respect to (non-trivial) rotations. Let us consider the row $P = 5409728136\text{AB}$ in Figure 7.3, its modification $P' = 4509728136\text{AB}$ in which the positions of the two first pitch classes have been exchanged and the rotations of both rows $r_{11}P = 409728136\text{AB}5$ and $r_{11}P' = 509728136\text{AB}4$. A comparison of of BADNESS OF SERIAL FIT values for the two rotated rows gives

$$BSF(P, P') = 2 \neq 132 = BSF(r_{11}P, r_{11}P'),$$

and therefore, BADNESS OF SERIAL FIT is not transformationally coherent under rotation. The transformational coherence is broken since the rotations "wrap around." Pitch classes 4 and 5 are adjacent in P and P' but non-adjacent in $r_{11}P$ and $r_{11}P'$. If the rows were treated as circular entities, BADNESS OF SERIAL FIT would be transformationally coherent under the rotations, but then the order of the pitch classes would not be defined.

We can use a similar argument to that used in Section 8.3 to show that BADNESS OF SERIAL FIT is not transformationally coherent under any other order-number operation than retrograde. If, for example, order-number operation $\boldsymbol{F}$ changes the unordered order-number interval between order numbers $\boldsymbol{p_0}$ and $\boldsymbol{p_1}$, it will change the number of pitch classes between the pitch classes at those order positions. Consequently, we can construct the following order-number rows:

$$\boldsymbol{X} = \boldsymbol{p_0p_1p_2p_3p_4p_5p_6p_7p_8p_9p_{10}p_{11}} \quad \text{and} \quad \boldsymbol{Y} = \boldsymbol{p_1p_0p_2p_3p_4p_5p_6p_7p_8p_9p_{10}p_{11}}.$$

The BADNESS OF SERIAL FIT value for rows $\boldsymbol{X}$ and $\boldsymbol{Y}$ and for rows $\boldsymbol{FX}$ and $\boldsymbol{FY}$ both depend only on the unordered order-number interval between order numbers $\boldsymbol{p_0}$ and $\boldsymbol{p_1}$. Since order-number operation $\boldsymbol{F}$ changes that unordered order-number interval, the BADNESS OF SERIAL FIT value for rows $\boldsymbol{X}$ and $\boldsymbol{Y}$ is not the same as its value for rows $\boldsymbol{FX}$ and $\boldsymbol{FY}$. Hence, BADNESS OF SERIAL FIT is not transformationally coherent under $\boldsymbol{F}$.

BADNESS OF SERIAL FIT is not transformationally coherent under the exchange operation either. Let rows $P = 5409728136\text{AB}$ and $r_{11}P = 409728136\text{AB}5$ be as above. The exchange transformations of these two rows are $EP = 2758109463\text{AB}$ and $Er_{11}P = 16470\text{B}83529\text{A}$. A comparison of the BADNESS OF SERIAL FIT values for rows P and $r_{11}P$ and their transformations using the exchange operation gives

$$BSF(P, r_{11}P) = 12 \neq 36032 = BSF(EP, Er_{11}P),$$

which confirms the observed dissimilarity.

The distribution of BADNESS OF SERIAL FIT

While the BADNESS OF SERIAL FIT values range from 1 to 479001600, the number of distinct values is considerably smaller, at 569573, but is several orders of magnitude larger than in the other twelve-tone row similarity measures.

It was noted above that the distribution of BADNESS OF SERIAL FIT values is extremely skewed. Let us now examine this in more detail. In general, at the lower end the values are densely and at the higher end sparsely distributed. On the one hand, all the integer values between 1 and 105946 are present, but on the other hand, while the largest is 479001600 the second largest value is only 239500800 – half of the maximum value. Of the values less than one in a hundred is larger than 1013760, and less than one in a thousand is larger than 3592512.

Figure 9.5 shows the distributions of BADNESS OF SERIAL FIT and LOGARITHMIC BADNESS OF SERIAL FIT (the values have been rounded and scaled to the interval $[0, 100]$ in order to facilitate comparison). When we compare values we must certainly take the scale into account. For example, it makes a significant difference whether the BADNESS OF SERIAL FIT value is 1 or 2: if it is 1 we have identical rows and if it is 2 one pair of adjacent pitch classes is reversed. However, it arguably makes no difference whether it is 1791647 or 1791648. Yet, in both cases the difference between them is 1 (that is, $|1 - 2| = 1 = |1791647 - 1791648|$). Hence, the resolution of BADNESS OF

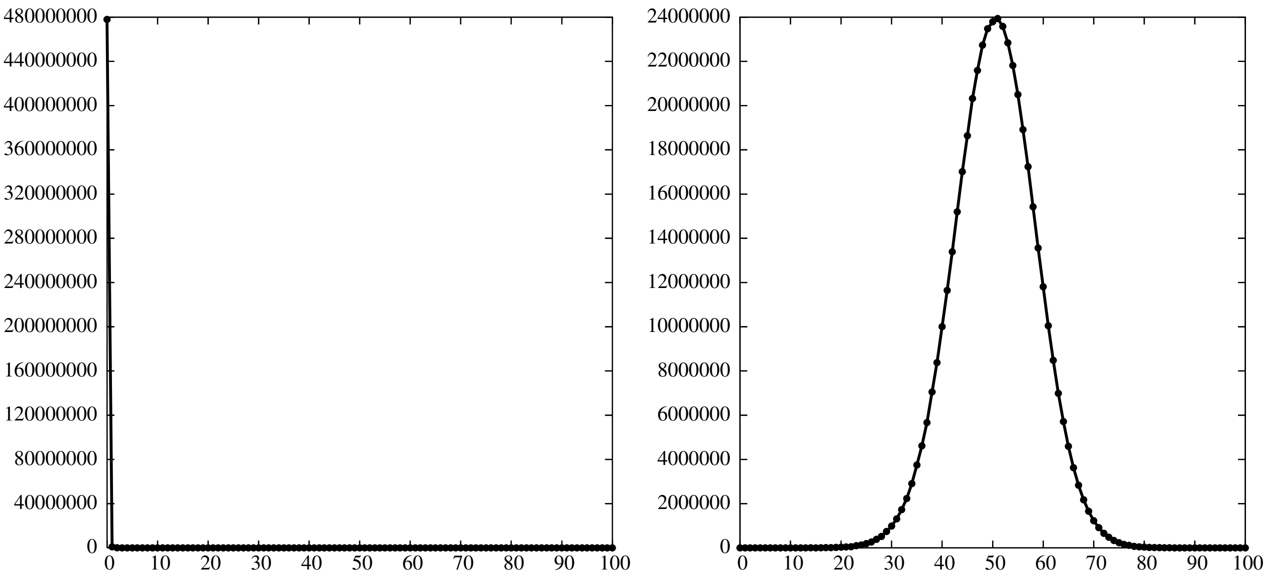


Figure 9.5: The distribution of BADNESS OF SERIAL FIT scaled and rounded to the interval $[0, 100]$ on the left, and the distribution of LOGARITHMIC BADNESS OF SERIAL FIT scaled and rounded to the interval $[0, 100]$ on the right. (The distribution of BADNESS OF SERIAL FIT goes almost along the axes and is therefore difficult to discern in the picture.) Both distributions were obtained by computing the distances defined by all 479001600 distinct transformations, and hence they are precise.

SERIAL FIT is considerably finer than a human can perceive.

It is rather extraordinary that David Lewin defines a similarity measure in his article, but does not give a single non-trivial example of calculating the BADNESS OF SERIAL FIT of two rows. The only examples he gives are a row measured against itself resulting in the value 1, and a row measured with its retrograde resulting in the value 479001600. The same applies to Daniel Starr and Robert Morris: Starr gives only the trivial values 1 and 479001600 (Starr 1984, 189), and Morris gives no examples (Morris 2001). John Ward enumerates the values of his own scaled variant of BADNESS OF SERIAL FIT for segments of sizes 2 to 6, and remarks that “there are limits to the feasibility of BADNESS OF SERIAL FIT in the large cardinalities, when potentially hundreds of millions of permutations must be examined” (Ward 1992, 100). Given the more technical vocabulary at my disposal in Chapter 13, I will show that computing the BADNESS OF SERIAL FIT of two arbitrary rows is a very difficult a task, but with an effective algorithm, it is nowhere close to being as hopeless as Ward implies. In particular, we certainly do not need to examine “hundreds of millions of permutations” if we have an efficient algorithm at our disposal.

9.4 Transformations in the ordered-pairs approach

The notion of a twelve-tone row as a set of ordered pairs is very suggestive with respect to transforming one row into another. A row can be gradually transformed into any other row by exchanging pitch classes successively. In the process rows gradually become more and more similar to the target row and more and more dissimilar to the original row. The pitch classes that are exchanged may be adjacent or non-adjacent. I will first discuss exchanging adjacent pitch classes in Section 9.4.1,

P:	5	4	0	9	7	2	8	1	3	6	10	11
	5	4	0	9	7	2	8	1	6	3	10	11
	5	4	0	9	7	2	8	6	1	3	10	11
	5	4	0	9	7	2	6	8	1	3	10	11
	5	4	0	9	7	6	2	8	1	3	10	11
	5	4	0	9	6	7	2	8	1	3	10	11
	5	4	0	6	9	7	2	8	1	3	10	11
	5	4	0	6	9	2	7	8	1	3	10	11
	5	4	0	6	9	2	8	7	1	3	10	11
	5	4	0	6	9	2	8	1	7	3	10	11
	5	4	0	6	9	2	8	1	3	7	10	11
	5	4	0	6	9	2	1	8	3	7	10	11
	5	4	0	6	9	1	2	8	3	7	10	11
S:	5	4	0	6	9	1	8	2	3	7	10	11

Figure 9.6: The transformation of row $P = 5409728136\mathrm{AB}$ into row $S = 5406918237\mathrm{AB}$ by exchanging adjacent pitch classes.

and then exchanging non-adjacent pitch classes in Section 9.4.2. The latter leads to a new similarity measure, CAYLEY DISTANCE, which I will discuss in detail in Section 9.5.

9.4.1 Exchanging adjacent pitch classes

Figure 9.6 shows a sequence in which row P in Figure 7.3 is gradually transformed into row S in Figure 7.3 by exchanging adjacent pitch classes. At the top of the figure is row $P = 5409728136\mathrm{AB}$, the second row $5409728163\mathrm{AB}$ is obtained by exchanging the adjacent pitch classes 3 and 6, the third row $5409728613\mathrm{AB}$ is obtained by exchanging the adjacent pitch classes 1 and 6, and so on. After thirteen exchanges we obtain the bottom row $S = 5406918237\mathrm{AB}$.

The order in which the adjacent pitch classes are exchanged is not unequivocal. For example, the first exchange in Figure 9.6 is between pitch classes 3 and 6 in row P, but it might as well be between pitch classes 7 and 2, for example.

The value of the ORDER INVERSIONS similarity measure for rows $P = 5409728136\mathrm{AB}$ and $S = 5406918237\mathrm{AB}$ is 13. Indeed, since in the process depicted in Figure 9.6 the number of order inversions with respect to the original row P is increased by one, it is obvious that the value of the similarity measure is precisely the number of exchanges of adjacent pitch classes needed to transform one row into another.

The idea of exchanging adjacent pitch classes is connected to the well-known sorting algorithm known as *bubble sort*: elements "bubble up" to their proper positions.[11] It can be proved that bubble sort is not particularly effective.[12] In terms of transformations of twelve-tone rows inefficiency means that the sequences in which the rows are transformed into other rows using exchanges of adjacent pitch classes tend to be lengthy. This is reflected in the high average value of ORDER INVERSIONS – on average it takes as many as 33 exchanges of adjacent pitch classes to transform one row into another.

I will return to these transformations in Chapter 13 in which graph theory is used to facilitate further analysis.

9.4.2 Exchanging non-adjacent pitch classes

Allowing only the exchange of adjacent pitch classes makes the transformation process slow. Exchanging non-adjacent pitch classes provides a potentially considerably quicker sequence of transformations. Indeed, in the case of rows $P = 5409728136\mathrm{AB}$ and $Q = 5406728139\mathrm{AB}$ we only need to exchange pitch classes 6 and 9 in order to transform row P into row Q. Since the value of the ORDER INVERSIONS similarity measure for rows P and Q is 11 we obtain with a single exchange of non-adjacent pitch classes the same result that would require 11 exchanges of adjacent pitch classes.

In terms of similarity, the adjacent exchanges introduce a smaller change in each step than the non-adjacent exchanges. However, transforming row P into row Q via a sequence of 11 adjacent exchanges is a lengthy process. Consequently, with respect to the similarity of the "endpoints" of the process, we must weight the non-adjacency of the exchange and the length of the transformational process. If the number of adjacent exchanges is particularly large compared to the non-adjacent exchanges, then the latter might be preferable. As an extreme example, it takes six exchanges to retrograde a row, whereas it takes as many as 66 adjacent exchanges.

In general, in order to discover a sequence of transformations that transform one row into another by exchanging (possibly non-adjacent) pitch classes, we need to examine the *cycle structure* of the operation that performs the transformation.[13] We could consider either the pitch-class operation or the order-number operation: the choice is immaterial, and in both cases we end up performing similar exchanges. Let us choose the pitch-class operation here, only because in that case we can describe the process in terms of the more familiar pitch classes.

It will be convenient to distinguish between an exchange of two elements and an operation in which two elements are exchanged, termed an *exchange*.[14] For easy reference, the formal definition

[11] For an introduction to the bubble sort algorithm and an analysis of its properties see, for example, Section 5.2.2 in Knuth (1998).

[12] In terms of computational complexity, the bubble sort algorithm is $O(N^2)$. This means that the average number of steps in it is proportional to the square of the number of elements. In the case of twelve-tone rows, the average number is $\frac{1}{2} \cdot \binom{12}{2} = 33$. If we were to double the length of the row, the average number of steps would grow to $\frac{1}{2} \cdot \binom{24}{2} = 138$, which is more than four times the average number in the case of twelve-tone rows.

[13] See Section A.3 in Appendix A for a discussion about cycles.

[14] Exchanges are usually called transpositions in the mathematical literature. I use the term "exchange" here since the word transposition is already in use.

of an exchange is given below.

DEFINITION 9.4 An exchange is a permutation that has one cycle of length 2 and the rest of the cycles are of length 1.

The exchange of two elements defines the cycle of length 2, and each of the cycles of length 1 denotes a fixed point.

For example, the permutation

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 0 & 1 & 2 & 3 & 4 & 5 & 9 & 7 & 8 & 6 & 10 & 11 \end{pmatrix}$$

exchanges two elements – 6 and 9 – and therefore it is an exchange. We can write this permutation conveniently as a single cycle $(6\ \ 9)$ – omitting (but implying) the cycles of length 1.

A well-known result in mathematics is that permutations can always be written as a product of exchanges.[15] For example, the permutation

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 0 & 2 & 1 & 3 & 4 & 5 & 7 & 9 & 8 & 6 & 10 & 11 \end{pmatrix}$$

can be written as a product of three exchanges: $(1\ \ 2)(7\ \ 9)(6\ \ 9)$.[16] Composing exchanges is no different from composing cycles in general; we only need to be careful always to proceed from the rightmost cycle towards the leftmost cycle. It should be noted that as the cycles are not distinct, the order is significant. If we reverse the order of the cycles $(7\ \ 9)$ and $(6\ \ 9)$ in the product $(1\ \ 2)(7\ \ 9)(6\ \ 9)$ so that our product of exchanges becomes $(1\ \ 2)(6\ \ 9)(7\ \ 9)$, we obtain a different permutation.

This product of exchanges is not a unique decomposition of the permutation into exchanges.[17] The product of exchanges $(1\ \ 2)(6\ \ 9)(6\ \ 7)$ results in the same permutation as the product of exchanges $(1\ \ 2)(7\ \ 9)(6\ \ 9)$. In fact, there is an arbitrary number of products of exchanges that result in a given permutation since we can always take one pair of pitch classes and flip them back and forth at will. However, there is always a minimum number of exchanges in each of these products. For example, the above permutation cannot be expressed with less than three exchanges. If a product of exchanges has the minimum number of exchanges we call it the shortest product of exchanges.

Let us consider rows $P = 5409728136\text{AB}$ and $S = 5406918237\text{AB}$ in Figure 7.3. According to the calculation performed in Section 7.2.3, the transformation that transforms row P into row S is $int(P, S) = SP^{-1} = 0213457986\text{AB}$. This transformation could be written as a product of cycles

[15] See, for example, Theorem 6 in Section 1.4 of Nicholson (1999).

[16] I would remind the reader that permutations are composed from right to left.

[17] As stated earlier, the decomposition of a permutation into disjoint cycles is unique except for the order of the cycles. However, as the cycles are not disjoint, the decomposition is not unique.

$$
\begin{array}{rcccccccccccc}
P\colon & 5 & 4 & 0 & 9 & 7 & 2 & 8 & 1 & 3 & 6 & 10 & 11 \\
 & 5 & 4 & 0 & 9 & 6 & 2 & 8 & 1 & 3 & 7 & 10 & 11 \\
 & 5 & 4 & 0 & 6 & 9 & 2 & 8 & 1 & 3 & 7 & 10 & 11 \\
S\colon & 5 & 4 & 0 & 6 & 9 & 1 & 8 & 2 & 3 & 7 & 10 & 11
\end{array}
$$

Figure 9.7: The transformation of row $P = 5409728136\mathrm{AB}$ into row $S = 5406918237\mathrm{AB}$ by exchanging (not necessarily adjacent) pitch classes.

as follows:

(9.4) $$(0)(1\ \ 2)(3)(4)(5)(6\ \ 7\ \ 9)(8)(10)(11) = (0)(1\ \ 2)(3)(4)(5)(6\ \ 9)(6\ \ 7)(8)(10)(11).$$

We can transform row $P = 5409728136\mathrm{AB}$ into row $S = 5406918237\mathrm{AB}$ with three exchanges. Let us begin with row P. Exchanging pitch classes 6 and 7 results in row $5409628137\mathrm{AB}$, then exchanging pitch classes 9 and 6 results in row $5406928137\mathrm{AB}$, and finally exchanging pitch classes 1 and 2 results in row $S = 5406918237\mathrm{AB}$. This process is illustrated in Figure 9.7.

Let us now derive a formula for counting the minimum number of exchanges that together generate a transformation. I will begin with two basic observations. First, a permutation can always be decomposed into a product of disjoint cycles. This decomposition is unequivocal (up to the ordering of the cycles (see Section A.3 in Appendix A). Secondly, a cycle of length n can be decomposed into a product of $n-1$ exchanges. For example, we could write the cycle $(a_1\ \ a_2\ \ldots\ a_{n-1}\ \ a_n)$ as a product of $n-1$ cycles $(a_1\ \ a_n)(a_1\ \ a_{n-1})\ldots(a_1\ \ a_3)(a_1\ \ a_2)$. Hence, we only need to decompose the transformation into disjoint cycles in order to be able to derive the minimum number of exchanges that together generate a transformation. Let us summarize these observations in the form of the following lemma.

LEMMA 9.2 The minimum number of exchanges that together generate transformation F is

$$\sum(|F_i| - 1)$$

in which F_i are cycles of F and $|F_i|$ denotes the length of the cycle.

This procedure was used in Formula 9.4 to create the sequence of exchanges that transforms row P into row S. Cycle $(1\ \ 2)$ is an exchange and cycle $(6\ \ 7\ \ 9)$ can be decomposed into a product of two exchanges.

The idea of exchanging pitch classes is connected to the family of sorting algorithms that could be given the moniker *sorting by exchanging*.[18] However, with sorting algorithms we usually do not know beforehand what elements are already in their correct positions, and consequently the pitch

[18]For a review of sorting algorithms based on exchanging see, for example, Section 5.2.2 in Knuth (1998).

classes that are already in their correct positions may be moved in the sorting process. Nevertheless, we could conceive of the process depicted in Figure 9.7 as the sorting of the pitch classes that are not in their correct positions.

9.5 CAYLEY DISTANCE

The transformational procedure for exchanging non-adjacent pitch classes leads directly to a method for measuring the similarity of twelve-tone rows: I call this CAYLEY DISTANCE because it is equivalent to a similarly named metric for permutations in mathematics, and will abbreviate it to C. Since this approach involves the decomposition of a transformation into exchanges, it is, obviously, a transformational similarity measure. Hence, I give the similarity measure only in transformational terms. Let us begin by defining CAYLEY DISTANCE formally in terms of the GIS of pitch-class rows.

DEFINITION 9.5 In the GIS of pitch-class rows, the value of the CAYLEY DISTANCE similarity measure for twelve-tone rows X and Y is the smallest number of exchanges in the decompositions of transformation $int(X, Y) = YX^{-1}$ into exchanges.

In order to calculate the CAYLEY DISTANCE value for rows X and Y we must first decompose pitch-class transformation YX^{-1} into exchanges (see Section A.3 in Appendix A); the CAYLEY DISTANCE value is then the number of exchanges.

We could also define CAYLEY DISTANCE in terms of the GIS of order-number rows, which would entail examining the order positions of the pitch classes that need to be exchanged and not the actual pitch classes. Both definitions result in the same similarity measure.

Let us utilize the same examples as in Section 9.4.2. First, the pitch-class transformation that maps row $P = 5409728136\text{AB}$ into row $Q = 5406728139\text{AB}$ is $QP^{-1} = 0123459786\text{AB}$. This transformation can be written as a single cycle $(6\ \ 9)$; consequently, the CAYLEY DISTANCE value for rows P and Q is 1. Secondly, the pitch-class transformation that maps row $P = 5409728136\text{AB}$ into row $S = 5406918237\text{AB}$ is $SP^{-1} = 0123459786\text{AB}$, and can be written as a product of three cycles $(1\ \ 2)(7\ \ 9)(6\ \ 9)$; consequently, the CAYLEY DISTANCE value for rows P and S is 3.

The minimum CAYLEY DISTANCE value is 0, denoting maximal similarity. It was proved in Lemma 9.2 that the minimum number of exchanges needed to transform one row into another is the sum of the lengths of the cycles in the transformation minus the number of cycles.[19] Therefore, the maximum CAYLEY DISTANCE value is obtained if the transformation contains only one cycle that is of length 12, in which case the value is 11. For example, pitch-class transformation T_1 contains one cycle that is of length 12. Since for any row A and T_1A, $C(A, T_1A) = 11$, two rows belonging to the same row class can be maximally dissimilar according to CAYLEY DISTANCE.

It is possible derive the following symmetry property of CAYLEY DISTANCE from these observations. First, as noted at the end of Section 5.4, the transformation that maps row X into row Y is the inverse of the transformation that maps row Y into row X. Secondly, note that the inverse

[19]Since each cycle F_i adds $|F_i| - 1$ to the sum, the CAYLEY DISTANCE value becomes the sum of the lengths of the cycles minus the number of cycles.

permutations have the same cycle structure.[20] Therefore, as CAYLEY DISTANCE is based on the cycle structure of transformations, it is the same from row X to row Y as from row Y to row X.

It is straightforward to show that CAYLEY DISTANCE defines a metric. First, the values are positive real values. Secondly, the four requirements of the metric are satisfied. (i) Trivially, the value of $C(X, X)$ is 0 for all rows: no exchanges are needed to keep a row unchanged. (ii) As discussed above, CAYLEY DISTANCE is symmetric and $C(X, Y) = C(Y, X)$. (iii) Triangle inequality $C(X, Z) \leq C(X, Y) + C(Y, Z)$ holds since if $C(X, Y) = m$ and $C(Y, Z) = n$, then it is possible to transform row X into row Y with a succession of m exchanges, and row Y into row Z with a succession of n exchanges. Consequently, by applying first the succession of m exchanges and then the succession of n exchanges, it is possible to transform row X into row Z with a succession of $n + m$ exchanges. (iv) Finally, if $C(X, Y) = 0$, no exchanges at all are needed transform row X into row Y, and therefore the two rows must be identical. Since all four requirements are satisfied, CAYLEY DISTANCE defines a metric.

CAYLEY DISTANCE for row classes

Again, in order to show that CAYLEY DISTANCE for row classes is well defined, we must show that it is transformationally coherent (see Corollary 6.1 in Section 6.2). The equation $C(X, Y) = C(FX, FY)$ clearly holds for all row operations: transpositions, inversion, retrograde, and their combinations. In fact, we can show that CAYLEY DISTANCE is transformationally coherent under all pitch-class operations and all order-number operations. First, all pitch-class operations are transformationally coherent since, with respect to CAYLEY DISTANCE, they constitute only a relabeling of the pitch classes. Hence, the cycle structures of pitch-class transformations $int(X, Y)$ and $int(FX, FY)$ are identical for all pitch-class operations F. Secondly, as noted above, CAYLEY DISTANCE could be defined in terms of order-number transformations. Therefore, it is transformationally coherent under all order-number operations are since, with respect to CAYLEY DISTANCE, they constitute only a relabeling of the order positions. Hence, the cycle structures of order-number transformations $int(\boldsymbol{X}, \boldsymbol{Y})$ and $int(\boldsymbol{FX}, \boldsymbol{FY})$ are identical for all order-number operations $\boldsymbol{F}$. Thus, as CAYLEY DISTANCE is a transformationally coherent metric under transposition, inversion, and retrograde, according to Corollary 6.1, it defines a metric for row classes.

Furthermore, it is transformationally coherent even under the exchange operation. In order to show this we need to examine the cycle structures of permutations. I have borrowed the following definition of the *type* of permutation from Bóna (2004, 79–80).

DEFINITION 9.6 Let p be a permutation of n elements with a_i cycles of length i. Then we say that p is of type $(a_1, a_2, \ldots, a_n)$.

The type of permutation is handy shorthand for describing the cycle structure of permutations.

[20]In other words, if permutation F has the cycle $(a_1\ a_2 \ldots a_{n-1}\ a_n)$ then its inverse permutation F^{-1} has the cycle $(a_n\ a_{n-1} \ldots a_2\ a_1)$.

For example, permutation

$$(0)(1\ \ 2)(3)(4)(5)(6\ \ 7\ \ 9)(8)(10\ \ 11)$$

is of type $(5, 2, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0)$ since it has five cycles of length 1, two cycles of length 2, and one cycle of length 3.

Two permutations σ and τ of the permutation group S_n are called *conjugates* if there exists permutation π such that $\sigma = \pi\tau\pi^{-1}$. We can use the conjugacy relation to prove the claim about transformational coherence under the exchange operation. Namely, the following lemma, borrowed from Bóna (2004, 80–81), links the conjugacy of permutations to their cycle structures.

LEMMA 9.3 Elements g and h of S_n are conjugates in S_n if and only if they are of the same type.

Proof. See Bóna (2004). □

If for a moment we could interpret the permutations representing rows as mappings instead of linear orderings, the exchange operation transforms row X into row X^{-1} and row Y into row Y^{-1}. The transformation that maps row X^{-1} into row Y^{-1} is $Y^{-1}X$, and we need to show that it is of the same type as transformation YX^{-1} that transforms row X into row Y.

First, note that $Y^{-1}X$ is of the same type as its inverse $(Y^{-1}X)^{-1} = X^{-1}Y$. Secondly, the equation

$$Y(X^{-1}Y)Y^{-1} = YX^{-1}(YY^{-1}) = YX^{-1}$$

shows that the inverse of $Y^{-1}X$ is a conjugate of YX^{-1} and therefore of the same type. The transformation $Y^{-1}X$ that maps row X^{-1} into row Y^{-1} is thus of the same type as the transformation YX^{-1} that transforms row X into row Y. Since the CAYLEY DISTANCE value depends on the cycle structure, that is on the type of transformation, CAYLEY DISTANCE is transformationally coherent even under the exchange operation.

The distribution of CAYLEY DISTANCE

Let us now turn to the distribution of values in CAYLEY DISTANCE. The distribution can be calculated by utilizing some well-known properties of permutations. I have borrowed the following lemma, giving the number of n-permutations of a given type, from Bóna (2004, 79–80).

LEMMA 9.4 If $(a_1 \cdot 1) + (a_2 \cdot 2) + \ldots (a_n \cdot n) = n$, then the number of n-permutations of type $(a_1, a_2, \ldots, a_n)$ is

$$\frac{n!}{a_1!a_2!\cdots a_n!1^{a_1}2^{a_2}\cdots n^{a_n}}.$$

Proof. See Bóna (2004). □

Now it is relatively straightforward to calculate the distribution of CAYLEY DISTANCE. Cases 0 and 1 are easy since there is only one permutation – the identity permutation – with no exchanges (value 0), and there are $\binom{12}{2} = 66$ ways to select two pitch classes for one exchange. Case 2 is

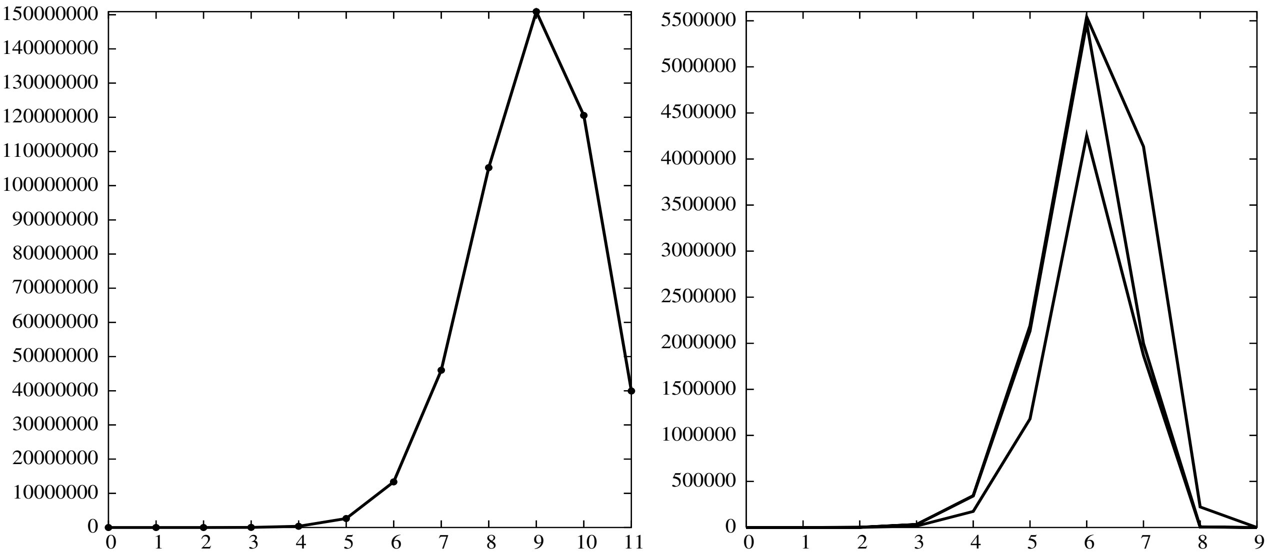


Figure 9.8: The distribution of CAYLEY DISTANCE on the left and the distribution of CAYLEY DISTANCE for row classes on the right. The former is obtained by computing the distances defined by all 479001600 distinct transformations, and hence it is precise; the latter is obtained by comparing 2000 random row classes to all other row classes. See Section 6.3 for a discussion on the three curves in the latter figure.

slightly more complicated since we need to consider two subcases: the permutation may be of type $(8, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$ (two cycles of length 2) or it may be of type $(9, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0)$ (one cycle of length 3); in both cases the total number of exchanges is 2.

According to Lemma 9.4, there are

$$\frac{12!}{8! \cdot 2! \cdot 1^8 \cdot 2^2} = 1485$$

permutations of type $(8, 2, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$ and

$$\frac{12!}{9! \cdot 1! \cdot 1^9 \cdot 3^1} = 440$$

permutations of type $(9, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0)$. Taking the two together amounts to $1485 + 440 = 1925$ transformations, giving a CAYLEY DISTANCE of 2. A similar procedure gives the other values. The distribution of CAYLEY DISTANCE is given in Figure 9.8. The average value is 8.9 and the mean is 9.

CAYLEY DISTANCE versus DERANGEMENT

At this point it would be useful to compare CAYLEY DISTANCE to DERANGEMENT. Both of these similarity measures have very similar scales and distributions, and both are transformationally coherent under all pitch-class operations and all order-number operations. While they both have closely related values, they are by no means identical. Consider the following two pitch-class transformations

$$(0)(1)(2)(3)(4)(5)(6\ 7\ 8)(9\ 10\ 11) \text{ and } (0)(1)(2)(3)(4)(5)(6\ 7)(8\ 9)(10\ 11).$$

Both have six fixed points; with the DERANGEMENT similarity measure, two rows related by either one of these transformations would have the value of 6. However, the former is a composition of four exchanges and the latter is a composition of three exchanges. Similarly, the two transformations

$$(0)(1)(2)(3)(4)(5)(6)(7)(8)(9\ 10\ 11) \text{ and } (0)(1)(2)(3)(4)(5)(6)(7)(8\ 9)(10\ 11)$$

can both be decomposed into two exchanges, but the former has a DERANGEMENT value of 3 and the latter has one of 4. Therefore, the CAYLEY DISTANCE value cannot be determined based on the DERANGEMENT value, and vice versa.

It is easy to see that the DERANGEMENT value is always equal to or greater than the CAYLEY DISTANCE value – the equality occurs only at the identity transformation. Enumerating all of the transformations and comparing the CAYLEY DISTANCE and the DERANGEMENT values produces, on average, a difference between the values of 2.1.

Two rows related by pitch-class transformation $(0\ 1)(2\ 3)(4\ 5)(6\ 7)(8\ 9)(10\ 11)$ constitute a case of maximum difference between DERANGEMENT and CAYLEY DISTANCE: it contains only six exchanges and the CAYLEY DISTANCE value is 6, yet the position of every pitch class is changed and therefore the DERANGEMENT value is 12.

Both the CAYLEY DISTANCE and the DERANGEMENT distributions are skewed. The average of the DERANGEMENT distribution is 11, which means that we can expect two random rows to have, on average, only one pitch class in the same order position. On the other hand, the average of the CAYLEY DISTANCE distribution is 8.9, which means that we can expect it to take approximately eight or nine exchanges to transform one random row into another.

9.6 CORRELATION COEFFICIENT

The CORRELATION COEFFICIENT similarity measure or CC was introduced by Robert Morris (1987, 120). As its name implies, it is borrowed from statistics, and is described in virtually all introductory books on statistics. Morris adapts the statistical measure to describe the degree to which the order numbers of two segments correlate.

Morris states his definition in transformational terms even if he does not discuss similarity in terms of transformations. Namely, his formula employs $\mathrm{OM_XY}$-type entities that denote the permutation that rearranges the pitch classes of segment Y so that the result is segment X. In the terminology of this study, $\mathrm{OM_XY}$ equals $int(\boldsymbol{Y}, \boldsymbol{X})$ in the GIS of order-number rows.

The following definition of CORRELATION COEFFICIENT is in the form in which Morris gives it, except that the length of segments is fixed at 12 and it is adapted to the notation for order-number transformations employed here.

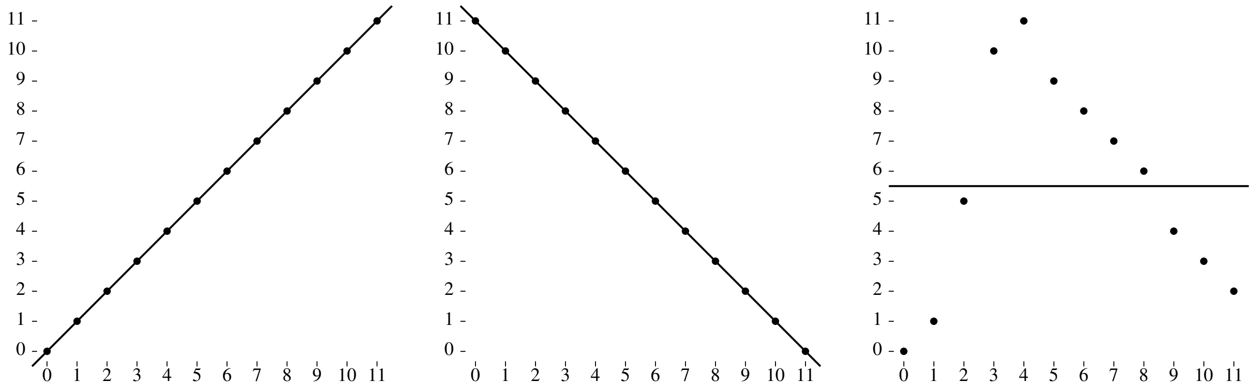


Figure 9.9: The first figure on the left depicts the correlation of order-number rows $\mathbf{0123456789AB}$ and $\mathbf{0123456789AB}$. Since the rows are identical, the correlation is maximal and the CORRELATION COEFFICIENT value is the maximum value 1. The line depicting the correlation is ascending. The middle figure depicts the correlation of order-number rows $\mathbf{0123456789AB}$ and $\mathbf{BA9876543210}$. Since the rows are retrograde related, the negative correlation is maximal and the CORRELATION COEFFICIENT value is the minimum value -1. The line depicting the correlation is descending. The figure on the right depicts the correlation of order-number rows $\mathbf{0123456789AB}$ and $\mathbf{015AB9876432}$. There is neither negative nor positive correlation between the rows and the CORRELATION COEFFICIENT value is 0. The line depicting the correlation is horizontal.

DEFINITION 9.7 The value of the CORRELATION COEFFICIENT similarity measure for twelve-tone rows X and Y is given by the formula

$$CC(X,Y) = FSUM(int(\mathbf{X},\mathbf{X}), int(\mathbf{Y},\mathbf{X}))/FSUM(int(\mathbf{X},\mathbf{X}), int(\mathbf{X},\mathbf{X}))$$

in which

$$FSUM(X,Y) = \left(\sum_{n=0}^{11} x_n \cdot y_n\right) - \frac{1}{12} \cdot \left(\sum_{n=0}^{11} n\right)^2.$$

In our case, the definition could be simplified. For example, for twelve-tone rows the value of $FSUM(int(\boldsymbol{X},\boldsymbol{X}), int(\boldsymbol{X},\boldsymbol{X}))$ is a constant since for any order-number row $\boldsymbol{X}$

$$FSUM(int(\boldsymbol{X},\boldsymbol{X}), int(\boldsymbol{X},\boldsymbol{X})) = \left(\sum_{n=0}^{11} n^2\right) - \frac{1}{12} \cdot \left(\sum_{n=0}^{11} n\right)^2 = 506 - 363 = 143.$$

In addition, including the identity transformation $int(\boldsymbol{X},\boldsymbol{X})$ as a parameter of the function $FSUM$ is redundant and we obtain the following streamlined formula for CORRELATION COEFFICIENT:

$$CC(\boldsymbol{X},\boldsymbol{Y}) = \frac{1}{143} \cdot \left(\sum_{n=0}^{11} n \cdot g_n\right) - \frac{363}{143} \tag{9.5}$$

in which g_n is the nth element of transformation $int(\mathbf{X},\mathbf{Y}) = \mathbf{Y}\mathbf{X}^{-1}$ interpreted as an integer.

The CORRELATION COEFFICIENT does not fit into our dichotomy of similarity measures and dissimilarity measures. It can display both positive and negative correlation and its values range from -1 to 1. The value 1 denotes maximal (positive) correlation and -1 denotes negative correlation. In our case, rows with maximal negative correlation are related by retrograde.

Let us consider an example. It was established in Section 7.2.3 that the order-number transformation that transforms row $P = 5409728136\mathrm{AB}$ into row $Q = 5406728139\mathrm{AB}$ is $\mathbf{0129456783AB}$. According to Formula 9.5 the CORRELATION COEFFICIENT value for rows P and Q is

$$\frac{0{\cdot}0+1{\cdot}1+2{\cdot}2+3{\cdot}9+4{\cdot}4+5{\cdot}5+6{\cdot}6+7{\cdot}7+8{\cdot}8+9{\cdot}3+10{\cdot}10+11{\cdot}11}{143}-\frac{363}{143}$$
$$=\frac{470}{143}-\frac{363}{143}=\frac{107}{143}\approx 0.75,$$

which denotes a relatively strong correlation.

Figure 9.9 illustrates the correlation measurement for three pairs of rows. In the first figure on the left rows $0123456789\mathrm{AB}$ and $0123456789\mathrm{AB}$ are identical, and the CORRELATION COEFFICIENT value is the maximum value of 1. In the second figure in the middle rows $0123456789\mathrm{AB}$ and $\mathrm{BA}9876543210$ are retrograde related and the CORRELATION COEFFICIENT value is the minimum value of -1. In the third figure on the right the value for rows $0123456789\mathrm{AB}$ and $015\mathrm{AB}9876432$ is 0. The points in the figures denote the order numbers and how they are transformed: the horizontal axis denotes the former and the vertical axis denotes the latter. For example, the order number **0** in the middle figure depicting rows $0123456789\mathrm{AB}$ and $\mathrm{BA}9876543210$ is transformed into $\mathbf{11}$, and therefore there is a point $(0, 11)$. The CORRELATION COEFFICIENT value is the slope of the line that has the smallest distance from all twelve points (according to the method known as least squares fitting). The distance is zero in the left and middle figures and positive in the one on the right.

I have acknowledged the usefulness of the metric in the analysis of similarity measures, but given the nature of CORRELATION COEFFICIENT, it is not useful here. Let me emphasize that this in no way diminishes the usefulness of the CORRELATION COEFFICIENT.

The distribution of the CORRELATION COEFFICIENT

Figure 9.10 depicts the distribution of the CORRELATION COEFFICIENT. It is perfectly symmetric owing to the equation

$$CC(X, Y) = -CC(X, RY). \tag{9.6}$$

The fact that this equation holds is easy to understand if we consider the dots and lines in Figure 9.9. Namely, the dots in a figure depicting rows X and Y would be “retrograded” in one depicting rows X and RY. For example, if there are dots

$$(0, y_0), (1, y_1), \ldots, (10, y_{10}), (11, y_{11})$$

in a figure depicting rows X and Y, then there would be dots

$$(0, y_{11}), (1, y_{10}), \ldots, (10, y_1), (11, y_0)$$

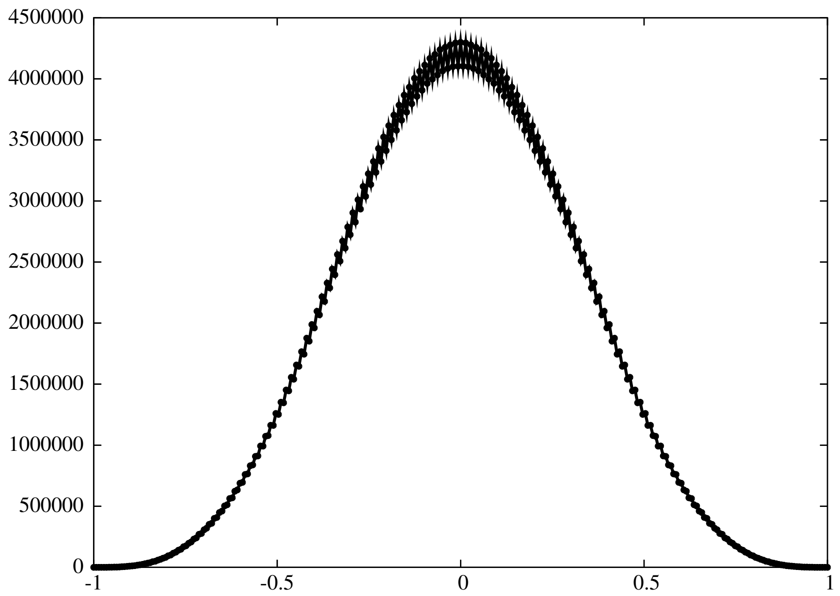


Figure 9.10: The distribution of the CORRELATION COEFFICIENT. It is obtained by computing the distances defined by all 479001600 distinct transformations, and hence it is precise.

in one depicting rows X and RY. Hence, if the dots define an ascending line (denoting a positive correlation of some degree) for rows X and Y, then they define a descending line (denoting a negative correlation of some degree) for rows X and RY.

The distribution of CORRELATION COEFFICIENT values resembles the bell curve. Hence, given a row, there are very few rows in which the order of pitch classes correlates positively or negatively with the given row, and most of the rows show neither positive nor negative correlation.

The CORRELATION COEFFICIENT and ordered pairs

It is not obvious how to classify the conception of a row that the CORRELATION COEFFICIENT similarity measure suggests. I have placed it in this chapter on ordered pairs, the justification being its correlation with the other similarity measures based on the conception of a row as a set of ordered pairs.

A comparison between this and all other similarity measures reveals varying degrees of correlation. Since the CORRELATION COEFFICIENT is based on the comparison of the order relations of a row it is obvious that it correlates only with measures that are also based on such a comparison. No significant correlation was found between the CORRELATION COEFFICIENT and the similarity measures based on subsegments, such as ULAM'S DISTANCE or DIVISIONS (see Sections 10.4 and 10.6, respectively). Hence, the two remaining options are the similarity measures based on the vector approach and those based on ordered pairs.

Figure 9.11 depicts the correlations of the CORRELATION COEFFICIENT and DISPLACEMENT similarity measures and of the CORRELATION COEFFICIENT and ORDER INVERSIONS similarity measures. In both cases there is correlation. However, in the latter case it seems to be stronger than in the former. In particular, the retrograde-related symmetry of both measures is reflected in their

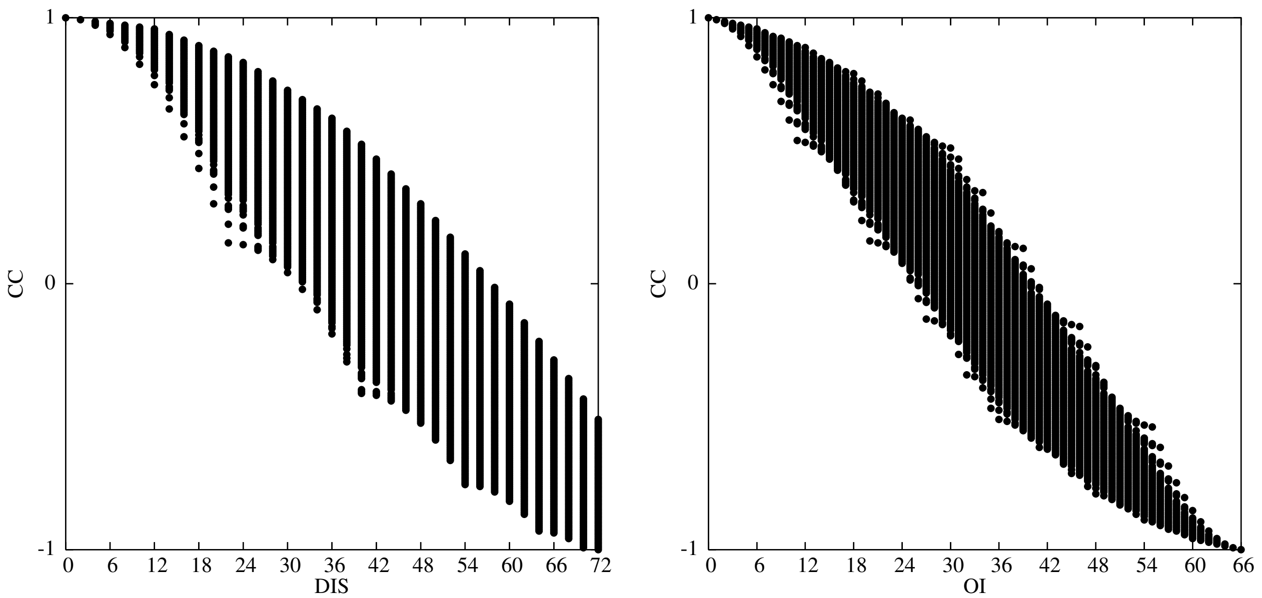


Figure 9.11: The correlation of the CORRELATION COEFFICIENT and DISPLACEMENT similarity measures on the left, and of the CORRELATION COEFFICIENT and ORDER INVERSIONS measures on the right.

distributions (compare Equation 9.3 on page 136 to Equation 9.6 on page 157).

Figure 9.12 shows the ORDER INVERSIONS and CORRELATION COEFFICIENT values for 67 rows compared to row $0123456789\mathrm{AB}$. The rows are arranged to begin with row $0123456789\mathrm{AB}$ and to end with its retrograde $\mathrm{BA}9876543210$. Consecutive rows differ only by one adjacent pair. The rows form a sequence in which one row is transformed step by step into its retrograde through the exchange of adjacent pitch classes. The ORDER INVERSIONS value increases by one at each step (compared to row $0123456789\mathrm{AB}$), as the CORRELATION COEFFICIENT value decreases. Figure 9.13 depicts the correlation of the CORRELATION COEFFICIENT and ORDER INVERSIONS values for the rows in Figure 9.12. Note that the correlation between the CORRELATION COEFFICIENT and ORDER INVERSIONS is strong but not perfect: the dots in Figure 9.13 do not form a straight line.

Finally, Figure 9.9 could be used to relate the CORRELATION COEFFICIENT and ordered pairs. There are twelve dots in each figure (denoting how the order numbers are transformed), and $12 \cdot 11 = 66$ pairs of dots. If a line drawn through a pair of dots is descending, those pairs constitute an order inversion. For example, in the figure on the right the line drawn through dots $(0, 0)$ and $(1, 1)$ is ascending, whereas that drawn through dots $(4, 11)$ and $(5, 9)$ is descending; consequently, order numbers **4** and **5** constitute an order inversion.

On the other hand, the line drawn through any pair of dots in the figure on the left is ascending: there are no order inversions. Correspondingly, the line drawn through any pair of dots in the middle figure is descending: there are 66 order inversions. All lines depicting the correlation will pass through point $(5.5, 5.5)$, thus each of the order inversions "nudge" the line clockwise.[21] The

[21]Point $(5.5, 5.5)$ is the average of the points on both the horizontal and vertical dimensions.

row	OI	CC	row	OI	CC
0123456789AB	0	$1,0000$	4563789AB210	33	$-0,2028$
1023456789AB	1	$0,9930$	4567389AB210	34	$-0,2308$
1203456789AB	2	$0,9790$	4567839AB210	35	$-0,2657$
1230456789AB	3	$0,9580$	4567893AB210	36	$-0,3077$
1234056789AB	4	$0,9301$	456789A3B210	37	$-0,3566$
1234506789AB	5	$0,8951$	456789AB3210	38	$-0,4126$
1234560789AB	6	$0,8531$	546789AB3210	39	$-0,4196$
1234567089AB	7	$0,8042$	564789AB3210	40	$-0,4336$
1234567809AB	8	$0,7483$	567489AB3210	41	$-0,4545$
1234567890AB	9	$0,6853$	567849AB3210	42	$-0,4825$
123456789A0B	10	$0,6154$	567894AB3210	43	$-0,5175$
123456789AB0	11	$0,5385$	56789A4B3210	44	$-0,5594$
213456789AB0	12	$0,5315$	56789AB43210	45	$-0,6084$
231456789AB0	13	$0,5175$	65789AB43210	46	$-0,6154$
234156789AB0	14	$0,4965$	67589AB43210	47	$-0,6294$
234516789AB0	15	$0,4685$	67859AB43210	48	$-0,6503$
234561789AB0	16	$0,4336$	67895AB43210	49	$-0,6783$
234567189AB0	17	$0,3916$	6789A5B43210	50	$-0,7133$
234567819AB0	18	$0,3427$	6789AB543210	51	$-0,7552$
234567891AB0	19	$0,2867$	7689AB543210	52	$-0,7622$
23456789A1B0	20	$0,2238$	7869AB543210	53	$-0,7762$
23456789AB10	21	$0,1538$	7896AB543210	54	$-0,7972$
32456789AB10	22	$0,1469$	789A6B543210	55	$-0,8252$
34256789AB10	23	$0,1329$	789AB6543210	56	$-0,8601$
34526789AB10	24	$0,1119$	879AB6543210	57	$-0,8671$
34562789AB10	25	$0,0839$	897AB6543210	58	$-0,8811$
34567289AB10	26	$0,0490$	89A7B6543210	59	$-0,9021$
34567829AB10	27	$0,0070$	89AB76543210	60	$-0,9301$
34567892AB10	28	$-0,0420$	98AB76543210	61	$-0,9371$
3456789A2B10	29	$-0,0979$	9A8B76543210	62	$-0,9510$
3456789AB210	30	$-0,1608$	9AB876543210	63	$-0,9720$
4356789AB210	31	$-0,1678$	A9B876543210	64	$-0,9790$
4536789AB210	32	$-0,1818$	AB9876543210	65	$-0,9930$
			BA9876543210	66	$-1,0000$

Figure 9.12: The ORDER INVERSIONS and CORRELATION COEFFICIENT values for 67 rows compared to row 0123456789AB. Each consecutive pair of rows differ only by one adjacent dyad.

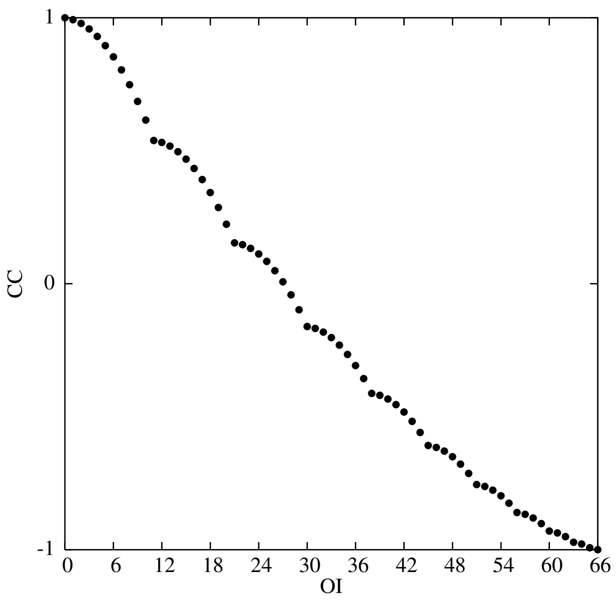


Figure 9.13: The ORDER INVERSIONS and CORRELATION COEFFICIENT values for the rows in Figure 9.12.

amount by which each inversion nudges the line depends on the position of the points representing the order inversion with respect to it. The CORRELATION COEFFICIENT could be described as a "summary" of these nudges.

CHAPTER 10

Similarity measures based on subsegments

The focus of the discussion in this chapter is on the measurement of the similarity of twelve-tone rows based on the conception of a twelve-tone row as a set of subsegments, which may be contiguous or non-contiguous. The concept is introduced in Section 10.1, and the six similarity measures based on it are discussed in more detail in the subsequent Sections 10.2, 10.3, 10.4, 10.5, 10.6, and 10.7. Finally, in Section 10.8 I will consider some transformational procedures that this approach suggests.

The subsegments of the first four measures discussed in this chapter could be non-contiguous. We could define variants of each of them that would require that them to be contiguous, but with the fifth and sixth measures, however, it is essential that the subsegments are contiguous.

10.1 The twelve-tone row as a set of subsegments

The notion of ordered pairs was discussed in Chapter 9. An ordered pair is a subsegment of a row of length 2. Hence, it is a natural avenue of development to consider the evaluation of the row similarity based on subsegments of a length other than 2. However, even if similarity measures based on ordered pairs have been historically significant, those based on larger segments are almost nonexistent. In fact, only one author, John Ward (1992), has explicitly formulated a similarity measure based on the subsegments (other than ordered pairs). In addition, in the context of combination matrices Daniel Starr and Robert Morris (1978) introduced the idea of *fragmentation*, their only requirement being that the rows are divided into segments with identical unordered pitch-class contents. In Section 10.7 I will develop their method of measuring the amount of fragmentation in a similarity measure based on the shared contiguous subsegments of rows.

Composers have used shared subsegments to relate twelve-tone rows. A well-known example is provided by the two rows in the fifth and sixth movements of Alban Berg's *Lyric Suite* (rows S and T in Figure 7.3), as the latter is composed of two subsegments of the former (and vice versa).

In addition, this process is explicitly demonstrated in the music (Headlam 1996, 282). Rows $E = 6958473\text{B}2\text{A}10$, $I_0E = 63748591\text{A}2\text{B}0$, $RE = 01\text{A}2\text{B}3748596$, and $RI_0E = 0\text{B}2\text{A}19584736$ of Anton Webern's song *Erlösung* op. 18 no. 2 provide another example. These rows are the only four row forms that Webern uses in the song, and they are very closely associated since rows E and RI_0E share the subsegments 958473 and B2A1, and rows I_0E and RE share the retrogrades of these subsegments: 374859 and 1A2B. Furthermore, these subsegments are contiguous in all rows.

A twelve-tone row contains a total of $2^{12} = 4096$ subsegments of different lengths: 66 ordered dyads, 220 subsegments of length 3, 495 subsegments of length 4, etc. In general, a row contains $\binom{12}{n}$ subsegments of length n. The set of 4096 subsegments includes the null subsegment of length 0 and the twelve trivial subsegments of length 1.

The number of contiguous subsegments is considerably smaller than the number of all subsegments. A twelve-tone row contains 11 contiguous subsegments of length 2, 10 contiguous subsegments of length 3, etc. The total number of contiguous subsegments is 79, which include the null segment of length 0 and the twelve trivial subsegments of length 1.

On the one hand, we could argue that as the number of subsegments decreases, their significance increases. For example, it is easier to observe in two rows that both contain the contiguous dyad FE than that in both rows pitch class F precedes pitch class E, and it is naturally easier to keep track of 11 contiguous dyads than of 66 dyads. On the other hand, it is easy to highlight shared subsegment contents in a composition even if the subsegments are not contiguous – the fifth and sixth movements of Alban Berg's *Lyric Suite* are a prime example.

There are a variety of approaches to the similarity of rows based on their subsegment contents. These include focusing on subsegments of some fixed length, or of any length, and finding the longest shared subsegment. Each approach has its merits and the values of the similarity measures based on them do not necessarily always coincide.

The sole focus in all of the approaches based on subsegments is on the mutual order of the pitch classes in them. A row is conceived of as a repository of segments. Again, the absolute positions of the pitch classes are immaterial, and we ignore the intervals between those that are adjacent. Naturally, we could derive the positions of pitch classes and other properties of rows but they are not the focus of this conception.

10.2 GENERALIZED ORDER INVERSIONS

The GENERALIZED ORDER INVERSIONS similarity measure was invented by the author. As its name implies, it is a generalization of the ORDER INVERSIONS similarity measure. As ORDER INVERSIONS is based on a comparison of the ordered dyads in two rows, a natural generalization would be to base the comparison on the ordered n-tuples in two rows. The n-tuples do not need to be contiguous.

There are at least two reasons for extending ORDER INVERSIONS to cover larger segments. First, if a piece of music is predominantly based on triplet segments, for example, it could give some insight into the rows of such a piece through consideration of their triplet contents. Secondly, since each row contains 66 ordered pairs, but 220 ordered triplets and 495 ordered quadruplets, etc., similarity measures based on the longer segments might give a better resolution. I will return to the issue of resolution in the discussion on the distribution of GENERALIZED ORDER INVERSIONS values below.

There is a fundamental difference between the ordered dyads and longer segments, however. We know that if a row does not contain an ordered dyad (i, j) then it must contain the retrograde (j, i). No such relation holds for the longer segments. For example, three pitch classes a, b, and c appear in six different orderings: (a, b, c), (a, c, b), (b, a, c), (b, c, a), (c, a, b), or (c, b, a). Therefore, if a row does not contain the segment (a, b, c) we cannot conclude that it contains the retrograde; we can only conclude that it contains one of the five other possible segments.

If a row does not contain the segment $(p_1, p_2, \ldots, p_n)$, then the order of at least one pair of the pitch classes is inverted. Hence, we are counting subsegments of some fixed length that contain at least one order inversion.

Let us define a family of similarity measures termed GENERALIZED ORDER INVERSIONS. In order to emphasize the genealogy of this family let us we label them OI_n, where n denotes the length of the segment. Under this definition, OI equals to OI_2. For the sake of completeness, let us include the segment lengths 0 and 1 even if the similarity measures based on them turn out to be trivial.

DEFINITION 10.1 If X and Y are twelve-tone rows and $0 \leq n \leq 12$, then the number of different tuples of length n in the two rows is given by the formula

$$OI_n(X, Y) = \#\{(a_1, a_2, \ldots, a_n) \mid (a_1, a_2, \ldots, a_n) \in X \text{ and } (a_1, a_2, \ldots, a_n) \notin Y\}$$

in which $a_1, a_2, \ldots, a_n$ are distinct pitch classes.

Thus the value of $OI_n(X, Y)$ is the number of subsegments $(a_1, a_2, \ldots, a_n)$ in row X that are not subsegments of row Y. The procedure for calculating the value of OI_n for rows X and Y is first to enumerate the subsegments of row X of length n and then to count the number of the subsegments that are not subsegments of row Y.

Let us consider GENERALIZED ORDER INVERSIONS and rows $P = 5409728136\text{AB}$ and $Q = 5406728139\text{AB}$ in Figure 7.3. It was shown in Section 9.2 that $OI(P, Q) = 11$. If we increase the length of the segments under scrutiny, the number of non-shared segments increases (except for length 12 – there is only one segment of length 12 in each row). For the segments of length 3, $OI_3(P, Q) = 80$: there are 80 subsegments of length 3 in row P that are not subsegments of row Q. For example, row P contains segment $(0, 9, 7)$ but row Q does not. For the segments of length 4, $OI_4(P, Q) = 265$. For example, row P contains segment $(0, 9, 7, 2)$ but row Q does not.

It is straightforward to show that GENERALIZED ORDER INVERSIONS defines a metric for all segment lengths $1 < n \leq 12$, and a pseudometric for segment lengths 0 and 1. First, the values are positive real values. Secondly, the four requirements of the metric are satisfied. (i) Trivially, the value of $OI_n(X, X)$ is 0 for all rows X, and there are no different subsegments between two identical rows. (ii) OI_n is symmetric since the number of segments in row X that are not in row Y equals the number of segments in row Y that are not in row X, and hence $OI_n(X,Y) = OI_n(Y,X)$. (iii) Triangle inequality $OI_n(X,Y)+OI_n(Y,Z) \geq OI_n(X,Z)$ holds, which can be shown as follows. If there are k segments that are in row X and not in row Y, then there are $\binom{12}{n} - k$ segments that rows X and Y share, whereas if there are l segments that are in row Y and not in row Z, then there are $\binom{12}{n} - l$ segments that rows Y and Z share. Consequently, rows X and Z share at least $\binom{12}{n} - k - l$ segments. This means that there are at most $k + l$ segments in row X that are not in row Z, and therefore triangle inequality holds. (iv) Finally, for segment lengths $1 < n \leq 12$, if $OI_n(X,Y) = 0$ then rows X and Y do not have any different segments so they must be identical. Since all four requirements of a metric are satisfied for segment lengths $1 < n \leq 12$, GENERALIZED ORDER INVERSIONS defines a metric for $1 < n \leq 12$, and since the three first requirements of a metric are satisfied for segment lengths 0 and 1, it also defines a pseudometric for $n = 0$ and $n = 1$.

The minimum value of GENERALIZED ORDER INVERSIONS is 0, denoting maximal similarity, and the maximum value is $\binom{12}{n}$ for segment lengths $n \geq 2$. For segment lengths 0 and 1 the maximum value is 0 – the same as the minimum value. Retrograde-related rows are maximally dissimilar for any segment length. Since for any row A and RA, $OI_n(A, RA) = \binom{12}{n}$ for segment lengths $n \geq 2$, two rows belonging to the same row class may be maximally dissimilar according to GENERALIZED ORDER INVERSIONS. For segment lengths $n \neq 2$ any row is maximally dissimilar to several other rows. For segment lengths $n > 2$, it follows from the above that if a row does not contain subsegment $(p_1, p_2, \ldots, p_n)$ it does not necessarily contain the retrograde. Hence, there are multiple ways of not having the same subsegments. Furthermore, somewhat paradoxically, for segment lengths 0 and 1 any row is both maximally dissimilar *and* maximally similar to all rows.

In the case of ordered dyads we could prove that if the ORDER INVERSIONS value for rows X and Y is n, then for rows X and RY it is $66 - n$. In other words, with this measure if rows X and Y are dissimilar, then rows X and RY would be similar and vice versa. No such property holds for GENERALIZED ORDER INVERSIONS if the length of the segment is longer than 2. Let us consider the OI_3 similarity measure and rows $C = 0123456789\mathrm{AB}$, $D = 643210\mathrm{BA}9875$, and $RD = 5789\mathrm{AB}012346$. The value of OI_3 for rows C and D is 220, which denotes maximal dissimilarity. Rows C and RD, however, are also very dissimilar since the value of OI_3 is 180. Hence, even if we know the value of OI_3 for rows X and Y we cannot predict its value for rows X and RY.

The transformational approach to GENERALIZED ORDER INVERSIONS

Since GENERALIZED ORDER INVERSIONS is a generalization of ORDER INVERSIONS, which is a transformational similarity measure, it requires only a small exercise in formal notation to show that these similarity measures have a transformational interpretation.

The concept of permutation inversion was applied in the definition of ORDER INVERSIONS in transformational terms (see Definition 9.2). Segments that contain at least one permutation inversion play the same role in GENERALIZED ORDER INVERSIONS as permutation inversions play in ORDER INVERSIONS. In more colloquial terms, we are now considering segments in which not all of the elements appear in their "natural" order. Of course, this notion (as in the case of ORDER INVERSIONS) is applied to the transformations between twelve-tone rows and not to the twelve-tone rows themselves.

DEFINITION 10.2 In the GIS of order-number rows, the value of the OI_n similarity measure for twelve-tone rows X and Y is

$$OI_n(X,Y) = \#\{(a_1, a_2, \ldots, a_n) \mid \text{ where } a_k > a_{k+1} \text{ for some } k\}$$

in which $(a_1, a_2, \ldots, a_n)$ is a segment of length n in transformation $int(\boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{Y}\boldsymbol{X}^{-1}$.

The OI_n similarity measure is thus defined as the number of segments of length n in order-number transformation $int(\boldsymbol{X}, \boldsymbol{Y})$ that contain at least one permutation inversion.[1]

Let us consider similarity measure OI_3 and rows $P = 5409728136\mathrm{AB}$ and $Q = 5406728139\mathrm{AB}$ in Figure 7.3. It was established in Section 7.2.3 that the order-number transformation in the GIS of order-number rows that transforms row P into row Q is $\mathbf{0129456783AB}$. We now need to find all segments of length 3 in transformation $\mathbf{0129456783AB}$ that contain at least one permutation inversion. For example, triplet $(1, 9, 4)$ contains one permutation inversion (since $9 > 4$), and triplet $(9, 6, 3)$ contains three (since $9 > 6$, $6 > 3$, and $9 > 3$). There are a total of 80 segments of length 3 in transformation $\mathbf{0129456783AB}$ that have at least one permutation inversion, hence $OI_3(P, Q) = 80$.

A similar argument as in the case of ORDER INVERSIONS shows that the transformational definition 10.2 of GENERALIZED ORDER INVERSIONS is equivalent to the non-transformational definition 10.1.

GENERALIZED ORDER INVERSIONS for row classes

In order to show that GENERALIZED ORDER INVERSIONS for row classes is well defined, we must show that it is transformationally coherent (see Corollary 6.1 in Section 6.2). It is transformationally coherent since the equation $OI_n(X, Y) = OI_n(FX, FY)$ clearly holds for all row operations: transpositions, inversion, retrograde, and their combinations. For transposition T_k, for every segment $(x_1, x_2, \ldots, x_n)$ in row X that is not a segment of row Y there is a corresponding segment $(T_k(x_1), T_k(x_2), \ldots, T_k(x_n))$ in row T_kX that is not a segment of row T_kY (and vice versa). Similarly, for inversion I_k, for every segment $(x_1, x_2, \ldots, x_n)$ in row X that is not a segment of row Y there is a corresponding segment $(I_k(x_1), I_k(x_2), \ldots, I_k(x_n))$ in row I_kX that is not a segment of row I_kY (and vice versa). Finally, for retrograde, for every segment $(x_1, x_2, \ldots, x_n)$ in row X that is not a segment of row Y there is a corresponding segment $(x_n, \ldots, x_2, x_1)$ in row RX that is not a segment of row RY (and vice versa). Thus, GENERALIZED ORDER INVERSIONS is a transformationally coherent metric and, according to Corollary 6.1, it defines a metric for row classes.

[1]Note that if $(a_1, a_2, \ldots, a_n)$ contains a permutation inversion it contains one in which the elements involved are adjacent.

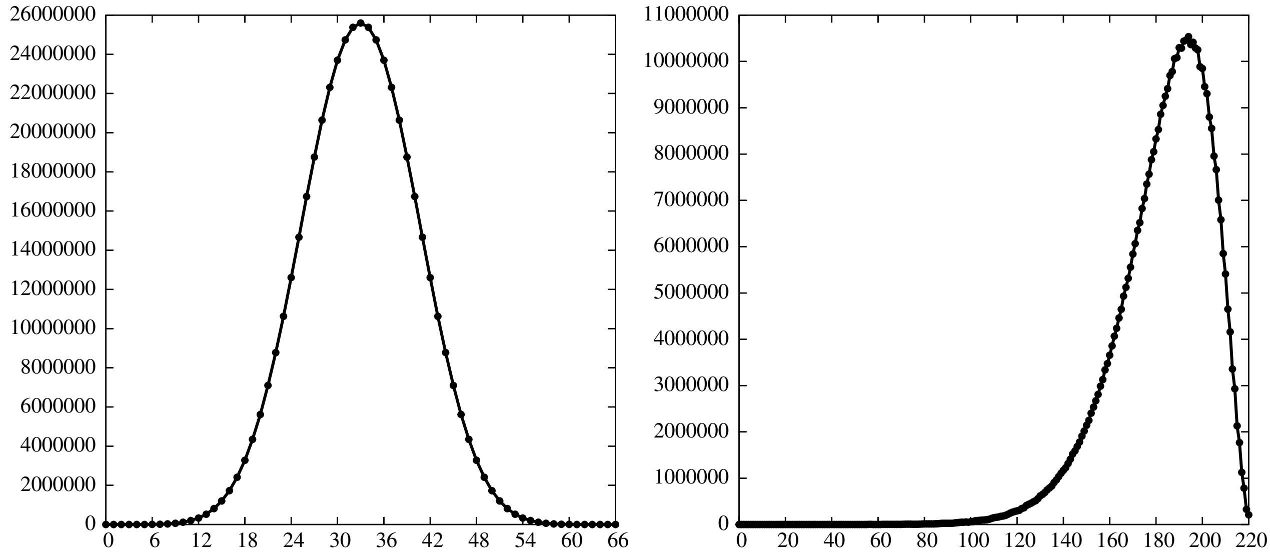


Figure 10.1: The distribution of the ORDER INVERSIONS similarity measure on the left and of GENERALIZED ORDER INVERSIONS with the segment length 3 on the right. The former is symmetrical while the latter is clearly skewed to the right.

The distribution of GENERALIZED ORDER INVERSIONS

The GENERALIZED ORDER INVERSIONS values obtained above for the sample rows tend to be large compared to the ORDER INVERSIONS value, for example – even accounting for the scales of the measures. Indeed, perhaps the largest disadvantage of similarity measures based on segments longer than a dyad is their skewed distribution. ORDER INVERSIONS has a symmetrical distribution, but that of GENERALIZED ORDER INVERSIONS becomes more and more skewed as the value of n increases. Figure 10.1 shows the distributions of similarity measures OI_2 and OI_3. The skewness increases as the length n of the segments grows until OI_{12} presents an extreme case since it has only two values 0 and 1: consequently, OI_{12} is maximally skewed since it gives the value 1 depicting maximum dissimilarity for all non-identical row pairs. Hence, OI_{12} defines a metric that is known as the *discrete metric*. This type of metric is too trivial to be useful in the comparison of twelve-tone rows.[2] Similarity measures OI_0 and OI_1 go to the other extreme: we consider the null subsegment of length 0 to be a subsegment of every row, and every row contains the twelve possible subsegments of length 1. Both measures thus give the value 0 for any pair of twelve-tone rows and therefore they only define a pseudometric.

In the case of ORDER INVERSIONS there are 67 possible values, from 0 denoting maximal similarity to 66 denoting maximal dissimilarity. Hence, every value between the maximum and the minimum is obtained at least once. In addition, as shown in Section 9.4.1, given two rows X and

[2]The discrete metric constitutes a special case of the metric that is regularly referred to in the mathematical literature (see, for example, Steen and Seebach (1978, 41)). However, this does not make it any more useful for the present purposes.

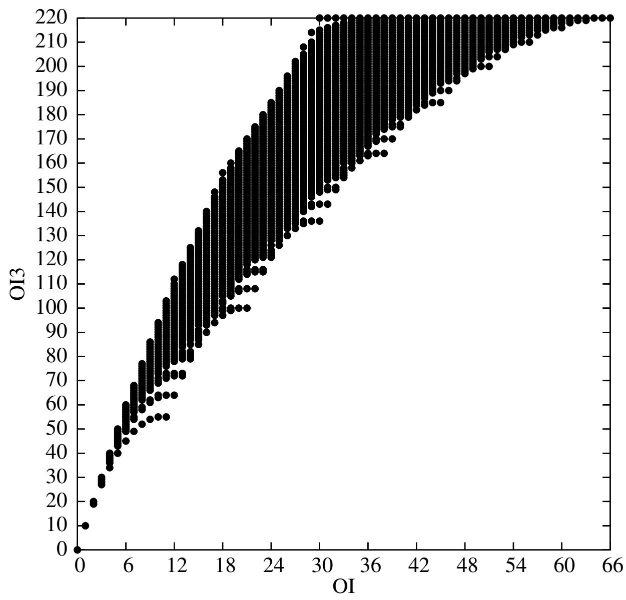


Figure 10.2: The correlation of the values of OI and OI_3. The horizontal axis denotes the OI values and the vertical axis the OI_3 values.

Y, it is always possible to form a sequence of rows

$$X = X_0, X_1, X_2, \ldots, X_{n-1}, X_n = Y$$

in which in each step we increase the number of order inversions with respect to row X by exactly one. Hence, it is possible that two rows differ only by one dyad. No such relation holds for similarity measures based on longer segments. For example, there are 220 ordered triplets but there are only 192 possible values between two rows. The smallest value is, of course, zero, and the smallest non-zero value is 10. Hence, no two rows differ by only one triplet: if two rows are not identical, they contain at least 10 different triplets.

In general, the values are sparse at the low end of the scale and denser at the high end. Therefore, while similarity measures based on segments longer than 2 do increase the resolution, we could argue that the increased resolution occurs in the wrong place: it would be preferable to have a good resolution at the low end of the scale than at the high end. Nevertheless, it is worth noting that a small change in a row results in big change in the triplet contents.

The correlation of the similarity measures OI and OI_3 is depicted in Figure 10.2. The figure shows that the two measures correlate more at the low end of the scale (similar rows) than at the high end (dissimilar rows). In general, we cannot predict with good accuracy the value of OI_3 when we know the value of OI and vice versa. As an extreme case, let us consider rows $C = 0123456789\mathrm{AB}$ and $D = 543210\mathrm{BA}9876$. These two rows share 36 ordered dyads, and 30 ordered dyads are inverted. Hence, according to OI, these two rows are slightly more similar than dissimilar. On the other hand, as these two rows do not share a single ordered triplet, according to OI_3 they are maximally dissimilar. While one might argue that OI_3 does a poor job in evaluating the similarity

	OI	OI_3
value of the similarity measure	11	80
value scaled to the interval $[0, 1]$	$0.1666\ldots$	$0.3636\ldots$
relative value	431886	112007

Table 10.1: Comparison of the absolute, scaled, and relative values of similarity measures OI and OI_3 for rows $P = 5409728136\text{AB}$ and $Q = 5406728139\text{AB}$.

of these rows, we should remember that a similarity measure measures only what it professes to measure: with respect to ordered triplets, rows C and D are maximally dissimilar.

It was established earlier that rows $P = 5409728136\text{AB}$ and $Q = 5406728139\text{AB}$ in Figure 7.3 have 80 different segments of length 3. At first sight, it seems that two rows with that many different segments of length 3 cannot be similar. In particular, it would seem that similarity measure OI captures the similarity of these two measures better since it gives the considerably smaller value of 11 – a value that is smaller even accounting for the different scales of the two measures. However, we obtain misleading results if we do not take into account the *distribution* of the values. Table 10.1 enumerates the absolute, scaled, and relative values of similarity measures OI and OI_3 for rows P and Q. If we scale the values of both measures to the interval $[0, 1]$, and round them to four decimal numbers, we obtain the scaled values of $OI(P, Q) = 0.1667$ and $OI_3(P, Q) = 0.3636$. The scaled value of OI is less than half of the scaled value of OI_3. However, if we take into account the distribution of values in both similarity measures we get the opposite picture. There are 431886 rows that, according to OI, are at most at a distance of 11 from row P, and there are only 112007 rows that, according OI_3, are at most at a distance of 80 from row P. If we take the distribution into account, rows P and Q are more similar according to OI_3 than according to OI. Hence, it might be that, after all, GENERALIZED ORDER INVERSIONS with a segment length of 3 captures the similarity of rows P and Q better than ORDER INVERSIONS.

Finally, it should be noted that similarity measure OI_6 has, at least in theory, the best resolution since the number of segments of length n is $\binom{12}{n}$ and the value of $\binom{12}{n}$ is highest (924) when $n = 6$. Hence, when $n > 6$ the resolution decreases as n increases until the range of OI_{12} encompasses only two possible values. In addition, the skewness of the distribution increases as n increases. Hence, the combination of a poor resolution and a skewed distribution suggests that OI_n is not very useful for larger values of n.

10.3 SUBSEGMENT CONTENT DIFFERENCE

The GENERALIZED ORDER INVERSIONS family of similarity measures was based on the consideration of subsegments of some fixed length n. A natural development then is to consider the subsegments of all lengths simultaneously. John Ward has devised a similarity measure based on this idea, which he describes as follows.[3]

[3]Ward follows Milton Babbitt's convention of using the word "set" for twelve-tone rows and the word "subsets" for subsegments.

> The author has devised the measure common subsets (CSS for segments and CSC for cycles) in response to the observation that, while OI and ORIN [similarity measure OI transformed from a dissimilarity measure into a similarity measure] provide sensitive and accurate measures of similarity, they consider only the ordered dyads shared by two ordered sets. They neglect the ordered subsets of any other cardinality that are also held in common by two ordered sets. (Ward 1992, 146)

Ward defines his measure not as a dissimilarity measure but as a similarity measure (see Section 3.5), which is consistent with his project to redefine existing dissimilarity measures also as similarity measures. As the project here is to analyze the properties of similarity measures in terms of the metric, I will redefine his similarity measure as a dissimilarity measure and rename it SUBSEGMENT CONTENT DIFFERENCE or SCD. First, I prefer the term "subsegment" rather than "subset" because I am referring to ordered entities. Secondly, since in order to demonstrate the relation of this similarity measure to the other similarity measures based on subsegments of rows, it is consistent to define it as a dissimilarity measure.

The GENERALIZED ORDER INVERSIONS family of similarity measures provides a convenient method for deriving SUBSEGMENT CONTENT DIFFERENCE. Since these measures take all subsegments (not only contiguous ones) into consideration, all subsegments are included in SUBSEGMENT CONTENT DIFFERENCE as well. The formal definition is given below.

DEFINITION 10.3 The value of the SUBSEGMENT CONTENT DIFFERENCE similarity measure for twelve-tone rows X and Y is given by the formula

$$SCD(X, Y) = \sum_{n=0}^{12} OI_n(X, Y).$$

Each of the similarity measures in the family returns the number of different subsegments of length n in two rows. Since the length of every subsegment of a row is between 0 and 12 inclusive SUBSEGMENT CONTENT DIFFERENCE is thus defined as the sum of the values of the thirteen similarity measures $OI_0, OI_1, OI_2, \ldots, OI_{12}$. The SUBSEGMENT CONTENT DIFFERENCE value for rows X and Y is thus the number of subsegments in row X that are not subsegments of row Y.

As noted in Section 10.2, a twelve-tone row has $\binom{12}{n}$ segments of length n. Using this information we can count the total number of segments in a twelve-tone row by summing the counts of each possible length, and we arrive at the following well-known formula:

$$\binom{12}{0} + \binom{12}{1} + \binom{12}{2} + \ldots + \binom{12}{10} + \binom{12}{11} + \binom{12}{12} = 4096 = 2^{12}.$$

Nevertheless, the maximum value of SCD is not 4096 but 4083, since 13 subsegments are shared by all 479001600 twelve-tone rows: these are the null segment of length 0 and the twelve segments of length 1 – which corresponds to the observation in Section 10.2 that similarity measures OI_0 and OI_1 return zero for any two rows. The minimum SUBSEGMENT CONTENT DIFFERENCE value is 0, denoting maximal similarity. The maximum value of 4083 is obtained by retrograde-related rows. Since for any row A and RA, $SCD(A, RA) = 4083$, two rows belonging to the same row class may be maximally dissimilar according to SUBSEGMENT CONTENT DIFFERENCE. Furthermore, for any row A, row RA is the only row at the maximum distance from row A.

Let us consider rows $P = 5409728136\mathrm{AB}$ and $Q = 5406728139\mathrm{AB}$ in Figure 7.3. The SUBSEGMENT CONTENT DIFFERENCE value for rows P and Q is 3008. A single exchange of two pitch classes results in a large number of different subsegments – almost three fourths (3008 of the total 4096) of them change due to this exchange. Similarly, if we exchange the first dyad in row $P = 5409728136\mathrm{AB}$ we obtain row $P' = 4509728136\mathrm{AB}$, and the SUBSEGMENT CONTENT DIFFERENCE value for these two rows is 1024: there are $2^{10} = 1024$ different segments that can be created from the last ten pitch classes of row P, and joining any of these with segment 54 of row P results in a segment that is not a subsegment of row P'.

It is straightforward to show that SUBSEGMENT CONTENT DIFFERENCE defines a metric. First, the values are positive real values. Secondly, the four requirements are satisfied. Some of the properties of SUBSEGMENT CONTENT DIFFERENCE follow directly from those of the constituent OI_n similarity measures, but for others we need to devise new proofs. (i) Trivially, the $SCD(X, X)$ value is 0 for all rows since $OI_n(X, X) = 0$ for every n and every row X so the sum is also zero. (ii) SUBSEGMENT CONTENT DIFFERENCE is symmetric since all similarity measures OI_n are symmetric. (iii) Triangle inequality can be shown to hold as follows. If there are k segments that are in row X and not in row Y, then there are $4096 - k$ segments that rows X and Y share. Assume then that there are l segments that are in row Y and not in row Z. Therefore, X and Z share at least $4096 - k - l$ segments. This means that there are at most $k + l$ segments in row X that are not in row Z and therefore triangle inequality holds. Alternatively, we could derive triangle inequality for SUBSEGMENT CONTENT DIFFERENCE from those of the thirteen similarity measures OI_n. (iv) Finally, if $SCD(X, Y) = 0$ then rows X and Y do not have any different segment so they must be identical. Since all four requirements of the metric are satisfied, SUBSEGMENT CONTENT DIFFERENCE defines a metric.

The transformational approach to SUBSEGMENT CONTENT DIFFERENCE

Since SUBSEGMENT CONTENT DIFFERENCE is derived from the family of transformational similarity measures OI_n it is straightforward to construct a transformational definition of it.

DEFINITION 10.4 In the GIS of order-number rows, the value of the SUBSEGMENT CONTENT DIFFERENCE similarity measure for twelve-tone rows X and Y is

$$SCD(X, Y) = \sum_{n=0}^{12} \#\{(a_1, a_2, \ldots, a_n) \mid \text{ where } a_k > a_{k+1} \text{ for at least one } k\}$$

where $(a_1, a_2, \ldots, a_n)$ is a segment of length n in the permutation $int(\boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{Y}\boldsymbol{X}^{-1}$.

The value is thus defined as the number of subsegments in the order-number transformation $int(\boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{Y}\boldsymbol{X}^{-1}$ that contain at least one permutation inversion. This definition is closely related to the transformational definition of the similarity measures OI_n. Hence, the proof that the transformational definition is equivalent to the non-transformational definition 10.3 is a simple corollary to the proof that the transformational formulations of constituent similarity measures OI_n are equivalent to their non-transformational formulations.

In the case of ORDER INVERSIONS we found a very concrete interpretation of the measure as a measurement of the complexity of a transformation that transforms one row into another since it could be described as the number of exchanges of adjacent pitch classes. In the case of SUBSEGMENT CONTENT DIFFERENCE we are no longer on such concrete ground. Nevertheless, the transformational approach has two advantages. First, picking up ascending segments in the transformation $int(\boldsymbol{X}, \boldsymbol{Y})$ may help in the search for the longest shared subsegment of two rows. Secondly, since the measure can be stated in transformational terms we know that every row has precisely the same network of distances to other rows, and hence the space defined by this similarity measure is perfectly symmetrical.

Finally, perhaps SUBSEGMENT CONTENT DIFFERENCE is more useful as an exploration of the network of subsegment relations between twelve-tone rows than as an actual similarity measure. What is most revealing is the fact that even exchanging two adjacent pitch classes changes one fourth of the subsegments.

SUBSEGMENT CONTENT DIFFERENCE for row classes

In order to show that SUBSEGMENT CONTENT DIFFERENCE for row classes is well defined we must show that it is transformationally coherent (see Corollary 6.1 in Section 6.2). SUBSEGMENT CONTENT DIFFERENCE is transformationally coherent since the equation $SCD(X, Y) = SCD(FX, FY)$ clearly holds for all row operations: transpositions, inversion, retrograde, and their combinations. For transposition T_k, for every segment $(x_1, x_2, \ldots, x_n)$ in row X that is not a segment of row Y there is a corresponding segment $(T_k(x_1), T_k(x_2), \ldots, T_k(x_n))$ in row T_kX that is not a segment of row T_kY. Similarly for inversion I_k, for every segment $(x_1, x_2, \ldots, x_n)$ in row X that is not a segment of row Y there is a corresponding segment $(I_k(x_1), I_k(x_2), \ldots, I_k(x_n))$ in row I_kX that is not a segment of row I_kY. Finally, for retrograde, for every segment $(x_1, x_2, \ldots, x_n)$ in row X that is not a segment of row Y there is a corresponding segment $(x_n, \ldots, x_2, x_1)$ in row RX that is not a segment of row RY. Thus, SUBSEGMENT CONTENT DIFFERENCE is a transformationally coherent metric and therefore, according to Corollary 6.1, it defines a metric for row classes.

The distribution of SUBSEGMENT CONTENT DIFFERENCE

As noted in Section 10.2, in the similarity measure OI_n the value distribution is skewed if n is larger than 2. In addition, the skewness increases as the length n of the segments increases. Since SUBSEGMENT CONTENT DIFFERENCE is the sum of the thirteen OI_n values it inherits the same behavior. It gives 1023 distinct values, the average being 3984.65 and the mean 4001. Figure 10.3 shows the distribution of the SUBSEGMENT CONTENT DIFFERENCE values.

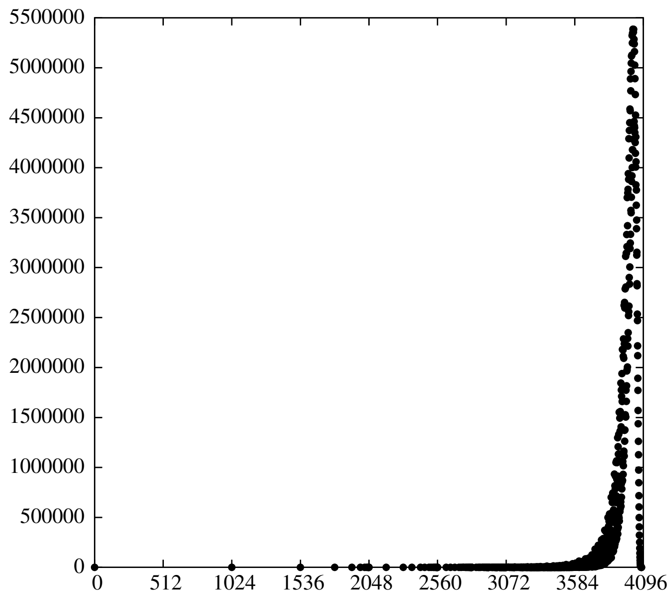


Figure 10.3: The distribution of the SUBSEGMENT CONTENT DIFFERENCE values.

An exchange of two adjacent pitch classes is the smallest change that can be introduced in a twelve-tone row. As noted above, an exchange of two adjacent pitch classes results in 1024 different subsegments. Hence, the smallest non-zero value obtained in SUBSEGMENT CONTENT DIFFERENCE is 1024.

Since the distribution is skewed, the distribution of SUBSEGMENT CONTENT DIFFERENCE for row classes is also skewed – which does not leave "room" for much variance in the distribution of values between row classes.

10.4 ULAM'S DISTANCE

The focus in the previous section was on the total subsegment contents of twelve-tone rows. Let us now take the opposite approach of merely finding a single shared (possibly non-contiguous) subsegment of two rows – the one that has the greatest length. This measurement of similarity is based on the assumption that two rows with a long shared subsegment should be similar. Again, as the aim is to define a dissimilarity measure, it is based on the number of pitch classes that do *not* belong to that longest shared subsegment; in this way the values decrease as the length of the shared subsegment increases. Let me now introduce a similarity measure known as ULAM'S DISTANCE, or U, which is named after a corresponding metric for permutations in mathematics. The formal definition is given below.

DEFINITION 10.5 The value of ULAM'S DISTANCE for two rows X and Y is 12 minus the length of the longest shared subsegment of X and Y.

For example, 54072813AB is the longest shared subsegment of rows $P = 5409728136\text{AB}$ and $Q = 5406728139\text{AB}$ in Figure 7.3. Since the length of that segment is 10, the value of ULAM'S

DISTANCE for these two rows is $12 - 10 = 2$.

This seems a natural and straightforward way of defining distance. The compositional relevance of ULAM'S DISTANCE is easy to demonstrate. For example, a shared subsegment in two rows can be highlighted by placing the pitch classes of the shared segment in one voice and the remaining pitch class in another voice. The two rows from Alban Berg's *Lyric Suite* discussed in Section 10.1 provide a good example.

There is one caveat to the idea of measuring similarity against the longest shared subsegment: the measure does not take into account the multiplicity of these subsegments. Consider rows $A = 0123456789\text{AB}$, $B = 6789\text{AB}012345$, and $C = \text{B0A}192837465$. The ULAM'S DISTANCE value for rows A and B is $12 - 6 = 6$ since the length of the longest shared subsegment (012345 or 6789AB) is 6. Similarly, the value for rows A and C is also $12 - 6 = 6$ since the length of the longest shared subsegment (012345) is again 6. There is a significant difference between the shared subsegments of rows A and B and those of rows A and C: rows A and B share two disjoint subsegments of length 6 but rows A and C share only one such subsegment. ULAM'S DISTANCE fails to distinguish between these two cases.[4]

It is straightforward to show that ULAM'S DISTANCE defines a metric. First, the values are positive real values. Secondly, the four requirements are satisfied. (i) Trivially, the value of $U(X, X)$ is 0 for all rows X since the longest shared subsegment is of length 12 and $12-12 = 0$. (ii) ULAM'S DISTANCE is obviously symmetric since it is calculated based on the length of the longest *shared* subsegment of two rows. (iii) Triangle inequality $U(X, Y) + U(Y, Z) \geq U(X, Z)$ holds and it can be shown as follows. Let us assume that $(x_1, x_2, \ldots, x_n)$ is the longest segment shared by rows X and Y. Hence, there are $12 - n$ pitch classes that do not belong to that segment. Let $(y_1, y_2, \ldots, y_m)$ then be the longest segment shared by rows Y and Z. Hence, there are $12 - m$ pitch classes that do not belong to that segment. We can now claim that $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_m)$ share at least $m+n-12$ pitch classes: if $m+n-12 \leq 0$ then the claim holds trivially, and if $m+n-12 > 0$ then the segments contain $m + n$ pitch classes, and as there are only twelve distinct pitch classes they must share $m + n - 12$ pitch classes. Now, the shared pitch classes must appear in the same order in both subsegments $(x_1, x_2, \ldots, x_n)$ and $(y_1, y_2, \ldots, y_m)$ since they are both subsegments of row Y. Therefore, rows X and Z have a shared subsegment of length $m + n - 12$, and thus triangle inequality holds. (iv) Finally, if $U(X, Y) = 0$ then rows X and Y have a shared subsegment of length 12 so they must be identical. Since all four requirements of the metric are satisfied, ULAM'S

[4]It is interesting to note that other similarity measures make a distinction between rows A and B and rows A and C but not necessarily in the same way. On the one hand, row A is more similar to row C than to row B according to, for example, the similarity measures DISPLACEMENT ($DIS(A, B) = 72$ versus $DIS(A, C) = 52$) and PITCH-CLASS DISPLACEMENT ($PCDIS(A, B) = 72$ versus $PCDIS(A, C) = 38$). On the other hand, row A is more similar to row B than to row C according to, for example, the similarity measures CAYLEY DISTANCE ($C(A, B) = 6$ versus $C(A, C) = 10$), BADNESS OF SERIAL FIT ($BSF(A, B) = 924$ versus $BSF(A, C) = 46080$), DIVISIONS ($DIV(A, B) = 1$ versus $DIV(A, C) = 11$), and ORDER-INTERVAL INVARIANT N-TUPLES ($OIIN(A, B) = 3969$ versus $OIIN(A, C) = 4081$). In addition, row A is equally similar to row C as to row B according to, for example, the similarity measures DERANGEMENT ($D(A, B) = 12 = D(A, C)$), SUBSEGMENT CONTENT DIFFERENCE ($SCD(A, B) = 3969 = SCD(A, C)$), and ORDER INVERSIONS ($OI(A, B) = 36 = OI(A, C)$). This provides some further support for the claim that similarity measures do not speak with a single extensional voice.

DISTANCE defines a metric.

The values of ULAM'S DISTANCE range from 0 to 11. The retrograde-related rows are the most dissimilar – the longest shared subsegment that such rows have is of length 1. Since for any row A and RA, $U(A, RA) = 11$, two rows belonging to the same row class may be maximally dissimilar, according to ULAM'S DISTANCE. Furthermore, for any row A, row RA is the only row at the maximum distance from row A – all other rows have a shared subsegment of at least length 2.

The transformational approach to ULAM'S DISTANCE

Let us now consider stating ULAM'S DISTANCE in GIS terms. Since it is an ordering relationship, let us select the GIS of order-number rows. As calculated in Section 7.2.3, the transformation in the GIS of order-number rows that maps row $P = 5409728136\text{AB}$ into row $Q = 5406728139\text{AB}$ is $int(\boldsymbol{P}, \boldsymbol{Q}) = \mathbf{0129456783AB}$. It is trivial to observe that the longest shared subsegment 54072813AB of rows P and Q corresponds to the longest ascending subsegment $\mathbf{01245678AB}$ of transformation $int(\boldsymbol{P}, \boldsymbol{Q})$.

ULAM'S DISTANCE has a more concrete transformational interpretation. However, I will defer this discussion to Section 10.8 in which I discuss transformations in general in using the conception of a twelve-tone row as a set of subsegments.

ULAM'S DISTANCE for row classes

ULAM'S DISTANCE is transformationally coherent since the equation $U(X, Y) = U(FX, FY)$ clearly holds for all row operations: transpositions, inversion, retrograde, and their combinations. For transposition T_k, segment $(x_1, x_2, \ldots, x_n)$ is the longest shared segment of rows X and Y if and only if $(T_k(x_1), T_k(x_2), \ldots, T_k(x_n))$ is the longest shared segment of rows T_kX and T_kY. Similarly for inversion I_k, segment $(x_1, x_2, \ldots, x_n)$ is the longest shared segment of rows X and Y if and only if segment $(I_k(x_1), I_k(x_2), \ldots, I_k(x_n))$ is the longest shared segment of rows I_kX and I_kY. Finally, segment $(x_1, x_2, \ldots, x_n)$ is the longest shared segment of rows X and Y if and only if segment $(x_n, \ldots, x_2, x_1)$ is the longest shared segment of rows RX and RY. Thus, ULAM'S DISTANCE defines a transformationally coherent metric and therefore, according to Corollary 6.1, it defines a metric for row classes.

The distribution of ULAM'S DISTANCE

No formula is known to produce the distribution of ULAM'S DISTANCE. The distribution for rows shown in Figure 10.4 was generated by a computer program using a brute-force algorithm. The average of the distribution is 7.15 and the mean is 7.

We can formally derive an upper limit of the distance for row classes using ULAM'S DISTANCE. Let us begin with the following classic result obtained by Erdös and Szekeres (1935).

THEOREM 10.1 (ERDÖS AND SZEKERES) A permutation of length $km + 1$ contains either an increasing subsequence of length $k + 1$ or a decreasing subsequence of length $m + 1$.

Let us now apply the theorem to order-number transformations. Setting $k = m = 3$ gives $3 \cdot 3 + 1 \leq 12$. Therefore, any order-number transformation contains either an increasing or a

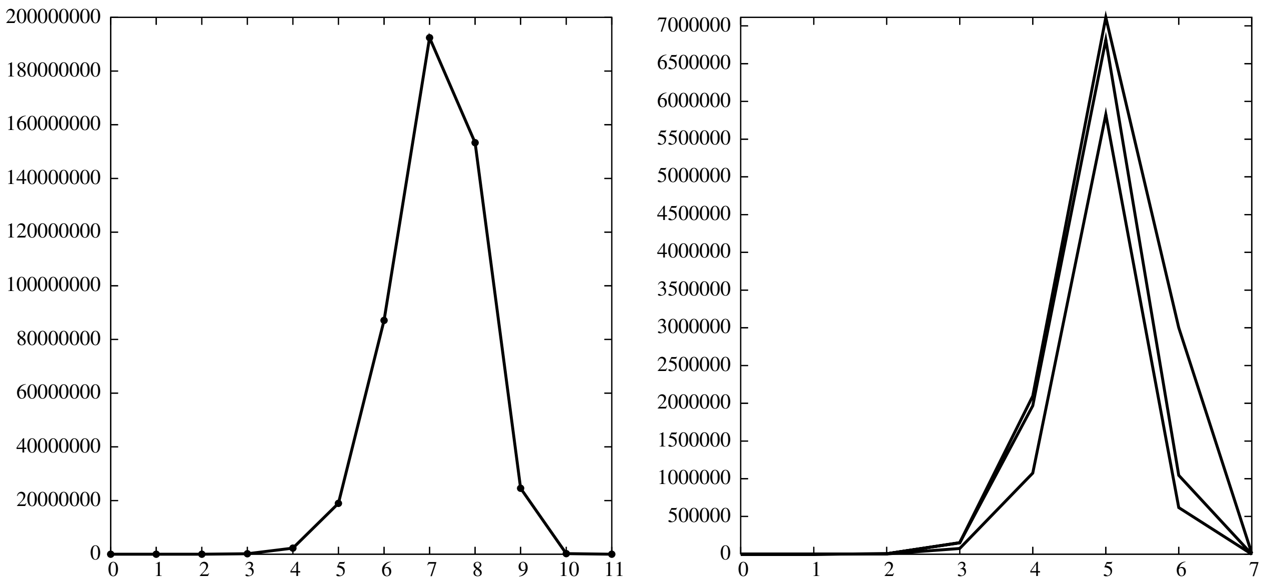


Figure 10.4: The distribution of ULAM'S DISTANCE on the left and the distribution of ULAM'S DISTANCE for row classes on the right. The former is obtained by computing the distances defined by all 479001600 distinct transformations, and hence is precise, and the latter by comparing 2000 random row classes to all other row classes. See Section 6.3 for a discussion on the three curves in the latter figure.

decreasing subsequence of length $3+1=4$. Hence, if we take the order-number transformation that transforms row A into row B, the result means that rows A and B either have a shared subsegment of length 4 or one that is retrograded in the other row. Consequently, for any rows A and B, either rows A and B or rows A and RB have a shared subsegment of length 4. Therefore, either the ULAM'S DISTANCE value for rows A and B is at most 8 or its value for rows A and RB is at most 8. It then follows that the ULAM'S DISTANCE value for two row classes cannot exceed 8.

In the sample, the maximum distance between two row classes is only 7. Consequently, we can always find rows in two row classes that contain a subsegment of length 5 – provided that 7 is also the absolute maximum and not only the maximum in the sample.

Variants of ULAM'S DISTANCE

As noted above, ULAM'S DISTANCE does not take into account the multiplicity of the longest shared subsegments. We could create a variant in which two rows are broken into shared disjoint subsetments. Let us consider again rows $A = 0123456789\mathrm{AB}$, $B = 6789\mathrm{AB}012345$, and $C = \mathrm{B0A192837465}$. Rows A and B share the disjoint subsegments 012345 and $6789\mathrm{AB}$. Similarly, rows A and C share the disjoint subsegments 012345, B, A, 9, 8, 7, and 6. I will introduce in Section 10.7 a method for weighting shared contiguous subsegments that could be adapted to the present case in which the shared subsegments are not necessarily contiguous. Such a similarity measure would deem rows A and B more similar than rows A and C.

Extending ULAM'S DISTANCE

ULAM'S DISTANCE could be easily extended to segments of any length. Furthermore, there could be pitch-class duplications and the two segments do not even need to contain the same pitch classes. The ULAM'S DISTANCE value would simply be the number of pitch classes that do not belong to the longest shared subsegment. If the two segments to be compared are not of equal length, it would be preferable to define this value as the number of pitch classes in the longer segment that do not belong to the longest shared subsegment since, for example, all pitch classes of segment 00 belong to the longest shared segment of 00 and 007.

10.5 ORDER-INTERVAL INVARIANT N-TUPLES

John Ward introduces ORDER-INTERVAL INVARIANT N-TUPLES, or $OIIN$, in his dissertation (Ward 1992). He describes this similarity measure as follows.

> The author has devised another similarity measure, Order-Interval Invariant n-tuples, which considers the exact positions of a subset's elements within a superset. (Ward 1992, 154)

This measure could be considered a refinement of SUBSEGMENT CONTENT DIFFERENCE: both measures compare the subsegment contents of rows, but ORDER-INTERVAL INVARIANT N-TUPLES also takes into account the order-number intervals between the entries in the shared subsegments. Naturally, two rows contain at least as many shared subsegments as they have shared subsegments with identical order-number intervals between the entries.

For example, rows $0123456789AB$ and $B0927456381A$ share subsegment $02468A$. In addition, while the pitch classes of the subsegment do not have the same order positions in the two rows they do have the same order-number intervals since in both rows we need to pick every second pitch class in order to obtain this subsegment.

Ward compares the SUBSEGMENT CONTENT DIFFERENCE and ORDER INVERSIONS similarity measures as follows.

> ORIN is concerned only with the relative order of the members of a dyad, irrespective of their distance from each other in any give ordered superset. [...] Furthermore, order-interval invariance is perhaps more perceptible, more marked for consciousness, than the mere absence of order inversion. It seems reasonable that instances of repeated utterances of an Order-Interval Invariant n-tuple may have an impact on the listener's consciousness. (Ward 1992, 154–155)

I compared ordered dyads and adjacent ordered dyads in Section 10.1, and suggested that it is easier to track adjacent dyads than non-adjacent dyads. Here we would need to keep track of all dyads and, in addition, the order intervals. I therefore have reservations about the "impact on the listener's consciousness," bearing in mind that keeping track of the exact order of the pitch classes and the order intervals seems to be a cognitively rather demanding task.

I derived SUBSEGMENT CONTENT DIFFERENCE from the GENERALIZED ORDER INVERSIONS family of similarity measures in Section 10.3. I will follow the same approach here and derive a modification of Ward's similarity measure from a family of measures. Ward defines his measure as a similarity measure: I will redefine it here as a dissimilarity measure.

I will begin by defining a family of similarity measures based on shared subsegments of length n that have identical order-number intervals between the entries. Let us refer to these measures as $OIIN_n$. As the aim is to define the ORDER-INTERVAL INVARIANT N-TUPLES similarity measure I will not examine the "constituent similarity measures" $OIIN_n$ in any detail. Let us extend the definition to segment lengths 0 and 1 (see Section 10.3).

DEFINITION 10.6 An n-tuple with order intervals is an ordered pair $(S_n, OINT(S_n))$ in which S_n is a segment of length k and $OINT(S_n)$ is the order-interval succession of S_n.

For example, rows 0123456789AB and B0927456381A both contain ordered pair (02468A, **22222**). Here 02468A denotes the shared subsegment and **22222** denotes the order-number intervals of the pitch classes in both rows.[5]

DEFINITION 10.7 If X and Y are twelve-tone rows and $0 \leq n \leq 12$, then the number of different tuples of length n with fixed order-number intervals in the two rows is given by the formula

$$OIIN_n(X, Y) = \#\{(S_n, OINT(S_n)) \mid (S_n, OINT(S_n)) \in X \text{ and } (S_n, OINT(S_n)) \notin Y\}$$

in which $(S_n, OINT(S_n))$ is an n-tuple with order intervals.

Definition 10.7 sets some segment length n, examines the segments S_n and the pertinent order-number intervals $OINT(S_n)$ in row X, and then counts the number of such pairs of segments and order-number intervals that are not found in row Y. From this it is straightforward to define the ORDER-INTERVAL INVARIANT N-TUPLES as the sum of similarity measures $OIIN_n$.

DEFINITION 10.8 If X and Y are twelve-tone rows then the number of different n-tuples with fixed successions of order-number intervals in the rows is given by the formula

$$OIIN(X, Y) = \sum_{n=0}^{12} OIIN_n(X, Y).$$

Figure 10.5 illustrates the calculation of ORDER-INTERVAL INVARIANT N-TUPLES for rows P = 5409728136AB and Q = 5406728139AB. For each segment length $0 \leq n \leq 12$ let us first calculate the number of subsegments of row P that are not subsegments of row Q, and then the number of shared subsegments of rows P and Q that have different order-number intervals. For instance, rows P and Q share segment 09 but the order-number interval is different since segment 09 is contiguous in row P but not in row Q.

[5] In order to emphasize that I am dealing with order-number intervals I have used boldface.

length	0	1	2	3	4	5	6	7	8	9	10	11	12
number of segments	1	12	66	220	495	792	924	792	495	220	66	12	1
shared segments	1	12	55	140	230	262	212	120	45	10	1	0	0
non-shared segments	0	0	11	80	265	530	712	672	450	210	65	12	1
diff. order-number intervals	0	0	10	20	20	10	2	0	0	0	0	0	0
$OIIN_n(P,Q)$	0	0	21	100	285	540	714	672	450	210	65	12	1

Figure 10.5: The calculation of $OIIN$ for rows $P = 5409728136\mathrm{AB}$ and $Q = 5406728139\mathrm{AB}$. The top row denotes the length of the segments, the second row denotes the total number of subsegments of a given length in a row, the third row denotes the number of segments that rows P and Q share, the fourth row denotes the number of segments that row P has but row Q does not, the fifth row denotes the number of segments that rows P and Q share but that have different order-number intervals, and the bottom row denotes the value of $OIIN_n(P,Q)$. The value of ORDER-INTERVAL INVARIANT N-TUPLES for rows P and Q is the sum of the values in the bottom row. The rows are related: the number of segments is the sum of shared segments and non-shared segments, and the value of $OIIN_n(P,Q)$ is the sum of non-shared segments and shared segments with different order-number intervals.

It is straightforward to show that ORDER-INTERVAL INVARIANT N-TUPLES defines a metric. First, the values are positive real values. Secondly, the four requirements of the metric are satisfied. (i) Trivially, the value of $OIIN(X,X)$ is 0 for all rows X. There are no different subsegments between two identical rows even accounting for the order-number intervals in the segments. (ii) ORDER-INTERVAL INVARIANT N-TUPLES is symmetric since the number of segments in X that are not in Y equals the number of segments in row Y that are not in row X, and hence $OIIN(X,Y) = OIIN(Y,X)$. (iii) Triangle inequality $OIIN(X,Y) + OIIN(Y,Z) \geq OIIN(X,Z)$ holds and it can be shown as follows (always taking the order-number intervals into account). If there are k segments that are in row X and not in row Y, then there are $4096 - k$ segments that rows X and Y share. Assume, then, that there are l segments that are in row Y and not in row Z. Therefore, rows X and Z share at least $4096-k-l$ segments. This means that there are at most $k+l$ segments in row X that are not in row Z and therefore triangle inequality holds. (iv) Finally, if $OIIN(X,Y) = 0$ then rows X and Y do not have any different segments; in particular the segments of length 12 are identical so rows X and Y must be identical. Since all four requirements of the metric are satisfied, ORDER-INTERVAL INVARIANT N-TUPLES defines a metric.

The minimum value of ORDER-INTERVAL INVARIANT N-TUPLES is 0 and the maximum value is 4083 (like that of SUBSEGMENT CONTENT DIFFERENCE). Any given row is maximally dissimilar to several rows, including its own retrograde. Since for any row A and RA, $OIIN(A,RA) = 4083$, two rows belonging to the same row class may be maximally dissimilar according to ORDER-INTERVAL INVARIANT N-TUPLES.

The transformational approach to ORDER-INTERVAL INVARIANT N-TUPLES

Let us now consider ORDER-INTERVAL INVARIANT N-TUPLES from the transformational perspective. It is now possible to utilize the transformational interpretation of SUBSEGMENT CONTENT DIFFERENCE. Recall from Section 10.3 that the shared subsegments of two rows correspond to the ascending sequences in the corresponding order-number transformation. In order to examine whether the order-number intervals are identical in shared subsegments we need to examine the "intervals" between the entries in the order-number transformation: the order-number intervals in a shared subsegment are identical if and only if the intervals between the transformed entries are equal to the intervals between the entries to be transformed.

Let us consider as an example a subsegment of length 2 in row A. Our task is to check whether another row B contains that segment and, if it does, whether the order-number interval is identical. The pitch classes of that subsegment are in row A at order positions a and $a+k$ in which $k > 0$. In the order-number transformation that transforms row A into row B, order number a is transformed into order number b and order number $a+k$ is transformed into order number $b+h$ for some h. Now, the two pitch classes are in the same order in both rows if and only if $h > 0$. Furthermore, if $h > 0$ then the order-number intervals are identical if and only if $h = k$.

Let us apply the above observation to a concrete case: rows $A = 0123456789\mathrm{AB}$ and $B = \mathrm{B}0927456381\mathrm{A}$. Let us first write the rows as order-number rows: $\boldsymbol{A} = \mathbf{0123456789AB}$ and $\boldsymbol{B} = \mathbf{1A38567492B0}$. The order-number transformation that transforms row $\boldsymbol{A}$ into row $\boldsymbol{B}$ is

$$\boldsymbol{BA}^{-1} = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \end{pmatrix} \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 1 & 10 & 3 & 8 & 5 & 6 & 7 & 4 & 9 & 2 & 11 & 0 \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 1 & 10 & 3 & 8 & 5 & 6 & 7 & 4 & 9 & 2 & 11 & 0 \end{pmatrix}.$$

Note in transformation $\boldsymbol{BA}^{-1}$ the ascending sequence of order numbers $\mathbf{135679B}$. In addition, order number $\mathbf{0}$ is transformed into order number $\mathbf{1}$ and order number $\mathbf{2}$ is transformed into order number $\mathbf{3}$: the order-number interval between the original order numbers $\mathbf{0}$ and $\mathbf{2}$ is 2 like that between the transformed order numbers $\mathbf{1}$ and $\mathbf{3}$. It is similar with the other order-number intervals between the entries in sequence $\mathbf{135679B}$ and, consequently, the pitch classes at order positions $\mathbf{024568A}$ in row A and those at order positions $\mathbf{135679B}$ in row B constitute an order-interval invariant n-tuple.

The two-line notation of permutations is convenient for the examination of order-interval invariant n-tuples in order-number transformations: the upper line denotes the original order numbers and the lower line denotes the transformed order numbers. The key to the transformational definition is that if an n-tuplet is order-interval invariant then every pitch class of it is either kept in place or displaced by the same number of order positions and in the same direction. Consequently, if we interpret the entries of the order-number transformation $int(\boldsymbol{X}, \boldsymbol{Y})$ as integers, the difference $int(\boldsymbol{X}, \boldsymbol{Y})(k) - k$ is constant for each order position of the order-interval invariant n-tuplet.

The "constant difference" phenomenon provides a quick way to discover the order-interval invariant n-tuples in rows X and Y: we need to look for identical differences $int(\boldsymbol{X}, \boldsymbol{Y})(k) - k$. For example, we could write the differences in transformation $\boldsymbol{BA}^{-1}$ as follows:

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ +1 & +9 & +1 & +5 & +1 & +1 & +1 & -3 & +1 & -7 & +1 & -11 \end{pmatrix}.$$

Note that seven times the difference $int(\boldsymbol{A}, \boldsymbol{B})(k) - k$ is $+1$: hence, there is an order-interval invariant n-tuple of length 7.

Given the observations above, we can now state similarity measures $OIIN_n$ (see Definition 10.7) in trasformational terms.

DEFINITION 10.9 In the GIS of order-number rows, the value of the $OIIN_n$ similarity measure for twelve-tone rows X and Y is

$$OIIN_n(X, Y) = \#\{(a_1, a_2, \ldots, a_n) \mid \text{ where } g_k - k \neq g_m - m \text{ for some } k \text{ and } m\}$$

in which g_k is the kth element of transformation $int(\boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{YX}^{-1}$ interpreted as an integer, g_m is the mth element of transformation $int(\boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{YX}^{-1}$ interpreted as an integer and $(a_1, a_2, \ldots, a_n)$ is a n-tuplet in transformation $int(\boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{YX}^{-1}$.

For example, $(4, 5)$ is an order-interval invariant n-tuplet of length 2 in the above transformation $\boldsymbol{BA}^{-1}$, since order number $\mathbf{4}$ is transformed into order number $\mathbf{5}$, order number $\mathbf{5}$ is transformed into order number $\mathbf{6}$, and $oint(\mathbf{4}, \mathbf{5}) = oint(\mathbf{5}, \mathbf{6})$. Correspondingly, $(0, 1)$ is not an order-interval invariant n-tuplet, since order number $\mathbf{0}$ is transformed into order number $\mathbf{1}$, order number $\mathbf{1}$ is transformed into order number $\mathbf{10}$, and $oint(\mathbf{0}, \mathbf{1}) \neq oint(\mathbf{1}, \mathbf{10})$.

Let us now define ORDER-INTERVAL INVARIANT N-TUPLES in transformational terms as the sum of the thirteen transformational similarity measures $OIIN_n$. This is almost identical to Definition 10.8, the only difference being that it is given in terms of the GIS of order-number rows.

DEFINITION 10.10 In the GIS of order-number rows, the value of the $OIIN_n$ similarity measure for twelve-tone rows X and Y is

$$OIIN(X, Y) = \sum_{n=0}^{12} OIIN_n(X, Y).$$

As discussed in Section 6.1, since ORDER-INTERVAL INVARIANT N-TUPLES measures the magnitude of the transformation, every row has precisely the same network of distances to the other rows.

ORDER-INTERVAL INVARIANT N-TUPLES for row classes

In order to show that ORDER-INTERVAL INVARIANT N-TUPLES for row classes is well defined, we must show that it is transformationally coherent (see Corollary 6.1 in Section 6.2). It is transformationally coherent since the equation $OIIN(X, Y) = OIIN(FX, FY)$ clearly holds for all row operations: transpositions, inversion, retrograde, and their combinations. For transposition T_k,

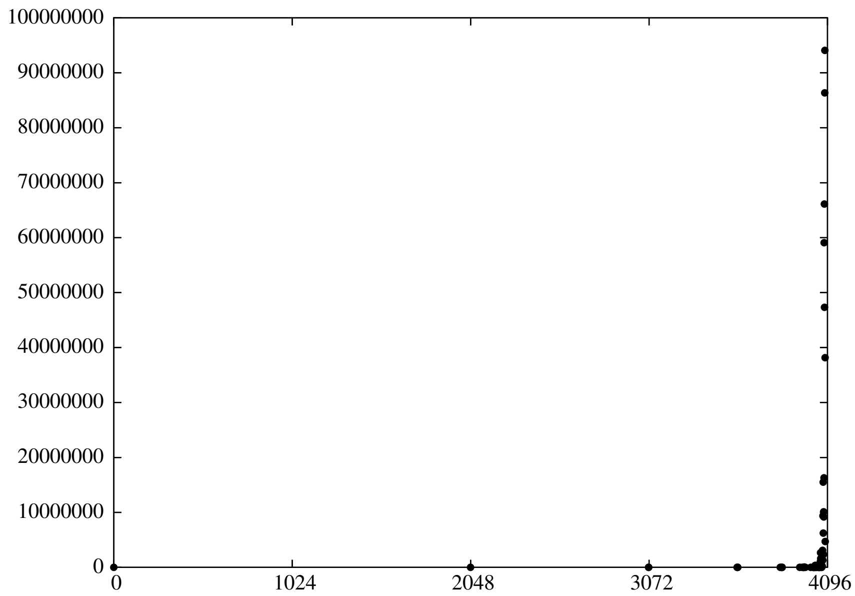


Figure 10.6: The distribution of the ORDER-INTERVAL INVARIANT N-TUPLES values. (Most of the dots denoting the distribution are on the very right edge of the picture and they are therefore difficult to discern.)

for every pair $(S_n, OINT(S_n))$ of row X that is not a pair of row Y there is a corresponding pair $(T_k(S_n), OINT(T_k(S_n)))$ of row T_kX that is not a pair of row T_kY (and vice versa). Similarly for inversion I_k, for every pair $(S_n, OINT(S_n))$ of row X that is not a pair of row Y there is a corresponding pair $(I_k(S_n), OINT(I_k(S_n)))$ of row I_kX that is not a pair of row I_kY (and vice versa). Finally for retrograde, for every pair $(S_n, OINT(S_n))$ of row X that is not a pair of row Y there is a corresponding pair $(R(S_n), OINT(R(S_n)))$ of row RX that is not a pair of row RY (and vice versa). Thus, ORDER-INTERVAL INVARIANT N-TUPLES is a transformationally coherent metric and therefore, according to Corollary 6.1, it defines a metric for row classes.

The distribution of ORDER-INTERVAL INVARIANT N-TUPLES

The ORDER-INTERVAL INVARIANT N-TUPLES values obtained above for the sample rows show a tendency to be large – even compared to the SUBSEGMENT CONTENT DIFFERENCE values, which were also notable for generating large values. For example, let us consider rows $P = 5409728136\mathrm{AB}$ and its variant $P' = 4509728136\mathrm{AB}$ in which the first two pitch classes are exchanged. The value of ORDER-INTERVAL INVARIANT N-TUPLES for rows P and P' is 3070, and that of SUBSEGMENT CONTENT DIFFERENCE for the same rows is only 1024. The smallest non-zero value of ORDER-INTERVAL INVARIANT N-TUPLES is 2047, and it is obtained if the rows are related by rotation r_1 or rotation r_{11}. For example, its value for rows $0123456789\mathrm{AB}$ and $123456789\mathrm{AB}0$ is 2047.

Figure 10.6 shows the distribution of the ORDER-INTERVAL INVARIANT N-TUPLES values. The average is 4078.04 and the mean is 4079. The distribution is extremely skewed: in particular, given the fact that the scale is from 0 to 4083, the average and the mean are very close to the maximum value.

Since this distribution is skewed, so is the distribution of ORDER-INTERVAL INVARIANT N-TUPLES for row classes. Given the skewness, there is not "room" for much variance in the distribution of values between row classes.

10.6 DIVISIONS

In most of the row measures discussed so far, rows $P = 5409728136\text{AB}$ and $r_6P = 8136\text{AB}540972$ have been deemed to be dissimilar rather than similar. For example, $D(P, r_6P) = 12$ (maximum value 12), $DIS(P, r_6P) = 72$ (maximum value 72), $PCDIS(P, r_6P) = 36$ (maximum value 72), $OI(P, r_6P) = 36$ (maximum value 66), $BSF(P, r_6P) = 924$ (maximum value 479001600), $C(P, r_6P) = 6$ (maximum value 11), $SCD(P, r_6P) = 3969$ (maximum value 4083), $U(P, r_6P) = 6$ (maximum value 11), and $OIIN(P, r_6P) = 3969$ (maximum value 4083). If we take the distributions of the measures into account, the measures BADNESS OF SERIAL FIT, CAYLEY DISTANCE, and ORDER-INTERVAL INVARIANT N-TUPLES indicate similarity. For example, according to CAYLEY DISTANCE only 0.63% of all twelve-tone rows are more similar to row P than row r_6P and according to ORDER-INTERVAL INVARIANT N-TUPLES only 0.02% of all twelve-tone rows are more similar to row P than row r_6P. Nevertheless, none of these measures seem to take into consideration the fact that these two rows are composed of the very same two contiguous segments 540972 and 8136AB.

These observations lead to the definition of a new similarity measure, DIVISIONS or DIV, invented by the author and based on dividing rows into shared contiguous subsegments. The fewer divisions we need to create, the more similar the rows are. For example, with only one division we can divide both rows $P = 5409728136\text{AB}$ and $r_6P = 8136\text{AB}540972$ into two contiguous subsegments as

$$540972|8136\text{AB} \text{ and } 8136\text{AB}|540972,$$

and hence the DIVISIONS value for these two rows is 1. Similarly, with four divisions we can divide rows $P = 5409728136\text{AB}$ and $Q = 5406728139\text{AB}$ into five contiguous subsegments as

$$540|9|72813|6|\text{AB} \text{ and } 540|6|72813|9|\text{AB},$$

and hence the DIVISIONS value for these two rows is 4.

I will defer the formal definition of DIVISIONS until I discuss it in terms of transformations.

It is straightforward to show that DIVISIONS defines a metric. First, the values are positive real values, and secondly, the four requirements are satisfied. (i) The value of $DIV(X, X)$ is 0 for all rows: since the rows are identical they do not need to be divided. (ii) DIVISIONS is clearly symmetric since it denotes the number of divisions needed to create *shared* subsegments: hence $DIV(X, Y) = DIV(Y, X)$. (iii) Triangle inequality $DIV(X, Y) + DIV(Y, Z) \geq DIV(X, Z)$ holds, as the following argument shows. Let X, Y, and Z be arbitrary twelve-tone rows. In order to obtain the same contiguous segments in rows X and Y we must insert n divisions into both. With these divisions in place, in order to obtain the same contiguous segments in rows Y and Z we must

	row X	row Y	row Z
step 1	012|345|678|9AB	012|678|345|9AB	
step 2		01|2|678|34|59|AB	01|78|59|2|6|34|AB
step 3		01|2|6|78|34|5|9|AB	
step 4	01|2|34|5|6|78|9|AB		01|78|5|9|2|6|34|AB

Figure 10.7: The triangle inequality of DIVISIONS for rows $X = 0123456789\text{AB}$, $Y = 0126783459\text{AB}$, and $Z = 0178592634\text{AB}$. The three appropriate divisions are inserted into rows X and Y in step 1, and the five appropriate divisions are inserted into rows Y and Z in step 2. The divisions inserted into row Y in the previous steps are combined in step 3, which results in a total of 7 divisions (not $3 + 5 = 8$ since two of those inserted in steps 1 and 2 coincide). The 7 divisions obtained in step 3 are inserted into rows X and Z in step 4. As a result, rows X and Z are divided into identical contiguous subsegments and the number of necessary divisions does not exceed the number of divisions between rows X and Y plus the number of divisions between rows Y and Z.

insert m divisions into both. If there is a division in row Y that is not in row X we add that division to row X. Similarly, if there is a division in row Y that is not in row Z we add that division to row Z. Consequently, there are at most $m + n$ divisions in rows X and Z that divide them into identical contiguous subsegments, and hence triangle inequality holds: Figure 10.7 illustrates the proof. (iv) Finally, if $DIV(X, Y) = 0$, then with no divisions at all the two rows consist of identical subsegments and therefore rows X and Y must be identical. Since all four requirements of the metric are satisfied, DIVISIONS defines a metric.

The minimum DIVISIONS value is 0, denoting maximal similarity. Inserting k divisions divides a row into $k + 1$ contiguous subsegments, hence inserting zero divisions keeps any row as a single segment. Correspondingly, inserting eleven divisions divides any row into twelve segments of length 1, and any two rows share at least these twelve segments. Consequently, the maximum DIVISIONS value is 11. Any given row is maximally dissimilar to several rows, including its own retrograde. Since for any row A and RA, $DIV(A, RA) = 11$, two rows belonging to the same row class may thus be maximally dissimilar.

The transformational approach to DIVISIONS

DIVISIONS has a natural interpretation as a measurement of the magnitude of a transformation. In order to illustrate this, let us examine the order-number transformation

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 0 & 1 & 2 & 9 & 4 & 5 & 6 & 7 & 8 & 3 & 10 & 11 \end{pmatrix}$$

that transforms row $P = 5409728136\text{AB}$ into row $Q = 5406728139\text{AB}$ (see Section 7.2.3). The three consecutive order numbers $\mathbf{0}$, $\mathbf{1}$, and $\mathbf{2}$ are transformed into three consecutive order numbers $\mathbf{0}$, $\mathbf{1}$, and $\mathbf{2}$, respectively. (It is immaterial here that they are the *same* order numbers.) The interpretation is that the contiguous subsegment 540 (at order positions $\mathbf{012}$) of row P is transformed into a contiguous subsegment 540 (also at order positions $\mathbf{012}$) of row Q. On the other hand, the four consecutive order numbers $\mathbf{0}$, $\mathbf{1}$, $\mathbf{2}$, and $\mathbf{3}$ are transformed into four order numbers $\mathbf{0}$, $\mathbf{1}$, $\mathbf{2}$, and $\mathbf{9}$, which are not consecutive. Hence, row P has the contiguous subsegment 5409 (at order positions $\mathbf{0123}$) but row Q does not.

We can find the contiguous segments of rows P and Q by examining the consecutive ascending pairs[6] in transformation $\boldsymbol{QP}^{-1}$. The number of pairs that are *not* consecutive ascending pairs equals the number of divisions. There are a total of eleven adjacent pairs in transformation $\boldsymbol{QP}^{-1} = \mathbf{0129456783AB}$: $\mathbf{01}$, $\mathbf{12}$, $\mathbf{29}$, $\mathbf{94}$, $\mathbf{45}$, $\mathbf{56}$, $\mathbf{67}$, $\mathbf{78}$, $\mathbf{83}$, $\mathbf{3A}$, and $\mathbf{AB}$. Of these, $\mathbf{01}$, $\mathbf{12}$, $\mathbf{45}$, $\mathbf{56}$, $\mathbf{67}$, $\mathbf{78}$, and $\mathbf{AB}$ are consecutive ascending pairs, and $\mathbf{29}$, $\mathbf{94}$, $\mathbf{83}$, and $\mathbf{3A}$ are not (pairs $\mathbf{29}$ and $\mathbf{3A}$ are ascending pairs but are not consecutive). Hence, the DIVISIONS value for these two rows is 4.

Let us now give a formal definition of DIVISIONS.

DEFINITION 10.11 In the GIS of order-number rows, the value of the DIVISIONS similarity measure for twelve-tone rows X and Y is the number of adjacent pairs in order-number transformation $int(\boldsymbol{X}, \boldsymbol{Y})$ that are *not* consecutive ascending pairs.

The transformational relation of rows could be conceptualized as a "shuffle" of segments. A row is first divided into segments (these segments do not need to be of equal length) and then it is transformed into another row by reordering them. DIVISIONS describes the number of segments a row is divided into.

DIVISIONS for row classes

In order to show that DIVISIONS for row classes is well defined, we must show that it is transformationally coherent (see Corollary 6.1 in Section 6.2). It is a transformationally coherent similarity measure since the equation $DIV(X, Y) = DIV(FX, FY)$ holds for all row operations: transpositions, inversion, retrograde, and their combinations. For transposition T_k, if two rows X and Y have the shared contiguous subsegment $p_1p_2 \ldots p_n$ then the transposed rows T_kX and T_kY have the shared contiguous subsegment $T_k(p_1)T_k(p_2) \ldots T_k(p_n)$ (and vice versa). Similarly for inversion I_k, if two rows X and Y have the shared contiguous subsegment $p_1p_2 \ldots p_n$ then the inverted rows I_kX and I_kY have the shared contiguous subsegment $I_k(p_1)I_k(p_2) \ldots I_k(p_n)$ (and vice versa). For the retrograde, if two rows X and Y have the shared contiguous subsegment $p_1p_2 \ldots p_{n-1}p_n$ then the retrograded rows RX and RY have the shared contiguous subsegment $p_np_{n-1} \ldots p_2p_1$ (and vice versa). Hence, the transformed rows have shared subsegments corresponding to the original rows and, given similar reasoning, the original rows have shared subsegments corresponding to the transformed rows. Since the rows have corresponding shared subsegments they have the same number of divisions. DIVISIONS is therefore transformationally coherent and it defines a metric for row classes.

If we consider the extended set of row operations, DIVISIONS is transformationally coherent under some of them but not all. In fact, it is transformationally coherent under all pitch-class operations, such as the M-operation, since with respect to DIVISIONS they constitute only a relabeling of the pitch classes. In general, DIVISIONS is not transformationally coherent under the order-number

[6] A consecutive ascending pair of a permutation is one in which k is followed by $k + 1$.

operations – retrograde is again the only exception. For example, it is not transformationally coherent under (non-trivial) rotation. Let us consider row $P = 5409728136\text{AB}$ in Figure 7.3, its modification $P' = 4509728136\text{AB}$ in which the positions of the two first pitch classes have been exchanged, and the rotations of both rows $r_9P = 9728136\text{AB}540$ and $r_9P' = 9728136\text{AB}450$. A comparison of the values for the two original rows and two rotated rows gives

$$DIV(P, P') = 2 \neq 3 = DIV(r_9P, r_9P'),$$

and therefore DIVISIONS is not transformationally coherent under rotation. The transformationality is broken since the rotations "wrap around."

DIVISION is not transformationally coherent under the exchange operation either. Let rows $P = 5409728136\text{AB}$ and $P' = 4509728136\text{AB}$ be as above. The exchange transformations of these two rows are $EP = 2758109463\text{AB}$ and $EP' = 2758019463\text{AB}$. A comparison of the DIVISION values for rows P and P' and for their transformations EP and EP' gives

$$DIV(P, P') = 2 \neq 3 = DIS(EP, EP'),$$

which shows that it is not transformationally coherent under the exchange operation.

Distribution of DIVISIONS

The above transformational interpretation of the DIVISIONS similarity measure gives quick access to its distribution, which is given in David, Kendall, and Barton (1966, 263) based on a recursive formula devised by David and Barton (1962, 168). The distribution is reproduced in Figure 10.8.

The distribution is highly skewed. The average is 10.08 and the mean is 10 – both of which are close to the maximum value 11. Any given row has no shared non-trivial contiguous subsegments with as many as 190899411 other rows.

The unordered interval-class contents of rows provides a method for examining some boundary conditions for the values of DIVISIONS for row classes. If two rows do not contain a single shared interval class, then the value for the respective row classes is necessarily 11. For example, rows 0123456789AB and $05\text{A}3816\text{B}4927$ do not contain a single shared interval class, and therefore no two row forms of their respective row classes contain a shared contiguous subsegment. I thus hypothesize that row classes with only a few distinct interval classes in the constituent rows have larger numbers of row classes at high distance than rows with all or most interval classes in the constituent rows. For example, the row class of row $2\text{B}871\text{A}903654$ (which contains only interval classes 1, 3, and 6) in the sample has the largest number of row classes at distance 11.

10.7 FRAGMENTATION

FRAGMENTATION is a new similarity measure invented by the author, and could be considered a variant of DIVISIONS. In order to justify proceeding with this measure, let us consider DIVISIONS and the three rows $A = 0123456789\text{AB}$, $B = \text{B}0123456789\text{A}$, and $C = 6789\text{AB}012345$. With

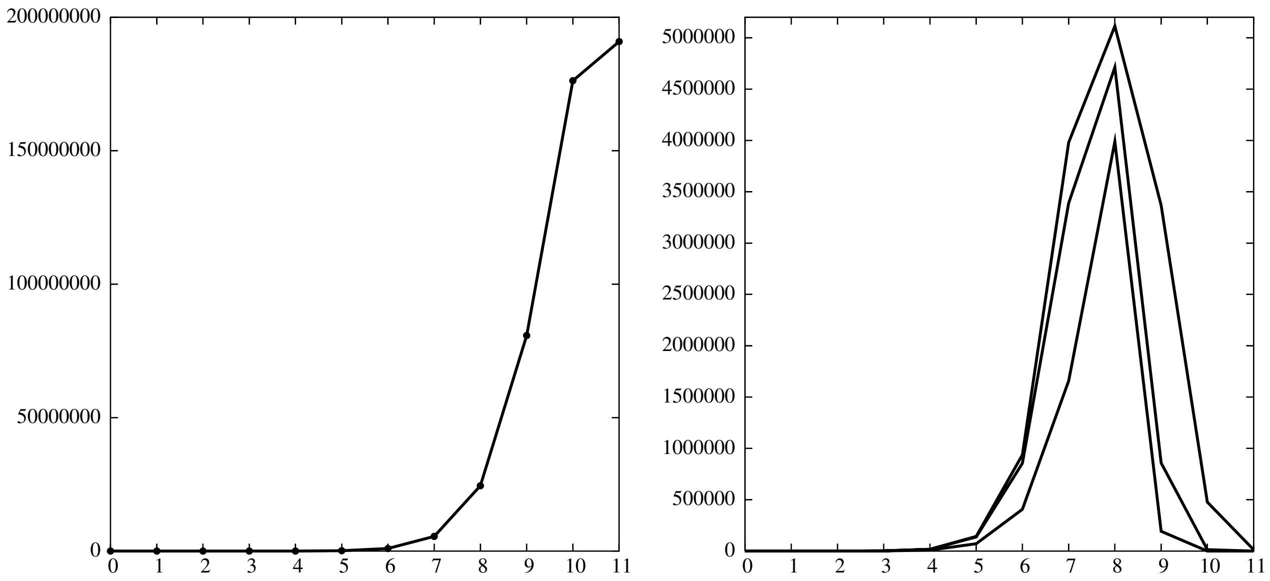


Figure 10.8: The distribution of DIVISIONS on the left and the distribution of DIVISIONS for row classes on the right. The former is obtained by means of a formula and hence it is precise, while the latter is obtained by comparing 2000 random row classes to all other row classes. See Section 6.3 for a discussion on the three curves in the latter figure.

only one division we can divide both rows A and B into two identical contiguous subsegments as

$$0123456789\mathrm{A}|\mathrm{B} \text{ and } \mathrm{B}|0123456789\mathrm{A},$$

and hence the DIVISIONS value for these two rows is 1. Similarly, with only one division we can divide both rows A and C into two identical contiguous subsegments as

$$012345|6789\mathrm{AB} \text{ and } 6789\mathrm{AB}|012345,$$

and hence the DIVISIONS value for these two rows is also 1. Therefore, the similarity of rows A and B is equal to that of rows A and C. The crucial question is whether we deem it to be significant that in the first case the longest shared subsegment is of length 11 while in the latter case it is only of length 6 (even if there are two shared subsegments of length 6). The FRAGMENTATION similarity measure is based on the premise that the length of the shared subsegments is significant: rows A and B should be more similar than rows A and C. Hence, we need a method for *weighting* the lengths of the shared segments – borrowed in this case from Starr and Morris (1978). While they base the division of rows into segments on combinatorial concerns, it is straightforward to extend their concepts to the similarity of rows.

Starr and Morris have extensively discussed the properties of combinatorial matrices (Starr and Morris 1977; Starr and Morris 1978). They introduce the concepts *fragmentation* and *association* to evaluate these properties, and describe their approach as follows.

> To begin with, we will consider the fragmentation of a single row broken into segments,

> which we can evaluate in terms of ordered pairs. If the PC k follows the PC j in some row-segment, we say that the segment contains the *ordered pair* j:k. A row-segment of n PCs will define n(n-1)/2 ordered pairs (henceforth 'OPs'). (Starr and Morris 1978, 73)

There are two steps in the application of this approach to the similarity of twelve-tone rows advocated here. The first is to divide the rows into shared contiguous segments, as in the case of the DIVISIONS similarity measure, and the second is to weight the segments based on the formula devised by Starr and Morris.

DEFINITION 10.12 The FRAGMENTATION value for twelve-tone rows X and Y is given by the formula

$$FRAG(X,Y) = 66 - \sum \frac{n_i(n_i-1)}{2}$$

in which n_i are the lengths of the shared subsegments of rows X and Y.

For example, the two shared subsegments of rows $A = 0123456789\mathrm{AB}$ and $B = \mathrm{B}0123456789\mathrm{A}$ have lengths 1 and 11, and therefore the FRAGMENTATION value is

$$FRAG(A,B) = 66 - \frac{11 \cdot 10}{2} - \frac{1 \cdot 0}{2} = 11.$$

Correspondingly, the two shared subsegments of rows $A = 0123456789\mathrm{AB}$ and $C = 6789\mathrm{AB}012345$ both have length 6, and therefore the FRAGMENTATION value is

$$FRAG(A,B) = 66 - \frac{6 \cdot 5}{2} - \frac{6 \cdot 5}{2} = 36.$$

The factor $n_i(n_i-1)/2$ in Definition 10.12 denotes the number of ordered pairs that are within the shared segments. Since each row contains 66 ordered pairs, the FRAGMENTATION value is the number of ordered pairs that are not in the same segment. For example, in the case of rows

$$A = 0123456789\mathrm{A}|\mathrm{B} \text{ and } B = \mathrm{B}|0123456789\mathrm{A}$$

there are

$$66 - \left(\frac{11 \cdot 10}{2} + \frac{1 \cdot 0}{2}\right) = 66 - 55 = 11$$

ordered pairs of row A – $(0,11)$, $(1,11)$, $(2,11)$, $(3,11)$, $(4,11)$, $(5,11)$, $(6,11)$, $(7,11)$, $(8,11)$, $(9,11)$, and $(10,11)$ – which are not in the same segment when we divide row A into the shared segments of rows A and B.

The minimum FRAGMENTATION value is 0, denoting maximal similarity. If two rows share only contiguous subsegments of length 1, then the value is 66, which is the maximum value. Any given row is maximally dissimilar to several rows, including its own retrograde. Since for any row A and RA, $FRAG(A,RA) = 11$, two rows belonging to the same row class may thus be maximally dissimilar.

It is straightforward to show that FRAGMENTATION defines a metric. First, the values are positive real values, and secondly, the four requirements of the metric are satisfied. (i) The value of $FRAG(X, X)$ is 0 for all rows. Since the rows are identical they do not need to be divided. (ii) Since FRAGMENTATION is based on shared segments, it is clearly symmetric and $FRAG(X, Y) = FRAG(Y, X)$. (iii) The following argument shows that triangle inequality

$$FRAG(X, Y) + FRAG(Y, Z) \geq FRAG(X, Z)$$

holds. Let X, Y, and Z be arbitrary twelve-tone rows. The left side of the above inequality denotes the sum of the ordered pairs that rows X and Y do not share (when they are divided into shared segments) and the ordered pairs that rows Y and Z do not share, and the right side denotes the ordered pairs that rows X and Z do not share. My claim is that if a pair appears on the right side it appears on the left side as well. If a pair appears on the right side it means that those pitch classes are not in a shared segment of rows X and Z, but if that pair does not appear on the left side it means that it is in a shared segment of rows X and Y and of rows Y and Z, which means that it is also in a shared segment of rows X and Z – which is a contradiction. Hence, the value on the left side must be greater than or equal to the value on the right side. (iv) Finally, if $FRAG(X, Y) = 0$, then with no divisions at all the two rows consist of identical subsegments and therefore rows X and Y must be identical. Since all four requirements of the metric are satisfied, FRAGMENTATION defines a metric.

The transformational approach to FRAGMENTATION

FRAGMENTATION has a natural interpretation as a measurement of a transformation. As observed in Section 10.6, the shared contiguous subsegments of rows X and Y correspond to the consecutive ascending segments in the order-number transformation $\boldsymbol{YX}^{-1}$.

Let us consider rows $A = 0123456789\text{AB}$, $B = \text{B}0123456789\text{A}$, and $C = 6789\text{AB}012345$. The order-number transformation that transforms row A into row B is $\mathbf{123456789AB0}$, which contains consecutive ascending segments $\mathbf{123456789AB}$ and $\mathbf{0}$. Correspondingly, the order-number transformation that transforms row A into row C is $\mathbf{6789AB012345}$, which contains consecutive ascending segments $\mathbf{6789AB}$ and $\mathbf{012345}$. As in the case of the DIVISIONS similarity measure, the shared contiguous segments of two rows correspond to consecutive ascending segments of the transformation that transforms one row into the other. Hence, the following transformational definition of FRAGMENTATION holds.

DEFINITION 10.13 In the GIS of order-number rows, the FRAGMENTATION value for twelve-tone rows X and Y is given by the formula

$$FRAG(X, Y) = 66 - \sum \frac{n_i(n_i - 1)}{2}$$

in which n_i are the lengths of the consecutive ascending segments of the transformation $\boldsymbol{YX}^{-1}$.

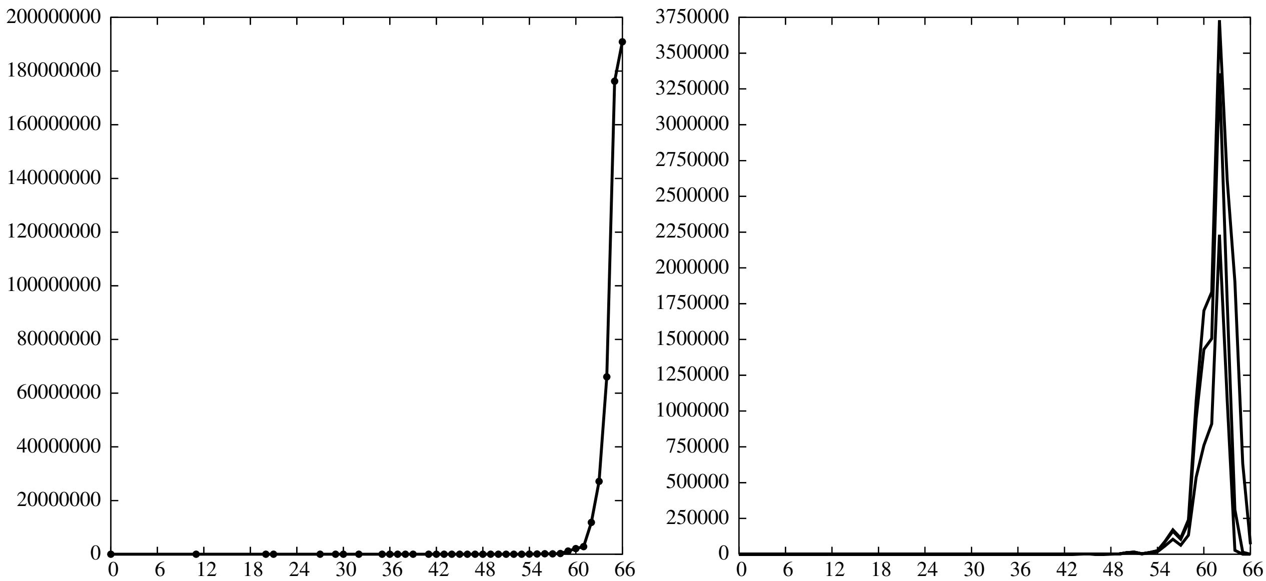


Figure 10.9: The distribution of FRAGMENTATION on the left and the distribution of FRAGMENTATION for row classes on the right. The former was obtained by computing the distances defined by all 479001600 distinct transformations, and hence is precise, while the latter was obtained by comparing 2000 random row classes to all other row classes. See Section 6.3 for a discussion on the three curves in the latter figure.

FRAGMENTATION for row classes

In order to show that FRAGMENTATION for row classes is well defined, we must show that it is transformationally coherent (see Corollary 6.1 in Section 6.2). FRAGMENTATION is a transformationally coherent similarity measure since the equation $FRAG(X, Y) = FRAG(FX, FY)$ holds for all row operations: transpositions, inversion, retrograde, and their combinations. The fact that this holds follows directly from the fact that DIVISIONS is transformationally coherent. It was shown in Section 10.6 that all row operations – transpositions, inversion, retrogression and their combinations – retain the number and length of the shared subsegments. Since the FRAGMENTATION value depends only on the number and length of the shared subsegments it is transformationally coherent and it defines a metric for row classes.

As far as the extended set of row operations is concerned, FRAGMENTATION is transformationally coherent under the same conditions as DIVISIONS. Hence, it is transformationally coherent under all pitch-class operations, such as the M-operation since, with respect to FRAGMENTATION, they constitute only a relabeling of the pitch classes. In general, FRAGMENTATION is not transformationally coherent under the order-number operations – retrograde being the only exception. In order to show that it is not transformationally coherent under rotation or the exchange operation we could use the same examples as given in Section 10.6.

The distribution of FRAGMENTATION

As illustrated in Figure 10.9, the distribution of FRAGMENTATION is highly skewed. The mean is 65 and the average is – curiously – $65\frac{1}{12!}$, both of which are very close to the maximum value of

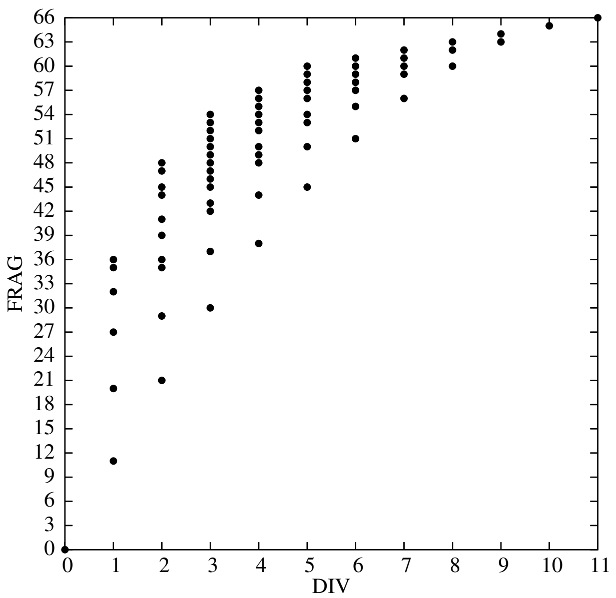


Figure 10.10: The correlation of DIVISIONS and FRAGMENTATION.

66. As noted in the discussion on DIVISIONS, any given row has no shared non-trivial contiguous subsegments with as many as 190899411 other rows. If two rows do not have a single shared non-trivial contiguous subsegment their shared contiguous subsegments do not contain a single ordered pair. Consequently, for any row, FRAGMENTATION returns the maximum value of 66 with as many as 190899411 other rows.

As with DIVISIONS, the unordered interval-class contents of rows provides a method for examining some boundary conditions for the FRAGMENTATION values for row classes. If two rows do not contain a single shared interval class, then the FRAGMENTATION value for the respective row classes is necessarily 66. For example, rows $0123456789\mathrm{AB}$ and $05\mathrm{A}3816\mathrm{B}4927$ do not contain a single shared interval class and therefore no two row forms of their respective row classes contain a shared contiguous subsegment.

Since the distribution of FRAGMENTATION is skewed, so is the distribution of FRAGMENTATION for row classes. As with the previous similarity measures, the distribution of the FRAGMENTATION values between row classes varies. However, as illustrated in Figure 10.9, due to the skewness the "room" for variance is limited.

FRAGMENTATION compared to DIVISIONS

The FRAGMENTATION similarity measure was devised as a weighted variant of the DIVISIONS similarity measure, hence it is to be expected that there is some correlation between them. Figure 10.10 depicts this correlation, which is obvious. We can, in fact, derive upper and lower limits for the FRAGMENTATION values from the DIVISIONS values. Given the factor $n_i(n_i-1)$ in Definition 10.12, the FRAGMENTATION value depends mostly on the longest shared segments of rows. Consequently, if there are k divisions in two rows, we can minimize the FRAGMENTATION value by making the longest shared segment as long as possible (the longest shared segment is of length

$12 - k$ and there are k segments of length 1) and, correspondingly, we can maximize it by making the longest shared segment as short as possible (making the lengths of all shared segments as close as possible).

For example, if the DIVISIONS value is 1, then the FRAGMENTATION value is low (11) when the shared subsegments are of lengths 11 and 1, and high (36) when they are both of length 6. Similarly, if the DIVISIONS value is 2, then the FRAGMENTATION value is low (21) when the shared subsegments are of lengths 10, 1, and 1, and high (48) when they are all of length 4.

10.8 Transformations in the subsegments approach

Transforming a row gradually into another one is a process in which in each step the rows resemble more and more the target row and less and less the original row. Given the focus on subsegments in this chapter, successive rows share more and more subsegments with the target row and the shared subsegments with the target row grow longer and longer.

Of the similarity measures discussed in this chapter, ULAM'S DISTANCE has a particularly interesting, even if somewhat surprising, interpretation as a transformation measurement. Since a transformation is the operation of transforming one row into another, it is possible to determine how many *moves* the most simple transformation entails. Consider rows $P = 5409728136\text{AB}$ and $S = 5406918237\text{AB}$ in Figure 7.3. The task is to transform row P into row S. Eight of the pitch classes – 5, 4, 0, 9, 2, 3, 10, and 11 – are already in the correct order (it is immaterial that these pitch classes are not in the same order positions in both rows). In other words, both rows P and S contain subsegment 540923AB. We need to move the remaining pitch classes to the right positions. Let us first interpose pitch class 6 from its current position in row P between pitch classes 0 and 9, then interpose pitch class 1 from its current position between pitch classes 9 and 7, then interpose pitch class 8 from its current position between pitch classes 1 and 7, and finally interpose pitch class 7 from its current position between pitch classes 3 and 10. We have thus transformed row P into row S in four moves. This process is depicted in Figure 10.11. Naturally, there are other sequences of moves that also transform row P into row S (for example, we might begin by moving pitch class 5 to the end of the row and then back again), but none is shorter than the one described above.

Let us now give a transformational definition of ULAM'S DISTANCE.

DEFINITION 10.14 In the GIS of order-number rows, the value of the ULAM'S DISTANCE similarity measure for rows X and Y is the minimum number of moves required to sort transformation $int(\boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{Y}\boldsymbol{X}^{-1}$.

The following lemma is borrowed from Diaconis (1988, 118). It connects the number of moves and the length of the shared subsegment and thereby shows that the non-transformational definition 10.5 and the transformational definition 10.14 of ULAM'S DISTANCE are equivalent.

LEMMA 10.1 The smallest number of moves to sort a permutation is the length of the permutation minus the length of the longest increasing subsequence in it.

P:	5	4	0	9	7	2	8	1	3	6	10	11
	5	4	0	6	9	7	2	8	1	3	10	11
	5	4	0	6	9	1	7	2	8	3	10	11
	5	4	0	6	9	1	8	7	2	3	10	11
S:	5	4	0	6	9	1	8	2	3	7	10	11

Figure 10.11: The transformation of row $P = 5409728136\text{AB}$ into row $S = 5406918237\text{AB}$ in four moves: first, interposing pitch class 6 between pitch classes 0 and 9 gives row 5406972813AB; second, interposing pitch class 1 between pitch classes 9 and 7 gives row 5406917283AB; third, interposing pitch class 8 between pitch classes 1 and 7 gives row 5406918723AB; finally, interposing pitch class 7 between pitch classes 3 and 10 gives row $S = 5406918237\text{AB}$.

Proof. See Diaconis (1988). □

The procedure described above is related to a family of sorting algorithms that could be put under the umbrella *insertion sort*.[7] There is a fundamental difference between this approach and the approach based on exchanges described in Section 9.4. Namely, if we reorder a sequence using exchanges, we always exchange the positions of two elements but the other elements remain in their positions. On the other hand, when we insert an element between other elements the positions of the other elements may change.

Let us consider again Figure 10.11. Pitch class 6 is at order position $\mathbf{9}$ and pitch class 9 is at order position $\mathbf{3}$ in row P. When we move pitch class 6 between pitch classes 0 and 9 the new order position of pitch class 6 is $\mathbf{3}$. In addition, the position of pitch class 9 also changes and it becomes $\mathbf{4}$. In fact, the positions of pitch classes 9, 7, 2, 8, 1, 3, and 6 change in this procedure.

Let us consider ULAM'S DISTANCE as moves in GIS terms. Since it is an ordering relationship, let us use the GIS of order-number rows. It was calculated in Section 7.2.3 that the interval from row P to row S in the GIS of order-number rows is $\boldsymbol{SP}^{-1} = \mathbf{0124976583AB}$. It is straightforward to see that the moves made above to transform row P into row S correspond directly to the moves that would sort the interval $\boldsymbol{SP}^{-1} = \mathbf{0124976583AB}$ into ascending order. The first move – moving pitch class 6 to between pitch classes 0 and 9 – corresponds to moving $\mathbf{3}$ between $\mathbf{2}$ and $\mathbf{4}$ in interval $\mathbf{0124976583AB}$: if we label the row obtained by moving pitch class 6 to between pitch classes 0 and 9 in row P as P', the interval in the GIS of order-number rows between rows $P' = 5406972813\text{AB}$ and $S = 5406728139\text{AB}$ is precisely $\mathbf{0123497658AB}$. All the remaining moves to transform row P into row S correspond similarly to the moves to sort the interval from

[7]This was not described as an exact algorithm, but was only given as a simple example: there are several algorithms that are related to the procedure. For a thorough description of the various algorithms, see Section 5.2.1 in Knuth (1998).

row P to row S in the GIS of order-number rows.

As noted in Section 10.4, the ULAM'S DISTANCE value is 11 for retrograde-related rows. Alternatively, we could now state that it takes a total of 11 moves to sort the order-number transformation $\mathbf{BA9876543210}$.

It is interesting to compare Figures 9.6, 9.7, and 10.11. All three give a sequence of transformations that transform row P into row S. Figure 9.6 shows a total of thirteen exchanges of adjacent pitch classes, all of which are small steps since the positions of only two pitch classes are changed at a time. However, pitch classes that are in their correct positions are exchanged, too. Figure 9.7 shows the exchange of (not necessarily adjacent) pitch classes – a procedure that is more clever since only the pitch classes that are not in their correct positions are changed. In Figure 10.11 there is no exchange of pitch classes at all – they are rather inserted. This procedure is efficient, but in each move the positions of several pitch classes are changed.

The above analysis help us to understand better the differences between the conceptions. Namely, if we consider the actual order positions of pitch classes to be more important than their mutual relations, then a transformational procedure that changes the positions of uninvolved pitch classes may not be the most appropriate. If, on the other hand, we consider only the mutual relations important, insertion algorithms provide an efficient transformational method.

It is now also possible to distinguish the measures that are truly linked to the transformations and those that are merely measurements of transformations. In other words, the values of certain similarity measures explicitly denote the number of steps required in order to transform a row step by step into another by using certain sorting algorithm. For example, the ORDER INVERSIONS value denotes the number of exchanges of adjacent pitch classes, the CAYLEY DISTANCE value denotes the number of exchanges of (not necessarily adjacent) pitch classes, and the ULAM'S DISTANCE value denotes the number of moves or insertions. The other similarity measures merely give an evaluation of how dissimilar two rows are, but do not offer a concrete algorithm of how to proceed with the transformation.

We can evaluate the effectiveness of the transformational procedures under ORDER INVERSIONS, CAYLEY DISTANCE, and ULAM'S DISTANCE by comparing their distributions. The average values of the three distributions are 33, 8.9, and 7.15, respectively. Hence, the quickest transformational procedure is via moves (ULAM'S DISTANCE) and the slowest is via exchanges of adjacent pitch classes (ORDER INVERSIONS).

Figure 10.12 depicts the correlation of the CAYLEY DISTANCE and ULAM'S DISTANCE values. The only correlation is that a large ULAM'S DISTANCE value implies at least a relatively large CAYLEY DISTANCE value. On the other hand, a small ULAM'S DISTANCE value does not guarantee a small CAYLEY DISTANCE value. Rotation r_1 is an extreme example. The ULAM'S DISTANCE value for rows related by rotation r_1 (for example rows $0123456789\mathrm{AB}$ and $\mathrm{B}0123456789\mathrm{A}$) is 1, but the CAYLEY DISTANCE value for these two rows is 11. In transformational terms, we obtain rotation r_1 with a single move (by moving the last pitch class as the first pitch class), whereas it takes a total of eleven exchanges of two pitch classes to obtain the same result.

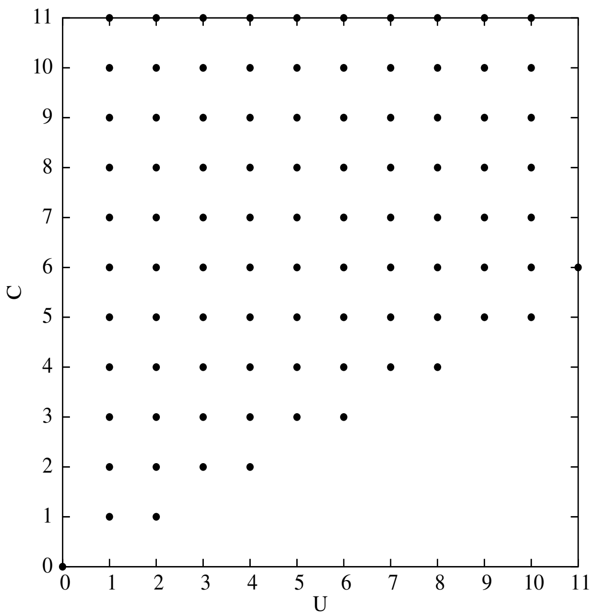


Figure 10.12: The correlation of the CAYLEY DISTANCE and ULAM'S DISTANCE values. The horizontal axis denotes the latter and the vertical axis the former values.

CHAPTER 11

Similarity measures based on subsets

The association of segments in two rows does not necessarily require that the pitch classes are precisely in the same order in the segments. A composer can create associations between rows by highlighting the subsegments that have identical (unordered) pitch-class contents.

Shared unordered subsets of two rows may be contiguous or non-contiguous. Non-contiguous unordered pitch-class sets are not particularly useful for the analysis of the similarity of twelve-tone rows since if we allow non-contiguous subsets, every twelve-tone row contains precisely the same unordered pitch-class sets. This type of research thus deals with the partitions of the (unordered) aggregate.[1] One way of applying this research to twelve-tone rows is through the concept of isomorphic partitioning, which was introduced by Haimo and Johnson (1984). According to this concept, related pitch-class sets are obtained by partitioning rows with the same partitioning scheme (a pattern of order positions). However, if we do not require the subsets to be contiguous, we are studying the properties of the aggregate rather than the properties of twelve-tone rows (although the order-position patterns may reveal something of the nature of the row structure).

The focus in this chapter is on the measurement of the similarity of twelve-tone rows based on the conception of a twelve-tone row as a set of (unordered) subsets. Naturally, it is a requirement that the subsets are contiguous.[2] Section 11.1 describes the conception in terms of a *nesting* that provides a method for addressing effectively and illustratively the shared subset contents of rows. The perspective here is broader than that of mere similarity since nesting is a relatively complex concept and its properties have not been fully addressed in the literature. This leads to the introduction in Section 11.2 of a similarity measure, NESTINGS, which is directly built on this concept. Section 11.3 concludes the chapter with a discussion of another similarity measure, SCATTERING,

[1] The first chapter of Alegant (1993) provides an excellent historical review of partitions.

[2] The most common approach to the analysis of rows based on unordered subset contents is that of combinatoriality. Other literature on combinatoriality includes Babbitt (1961b), Martino (1961), Babbitt (1974), Starr and Morris (1977), and Starr and Morris (1978). However, it is not similarity.

which involves the unordered adjacent dyad contents of rows.

The transformational approach has a minor role in this chapter. While it is possible to interpret both NESTINGS and SCATTERING in transformational terms, neither of them is directly linked to any transformational procedure.

11.1 Nestings

11.1.1 Background

David Lewin presents "segmental association" as a means of relating twelve-tone rows (Lewin 1962). Segments[3] with identical pitch-class contents provide a means of creating associations between twelve-tone rows. The locus classicus of such an association between two distinct row forms is the opening of Anton Webern's *Concerto for Nine Instruments* Op. 24 in which the three-note segments of the first two row forms $\mathrm{BA2376845019}$ and $\mathrm{2AB673548910}$ have identical pitch-class contents.[4]

Lewin motivates the study of the segmental association of rows as follows.

> Compositionally, we have observed that segments in common between two row forms can supply a natural basis for connecting the musical presentations of those row forms by "associative harmony." Structurally, we have observed that certain aspects of the internal structure of a row may be manifested through segmental relations with various other row forms; conversely, every segmental relation between two twelve-tone related row forms may be viewed as a manifestation of some properties of that internal structure which they, and all other rows related to them through any twelve-tone operation, share. (Lewin 1962, 96)

Lewin discusses relating two rows by a *nesting*, which consists of the shared contiguous subsets of two rows. A twelve-tone row contains a total of $1 + 2 + \ldots + 11 + 12 = 78$ unordered nonempty subsets (one subset of size 12, two subsets of size 11, etc.); the empty set is, of course, also a subset of any row, but as a trivial subset it is not of interest here.

Figure 11.1 reproduces part of Lewin's example 21 (Lewin 1962, 106). It shows a nesting created from rows $O_3 = \mathrm{01627934AB58}$ and $I_3 = \mathrm{54B3A8217609}$ from Schoenberg's *Concerto for Violin and Orchestra* op. 36. For example, row $\mathrm{01627934AB58}$ contains segment 1627 and row $\mathrm{54B3A8217609}$ contains segment 2176, and therefore the nesting defined by these two rows contains the unordered pitch-class set $\{1, 2, 6, 7\}$.

I have changed some of the notation in this example. In Lewin's original notation the expression $(D\flat\ G\flat\ D\ G)$ denotes an unordered pitch-class set, whereas I use the standard notation $\{1, 2, 6, 7\}$. In addition, the pitch-class sets "grow" upwards in the above example (through the tetrachord up

[3] Lewin uses the term segment to refer to an unordered collection of pitch classes. In the following I will use the term *segment* to refer to an ordered collection of pitch classes that are consecutive in a row, and the term *subset* to refer to the pitch classes of a segment as an unordered collection.

[4] In fact, the three-note segments are even more closely associated than by mere pitch-class contents since the corresponding three-note segments are related by retrograde.

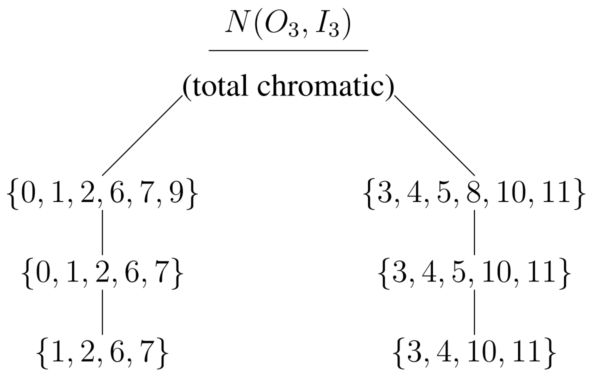


Figure 11.1: A nesting after Lewin's example 21. Here row O_3 is $01627934AB58$ and row I_3 is $54B3A8217609$ from Schoenberg's *Concerto for Violin and Orchestra* op. 36.

to the level of the same combinatorial hexachords in the two rows):[5] in Lewin's original example, pitch-class sets "grow" downwards. The reason for the change is that my example is a standard *Hasse diagram* – a well-known representation of partial orders – and I use the theory of partial orders in my treatment of nestings.

In the following I will define nestings formally in terms of partial orders and graph theory. The necessary concepts of partial orders are discussed in Appendix C and the pertinent concepts of graph theory are introduced in Appendix D.

My approach is more formal than that of Lewin: it offers answers to some questions that Lewin poses but leaves unresolved. Some of the answers given could also be obtained by employing the more visual approach adopted by Morris (2001), who discusses nestings in terms of invariance matrices.

Section 11.1.2 begins with a description of the general properties of nestings and how they can be created from rows. Sections 11.1.3 and 11.1.4 will follow the opposite path and discuss how nestings define rows and how uniquely they do so, and in Section 11.1.5 the discussion moves on to the creation of nestings from more than two rows.

11.1.2 Creating a nesting from rows

Before introducing the formal definition of a nesting we need to consider retrograde-related rows. Lewin notes that a row and its retrograde have the same subsets (Lewin 1962, 101). Consequently, we cannot distinguish between retrograde-related rows only by their subset contents. On the other hand, no two distinct rows that are not retrograde related have the same subset contents. This observation is worth stating as a simple theorem.

THEOREM 11.1 If two rows P and Q have the same subset contents, then they are either the same row or are retrograde related.

[5] It should be noted that even if nestings are "tree-like" and my initial example resembles a binary tree, nestings in general are not binary trees or even trees: they may have loops, and a "parent" may have more than two "children."

Proof. Assume that two rows P and Q have the same subset contents. Let us write

$$P = p_0p_1p_2p_3p_4p_5p_6p_7p_8p_9p_{10}p_{11}.$$

First, note that the rows have two subsets of size 11 that are composed of adjacent pitch classes: one that excludes the last pitch class and one that excludes the first pitch class. Since the rows have the same subsets of size 11 the two excluded pitch classes must be the same. Therefore there are two alternatives for row Q,

$$Q = p_0 \cdots\cdots\cdots p_{11} \quad \text{or} \quad Q = p_{11} \cdots\cdots\cdots p_0,$$

where the dots denote as yet unspecified pitch classes. Now, since row P has a subset $\{p_0, p_1\}$, row Q must have that subset. Similarly, since row P has a subset $\{p_{10}, p_{11}\}$, row Q must have that subset. Therefore, there are again two alternatives for row Q,

$$Q = p_0p_1 \cdots\cdots\cdots p_{10}p_{11} \quad \text{or} \quad Q = p_{11}p_{10} \cdots\cdots\cdots p_1p_0.$$

By continuing along the same lines we can show that rows P and Q are either identical or they are retrograde related. □

I will use the expression "modulo retrograde" to denote a row or row segment that is unequivocally defined except for the retrograde aspect; instead of row or row segment P we can have its retrograde RP.

The approach in the proof of the above lemma is used throughout the discussion. Typically, we have an existing segment and we need to discover what other pitch classes or segments must be adjacent to it.

Let us now define nestings formally.

DEFINITION 11.1 A nesting of two rows is a graph in which the nodes or vertices are the intersection of the nonempty subsets of the rows. Two nodes are connected if and only if one is a proper subset of the other and there is not an intermediate node – one that is a proper subset of the larger node and a proper superset of the smaller node.

We could use the notion of *partially ordered sets* or *partial orders* to analyze nestings (see Appendix C for an introduction to partially ordered sets). The partial order induced by the subset relation $\subseteq$ is well known in mathematics (van Lint and Wilson 2001, 53):[6] a proof of this is provided in Lemma C.5 in Appendix C.

I have defined nesting as a graph in which the vertices are connected if and only if one is a proper subset of the other and there is not an intermediate node. In the language of partial orders, the nodes of the graph are the members of the partial order (that is, the intersection of the subsets of

[6]I would remind the reader that in this work the symbol $\subset$ denotes a proper subset and the symbol $\subseteq$ allows for identity.

the rows) and the edges are the *cover relations*. Since the vertices of a nesting are the intersection of the subsets of the rows, it follows that the nesting of rows A and B is identical to that of rows B and A.

It is now convenient to introduce the following terminology. If two nodes are connected, the larger of them is called a *parent* and the smaller is called a *child*. Hence, a child is always a proper subset of its parent(s), children of the same parent are *siblings*, and a subset of cardinality 1 is called a *singleton*. For example, in Figure 11.1 pitch-class sets $\{0, 1, 2, 6, 7, 9\}$ and $\{3, 4, 5, 8, 10, 11\}$ are siblings and they are both children of the total chromatic.

Let us examine a concrete example of creating a nesting from the subsets of two rows. Table 11.1 enumerates the subsegments of rows $T_5Q =$ A95B07168234 and $r_4I_8Q =$ 5BA934821607 from the third movement of Alban Berg's *Lyric Suite*. The first step is to identify the shared subsets that the subsegments define (we obtain a subset from a subsegment by simply ignoring the order of the pitch classes in it). The shared subsets of rows T_5Q and r_4I_8Q are written in bold in Table 11.1. The next step is to arrange the subsets hierarchically and connect them to their supersets. Figure 11.2 shows the result as a Hasse diagram.

Some notes on Figure 11.2 are due. First, omitting the singletons would avoid extraneous clutter in the graph. Their inclusion does not bring any new information – here they are included only for the sake of completeness. Even if they are excluded from a nesting, they are implied. Secondly, for the sake of perspicuity the nodes of the nesting are written as segments from the two rows, thereby demonstrating how these segments were arrived at. They represent nevertheless unordered pitch-class sets.

Let us now examine some properties of nestings and formulate them as small lemmas. The first observation is almost trivial: a nesting is connected, in other words there is a sequence of edges that forms a path between any two vertices of the graph.

LEMMA 11.1 A nesting is connected.

Proof. The proof follows immediately from the observation that all nestings contain the aggregate as a node. If we can show that from every node there is a path to the aggregate, we have shown that the graph is connected.

Now, for a non-aggregate pitch-class set X in the graph there is always (since we are dealing with a finite number of pitch-class sets) at least one pitch-class set X' such that $X \subset X'$, and that for no pitch-class set Z the relation $X \subset Z \subset X'$ holds. Thus, pitch-class set X is connected to pitch-class set X'. If pitch-class set X' is the aggregate we have found the path, otherwise we continue to find a (proper) superset X'' to which X' is connected. Continuing similarly we finally reach the aggregate since X' is larger than X, X'' is larger than X', etc., and therefore the chain of supersets must end at the aggregate. Hence, we have found a path from an arbitrary non-aggregate subset to the aggregate, which shows that the nesting is connected. □

The nesting in Figure 11.2 is very orderly. On the informal level, there seems to be some regularity in the number of parents and in how the network is formed. While not all nestings are

size	subsets of row T_5Q = A95B07168234	subsets of row r_4I_8Q = 5BA934821607
12	**A95B07168234**	**5BA934821607**
11	A95B0716823, 95B07168234	5BA93482160, BA934821607
10	A95B071682, 95B0716823, 5B07168234	5BA9348216, BA93482160, A934821607
9	A95B07168, 95B071682, 5B0716823, B07168234	5BA934821, BA9348216, A93482160, 934821607
8	A95B0716, 95B07168, 5B071682, B0716823, **07168234**	5BA93482, BA934821, A9348216, 93482160, **34821607**
7	A95B071, 95B0716, 5B07168, B071682, 0716823, 7168234	5BA9348, BA93482, A934821, 9348216, 3482160, 4821607
6	A95B07, 95B071, 5B0716, B07168, **071682**, 716823, **168234**	5BA934, BA9348, A93482, 934821, **348216**, 482160, **821607**
5	A95B0, 95B07, 5B071, B0716, 07168, 71682, 16823, 68234	5BA93, BA934, A9348, 93482, 34821, 48216, 82160, 21607
4	**A95B**, 95B0, 5B07, B071, **0716**, 7168, **1682**, 6823, **8234**	**5BA9**, BA93, A934, 9348, **3482**, 4821, **8216**, 2160, **1607**
3	A95, 95B, 5B0, B07, 071, 716, 168, 682, 823, 234	5BA, BA9, A93, 934, 348, 482, 821, 216, 160, 607
2	**A9**, 95, **5B**, B0, **07**, 71, **16**, 68, **82**, 23, **34**	**5B**, BA, **A9**, 93, **34**, 48, **82**, 21, **16**, 60, **07**
1	**A**, **9**, **5**, **B**, **0**, **7**, **1**, **6**, **8**, **2**, **3**, **4**	**5**, **B**, **A**, **9**, **3**, **4**, **8**, **2**, **1**, **6**, **0**, **7**

Table 11.1: Subsegments of rows T_5Q = A95B07168234 and r_4I_8Q = 5BA934821607 from the third movement of Alban Berg's *Lyric Suite*. The subsegments whose unordered pitch-class contents are the same in both rows are written in bold.

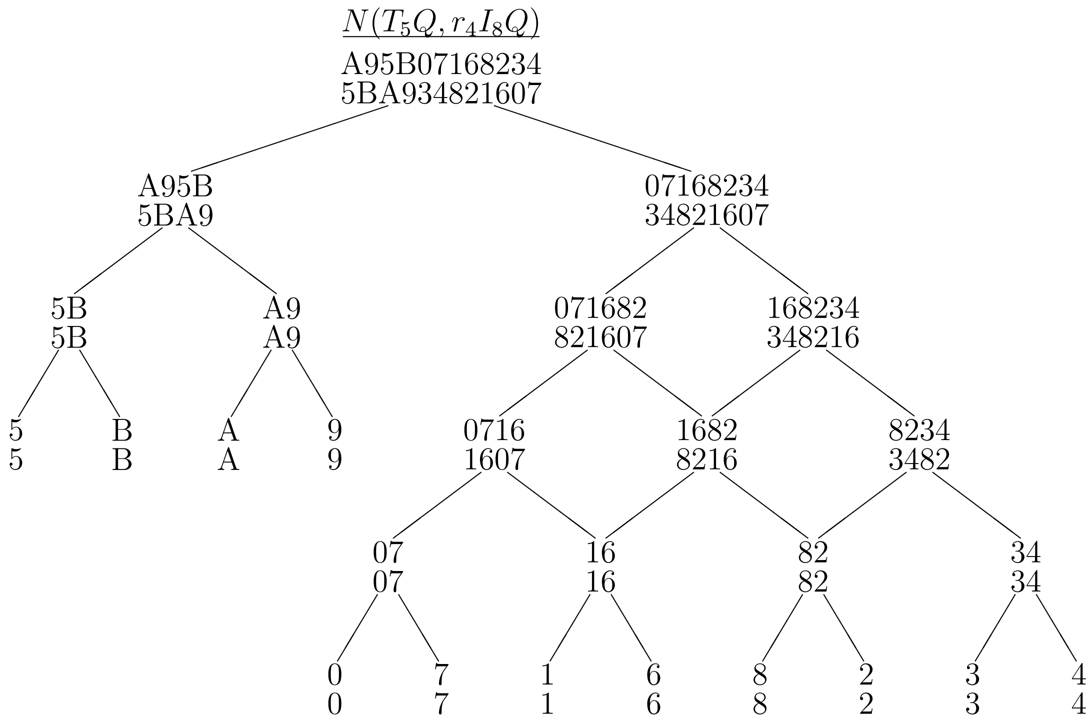


Figure 11.2: Nesting defined by two row forms from the third movement of Alban Berg's *Lyric Suite*: $T_5Q =$ A95B07168234 and $r_4I_8Q =$ 5BA934821607.

this "unsurprising," the following lemmas concern the relation of a parent and its children, and they explicate what we mean by orderliness and show that the orderliness of this particular nesting is not a coincidence.

LEMMA 11.2 If the parent is not a singleton, then the union of the children equals the parent.

Proof. Since the children are by definition proper subsets of their parent, it is evident that the union of children is a subset of the parent.

We should keep in mind the recursive structure of the nesting throughout the discussion: the top level is the aggregate. A child is "a new smaller aggregate." Therefore, it suffices to show that the union of the children of the aggregate equals the aggregate.

Assume now that there is a set of pitch classes C that do not belong to any of the children of the aggregate. Then, if a pitch class c is a member of set C then the singleton $\{c\}$ is a member of the nesting and, according to Lemma 11.1, it is connected to the aggregate. Then again, some child of the aggregate contains pitch class c, which cannot therefore belong to set C, against the assumption. Hence, set C is empty and the union of the children of the aggregate equals the aggregate. $\square$

A number of properties can be proved concerning the possible number of children. Let us first prove that all nodes except singletons have at least two children: I will return later to the case of nodes with more than two children.

LEMMA 11.3 A singleton has no children and the other nodes have at least two children.

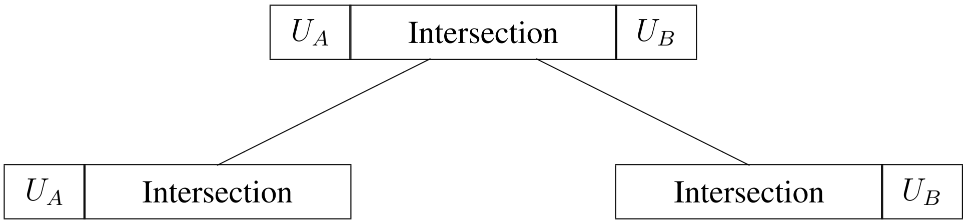


Figure 11.3: The mutual locations of the pitch classes in a row when a parent has two children whose intersection is not empty.

Proof. Both of the claims in this lemma follow from the fact that a child is always a proper subset of the parent. Hence, the size of the child is always smaller than the size of the parent. A singleton cannot have children since the only proper subset of a singleton is the empty set and this is not included in the graph. Since the children are smaller than the parent at least two children are needed to make the union of the children equal to the parent. □

LEMMA 11.4 If the intersection of any two children is not empty, then the parent has exactly two children, and the union of the two children equals the parent.

Proof. Assume that a parent has two children A and B whose intersection is not empty. Let us consider one of the rows generating the nesting and how the pitch classes are located within it. Let us divide the pitch classes into three disjoint sets: the (non-empty) intersection I, the unique pitch classes in one of the children U_A, and the unique pitch classes of the other children U_B.

First we claim that the pitch classes of the intersection are contiguous in both rows. The case $\#I = 1$ is trivial since the pitch classes in any segment of length 1 are contiguous. Let us then consider the case $\#I > 1$. The pitch classes $A = I \cup U_A$ are contiguous in A since they constitute a subset. Assume that the pitch classes of I are not contiguous: then in the row there is an ordered sequence of pitch classes $i_1\ x\ i_2$ where $i_1, i_2 \in I$ and $x \notin I$. Now, since pitch classes i_1 and i_2 belong to the intersection they belong to both A and B. But then, as A and B are contiguous row segments they must also contain x. This is a contradiction and therefore the pitch classes of the intersection are contiguous in both rows.

Since I, $I \cup A$ and $I \cup B$ are all subsets of the rows, the intersection must be in the middle, the unique pitch classes of A on one side of the intersection and the unique pitch classes of B on the other. This is illustrated in Figure 11.3. Hence, the pitch classes of $I \cup U_A \cup U_B$ are contiguous in both rows.

It is now straightforward to prove the claim in the lemma. I have shown above that $I \cup U_A \cup U_B$ is a subset of both rows. Hence, $I \cup U_A$ and $I \cup U_B$ are children of $I \cup U_A \cup U_B$ and not of any larger set. Furthermore, since the parent has two children and it is not a singleton, therefore according to Lemma 11.2 the union of the two children equals the parent. □

Lemma 11.4 formalizes and Figure 11.3 illustrates a feature of nestings that is prominent in Figure 11.2. We obtain the children by taking the "end segments" of the parent. For example, the

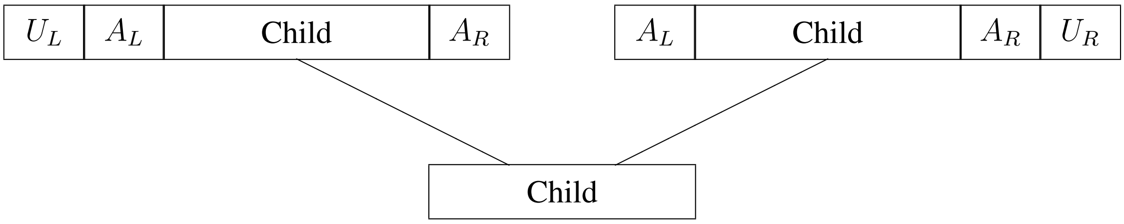


Figure 11.4: The positions of the pitch classes in a row segment when a child in the nesting has two parents.

two children of the top node of the nesting in Figure 11.2 are the four leftmost pitch classes and the eight rightmost pitch classes of rows A95B07168234 and 5BA934821607.

LEMMA 11.5 If a child has two parents it is the intersection of its parents.

Proof. Let us first assume that child C has two distinct parents that we label conveniently L and R (for left and right, respectively). Let us again consider the rows generating the nesting. Both parents contain the child, possibly some additional elements A_L and A_R on the left and right sides of the pitch classes of the child in one of the parent rows, and some elements U_L and U_R that are unique to the parents L and R. By definition, neither $L \subset R$ nor $R \subset L$, and hence sets U_L and U_R are nonempty.

Let us now proceed to prove that the pitch classes must be positioned in a row as depicted in Figure 11.4. First, the child must appear as a contiguous segment in both parents. Secondly, let us divide the possible additional elements into those that are to the left of the child A_L and those that to the right of the child A_R. Both parents have these. We now need to show that the unique elements U_L of L are to the left of A_L and the unique elements U_R of R are to the right of A_R (naturally, the choice of left and right is arbitrary). This follows naturally from the observation that the elements of L and the elements of R are contiguous: the only possibility is that the shared elements of L and R are contiguous and the unique elements U_L and U_R are on the left and right sides of the shared elements, respectively.

Finally, parents L and R cannot contain any additional shared elements A_L and A_R that are not in the child, since the shared elements are contiguous in L and R and therefore a node in the nesting. In this case, neither L nor R would be a parent of the child: their shared elements would be a node in the nesting and C would be a child of that node. Hence, $A_L = \emptyset = A_R$ and the child is the intersection of L and R. □

A note on the proof of Lemma 11.5 is in order. The nodes of a nesting are the shared subsets of *two* rows. Naturally, all nodes in a partition are the subsets that are shared in both rows. In the proof reference was made only to the subsegments of one row. However, what is claimed of the one row also applies to the other. For example, it was derived that the shared elements of L and R are contiguous in the row in question since L and R are nodes of a nesting. It thus follows that the shared elements of L and R are contiguous in the two rows of which the nesting is composed.

COROLLARY 11.1 A child has at most two parents.

Proof. Lemma 11.5 showed that if a child has two parents it is the intersection of the parents L and R, and the unique elements of the parents are to the left and right of the elements of the child in the row. If there were a third parent T we would have either $L \cap T$ or $R \cap T$ as a node in the nesting, in which case either L or R would not be a parent of the child, against the assumption. □

LEMMA 11.6 If a child has two parents then both have exactly two children.

Proof. From the proof of Lemma 11.5 we know that a child C is an intersection of its two parents L and R. Let us consider the right parent R. Since R is not a singleton it has at least two children and one of them is C. Let us assume that R has more than two children. According to Lemma 11.4, the children must be disjoint. However, as C is one child and the unique elements of R are contiguous in R, the assumption leads to a contradiction since the unique elements of R form a node in the nesting and therefore R has only two children. □

LEMMA 11.7 No parent has three children.

Proof. We know from Lemma 11.4 that if a node has more than two children, the children are disjoint. Assume then that a parent has three children A, B, and C. The children are contiguous in the parent and the union of the children equals the parent. Let A_1, B_1, and C_1 be the three segments corresponding to the three children in one of the rows generating the nesting. It can be assumed that the row is $A_1B_1C_1$. Let A_2, B_2, and C_2 be the three segments corresponding to the three children in the second row generating the nesting. There are six possible orderings of these three segments. If the second row is $A_2B_2C_2$ or $C_2B_2A_2$ then the parent has children AB and BC; if it is $A_2C_2B_2$ or $B_2C_2A_2$ then the parent has children BC and A; and if it is $B_2A_2C_2$ or $C_2A_2B_2$ then the parent has children AB and C. Thus, in none of the six orderings does a parent have three children. □

COROLLARY 11.2 A node of size three has at least one child of size two.

Proof. The corollary follows directly from the previous lemma if we set the subsets A, B, and C to be singletons. □

LEMMA 11.8 It is possible to construct such a pair of rows that in the nesting they define there is a node of a size larger than three that has only singletons as children.

Proof. The following algorithm generates examples of pairs of segments of any cardinality that have only singletons as children. Let the cardinality be $n > 3$. Then the only shared proper subsets of the two rows

$$a_1\ a_2 \ldots a_{n-1}\ a_n \quad \text{and} \quad \ldots a_{n-3}\ a_{n-1}\ a_1\ a_n\ a_{n-2}\ a_{n-4} \ldots$$

are singletons. This is because in the first row a non-singleton subset contains a pair a_ka_{k+1} but in the lower row such a pair must always contain a_1 and a_n, and hence a subset containing more than one pitch class must the whole set. □

The algorithm in Lemma 11.8 can also be used to organize four or more pairs of subsegments. If $n \geq 4$ and subsegments X_i and Y_i have the same unordered pitch-class contents for $1 \leq i \leq n$, then the node defined by the segments

$$X_1\ X_2 \ldots X_{n-1}\ X_n \quad \text{and} \quad \ldots Y_{n-3}\ Y_{n-1}\ Y_1\ Y_n\ Y_{n-2}\ Y_{n-4} \ldots$$

has n disjoint children.

Figure 11.2 serves to illustrate a recurring pattern in nestings. The two children of the aggregate are disjoint and the nesting divides into two branches down from the aggregate. In the left branch every node has only one parent whereas in the right branch some of them have two. However, even in the right branch every non-singleton node that has only one parent also has at least one child that has only one parent. It follows, then, that in every "generation" of the right branch there is at least one node with only one parent.

In the following, nodes that either do not have siblings or do not share pitch classes with any of their siblings will play a prominent role. It is convenient to label such nodes *non-intersecting*, and thus all nodes in the left branch of Figure 11.2 are non-intersecting.

The following results formalize the above observations.

LEMMA 11.9 Every non-singleton node that has only one parent has at least one child that has only one parent.

Proof. The lemma can be proved by induction starting from the aggregate. First, the children of the aggregate have, by definition, only one parent. If these children are disjoint, then their children have only one parent. If they are not, then according to Lemma 11.4 the aggregate has two children, both of which contain an end segment of the rows that form the nesting. In particular, they contain an end segment of a segment forming a non-intersecting node (since the aggregate does not have any siblings).

The end segment of a segment forming a non-intersecting node has a special property with respect to the parent count. Let us assume that segment $x_1x_2x_3 \ldots x_n$ is a segment defining such a node. Let us consider child of that node that contains pitch class x_1. This child is defined by segment $x_1 \ldots x_k$ (an end segment of $x_1x_2x_3 \ldots x_n$). Since x_1 is at one end of segment $x_1x_2x_3 \ldots x_n$ and the child is defined by some of its subsegment, it ensues that the child defined by segment $x_1 \ldots x_k$ necessarily has only one parent.

In general, if node S_1 shares pitch classes with its sibling S_2 and contains an end segment of a segment that forms a non-intersecting node, then it follows that S_1 has two children: the child that contains the end segment has only one parent (node S_1) and the other child has two (nodes S_1 and S_2). Thus, whenever there is a node with only one parent we know that it is either a non-intersecting node or that it contains an end segment of a segment that forms a non-intersecting node.

Assume now that there is a non-singleton node S that has only one parent. If the children of S are disjoint, all of them have only one parent. If the children of S are not disjoint, then the argument depends on S: either it is a non-intersecting node or it contains an end segment of a segment that forms a non-intersecting node. In the first case the same argument that was applied to the aggregate can be applied to node S. In the second case the parent of S has two children and of them at least node S contains an end segment of a segment that forms a non-intersecting node (whether the sibling of S contains an end segment of a segment that forms a non-intersecting node depends on whether the parent of S is a non-intersecting node or not). Of the children of node S one has only one parent and contains the end segment of S, and the second has two parents. Thus, in all cases a non-singleton node with only one parent has been shown to have a child with only one parent, which proves the lemma. □

DEFINITION 11.2 An *extended set of siblings* is a set of nodes formed by selecting one node, adding its siblings, adding the siblings of the newly added siblings, and continuing adding siblings until no new ones can be added.

For example, in the nesting in Figure 11.2 we can construct an extended set of siblings by starting with node $\{0\}$ and adding its sibling $\{7\}$. As node $\{7\}$ does not have siblings that are not siblings of node $\{0\}$, the two nodes $\{0\}$ and $\{7\}$ form an extended set of siblings. Similarly, we can construct an extended set of siblings by starting with node $\{0, 7\}$ and adding its sibling $\{1, 6\}$. We continue by adding node $\{2, 8\}$ (sibling of $\{1, 6\}$) and node $\{3, 4\}$ (sibling of $\{2, 8\}$). The extended set of siblings thus contains four nodes: $\{0, 7\}$, $\{1, 6\}$, $\{2, 8\}$, and $\{3, 4\}$.

It follows from the way the extended sets of siblings are constructed that belonging to the same extended set of siblings is a reflexive, symmetric and transitive relation. Thus, the extended sets of siblings define a partition of the nesting.

LEMMA 11.10 Any extended set of siblings contains a node with a single parent.

Proof. Let us assume that there is an extended set of siblings S_1 in which all nodes have two parents and show that this assumption leads to contradiction. The parents of the extended set of siblings S_1 also form an extended set of siblings, let it be S_2. Since the parents of an extended set of siblings always form an extended set of siblings, there is a sequence of extended sets of siblings $S_1, S_2, \ldots, S_k$, in which S_k is the aggregate. Since no node in set S_1 has only one parent it follows from Lemma 11.9 that no node in set S_2 has only one parent. Indeed, it follows that none of the nodes in sets $S_1, S_2, \ldots, S_{k-1}$ has only one parent, which is a contradiction since the aggregate is the only parent of the nodes in set S_{k-1}. Hence, the assumption that there is an extended set of siblings S_1 in which all nodes have two parents leads to contradiction, which proves the lemma. □

11.1.3 Creating rows from a nesting

Lewin discusses two aspects of a nesting. First, it shows the common subsets of two rows, and secondly it represents a “harmonic idea” (Lewin 1962, 100). A harmonic idea defines in the form of a graph a set of relations between subsets that two rows may share: hence, despite its generality, I will use the concept of a harmonic idea as a special term.

I defined a nesting as a graph in which the nodes are the intersection of subsets of the rows and the edges are the cover relations of the partial order induced by the subset relation (see Defini-

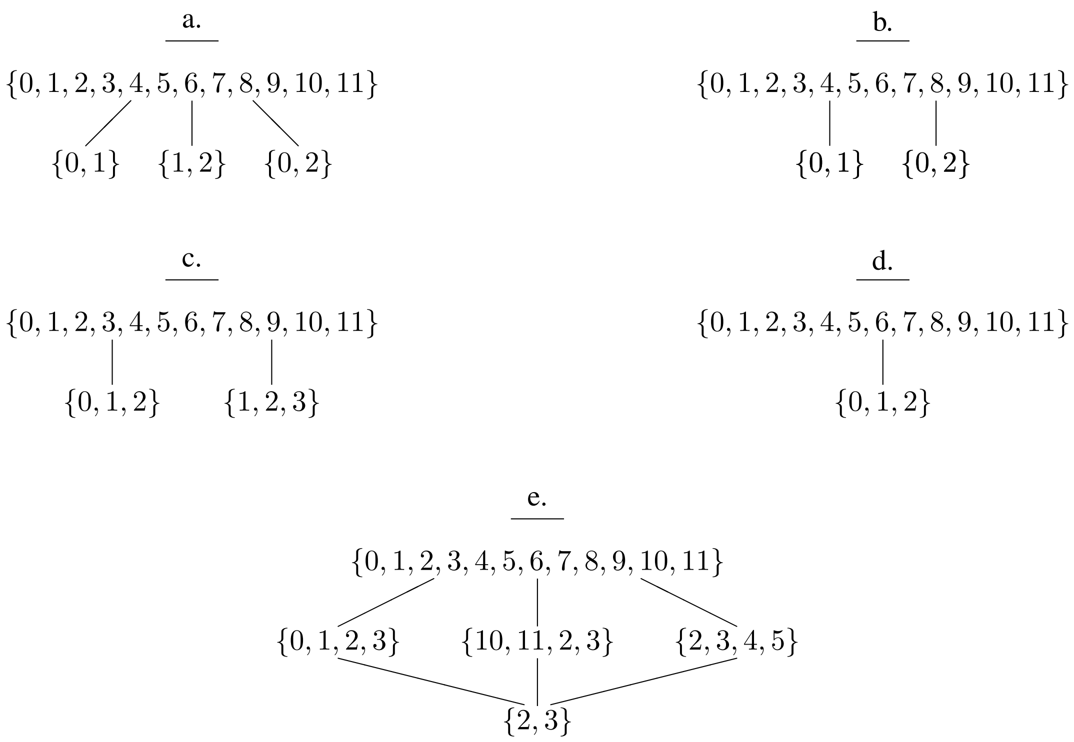


Figure 11.5: "Impossible" nestings from Lewin's example 23.

tion 11.1). A harmonic idea is a similar structure, except that the nodes can be chosen arbitrarily. The formal definition of a harmonic idea is very similar to that of a nesting, the only difference being in how the nodes are selected.

DEFINITION 11.3 A harmonic idea is a graph in which the vertices are (arbitrarily selected) pitch-class sets. Two vertices are connected if and only if one is a proper subset of the other and there does not exist an intermediate vertex (a pitch-class set that is a proper subset of the larger pitch-class set and a proper superset of the smaller pitch-class set).

The distinction between a harmonic idea and a nesting raises the question of under what conditions there exist two rows with shared subsets that are precisely those of a given harmonic idea. In other words, under what conditions is a harmonic idea a nesting? Lewin asks this question, but he is not able to give a definite answer. Instead, he provides five examples of harmonic ideas that cannot be formed from any two rows, and follows this with a discussion on why they do not define nestings (Lewin 1962, 110). Figure 11.5 reproduces these nestings as Hasse diagrams.[7]

For the sake of clarity, let us in the following denote by a harmonic idea a graph of related subsets, and by a nesting a harmonic idea that is satisfied by at least one pair of rows. Hence, a nesting is also a harmonic idea but a harmonic idea is not necessarily a nesting. In addition, a harmonic idea need not be complete.

[7]"Impossible nestings" are a phenomenon similar to the "impossible melodies" discussed by Polansky and Bassein (1992).

Some necessary conditions of nestings were set out in Section 11.1.2. Later on I propose later that these conditions are also sufficient. Let us now use them to examine Lewin's examples of harmonic ideas that are not nestings.

First, all five harmonic ideas fail to satisfy the condition concerning non-singleton nodes stated in Lemma 11.2 – i.e. that the union of the children should equal the parent. If we allow the omission of singletons in the graph, then harmonic idea d does not flout this condition. However, we might consider whether some of the harmonic ideas could be considered "incomplete" nestings, in other words whether by adding some nodes they could be turned into nestings. This might be desirable if our interest is only in establishing some relations and leaving other details undetermined.

Harmonic idea a fails to satisfy another condition: the intersection of two children, for example $\{0, 1\}$ and $\{1, 2\}$, is not empty, and therefore, according to Lemma 11.4, the parent should have only two children whereas in harmonic idea a it has three.

Let us consider the general case in which the mutual intersections of three children are all non-empty. Let us pick two children and label them A and B. Since their intersection is not empty we have the situation depicted in Figure 11.3. However, if the third child is to have common pitch classes with both A and B it has to be a subset of the intersection $A \cap B$, but then the third one is not a sibling of A and B, but it is their child.

Harmonic idea b only fails to satisfy the condition that the union of the children should equal the parent. Hence, harmonic idea b could be completed to create a nesting without removing any of the present nodes. In fact, the mere addition of a node $\{0, 1, 2\}$ and nine singletons would do the trick: the nodes $\{0, 1\}$ and $\{0, 2\}$ would then, of course, be children of the node $\{0, 1, 2\}$ and not of the aggregate. For example, rows 102A864B3579 and B1023456789A define such a nesting.

Harmonic ideas c and d both fail to satisfy the condition stated in Corollary 11.2 that nodes of cardinality 3 have at least one child of cardinality 2. In harmonic idea c the child would be $\{1, 2\}$. Both harmonic ideas could be completed to create a nesting. In addition, in harmonic idea c the parent of nodes $\{0, 1, 2\}$ and $\{1, 2, 3\}$ should be $\{0, 1, 2, 3\}$ and not the aggregate.

Finally, harmonic idea e has the same fault as harmonic idea a: it has three children but the intersection of at least two of them is not empty. While in harmonic idea a the intersection of the three nodes is empty but the intersection of any two children is not, in harmonic idea e the intersection of the three nodes is not empty.

In general, if some pitch class resides in three (or more) siblings, then the harmonic idea cannot be a nesting (assuming that none of the children is a subset of some other child). Informally, we could explain this by referring to Figure 11.3. The unique pitch classes of one child reside in a row on the left side of the intersection and the unique pitch classes of another child reside in a row on the right side. Where, then, could the unique pitch classes of the third child reside?

In the previous section a sequence of lemmas proved some necessary conditions of nestings. Let us now prove that these conditions are also sufficient.

THEOREM 11.2 It is possible to find two rows that define a given harmonic idea (that thus will be a nesting) if it satisfies the following criteria: (i) it is well formed (all proper subset relations are present in it, the children are proper subsets of their parents, and all nodes are distinct); (ii) the top node is the aggregate; (iii) it contains all twelve singletons; (iv) the harmonic idea satisfies Lemma 11.2 (if the parent is not a singleton, then the union of the children equals the parent), Lemma 11.3 (a singleton has no children, and the other nodes have at least two children), Lemma 11.4 (if the intersection of any two children is not empty, then the parent only has exactly two children, and the union of the two children equals the parent), Lemma 11.5 (if a child has two parents it is the intersection of its parents), Lemma 11.6 (if a child has two parents then both have exactly two children), Lemma 11.7 (no parent has three children), and Lemma 11.9 (every non-singleton node that has only one parent has at least one child that has only one parent).

Proof. Let us assume that we have an arbitrary harmonic idea that satisfies the four conditions of the theorem. The task is now to prove that this harmonic idea is a nesting, that is, that there are two such twelve-tone rows that the nesting they define is this harmonic idea. Conditions (i), (ii) and (iii) state that the harmonic idea is well formed and contains both the aggregate and the singletons. The proof is based on progressing step by step from the singletons towards the aggregate by identifying subnetworks and finding subsegments that define them.

Since the network is well formed, each node at a distance of n from the aggregate is connected to one or two nodes at a distance of $n-1$ from the aggregate. Every node except the aggregate has siblings (since according to Lemma 11.3 every node except the singletons have at least two children). Let us begin at the nodes that have the longest distance from the aggregate and start "pruning" the network. It will be characteristic of this process that the sets of nodes from which we commence the pruning will always be disjoint: the process starts from the singletons that are, by definition, disjoint and – as will be demonstrated below – the process will make it sufficient to start pruning subnetworks in which the siblings at the largest distance from the aggregate are disjoint.

In the process we will construct the rows subsegment pair by subsegment pair. Two issues need to be taken into account. First, we must ensure that the nesting that the rows (and their subsegments) define contains all the nodes in the harmonic idea. This will be accomplished by the way the segments are defined. Secondly, we must ensure that the nesting that the rows (and their subsegments) define contains no other nodes but those in the harmonic idea. Extra nodes might appear when two or more disjoint segments are joined without taking extra care. However, the process encompasses a method for guaranteeing that this will not happen: either by using the algorithm in Lemma 11.8 or by joining in such a way that the segments have the same subset at one end (since retrograde-related segments and rows have the same subset contents, it is immaterial at which end the shared subset is).

In the following, the subnetworks are selected as follows. We first pick one node S_1 that has the greatest distance from the aggregate. Then we find the extended set of siblings $S_1, S_2, \ldots, S_k$ to which node S_1 belongs.

For each node at a distance of $n > 0$ from the aggregate there are two possibilities with regard to its parent: (i) either all children of the parent have only one parent or (ii) some of the children have more than one parent. Both cases must be considered.

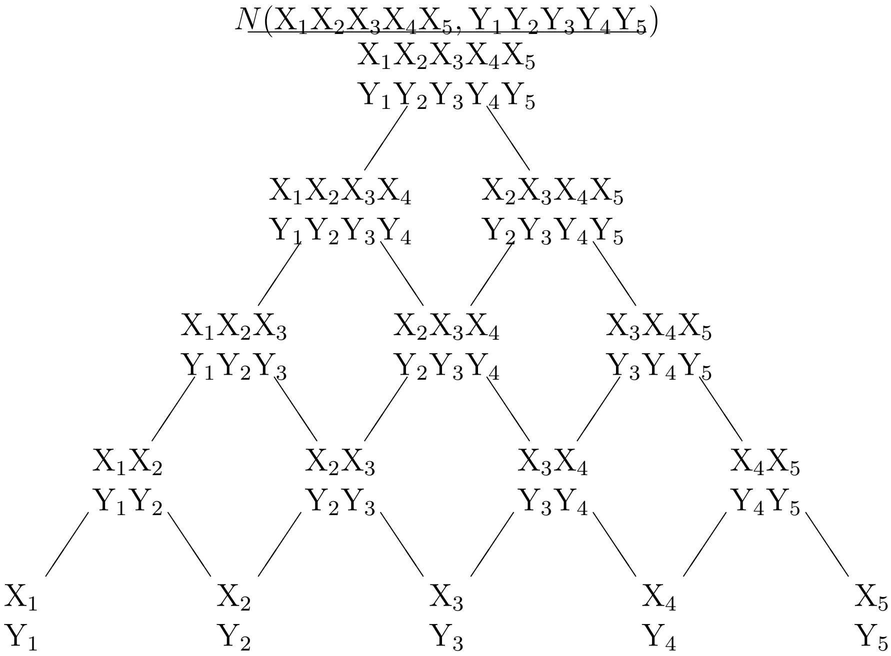


Figure 11.6: A harmonic idea in which the intersections of the siblings are nonempty.

Let us consider the first case. The nodes $S_1, S_2, \ldots, S_k$ have the same parent and they all have only one parent. Consequently, the parent is non-intersecting, that is, it does not share any pitch classes with its siblings (if node S_i had two parents then, according to Lemma 11.5, its parents would be siblings that share the pitch classes of node S_i).

It can be assumed that segments X_i and Y_i that define the nodes S_i for $1 \leq i \leq k$ have been found: they are either singletons or segments defined using the process. It can also be assumed that either the segments X_i and Y_i have the same subset at one end or they have been constructed by using the algorithm in Lemma 11.8. In the former case it is assumed that the shared subset of X_i and Y_i is at the left end of X_i and at the right end of Y_i.

Since the nodes $S_1, S_2, \ldots, S_k$ are children of a single parent, according to Lemma 11.7 either $k = 2$ or $k \geq 4$. If $k = 2$, then the parent is $S_1 \cup S_2$ and the segments X_1X_2 and Y_1Y_2 define the desired subnetwork. They have the same subsets at the (left) end and they do not contain additional shared subsets neither here nor in the next iteration of the process (neither when there are 2 disjoint nodes to join nor when there are 4 or more disjoint nodes to join). If $k \geq 4$, the algorithm in Lemma 11.8 provides a method for generating a segment pair that does not contain shared subsets except those of $S_1, S_2, \ldots, S_k$.

Let us then consider the second case: at least one of the nodes $S_1, S_2, \ldots, S_k$ has two parents (nodes $S_1, S_2, \ldots, S_k$ are all disjoint). In this case, the subnetwork will necessarily be like that in Figure 11.6. Let us examine why.

It follows from Lemmas 11.9 and 11.10 that at least one of nodes $S_1, S_2, \ldots, S_k$ has only one parent. Let node S_1 be such a node. The parent of S_1 has at least one other child, let it be S_2. Now, if the parent does not share subsets with any of its siblings, then there are only two nodes S_1 and S_2 in the extended set of siblings and we could proceed as in the first case. Hence, since we assumed that at least one of the nodes $S_1, S_2, \ldots, S_k$ has two parents, it follows that S_2 must have two parents. The nodes $S_1, S_2, \ldots, S_k$ are all disjoint, and according to Lemma 11.2 the parents are unions of their children and according to Lemma 11.5 the children with two parents are intersections of their parents. Hence, the parents of S_2 are $S_1 \cup S_2$ and $S_2 \cup S_3$ (and node S_2 is the intersection of its parents $S_1 \cup S_2$ and $S_2 \cup S_3$. By proceeding similarly, we can show that the network will necessarily be like that in Figure 11.6: if there are k nodes $S_1, S_2, \ldots, S_k$ at a distance of n from the aggregate, there are $k-1$ nodes $S_1 \cup S_2, S_2 \cup S_3, \ldots, S_{k-1} \cup S_k$ at a distance of $n-1$ from the aggregate and so on, until we reach the node $S_1 \cup S_2 \cup S_3 \ldots \cup S_k$ at a distance of $n-k+1$ from the aggregate.

The necessary segments are X_1X_2 and Y_1Y_2 and they have the same subset at one end. Moreover $X_1X_2 \ldots X_k$ and $Y_1Y_2 \ldots Y_k$ have the same segment at one end. If segments X_i are joined in such a way that the subsets are to the left, and segments Y_i are joined in such a way that the subsets are to the right, then the combined segments $X_1X_2 \ldots X_k$ and $Y_1Y_2 \ldots Y_k$ will not have extra subsets.

The two cases have now been verified and the theorem is thus proved. □

Another question Lewin poses concerns the conditions under which it is possible to find transpositionally or inversionally related rows that define a given nesting (Lewin 1962, 112). Naturally, the nesting must satisfy the conditions of Theorem 11.2. We could also impose some further necessary conditions.

There is some guidance in this in some of the formulas that Lewin provides (Lewin 1962, 102–103):

$$
\begin{aligned}
&T(N(O,O')) = N(T(O),T(O'))\\
&I(N(O,O')) = N(I(O),I(O'))\\
&I(N(O,I(O))) = N(O,I(O))\\
&T_6(N(O,T_6(O))) = N(O,T_6(O)).
\end{aligned}
$$

These formulas concern transposed or inverted nestings. The first two could be read as stating that a transposed or inverted nesting equals the nesting composed of transposed or inverted rows. The third and fourth are a direct application of the first two, and rely on the fact that transposition T_6 and all inversions are involutions: applying transposition T_6 to the unordered pair of rows O and T_6O gives the very same rows, and applying inversion I_k to the unordered pair of rows O and I_kO gives the very same rows.

It turns out that it is easier to formulate a condition for inversionally related rows. Let us start by examining the shared subsets of rows $T_5Q = \mathrm{A95B07168234}$ and $I_8Q = \mathrm{348216075BA9}$ shown in Table 11.2. An examination of the shared pitch-class sets reveals that either they are invariant under I_1 or they come in pairs that are related by I_1. For example, of the shared tetrachords $\{0,7,1,6\}$ is invariant under I_1, tetrachords $\{\mathrm{A},9,5,\mathrm{B}\}$ and $\{8,2,3,4\}$ are related by I_1 and so are tetrachords $\{5,\mathrm{B},0,7\}$ and $\{1,6,8,2\}$.

This is not a coincidence. Assume that rows P and I_kP share the subset $\{p_0, p_1, \ldots, p_n\}$. Then

segment of $T_5Q =$ A95B07168234	segment of $I_8Q =$ 348216075BA9
A95B07168234	348216075BA9
A95B071682, 5B07168234	348216075B, 8216075BA9
A95B0716, 5B071682, 07168234	34821607, 8216075B, 16075BA9
A95B07, 5B0716, 071682, 168234	348216, 821607, 16075B, 075BA9
A95B, 5B07, 0716, 1682, 8234	3482, 8216, 1607, 075B, 5BA9
A9, 5B, 07, 16, 82, 34	34, 82, 16, 07, 5B, A9

Table 11.2: Shared nontrivial subsets of the two inversionally related rows $T_5Q =$ A95B07168234 and $I_8Q =$ 348216075BA9.

row P has those pitch classes in some order in order positions $\{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_n\}$, and I_kP has them in some order in order positions $\{\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_n\}$. If the two sets of order positions are identical it follows that the subset $\{p_0, p_1, \ldots, p_n\}$ is invariant under I_k. If the two sets of order numbers are not identical, then row P has some pitch classes $\{q_0, q_1, \ldots, q_n\}$ in some order in order positions $\{\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_n\}$ and I_kP has those pitch classes in some order in order positions $\{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_n\}$. It follows now that the pitch-class sets $\{p_0, p_1, \ldots, p_n\}$ and $\{q_0, q_1, \ldots, q_n\}$ are related by I_k, and both are members of the nesting.

The above reasoning relied on the fact that inversions are involutions, in other words for any inversion I_k, $I_kI_k = T_0$. A similar reasoning applies to nestings composed of rows related by transposition T_6. However, less can be said about the other transpositions. Assume that rows P and T_kP share the subset $\{p_0, p_1, \ldots, p_n\}$. Then row P has those pitch classes in some order in order positions $\{\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_n\}$, and T_kP in order positions $\{\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_n\}$. However, if the transposition interval is not 6 (or 0) nothing can be said about whether the pitch classes of P in order positions $\{\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_n\}$ are shared or not.

Nevertheless, it is possible to comment on the nesting of a row and transposition T_k, and on the nesting of a row and its complementary transposition T_{12-k}. Let us first take an example. Figure 11.7 shows the nesting of rows T_5Q and $T_{10}Q$ and of rows T_5Q and Q. The two nestings have identical structures. That this must be so becomes obvious given that all subsets of the nesting on the right are obtained by transposing subsets of the nesting on the left by T_7. Furthermore, if we transpose rows T_5Q and $T_{10}Q$ by T_7 we obtain rows T_0Q and T_5Q. All this is based on the simple observation that if pitch-class set A is a subset of pitch-class set B then pitch-class set T_nA is a subset of pitch-class set T_nB.

The following lemma summarizes the relation between a pair of rows and a nesting with the rows transformed by some pitch-class operation. It follows directly from Lewin's observation that a transformation of a nesting equals the nesting of transposed or inverted rows (see page 212).

LEMMA 11.11 The nesting defined by rows A and B is isomorphic to the nesting defined by rows FA and FB where F is any pitch-class operation.

It should be noted that the nesting defined by rotated rows is not necessarily isomorphic (structurally identical) to the original nesting. Figure 11.8 shows the nesting defined by rows $r_2T_5Q =$

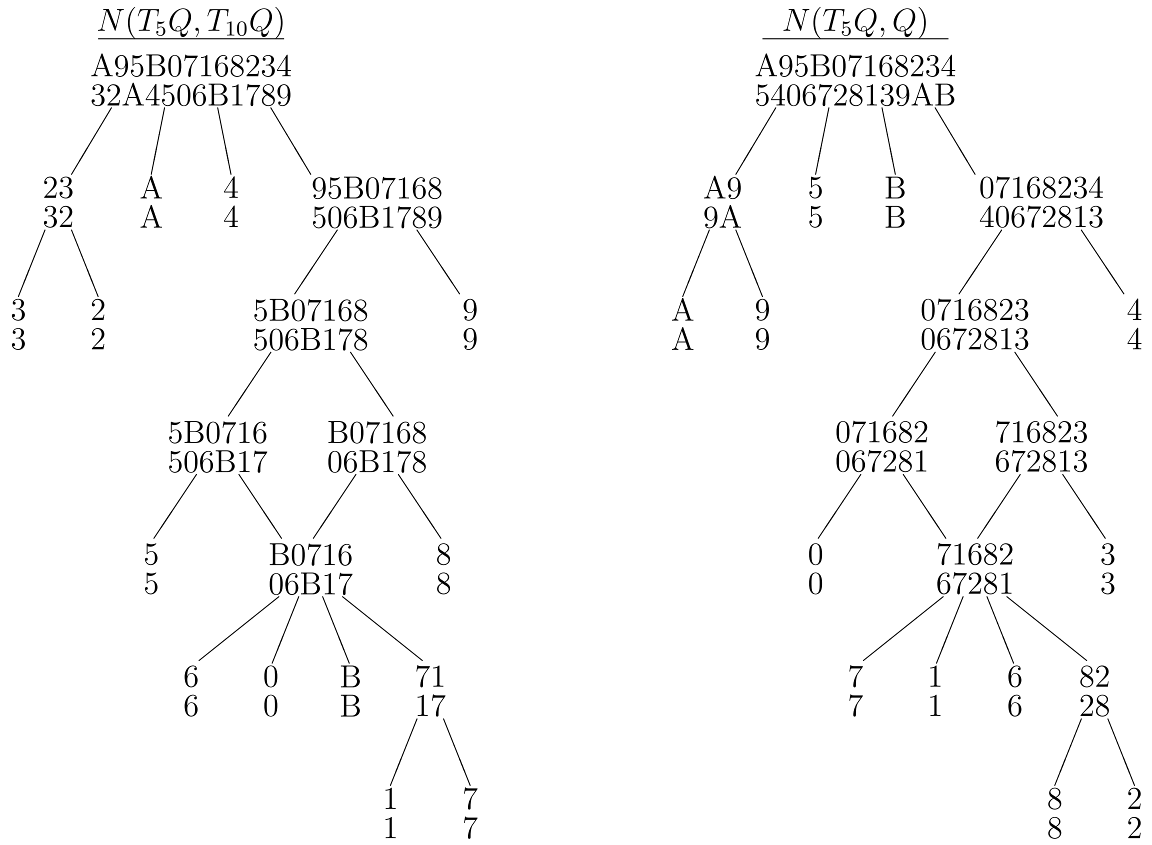


Figure 11.7: Two transpositionally related nestings. The nesting on the left is that of rows Q and T_5Q and the nesting on the right is that of rows Q and T_7Q.

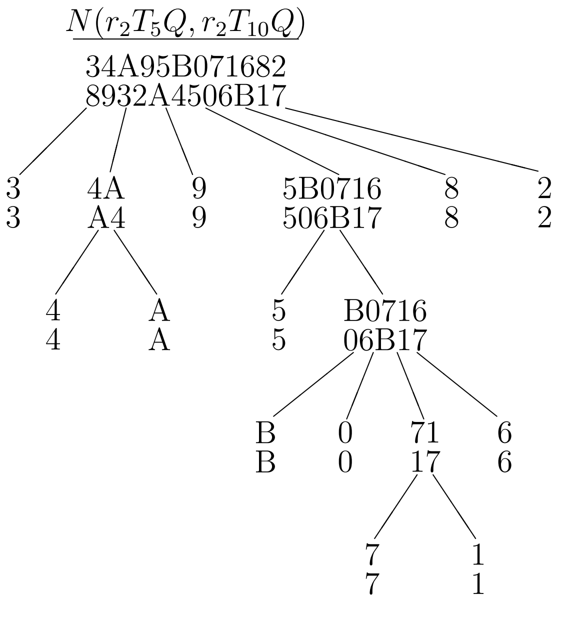


Figure 11.8: The nesting of rows $r_2T_5Q = 34\text{A}95\text{B}071682$ and $r_2T_{10}Q = 8932\text{A}4506\text{B}17$.

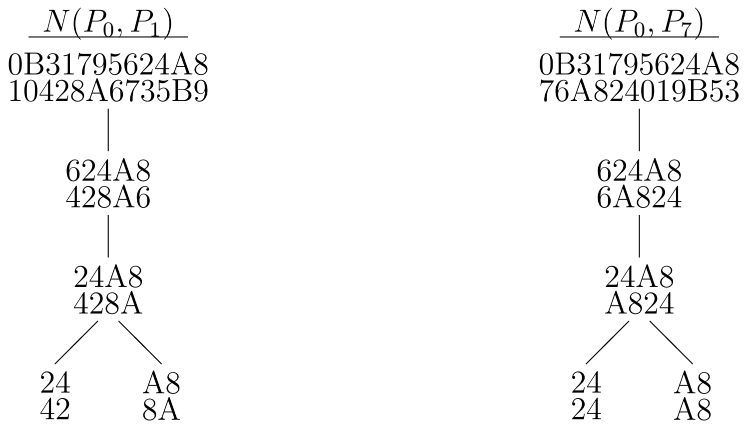


Figure 11.9: Two nestings of two pairs of rows from Schoenberg's *Fantasy for Violin and Piano* op. 47: the singletons have been omitted.

34A95B071682 and $r_2T_{10}Q = 8932\text{A}4506\text{B}17$. By comparing the nesting defined by the rotated rows r_2T_5Q and $r_2T_{10}Q$ to that defined by the unrotated rows T_5Q and $T_{10}Q$ in Figure 11.7 it is easy to see that the nestings are not isomorphic.

11.1.4 Nesting and uniqueness

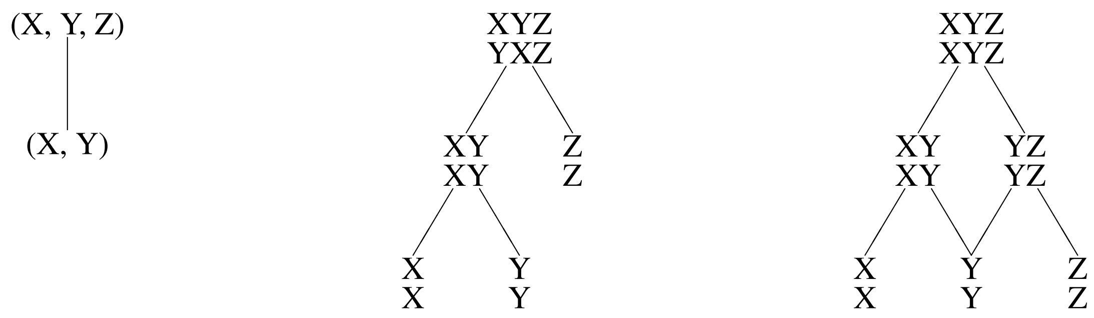
A further question arises concerning the conditions under which a nesting defines a pair of rows unequivocally. As already noted in Theorem 11.1, two rows related by retrograde are undistinguishable since they have precisely the same subsets. Hence, we need to consider under what conditions a nesting defines a pair of rows unequivocally modulo retrograde.

Figure 11.9 provides an example in which a nesting does not unequivocally define a pair of rows. Furthermore, even within the same row class there are two pairs of rows that define the very same nesting: the rows in Figure 11.9 are all transpositions of the row of Schoenberg's *Fantasy for Violin and Piano* op. 47.[8]

Given the uniqueness of nestings it is also crucial to be aware that a nesting also defines what is *not* in it. What I have called a harmonic idea may omit something. Therefore it makes a major difference whether or not a harmonic idea defines a pair of rows uniquely, and whether or not a nesting defines a pair of rows uniquely.

Figure 11.10 illustrates the difference between a harmonic idea and a nesting. Figure (a) shows a three-note harmonic idea or a nesting with the singleton nodes omitted. On the one hand, if we consider it as a harmonic idea, both nestings (b) and (c) are examples of pairs of rows that can be constructed from it. Clearly then, figure (a) as a harmonic idea does not uniquely define the rows. On the other hand, if we consider it as a nesting it does define a pair of rows uniquely (modulo retrograde). The nesting in figure (c) contains an extra node (Y, Z); figure (a) as a nesting excludes

[8]The rows 0162A954B378 from Schoenberg's *A Survivor from Warsaw* op. 46, 0B674589A123 from Webern's *Three Traditional Rhymes* op. 17 no 1, 0987BA456321 from Webern's *Three Traditional Rhymes* op. 17 no 3, and 0B58A9341726 from Webern's *Three Songs* op. 18 no 1 provide further examples of cases in which nestings do not unequivocally define a pair of rows within a row class.



(a) (b) (c)

Figure 11.10: Harmonic idea (a), which could also be interpreted as a nesting, and two nestings (b) and (c). In addition to the singletons, nesting (b) contains only the nodes that are in (a), and nesting (c) contains an additional node (Y, Z).

that node and, for that matter, any other node not in it (except the omitted singletons). In sum, the two three-note rows in figure (b) are the only two rows (modulo retrograde) that define the *nesting* in figure (a), but the row pair in figure (b) and in figure (c) could be generated from the *harmonic idea* in figure (a).

Let us next prove some lemmas on uniqueness. We are constantly operating "modulo retrograde." If some parent in a nesting does not uniquely define the row segments that define its children, then the pair of twelve-tone rows that define the complete nesting is not unique.

First of all, let us assume that there is a nesting. Consequently, it has to satisfy the conditions enumerated in Theorem 11.2: in particular, a parent of cardinality 3 must have at least one child of cardinality 2. If we have a parent of cardinality 3 without the children being specified, we have a harmonic idea and not a nesting.

LEMMA 11.12 If a parent has two disjoint non-singleton children A and B then the nesting does not define the rows unequivocally.

Proof. Let A_1 and A_2 be two orderings of A, and let B_1 and B_2 be two orderings of B. Assume that the pair of rows A_1B_1 and A_2B_2 generates the pertinent nesting. Observe first that orderings A_1 and A_2 cannot be identical and orderings B_1 and B_2 cannot be identical. We could then conclude that the pair of rows $A_1R(B_1)$ and $A_2R(B_2)$ generates the same nesting as A_1B_1 and A_2B_2. Hence, we have found two distinct pairs of rows that generate the same nesting, which proves the lemma. □

Let us take an example. Let the parent be $\{a, b, c, d\}$ and let the children be $\{a, b\}$ and $\{c, d\}$. Then rows $abcd$ and $badc$ generate the nesting, but so do rows $abdc$ and $bacd$. None of these rows are related by retrograde. In addition, the segments defining the children are the same modulo retrograde: one child is composed of segments A and $R(A)$ and the other is composed of segments B and $R(B)$.

The requirement that the children are not singletons is essential since, trivially, the retrograde of a singleton equals the prime form. Furthermore, the proof relies on the observation that A_1 and A_2

cannot be identical and B_1 and B_2 cannot be identical. If, for example, $A_1 = A_2 = a_1a_2 \ldots a_n$, then the children of rows $A_1B_1 = a_1a_2 \ldots a_nB_1$ and $A_1B_2 = a_1a_2 \ldots a_nB_2$ would not be disjoint: one of them would be the node defined by segments $a_2 \ldots a_nB_1$ and $a_2 \ldots a_nB_2$ and the other would contain at least the pitch classes $a_1a_2 \ldots a_n$.

If a parent has two disjoint children A and B then the two segments of which A is generated must be such that if a node contains any of the pitch classes of A and of B then all pitch classes of A must be in the node.

LEMMA 11.13 If a nesting has a parent with four (or more) singleton children then it does not define a pair of rows unequivocally.

Proof. Let us consider the case of four singleton children. If the parent has children a, b, c, and d then the row pair $abcd$ and $bdac$ has four singleton children and so does the row pair $cbad$ and $bdca$. A corresponding example can be constructed in the case of more than four singleton children. □

In general, the more nodes a nesting has the fewer the pairs of rows that generate it. The issue of uniqueness is not a mere function of the number of nodes, however, but also includes their configuration. For example, rows $Q = 5406728139\text{AB}$ and $I_5Q = 015\text{BA}3942876$ define a nesting with 21 nodes (including the singletons), and rows $Q = 5406728139\text{AB}$ and $I_3Q = \text{A}95\text{B}07168234$ define one with 20 nodes (including the singletons). The former is a larger nesting and therefore we expect it to be more "selective." However, there are 264960 distinct pairs of rows (5760 distinct rows) that generate the former nesting but only 36864 distinct pairs (4608 distinct rows) that generate the latter.

Correspondingly, the same number of nodes in a nesting does not guarantee that the same number of row pairs define it. To take an extreme example, rows 0123456789AB and $6723\text{AB}014589$ define a nesting with 19 nodes (including the aggregate and the singletons), as do rows 0123456789AB and $0\text{A}352684971\text{B}$. A total of 16957440 distinct pairs of rows (different combinations of 23040 distinct rows) generate the former, while a total of 117976320 distinct pairs of rows (different combinations of 80640 distinct rows) generate the latter.

11.1.5 Nestings of more than two rows

Since the concept of a nesting was formalized in terms of graph theory and the inclusion relation between pitch-class sets, it is easy to make a generalization concerning the nesting of any number of rows. Given any number of rows, the nesting defined by those rows has the shared subsets as its nodes, and the vertices are the cover relations of the inclusion relation. For example, Figure 11.11 shows the nesting defined by four rows: $Q = 5409728136\text{AB}$, $r_3T_7Q = 5406728139\text{AB}$, $r_9T_{10}Q = 5406918237\text{AB}$, and $r_4I_8Q = 5\text{BA}934821607$ from Alban Berg's *Lyric Suite*. The motivic status of the tetrachord $\{5, 9, 10, 11\}$ in the third movement is well represented by this nesting, which shows that, in general, the four rows share very few segments.[9] However, the tetrachord $\{5, 9, 10, 11\}$ and its complement octachord $\{0, 1, 2, 3, 4, 6, 7, 8\}$ divide the nesting into two distinct branches.

[9] See, for example, Headlam (1996, 262) for a discussion on the motivic status of the tetrachord $\{5, 9, 10, 11\}$.

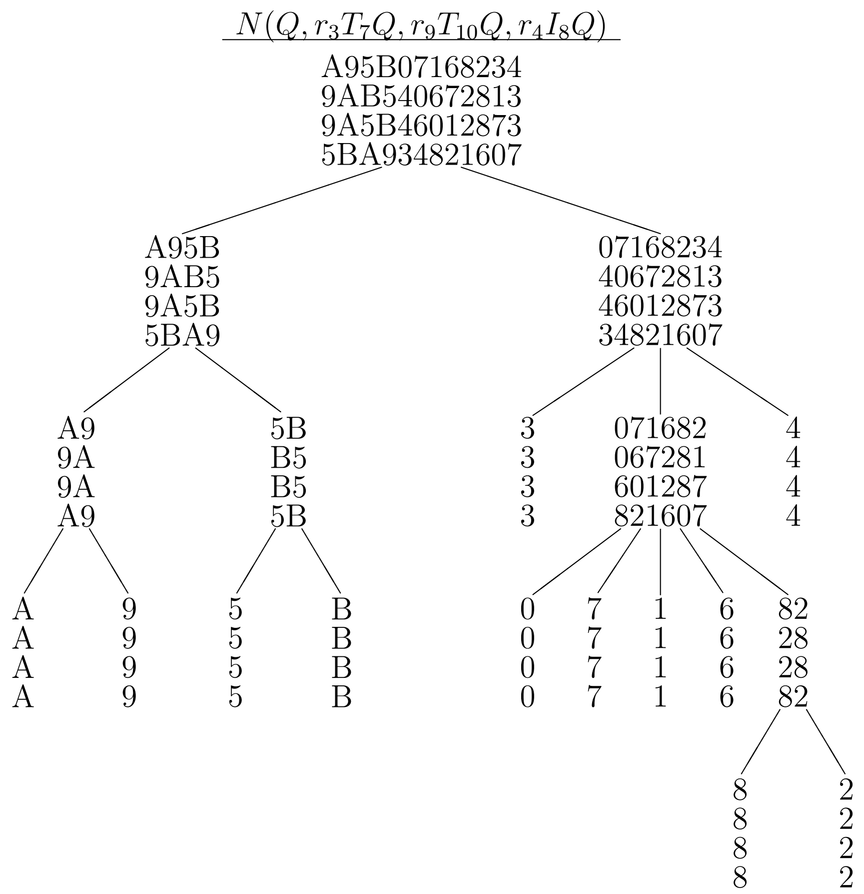


Figure 11.11: The nesting defined by rows T_5Q = A95B07168234, r_3Q = 9AB540672813, r_9T_3Q = 9A5B46012873, and r_4I_8Q = 5BA934821607, where Q = 5406728139AB is the row of the third movement of Alban Berg's *Lyric Suite*.

Naturally, some of the lemmas proved with regard to nestings generated by two rows do not apply to those generated by three or more rows. For example, the nesting generated by the three-note rows abc, bca, and cab has the node $\{a, b, c\}$, which has three singletons as its children.

Nestings of several rows could be used to examine how closely related a set of rows is. For example, *Lyric Suite* uses four different rows, as shown in Figure 7.3. The nesting of the first three rows has 34 nodes (of which 12 are singletons and one is the aggregate). If we add the fourth row the nesting contains only 16 nodes (of which 12 are singletons and one is the aggregate). Hence, the first three rows contain relatively many shared subsets whereas all four rows contain very few.

11.2 The NESTINGS similarity measure

Lewin notes that the discussion could be extended to the analysis of the relations between arbitrary rows even if it mostly involves rows belonging to a single row class. Hence, even if he does not discuss the segmental association of nestings in terms of similarity, it is a natural consideration.

Given the conception of a twelve-tone row as a set of subsets it is natural to measure the similarity of rows based on the number of subsets they share. The more subsets two rows share, the larger is the nesting that they define, and the more similar the two rows are (with respect to subset contents). The formal definition of the NESTINGS similarity measure, or $NEST$, is given below.

DEFINITION 11.4 The value of the NESTINGS similarity measure for twelve-tone rows X and Y is

$$NEST(X,Y) = 78 - \#N(X,Y)$$

in which $N(X,Y)$ is the nesting defined by rows X and Y and $\#N(X,Y)$ is the number of nonempty nodes in it (including the nodes of size 1).

A nesting contains all the shared contiguous unordered subsets of two rows (except the empty set); there are 78 such sets in a row. Consequently, NESTINGS is defined simply as the number of contiguous unordered subsets in one row that are not contiguous unordered subsets in the other. Hence, we could also write NESTINGS as

$$NEST(X,Y) = \#\{S(X) \setminus S(Y)\}$$

where function $S(X)$ denotes the set of contiguous unordered subsets in row X.

The more two rows share subsets, the smaller the NESTINGS value is. The minimum value is zero for identical rows or rows related by retrograde, and the maximum value is 65 since the 78 non-empty subsets of all rows include the twelve singletons and the aggregate.

Figure 11.2 shows the nesting defined two rows from the third movement of Alban Berg's *Lyric Suite*: T_5Q = A95B07168234 and r_4I_8Q = 5BA934821607. It contains 26 nodes. Consequently, the NESTINGS value for rows T_5Q and r_4I_8Q is $78 - 26 = 52$.

In all the similarity measures discussed so far, two rows belonging to the same row class may be maximally dissimilar. In the case of NESTINGS this depends on the row class. For example, the NESTINGS value for rows A = 012345796AB8 and $I_{10}A$ = A987653140B2 is 65, and they are maximally dissimilar. On the other hand, the largest value between the rows of the row class of row 0123456789AB is 35, and it is obtained, for example, for rows 0123456789AB and 6789AB012345.

It is straightforward to prove that NESTINGS defines a pseudometric. First, the values are positive real values ($78 - \#N(X,Y) \geq 0$ for all rows X and Y since the maximum size of a nesting is 78). Secondly, the first three requirements of the metric are satisfied. (i) The value of $NEST(X,X)$ is 0 for all rows X: two identical rows have precisely the same subsets and therefore the size of the nesting $N(X,X)$ is 78 for all rows X. (ii) Since NESTINGS is based on the *shared* subsets of rows it it clearly symmetric, and hence $NEST(X,Y) = NEST(Y,X)$. (iii) Triangle inequality $NEST(X,Y) + NEST(Y,Z) \geq NEST(X,Z)$ holds, which can be shown as follows. Row X shares $78 - NEST(X,Y)$ subsets with row Y. Since row Y shares $78 - NEST(Y,Z)$ subsets with row Z it follows that row X shares at least $78 - NEST(X,Y) - NEST(Y,Z)$ subsets with row Z. Consequently, $78 - NEST(X,Z) \geq 78 - NEST(X,Y) - NEST(Y,Z)$, and by moving terms we obtain $NEST(X,Y) + NEST(Y,Z) \geq NEST(X,Z)$. The fourth requirement of the metric does not hold since $NEST(X,RX) = 0$ for all rows X. Since the three first requirements are satisfied, NESTINGS defines a pseudometric.

Transformational approach to NESTINGS

The NESTINGS similarity measure can be expressed in terms of order-number transformations based on the following observation. By definition, the order numbers of a contiguous subset form a set of "contiguous" order numbers, that is, a set of order numbers of the type $\{\boldsymbol{p}, \boldsymbol{p+1}, \ldots, \boldsymbol{p+k}\}$. If these pitch classes form a contiguous subset when a row is transformed by an order-number transformation, then the order numbers must be transformed into another set of contiguous order numbers $\{\boldsymbol{q}, \boldsymbol{q+1}, \ldots, \boldsymbol{q+k}\}$. Hence, we must look for contiguous subsets in the transformation.

Let us define NESTINGS transformationally in much the same way as we defined the similarity measures SUBSEGMENT CONTENT DIFFERENCE and ORDER INTERVAL INVARIANT N-TUPLES. Let us first define a measure restricted to a fixed length of subsets, and then define NESTINGS as the sum of these individual measures. Hence, the first step is to define a family of similarity measures that indicate the number of subsets of cardinality n that are transformed into non-contiguous subsets, which are labeled $NEST_n$. Since the empty set was omitted in the nestings, $NEST_n$ is defined for $1 \leq n \leq 12$.

DEFINITION 11.5 In the GIS of order-number rows, the value of the $NEST_n$ similarity measure for twelve-tone rows X and Y is

$$
\begin{aligned}
&NEST_n(X, Y) \\
&\qquad = \#\{\{g_k, \ldots, g_{k+n-1}\} \mid \max\{g_k, \ldots, g_{k+n-1}\} - \min\{g_k, \ldots, g_{k+n-1}\} > n-1\}
\end{aligned}
$$

in which $1 \leq n \leq 12$ and g_k is the kth element of the order-number transformation $int(\boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{Y}\boldsymbol{X}^{-1}$ interpreted as an integer.

Admittedly, this definition is slightly awkward. The idea is to find the sets of (transformed) order numbers that form a contiguous set. The n order numbers $\{\boldsymbol{k}, \boldsymbol{k+1}, \ldots, \boldsymbol{k+n-1}\}$ are contiguous and they are transformed into a contiguous set of order numbers if and only if the maximum difference of the transformed order numbers is $n-1$. Hence, the idea in Definition 11.5 is to examine whether the sets of transformed order numbers are contiguous by examining the maximum difference within them.

Let us consider rows T_5Q = A95B07168234 and r_4I_8Q = 5BA934821607 in Figure 11.2. The order-number transformation that transforms row T_5Q into row r_4I_8Q is

$$
\begin{pmatrix}
0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\
2 & 3 & 0 & 1 & 10 & 11 & 8 & 9 & 6 & 7 & 4 & 5
\end{pmatrix}.
$$

Let us now consider, for example, how contiguous sets of order numbers of size four are transformed. There are nine such sets: $\{\boldsymbol{0,1,2,3}\}$, $\{\boldsymbol{1,2,3,4}\}$, $\{\boldsymbol{2,3,4,5}\}$, $\{\boldsymbol{3,4,5,6}\}$, $\{\boldsymbol{4,5,6,7}\}$, $\{\boldsymbol{5,6,7,8}\}$, $\{\boldsymbol{6,7,8,9}\}$, $\{\boldsymbol{7,8,9,10}\}$, and $\{\boldsymbol{8,9,10,11}\}$, which are transformed into the following sets: $\{\boldsymbol{2,3,0,1}\}$, $\{\boldsymbol{3,0,1,10}\}$, $\{\boldsymbol{0,1,10,11}\}$, $\{\boldsymbol{1,10,11,8}\}$, $\{\boldsymbol{10,11,8,9}\}$, $\{\boldsymbol{11,8,9,6}\}$, $\{\boldsymbol{8,9,6,7}\}$, $\{\boldsymbol{9,6,7,4}\}$, and $\{\boldsymbol{6,7,4,5}\}$. Of these, four are contiguous and five are not. Consequently, the $NEST_4$ value for rows T_5Q and r_4I_8Q is 5.

We can now define NESTINGS in transformational terms as the sum of the twelve transformational similarity measures $NEST_n$.

DEFINITION 11.6 In the GIS of order-number rows, the NESTINGS value for twelve-tone rows X and Y is

$$NEST(X,Y) = \sum_{n=1}^{12} NEST_n(X,Y).$$

In transformational terms, calculating the NESTINGS value for rows X and Y thus amounts to examining "contiguous segments" of transformation $\boldsymbol{YX}^{-1}$ and counting the number of those that do not form contiguous (unordered) sets of order numbers.

NESTINGS for row classes

In order to show that NESTINGS for row classes is well defined, we must show that it is transformationally coherent (see Corollary 6.1 in Section 6.2). It is transformationally coherent since the equation $NEST(X,Y) = NEST(FX,FY)$ clearly holds for all row operations: transpositions, inversion, retrograde, and their combinations. For transposition T_k, for every subset $\{p_1, p_2, \ldots, p_n\}$ of row X that is not a subset of row Y there is a corresponding subset $\{T_k(p_1), T_k(p_2), \ldots, T_k(p_n)\}$ of row T_kX that is not a subset of row T_kY (and vice versa). Similarly for inversion I_k, for every subset $\{p_1, p_2, \ldots, p_n\}$ of row X that is not a subset of row Y there is a corresponding subset $\{I_k(p_1), I_k(p_2), \ldots, I_k(p_n)\}$ of row I_kX that is not a subset of row I_kY (and vice versa). Finally for retrograde, as a row has precisely the same subsets as its retrograde, $NEST(X,Y) = NEST(RX,Y) = NEST(X,RY) = NEST(RX,RY)$. Thus, NESTINGS is a transformationally coherent metric and therefore, according to Corollary 6.1, it defines a metric for row classes. Curiously, since it gives the value 0 only for identical rows or rows related by retrograde, it only defines a pseudometric for twelve-tone rows, but it defines a metric for the row classes.

The distribution of NESTINGS

No formula is known to produce the distribution of NESTINGS: the distribution in Figure 11.12 was obtained by computer using a brute-force algorithm. It is highly skewed: the average is 61.78 and the mean is 62, both of which are close to the maximum value 65. This means that we can expect two randomly selected twelve-tone rows to contain only three non-trivial shared subsets on average.

Since the distribution of NESTINGS is skewed, the distribution of NESTINGS for row classes is also skewed. As with the previous similarity measures, the distribution of NESTINGS values between row classes varies, although due to the skewness there is not "room" for much variance.

Extensions of NESTINGS

NESTINGS could be easily extended to segments of any length as long as the segments contain the same pitch classes and the pitch classes are distinct. If pitch classes are duplicated, we can use multisets. If the segments do not contain the same pitch classes, or have different lengths, we can evaluate their similarity by counting the number of subsets that are in one segment but not in

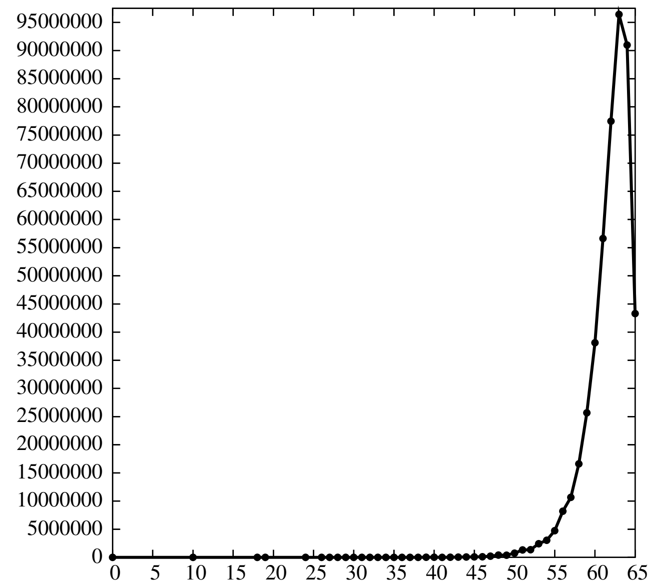


Figure 11.12: The distribution of NESTINGS. It was obtained by computing the distances defined by all 479001600 distinct order-number transformations, and hence it is precise.

the other. However, if there are pitch-class duplications or the segments have different pitch-class contents, some of the results on nestings reported in Section 11.1 no longer hold.

In order to take into account not only the number of the shared subsets of rows but also their sizes, a weighted variant of NESTINGS could be created in the same way as the FRAGMENTATION similarity measure was defined as a weighted variant of the DIVISIONS similarity measure (see Sections 10.6 and 10.7). Let us consider rows $A = 0123456789\mathrm{AB}$, $B = 2405318\mathrm{A}6\mathrm{B}97$, and $C = 10468\mathrm{A}2\mathrm{B}9753$. The nontrivial (unordered) subsets of rows A and B are $\{0, 1, 2, 3, 4, 5\}$ and $\{6, 7, 8, 9, 10, 11\}$ and those of rows A and C are $\{0, 1\}$ and $\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11\}$. The NESTINGS value for both pairs is the same, $NEST(A, B) = 63 = NEST(A, C)$. A weighted variant would make a difference in the similarities of these two row pairs, however.

11.3 SCATTERING

The SCATTERING similarity measure was invented by Robert Morris (1987). In addition, John Ward devoted some twenty pages to the analysis of it in his dissertation (Ward 1992). My aim in the following it to deepen our understanding of SCATTERING by examining it from the transformational perspective, discussing the issue of the metric, analyzing its distribution, defining SCATTERING for row classes, and introducing some variants.

Morris motivates SCATTERING as follows.

> The following similarity measure is motivated in part by the observation that the OI(P,Q) shows minimal similarity for a segment and its retrograde. Yet every pc in RP has exactly the same neighbor pcs as in P. (Morris 1987, 119)

SCATTERING takes the adjacent pitch classes in one row as its starting point and measures how far

$5409728136\mathrm{AB}$	$5406918237\mathrm{AB}$	intervening pitch classes	scattering
54	54		0
40	40		0
09	069	9	1
97	918237	$1, 8, 2, 3$	4
72	237	3	1
28	82		0
81	18		0
13	1823	$8, 2$	2
36	691823	$9, 1, 8, 2$	4
$6\mathrm{A}$	$6918237\mathrm{A}$	$9, 1, 8, 2, 3, 7$	6
AB	AB		0

Table 11.3: Scattering between rows $P = 5409728136\mathrm{AB}$ and $S = 5406918237\mathrm{AB}$ in Figure 7.3. The first column depicts the adjacent pitch classes of row P, the second depicts the corresponding segments of S that are delimited by the pitch classes of the first column, the third enumerates the intervening pitch classes, and the last column depicts the amount of scattering of these pitch classes, that is, the number of intervening pitch classes.

apart they have strayed in the other row. The procedure for calculating SCATTERING from row X to row Y can be expressed informally as follows. For every adjacent dyad pq in row X, calculate the number of pitch classes that are between p and q in row Y. The SCATTERING between two rows is the sum of these "individual scatterings."

Let us consider rows $P = 5409728136\mathrm{AB}$ and $S = 5406918237\mathrm{AB}$ in Figure 7.3. Figure 11.3 enumerates the adjacent pairs of row P and how they have strayed in row S. For example, pitch classes 5 and 4 are adjacent in both rows, and pitch classes 9 and 7 are adjacent in row P, but there are four intervening pitch classes between them in row S. The mutual order of the pitch classes is not taken into account. Therefore, pitch classes 2 and 8 are adjacent in both rows P and S even if pitch class 2 precedes pitch class 8 in row P and pitch class 8 precedes pitch class 2 in row S.

My justification for placing SCATTERING under the moniker "Similarity measures based on subsets" is that the order of the pitch classes is immaterial, and only their adjacency matters. Consequently, the SCATTERING value between retrograde-related rows is zero since such rows contain identical adjacent (unordered) pairs of pitch classes. It should be noted, however, that the dyads that are considered in the similarity measure are adjacent in one row but not necessarily in the other.

Let us now formally define SCATTERING. I will adopt Morris' convention of labeling it $SCAT$, and will give an equivalent but slightly differently formulated definition. Furthermore, I will only give a transformational definition. My rationale here is that by formalizing SCATTERING in terms of the GIS of order-number rows it is possible to prove some of its properties easily and to connect it to other similarity measures.

DEFINITION 11.7 In the GIS of order-number rows, the SCATTERING value for twelve-tone rows

X and Y is

$$SCAT(X,Y) = \sum_{1}^{11} |g_n - g_{n-1}| - 11$$

where g_n is the nth element of transformation $int(\boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{Y}\boldsymbol{X}^{-1}$ interpreted as an integer.

The SCATTERING value can be derived from the "intervals" in the transformation mapping one row into the other. There is a total of eleven pairs of adjacent pitch classes in a row. Since the order-number interval between any two order numbers is at least 1 we need to subtract $11 \cdot 1 = 11$ from the total sum of intervals in the formula in order to guarantee that the SCATTERING value is zero between two identical rows.

The range of SCATTERING contains 61 distinct values. The minimum value is 0 denoting maximum similarity, and the maximum is 60 denoting maximum dissimilarity.

In order to illustrate why the formula in Definition 11.7 gives the correct value let us consider rows $P = 5409728136\text{AB}$ and $S = 5406918237\text{AB}$. As calculated in Section 7.2.3, the order-number transformation

$$int(\boldsymbol{P}, \boldsymbol{S}) = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 0 & 1 & 2 & 4 & 9 & 7 & 6 & 5 & 8 & 3 & 10 & 11 \end{pmatrix}$$

transforms row P into row S. The upper row in $int(\boldsymbol{P}, \boldsymbol{S})$ denotes the order numbers in row P and the lower row denotes how these order numbers are transformed in order to transform row P into row S. For example, since order numbers $\mathbf{0}$, $\mathbf{1}$ and $\mathbf{2}$ are mapped into themselves in $int(\boldsymbol{P}, \boldsymbol{S})$, the interpretation is that the pitch classes at order positions $\mathbf{0}$, $\mathbf{1}$ and $\mathbf{2}$ are the same in rows P and S. Since order number $\mathbf{3}$ is mapped into order number $\mathbf{4}$ in $int(\boldsymbol{P}, \boldsymbol{S})$, the interpretation is that the pitch class that is at order position $\mathbf{3}$ in row P is at order position $\mathbf{4}$ in row S.

Now, since the order positions in the upper row of $int(\boldsymbol{P}, \boldsymbol{S})$ are listed in ascending order, they are adjacent. However, those in the lower row are not necessarily adjacent. In transformation $int(\boldsymbol{P}, \boldsymbol{S})$ the pitch classes at order positions $\mathbf{3}$ and $\mathbf{4}$ are moved to order positions $\mathbf{4}$ and $\mathbf{9}$, respectively. Therefore, the amount of scattering with respect to these two pitch classes is $|9-4| - 1 = 4$. (We need to subtract 1 so that the amount of scattering of adjacent pitch classes is 0.) The total amount of scattering is the sum of the scatterings of the 11 pairs of adjacent pitch classes. Therefore, the SCATTERING value for rows P and S is

$$\begin{aligned} SCAT(P,S) = |1-0| + |2-1| + |4-2| - |9-4| + |7-9| + |6-7| + |5-6| + \\ + |8-5| + |3-8| + |10-3| + |11-10| - 11 = 18. \end{aligned}$$

The usefulness of SCATTERING is somewhat diminished, however, due to the fact that it is not symmetric. In fact, it is the only similarity measure discussed in this work that is not symmetric. In order to show why it is not symmetric let us consider it in transformational terms. It was shown in Section 7.2.3 that the order-number transformation that transforms order-number row $\boldsymbol{P} = \mathbf{2758109463AB}$ into order-number row $\boldsymbol{S} = \mathbf{2578103964AB}$ is $\boldsymbol{S}\boldsymbol{P}^{-1} = \mathbf{0124976583AB}$.

The order-number transformation that transforms order-number row $\boldsymbol{S}$ into order-number row $\boldsymbol{P}$ is the inverse of $\boldsymbol{SP}^{-1}$, that is, $(\boldsymbol{SP}^{-1})^{-1} = \boldsymbol{PS}^{-1} = \mathbf{0129376584AB}$. Symmetric SCATTERING would require that sum of the "intervals" of inverse permutations was always the same. This is not the case, however. Indeed, we can use rows P and S as a counterexample. The SCATTERING value for rows S and P is 24, which is not equal to the value for rows P and S, calculated above to be 18.[10]

In addition to not being symmetric, SCATTERING does not satisfy triangle inequality, either. Consequently, it does not define a metric.

In order to give an example of SCATTERING failing to satisfy triangle inequality let us examine rows $A = 0123456789\text{AB}$, $B = 01234568\text{AB}79$, and $C = 68\text{A}453210\text{B}79$. $SCAT(A, B) = 10$ and $SCAT(B, C) = 10$, and consequently row A is relatively similar to row B and row B is relatively similar to row C (since if we take the distribution into account, in only 0.0066% of all row pairs we obtain a SCATTERING value that is lower than 10). On the other hand, $SCAT(A, C) = 44$, and consequently row A is relatively dissimilar to row C (since in as many as 80.30% of all row pairs we obtain a SCATTERING value that is lower than 44). Hence, taking a small step away from row A brings us to row B, and another small step brings us to row C, but the distance between rows A and C is larger than the two small steps would imply.

Rows $P = 5409728136\text{AB}$ and $P' = 287193064\text{A}5\text{B}$ present an extreme case of the non-symmetry of SCATTERING. Row P' is constructed as follows: let us start by taking the two hexachords $H_1 = 540972$ and $H_2 = 8136\text{AB}$ of row P, then hexachord H_1 is retrograded and we obtain $RH_1 = 279045$. Finally, the two hexachords RH_1 and H_2 are interleaved and we obtain $P' = 287193064\text{A}5\text{B}$. As a result, $SCAT(P, P') = 10$ but $SCAT(P', P) = 55$, and the difference between these two values is 45.

Rows $P = 5409728136\text{AB}$ and $P'' = 49216\text{B}50783\text{A}$ present another extreme case of the non-symmetry of SCATTERING. Row P'' is constructed as follows: let us start by dividing the row into two interleaved hexachords $H_1 = 50783\text{A}$ and $H_2 = 49216\text{B}$ and then simply catenate H_2 and H_1. As a result, $SCAT(P, P'') = 60$ (which is also the maximum SCATTERING value), but $SCAT(P'', P) = 20$, and the difference between the two values is 40.

Figure 11.13 shows the correlation of the SCATTERING values for symmetric row pairs, in other words the values of $SCAT(X, Y)$ and $SCAT(Y, X)$. If SCATTERING were a symmetric similarity measure, all points would along in the diagonal $x = y$. The farther away from the diagonal they are, the less correlation there is. As the figure shows, the correlation between the $SCAT(X, Y)$ and $SCAT(Y, X)$ values is very weak. Hence, the closeness of the adjacent pitch classes of row X in row Y is not a reliable indicator or prediction of the closeness of adjacent pitch classes of row Y in row X. In other words, even if pairs of pitch classes that are adjacent in row X are close to each other in row Y, it does not necessarily mean that those that are adjacent in row Y are close to each other in row X.

I will now show why SCATTERING does not satisfy triangle inequality. For the sake of simplicity,

[10] As SCATTERING is not symmetric, the expressions "P and S" and "S and P" do not have the same meaning here.

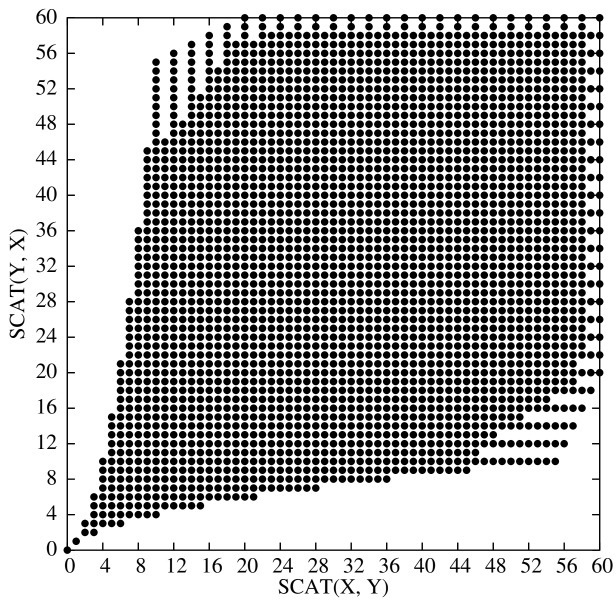


Figure 11.13: The correlation of the values of $SCAT(X, Y)$ and $SCAT(Y, X)$. The horizontal axis denotes the former and the vertical axis denotes the latter.

0123	0132	scattering
01	01	0
12	132	1
23	32	0

0132	1302	scattering
01	130	1
13	13	0
32	302	1

0123	1302	scattering
01	130	1
12	1302	2
23	302	1

Figure 11.14: Scattering between the four-tone rows 0123, 0132, and 1302. The first column in each table shows the adjacent dyad of the first of the rows, the second column shows the pitch classes of the dyads in the second row with possible intervening pitch classes, and the third column shows the amount of scattering that this dyad causes.

let us consider the four-note segments $A = 0123$, $B = 0132$ and $C = 1302$. Figure 11.14 shows the scatterings of each dyad in row pairs A and B, B and C, and A and C. As pitch class 3 "travels" towards the beginning of the segment (order position **3** in row A, order position **2** in row B, and order position **1** in row C), it adds only small local scatterings. However, with respect to the original segment A, the amount of scattering it adds is larger than the sum of the small local scatterings.

SCATTERING for row classes

SCATTERING is transformationally coherent under all pitch-class operations, since with respect to SCATTERING, they constitute only a relabeling of the pitch classes. It is also transformationally coherent under retrograde since all the adjacencies are retained in retrograde. Consequently, it is well defined as a distance. Nevertheless, since it is not symmetric, it is not symmetric when applied to row classes either, and hence it does not define a metric for row classes.

In order to calculate the SCATTERING value for two row classes we need to find the representatives of the row classes that have the smallest value. For example, it gives the value 38 for rows $P = 5409728136\text{AB}$ and $W = 3\text{B}215476\text{A}908$ in Figure 7.3, but for the corresponding row classes

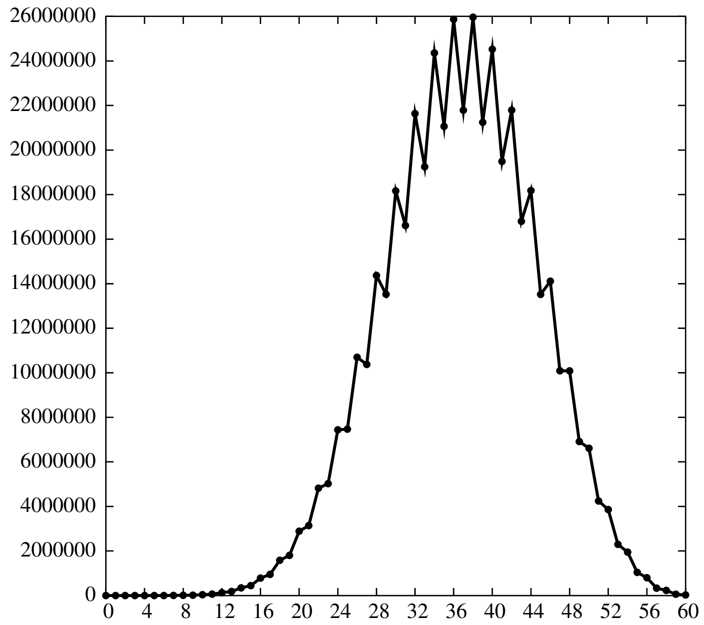


Figure 11.15: The distribution of SCATTERING. It was obtained by computing the distances defined by all 479001600 distinct order-number transformations, and hence it is precise.

$[P]$ and $[W]$ we obtain the value 28 since SCATTERING gives this value, for example, for rows $P = 5409728136\text{AB}$ and $T_2W = 514376980\text{B}2\text{A}$. Since SCATTERING is not symmetric, the value for row classes $[P]$ and $[W]$ is not the same as that for row classes $[W]$ and $[P]$. The SCATTERING value for rows W and P is 29, which is also the value for the corresponding row classes $[W]$ and $[P]$.

The distribution of SCATTERING

No formula is known to produce the distribution of SCATTERING. The distribution of SCATTERING for rows shown in Figure 11.15 was obtained by computer using a brute-force algorithm. The mean is 37 and the average is $36\frac{2}{3}$. The distribution defines a jagged curve, but its overall shape resembles the bell curve. Hence, we can say that there are relatively few very similar and very dissimilar rows, and most distances denote neither similarity nor dissimilarity.

Variants of SCATTERING

It would be straightforward to generalize SCATTERING to contiguous subsegments larger than dyads as follows. Let us consider subsegments of length n in one row. We could then compare the "spans" of those subsegments in the other row. Table 11.4 shows the contiguous triplets of row P and the corresponding segments of row Q. Both rows contain the triplet 540. The shortest segment of row Q that contains the pitch classes of the triplet 409 is 4069, which contains one extra pitch class, 6. By counting the intervening pitch classes of each of the triplets we obtain the value of this variant of SCATTERING. Hence, the value for rows P and Q is

$$0 + 1 + 5 + 3 + 1 + 0 + 1 + 3 + 5 + 6 = 25.$$

5409728136AB	5406918237AB	intervening pitch classes	scattering
540	540		0
409	4069	6	1
097	06918237	6, 1, 8, 2, 3	5
972	918237	1, 8, 3	3
728	8237	3	1
281	182		0
813	1823	2	1
136	691823	9, 8, 2	3
36A	6918237A	9, 1, 8, 2, 7	5
6AB	6918237AB	9, 1, 8, 2, 3, 7	6

Table 11.4: The scattering of triplets between rows $P = 5409728136\mathrm{AB}$ and $S = 5406918237\mathrm{AB}$. The first column shows the adjacent triplets of row P, the second shows the corresponding segments of S that contain the pitch classes of the triplet of the first one, the third column enumerates the intervening pitch classes, and the last column is the amount of scattering of these pitch classes, that is, the number of intervening pitch classes.

These generalizations of SCATTERING are not symmetric either. For example, using triplets gives the value 25 for rows P and Q, and the value 28 for rows Q and P.

CHAPTER 12

Similarity measures based on intervals

So far this discussion on similarity measures has given no attention to the intervals between the adjacent pitch classes in twelve-tone rows. The focus in this chapter is thus on the measurement of similarity based on these intervals.

There are two approaches here – the ordered and the unordered interval contents of rows. First, the succession of ordered pitch-class intervals is unique to every row modulo transposition. As is well known, inversion and retrograde create mirrors of this succession. Josef Rufer refers to this phenomenon as the "principle of intervals."

> Das Prinzip der Intervalle gilt für alle vier Erscheinungsformen, in welchen die Reihe auftreten kann: 1. die Grundform (G), 2. ihre Umkehrung (U), die aus der Umkehrung aller Intervalle der Grundform entsteht, 3. ihre Krebsform (K), die mit dem 12. Ton beginnt und rückläufig zum 1. Ton der G-reihe führt, und 4. die Umkehrung der Krebsform (KU).[1] (Rufer 1966, 78)

In a similar vein, Anton Webern formulated the classic row operations as follows in a lecture originally given on February 26th, 1932.

> Die Reihe ist nun da. — Sofort setz die Umbildung, Entwicklung ein. — Wie wurde nun das System ausgebaut? — Die kombinierende Phantasie fand die folgenden Formen: Krebs, Umkehrung, Umkehrung des Krebses. — Also vier Formen. Andere gibt es nicht. Trotz aller Bemühungen der Theoretiker.[2] (Webern 1960, 58)

[1]The principle of intervals applies to all four forms in which a row can appear: 1. the prime form, 2. its inversion which results from the inversion of all intervals of the prime form, 3. its retrograde in which we begin with the last tone and proceed backwards towards the first tone of the prime form, and 4. the inversion of the retrograde.

[2]So the row is there. At once re-casting, development starts. However is the system now built up? Our inventive resourcefulness discovered the following forms: cancrizan, inversion, inversion of the cancrizan. Four forms altogether. There aren't any others. However much the theorists try. (Webern 1963, 54)

Webern does not include transpositions in his list of four row forms, thus we might conclude that the transpositionally related rows were conceived as one form. Naturally, it is precisely the ordered succession of ordered pitch-class intervals that transpositionally related rows share.

Secondly, the unordered INT contents constitute a property of twelve-tone rows that has traditionally been considered important. An early instance of this was Alban Berg's *Lyric Suite*, in which he used an all-interval row derived from a "Mutter Akord" invented by Fritz Heinrich Klein (Headlam 1992). Alban Berg wrote to Arnold Schoenberg (incorrectly) that the row was the only such row (Brand, Hailey, and Harris 1987, 351). Nevertheless, the example shows the relevancy of the unordered INT contents for the early composers. Other famous all-interval rows include that in Luigi Nono's *Il Canto Sospeso* and the so-called Mallalieu row,[3] and since then all-interval rows have been studied carefully (Bauer-Mengelberg and Ferentz 1965; Morris and Starr 1974; Mead 1988).

Even if the measurement of the similarity of twelve-tone rows based on their intervallic succession is conceptually rather obvious, and despite the importance of intervals in the literature on twelve-tone music, the intervallic approach has not been popular in the literature concerning the similarity of twelve-tone rows: in fact, only one similarity measure has been proposed.

Section 12.1 focuses on the conception of a twelve-tone row as an ordered succession of ordered pitch-class intervals, and Section 12.2 concerns the INTERVALLIC DISTANCE similarity measure, which is based on this conception. Section 12.3 briefly outlines some characteristics of similarity measurement based on the unordered INT contents of rows, and in Section 12.4 the discussion turns to how we can measure the similarity of rows with identical unordered INT contents by measuring how the intervals are displaced.

12.1 The twelve-tone row as a succession of ordered pitch-class intervals

One of the fundamental differences between the conceptions discussed so far and the conception of a twelve-tone row as a succession of ordered pitch-class intervals is that the latter is invariant under transposition. Consequently, such a succession does not unequivocally define a twelve-tone row; in addition to recognizing the intervals we also need to know the opening pitch class (or, equivalently, the pitch class at any other order position). The conception of the twelve-tone row as a set of subsets (see Chapter 11) is analogous in that it does not enable us to distinguish between related-retrograde rows.

Perhaps an even more fundamental difference from the other conceptions (even that of the twelve-tone row as a set of subsets), however, is that by focusing on the intervals we abandon the permutational approach, according to which we always permute the same elements. When the focus is on the intervals, two rows do not necessarily contain the same intervals at all.

By abandoning the permutational approach we must abandon the transformational approach as well. We cannot interpret the differences between intervallic successions in terms of pitch-class transformations or order-number transformations. We could design a new set of transformations that

[3]Number 8 of volume 2 of the journal *In Theory Only* contains contributions by several authors concerning the properties of this row, and Andrew Mead (1989) provides further discussion.

transform successions of ordered pitch-class intervals into other successions of ordered pitch-class intervals, but it is difficult to see how we could *meaningfully* construct such a set of transformations in terms of a GIS. Namely, speaking solely in terms of intervals it is difficult to imagine how we could apply the same set of transformations to succession $\langle 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\rangle$ of row $0123456789\mathrm{AB}$ and succession $\langle 11\ \ 8\ \ 9\ \ 10\ \ 7\ \ 6\ \ 5\ \ 2\ \ 3\ \ 4\ \ 1\rangle$ of row $5409728136\mathrm{AB}$, for example. Consequently, we lose the result that every row has the same set of distances to the other rows.

In principle, if we were to stick to the idea of the isomorphism of order numbers and pitch classes, we could define similarity measures based on order-number intervals. For example, row P in Figure 7.3 is as an order-number row $\boldsymbol{P} = \mathbf{2758109463AB}$. The order-number interval between the first two order numbers $\mathbf{2}$ and $\mathbf{7}$ is $\mathbf{5}$, which means that the order-number interval between the positions of pitch classes C and $D\flat$ is $\mathbf{5}$. However, it does not seem a musically viable approach to conceptualize rows based on the order-number intervals of the pairs of pitch classes $(0,1)$, $(1,2)$, …, $(10,11)$. Adjacency in the pitch-class space does not seem to be as fundamental to our conception of rows as adjacency in the order-number space. Consequently, I will not follow this avenue of development.

12.2 INTERVALLIC DISTANCE

John Roeder developed a method for measuring the similarity between two ordered series (sequences of pitch classes) based on their intervals (Roeder 1987). I have adapted his idea to the comparison of the successions of ordered pitch-class intervals that the twelve-tone rows define. The similarity of rows is defined as the similarity of these successions. Roeder terms his similarity measure the *similarity index*. In order to avoid confusion with Teitelbaum's similarity index for set classes I have renamed the similarity measure described in this section INTERVALLIC DISTANCE or ID. For the sake of brevity I will henceforth refer to the successions of ordered pitch-class intervals that rows or segments define simply as the INT of the row or segment (see Definitions).

INTERVALLIC DISTANCE sums the distances between the corresponding ordered pitch-class intervals of rows. These ordered pitch-class intervals are conceived of in a modular space. The distance between them is measured like the distance between pitch classes, and corresponds to the unordered pitch-class interval:[4] hence it ranges from 0 to 6. For example, the distance between pitch-class intervals 1 and 11 in this modular space is not 10 but 2. Let us notate the distance between two pitch-class intervals as ic.

Following these preliminary remarks let us now give the formal definition of INTERVALLIC DISTANCE.

DEFINITION 12.1 The INTERVALLIC DISTANCE value for twelve-tone rows X and Y is

$$ID(X,Y) = \sum_{n=0}^{10} ic(int(x_n, x_{n+1}), int(y_n, y_{n+1}))$$

[4]As stated in Chapter 4, distance is conceived of as non-directional.

in which $int(x_n, x_{n+1})$ is the ordered pitch-class interval between the pitch classes x_n and x_{n+1} of row X, $int(y_n, y_{n+1})$ is the ordered pitch-class interval between the pitch classes y_n and y_{n+1} of row Y, and $ic(int(x_n, x_{n+1}), int(y_n, y_{n+1}))$ is the distance between the two ordered pitch-class intervals interpreted as an integer.

INTERVALLIC DISTANCE sums the distances between the eleven corresponding ordered pitch-class intervals of rows using modular arithmetic. For example, the INT of row $P = 5409728136\text{AB}$ is $\langle 11 \quad 8 \quad 9 \quad 10 \quad 7 \quad 6 \quad 5 \quad 2 \quad 3 \quad 4 \quad 1 \rangle$: the pitch-class interval between the first and the second pitch classes is 11, the pitch-class interval between the second and the third pitch classes is 8, etc. Similarly, the INT of row $Q = 5406728139\text{AB}$ is $\langle 11 \quad 8 \quad 6 \quad 1 \quad 7 \quad 6 \quad 5 \quad 2 \quad 6 \quad 1 \quad 1 \rangle$. We obtain $ID(P, Q)$ by summing the distances of the corresponding ordered pitch-class intervals in the INTs of these two rows as follows:

$$\begin{aligned} ID(P, Q) = ic(11, 11) + ic(8, 8) + ic(6, 9) + ic(1, 10) + ic(7, 7) + ic(6, 6) + \\ + ic(5, 5) + ic(2, 2) + ic(6, 3) + ic(1, 4) + ic(1, 1) = 12. \end{aligned}$$

As noted by Roeder, INTERVALLIC DISTANCE defines a metric on the set of INTs. The metric could be described as a modular taxicab metric: it is not a "normal" taxicab metric since the minimum of the values $|y - x|$ and $12 - |y - x|$ was selected as the distance between the two pitch-class intervals x and y.[5]

Transpositionally related rows have the same INT. Hence, $ID(X, T_n X) = 0$ for all rows X and all transpositions T_n. Therefore, INTERVALLIC DISTANCE defines pseudometric on the set of twelve-tone rows, not a metric.

We cannot define intervallic distance in terms of transformations. One way to illustrate this is that, as will be explained below, rows have different distributions of distances to the other rows. If intervallic distance could be defined in terms of transformations, all rows should have the same distribution.

INTERVALLIC DISTANCE for row classes

INTERVALLIC DISTANCE is a transformationally coherent measure since the equation $ID(X, Y) = ID(FX, FY)$ clearly holds for all row operations: transpositions, inversion, retrograde, and their combinations. Trivially, for transpositions $ID(X, Y) = ID(T_k X, T_k Y)$ since transpositionally related rows have the same INT. For inversions $ic(x, y) = ic(I(x), I(y))$ and therefore $ID(X, Y) = ID(IX, IY)$. The INT of a retrograde is a retrograded INT of the inversion form, and similarly, the INT of a retrograde inversion is a retrograded INT of the prime form. Consequently, $ID(X, Y) = ID(RX, RY)$ and $ID(X, Y) = ID(RIX, RIY)$. INTERVALLIC DISTANCE therefore defines a metric for row classes – even if it only defines a pseudometric for twelve-tone rows.

In order to calculate the INTERVALLIC DISTANCE value for row classes we only need to make

[5]Naturally, if $x = y$ then $|y - x| = 0$ and does not matter that $12 - |y - x| = 12$. If $x \neq y$, then $0 < 12 - |y - x| < 12$.

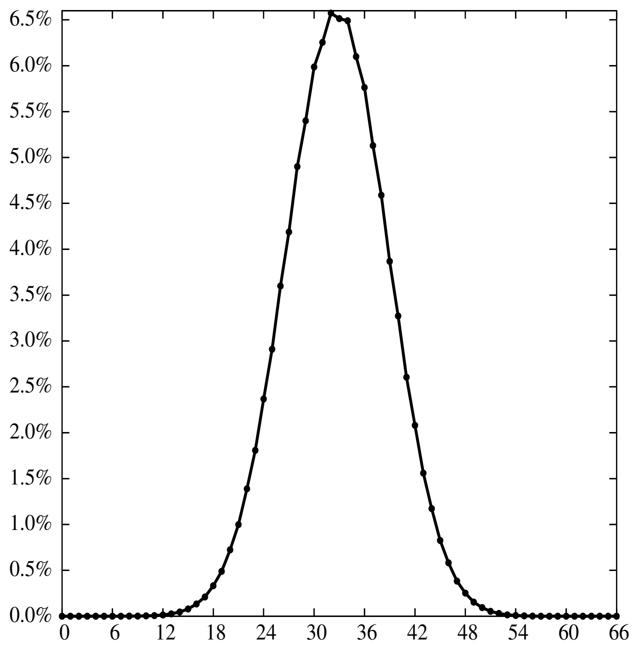


Figure 12.1: The distribution of INTERVALLIC DISTANCE. It was obtained by calculating the INTERVALLIC DISTANCE value for 10000000000 pairs of randomly generated twelve-tone rows.

four comparisons:

$$ID([X],[Y]) = \min\{ID(X,Y), ID(X,IY), ID(X,RY), ID(X,RIY)\}.$$

For example, INTERVALLIC DISTANCE gives the value 32 for rows $P = 5409728136\text{AB}$ and $V = \text{A}463592178\text{B}0$, which is also the value it gives for the corresponding row classes $[P]$ and $[V]$.

The distribution of INTERVALLIC DISTANCE

A twelve-tone row contains eleven ordered pitch-class intervals. The maximum distance between two ordered pitch-class intervals is 6, therefore the maximum INTERVALLIC DISTANCE value is $11 \cdot 6 = 66$. This value is obtained between rows 0123456789AB and $07294\text{B}6183\text{A}5$, for example. Not every row has other rows at a distance of 66, however. (This is the quickest method of showing that INTERVALLIC DISTANCE is not transformational since in transformational similarity measures every twelve-tone row has the same network of distances to the other rows.) No twelve-tone row with a tritone has a row at a distance of 66 since the distance between a tritone and any non-zero pitch-class interval is less than 6. For example, the maximum distance from row $P = 5409728136\text{AB}$ is 60 and it is obtained by row $r_6P = 8136\text{AB}540972$.

Figure 12.1 depicts the distribution of INTERVALLIC DISTANCE values. The average of the distribution is 32.73 and the mean is 33, and the shape resembles the bell curve. Hence, given a row, there are very few rows that are very similar to it, relatively few rows that are very dissimilar to it, and the majority are neither particularly similar nor dissimilar to it.

The distribution is not perfectly symmetrical, however. This follows from the observation that

rows		ID	$PCDIS$
0123456789AB	07294B6183A5	66	36
0123456789AB	061728394A5B	60	30
0123456789AB	012389AB4567	8	32

Figure 12.2: Some rows in which the INTERVALLIC DISTANCE and PITCH-CLASS DISPLACEMENT values do not correlate even in comparisons of the row classes of transpositionally related rows.

every row has exactly twelve rows at the minimum distance of 0; these are the twelve transpositions of the row. However, some but not all rows have rows at the maximum distance of 66, and hence the distribution cannot be symmetrical.

INTERVALLIC DISTANCE compared to PITCH-CLASS DISPLACEMENT

We can relate INTERVALLIC DISTANCE to the pitch-class domain even if it is not explicitly defined in it. Of the similarity measures discussed in this work, the only one pronouncedly defined in the pitch-class domain is PITCH-CLASS DISPLACEMENT (DERANGEMENT and CAYLEY DISTANCE could be defined equally well in the pitch-class domain and in the order-number domain). Both INTERVALLIC DISTANCE and PITCH-CLASS DISPLACEMENT measure, in their own ways, the similarity of the "curves" that the rows draw (see Section 8.4). Hence, it is appropriate to find out if there is any correlation between the two measures.

First, note that the INTERVALLIC DISTANCE and PITCH-CLASS DISPLACEMENT distributions are very similar (compare Figures 8.5 and 12.1): both resemble the bell curve. The maximum INTERVALLIC DISTANCE value is 66 and the maximum PITCH-CLASS DISPLACEMENT value is 72. Hence, the values are easy to compare.

It turns out, however, that the INTERVALLIC DISTANCE and PITCH-CLASS DISPLACEMENT values do not necessarily coincide. The most obvious case is in the comparison of transpositionally related rows, which have the same INT but the PITCH-CLASS DISPLACEMENT value varies according to the transposition level: the maximum value of 72 is obtained between rows X and T_6X for any row X.

We could bypass the problem of transposition by comparing classes of transpositionally related rows. Even for these we find instances in which the INTERVALLIC DISTANCE and PITCH-CLASS DISPLACEMENT values do not correlate. Figure 12.2 gives some examples.

The difference between the behavior of the INTERVALLIC DISTANCE and PITCH-CLASS DISPLACEMENT similarity measures arises from the way that the displacement of one pitch class affects the measure. Namely, if we begin with segment abc and transpose the middle pitch class b we obtain segment axc. With respect to PITCH-CLASS DISPLACEMENT, only one pitch class has been displaced, but with respect to INTERVALLIC DISTANCE both the interval between pitch classes a and b and that between pitch classes b and c in the original segment have been changed.

The value of INTERVALLIC DISTANCE for rows 0123456789AB and 07294B6183A5 in Figure 12.2 is the maximum 66. However, with respect to PITCH-CLASS DISPLACEMENT, half of the pitch classes have the same order positions in the two rows, and hence the rows are not deemed either dissimilar or similar. Rows $0123456789\mathrm{AB}$ and $061728394\mathrm{A}5\mathrm{B}$ have a similar difference. In contrast, the INTERVALLIC DISTANCE value of 8 is relatively small for rows $0123456789\mathrm{AB}$ and $012389\mathrm{AB}4567$, thus indicating similarity, while the PITCH-CLASS DISPLACEMENT value of 32 is clearly larger. Here only two of the pitch-class intervals are different in the two rows, but as eight of the twelve pitch classes have been displaced, the PITCH-CLASS DISPLACEMENT value grows high enough to indicate that the two rows are not similar.

In sum, judging from the rows given in Figure 12.2, we could argue that INTERVALLIC DISTANCE does a better job than PITCH-CLASS DISPLACEMENT in evaluating the similarity of the contours or the "curvatures" of rows. Namely, the zig-zag contours of rows $07294\mathrm{B}6183\mathrm{A}5$ and $061728394\mathrm{A}5\mathrm{B}$ are clearly different form the smoothly ascending contour of row $0123456789\mathrm{AB}$.

Extending INTERVALLIC DISTANCE to segments

INTERVALLIC DISTANCE could easily be extended to segments, as long as the segments are of the same length. As noted before, it is not permutational. Since only the intervals between adjacent pitch classes affect this measure, we do not need to require that the segments contain the same pitch classes, or that there are no pitch-class duplications in them. Indeed, John Roeder's (1987) original definition of the measure is for segments.

If we wanted to compare two segments of different lengths, we might define a "penalty" for such a situation along the lines outlined in Section 8.3.

12.3 Similarity measures based on unordered interval contents

The interval contents are a prominent property of twelve-tone rows. For example, the rows used by Anton Webern are characterized by an almost complete lack of whole-tone steps and tritones.[6] While a composer may introduce intervals that are not present in the INT of the row, the intervals of the row are usually reflected in the musical surface.[7]

Twelve-tone rows can be compared with respect to their unordered interval contents. The unordered INT contents of rows is a variant of Allen Forte's concept of reducing a succession of notes to a *Basic Interval Pattern*, or BIP (Forte 1973a). A BIP is the unordered collection of unordered pitch-class intervals between the adjacent pitch classes of a row or segment arranged in ascending order. Alternatively, we might create a similar construct by arranging the unordered collection of ordered pitch-class intervals between the adjacent pitch classes of a row or segment in ascending order.

For example, the INT of row $V = \mathrm{A}463592167\mathrm{B}0$ is $\langle 6 \quad 2 \quad 9 \quad 2 \quad 4 \quad 5 \quad 11 \quad 6 \quad 1 \quad 3 \quad 1 \rangle$. If we interpret these ordered pitch-class intervals as unordered pitch-class intervals we obtain the

[6] If we take Webern's *Klavierstuck* (1925), *Kinderstuck* (1924) and the works from op. 17 to op. 31 as our body of rows, we find a total of 21 rows (opuses 17 and 18 both contain three compositions with three distinct rows). In the $11 \cdot 21 = 231$ intervals in these rows we find only nine whole-tone steps and twelve tritones.

[7] For example, in the fifth variation of his *Variations for Orchestra* op. 31 Arnold Schoenberg divides row A463592178B0 into six semitone dyads (A9, 43, 65, 21, 78, and B0), even if it contains only three adjacent semitone dyads.

succession $\langle 6\ \ 2\ \ 3\ \ 2\ \ 4\ \ 5\ \ 1\ \ 6\ \ 1\ \ 3\ \ 1\rangle$, and we obtain a representation of these unordered pitch-class intervals as a BIP by arranging them in ascending order as follows 11122334566. Similarly, by enumerating the members of the unordered collection of ordered pitch-class intervals we obtain the BIP-like construct $1122345669\mathrm{B}$. The same information could be conveyed by using a notation similar to the interval-class vector of set classes. The vector of unordered pitch-class intervals of row V would be [322112], and that of ordered pitch-class intervals would be [22111200101].

Forte devotes a major part of his article to a discussion on how different set classes generate different amounts of basic interval patterns. In the present case, however, every twelve-tone row contains all the twelve pitch classes, and hence the interest is only in the basic interval patterns that are generated by set class 12-1[0123456789AB]. Curiously enough, with the exception of all-interval rows, the unordered interval contents of twelve-tone rows has not been studied.[8]

It would be easy to create measures for comparing the similarity of twelve-tone rows based on the unordered collections of either ordered or unordered pitch-class intervals. In fact, all set-class similarity measures based on the interval-class contents of set classes could be thus adapted with little effort. Furthermore, such similarity measures could be extended to segments of other cardinalities, to segments with different cardinalities, and to segments containing pitch-class duplications. The literature on the similarity of set classes contains guidelines on how to cope with different cardinalities.

We could define row measures based on the ordered pitch-class intervals of rows. However, in that case rows $0123456789\mathrm{AB}$ and $\mathrm{BA}9876543210$ would be maximally dissimilar, for example, since the former contains only pitch-class intervals 1 and the latter contains only pitch-class intervals 11. Consequently, it might be more appropriate to base the measures on the interval classes (as with set classes). In addition, since all rows in a row class have identical unordered interval-class contents, these would define measures in which they are all maximally similar to each other.

An inevitable feature of row measures based on unordered interval contents is that they have a poor resolution: the 479001600 rows have only 301666 different unordered INT contents and only 4129 unordered interval-class contents. Furthermore, as many as 3856 twelve-tone rows have identical unordered INT contents, and as many as 162888 have identical unordered interval-class contents. In Section 1.2 I suggested placing the relations of twelve-tone rows in three categories: by operation, by property, and by similarity. Given the issues outlined above, it seems that the unordered INT contents of rows is better suited for defining relations by property than relations by similarity. On the other hand, we might define similarity relations (see Section 3.2) in which two twelve-tone rows are related if and only if they have identical unordered INT contents or identical unordered interval-class contents.

12.4 INTERVAL DISPLACEMENT

In the approach described in the previous section the focus is on the unordered INT contents of rows, but not on the order in which the intervals occur in them. Two rows with identical unordered

[8]Forte only considers basic interval patterns derived from set classes of cardinalities 3 to 7, presumably due to the amount of computation needed and the limited capacity of computers at the time of writing.

INT contents might be more or less similar depending on the order in which the intervals occur. This section charts the development of a similarity measure in which the order is also significant. For the sake of brevity, I will henceforth in this section refer to ordered pitch-class intervals simply as intervals.

The INTERVAL DISPLACEMENT similarity measure was suggested to the author by Robert Morris in private communication. It is based on the idea that two rows with identical unordered INT contents contain the same intervals but in a different order. Just as DISPLACEMENT measures how far the pitch classes have strayed, INTERVAL DISPLACEMENT measures how far the intervals have strayed.

Let us consider row $P = 5409728136\text{AB}$ and its retrograde inversion $RI_3P = 4590271863\text{BA}$. Both rows are all-interval rows – the successions of the intervals are $\langle 11 \ \ 8 \ \ 9 \ \ 10 \ \ 7 \ \ 6 \ \ 5 \ \ 2 \ \ 3 \ \ 4 \ \ 1\rangle$ and $\langle 1 \ \ 4 \ \ 3 \ \ 2 \ \ 5 \ \ 6 \ \ 7 \ \ 10 \ \ 9 \ \ 8 \ \ 11\rangle$, respectively. It is a well-known fact that in rows related by retrograde inversion the succession of intervals is retrograded: the first interval 11 in row P is the last interval in row RI_3P, the second interval 8 in row P is the next to the last interval in row RI_3P, etc. Hence, in order to turn the succession of intervals in row P into that in row RI_3P, we might rearrange them as shown in Figure 12.3. The first interval is displaced by 10 positions, the second by 8 positions, etc. Consequently,

$$10 + 8 + 6 + 4 + 2 + 0 + 2 + 4 + 6 + 8 + 10 = 60$$

is the total sum of displacements.

Let us then consider rows $T_5Q = \text{A95B07168234}$ and $RI_8Q = 9\text{AB570612843}$. Neither of them is an all-interval row but they both have the same unordered INT contents. The successions of intervals are $\langle 11 \ \ 8 \ \ 6 \ \ 1 \ \ 7 \ \ 6 \ \ 5 \ \ 2 \ \ 6 \ \ 1 \ \ 1\rangle$ and $\langle 1 \ \ 1 \ \ 6 \ \ 2 \ \ 5 \ \ 6 \ \ 7 \ \ 1 \ \ 6 \ \ 8 \ \ 11\rangle$, respectively. We might rearrange the former succession in a similar fashion as above in order to obtain the latter. This is depicted in Figure 12.4. Again,

$$10 + 8 + 6 + 4 + 2 + 0 + 2 + 4 + 6 + 8 + 10 = 60$$

is the total sum of displacements. There is another possibility, however. We might rearrange the succession of intervals as depicted in Figure 12.5. In this,

$$10 + 8 + 0 + 3 + 2 + 0 + 2 + 4 + 0 + 8 + 3 = 40$$

is the total sum of displacements.

The crucial observation here is that there may be several different ways in which to rearrange the intervals. In fact, the all-interval rows are the only ones in which such rearrangement is unequivocal. Consequently, we have to decide on which transformation the measurement of INTERVAL DISPLACEMENT should be based.

I discussed the similarity of row classes in Section 6.2, my guiding principle being that the dis-

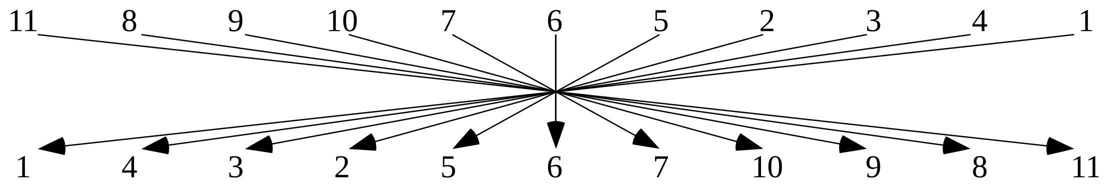


Figure 12.3: Displacing the intervals in rows $P = 5409728136\mathrm{AB}$ and $RI_3P = I_9P = 4590271863\mathrm{BA}$ by retrograding the interval succession.

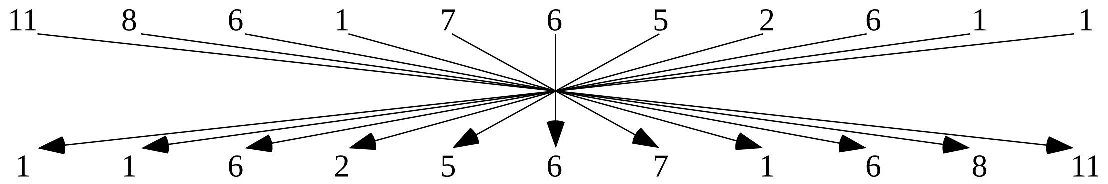


Figure 12.4: Displacing the intervals in rows $T_5Q = \mathrm{A95B07168234}$ and $RI_8Q = \mathrm{9AB570612843}$ by retrograding the interval succession.

tance between two classes is the distance between the closest rows in them. In a similar vein, of the ways of rearranging the intervals, the INTERVAL DISPLACEMENT value is based on the rearrangement that yields the lowest value. For example, in the case of rows $T_5Q = \mathrm{A95B07168234}$ and $RI_8Q = \mathrm{9AB570612843}$, the rearrangement depicted in Figure 12.5 is preferable to that depicted in Figure 12.4.

In more formal terms, the INT of a twelve-tone row is a permutation of its unordered INT contents, just as a twelve-tone row is a permutation of the twelve pitch classes. There are two differences compared to twelve-tone rows, however. First, the unordered INT contents of a row may be a *multiset*[9] since some intervals may be duplicated. Secondly, not all permutations of the unordered INT contents necessarily represent the succession of intervals in any twelve-tone row: some permutations may lead to pitch-class duplication. I will return to this issue below.

The INTERVAL DISPLACEMENT similarity measure can be applied to two rows if and only if they have the same unordered INT contents. In such a case, we would permute the INT (or "re-permute" the unordered INT contents), which could be formalized as a permutation acting on the INT. This formalization is analogous to the idea of pitch-class operations or order-number operations acting on the twelve-tone rows. However, I will not develop the full formalization of the pertinent group of permutations and its action on the set of successions of intervals here. Suffice it note that the action of the permutations on the set of INTs is analogous to the action of the permutations (order-number transformations) on the order-number rows in the GIS of order-number rows, and to that (pitch-class transformations) on the pitch-class rows in the GIS of pitch-class rows. Consequently, let us define the INTERVAL DISPLACEMENT value in terms of the complexity of the permutation.

[9]A multiset is like a set in the sense that the order of elements is immaterial, but different in the sense that their multiplicity is significant. For example, $\{0, 1, 2\}$ and $\{0, 0, 1, 2\}$ are identical as sets but not as multisets.

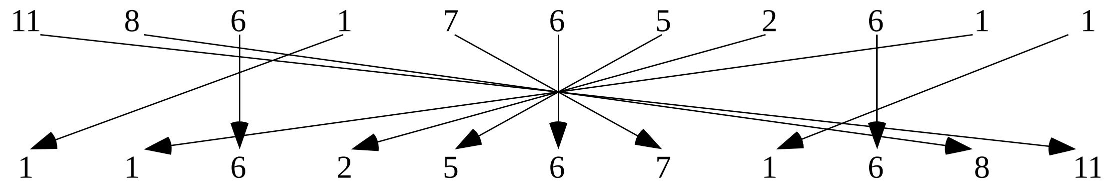


Figure 12.5: Displacing the intervals in rows $T_5Q = \text{A95B07168234}$ and $RI_8Q = \text{9AB570612843}$ but not by retrograding the interval succession.

DEFINITION 12.2 If X and Y are twelve-tone rows with identical unordered INT contents, then permutation F transforms INT of row X into that of row Y if $F(INT(X)) = INT(Y)$.

The next step it to define the INTERVAL DISPLACEMENT value for each permutation. Let us follow the formula used in the transformational definition of the DISPLACEMENT similarity measure. Note, however, that we could define any additional similarity measure using any formula that was used to define similarity measures based on the GIS of order-number rows.

DEFINITION 12.3 The INTERVAL DISPLACEMENT value for permutation G is given by the formula

$$\mathit{INTDIS}(G) = \sum_{n=1}^{11} |g_n - n|$$

in which g_n is the nth interval of permutation G interpreted as an integer.

The interpretation of Definition 12.3 is that permutation G moves the interval at position n to position g_n in the succession of intervals: the INTERVAL DISPLACEMENT value for permutation G is the sum of the individual displacements.

Given these preliminary considerations let us now define INTERVAL DISPLACEMENT formally.

DEFINITION 12.4 The INTERVAL DISPLACEMENT value for two twelve-tone rows X and Y is given by the formula

$$\mathit{INTDIS}(X,Y) = \min\{\mathit{INTDIS}(G) \mid G(INT(X)) = INT(Y)\}$$

where G is a permutation acting on the INT of row X.

Definition 12.3 defines the magnitude of a permutation that transforms the INT of row X into that of row Y, and the INTERVAL DISPLACEMENT value is thus the minimum of the magnitudes of all permutations that transform the INT of row X into that of row Y (Definition 12.4).

INTERVAL DISPLACEMENT is defined in terms of the permutation acting on the successions of intervals. Since the formula is almost identical to that used to define DISPLACEMENT (see Definition 8.6), many of the properties of INTERVAL DISPLACEMENT can be derived in a similar fashion. Consequently, let us briefly note that INTERVAL DISPLACEMENT defines a metric on every set of rows that have identical unordered INT contents (but not on the set of all twelve-tone rows). The

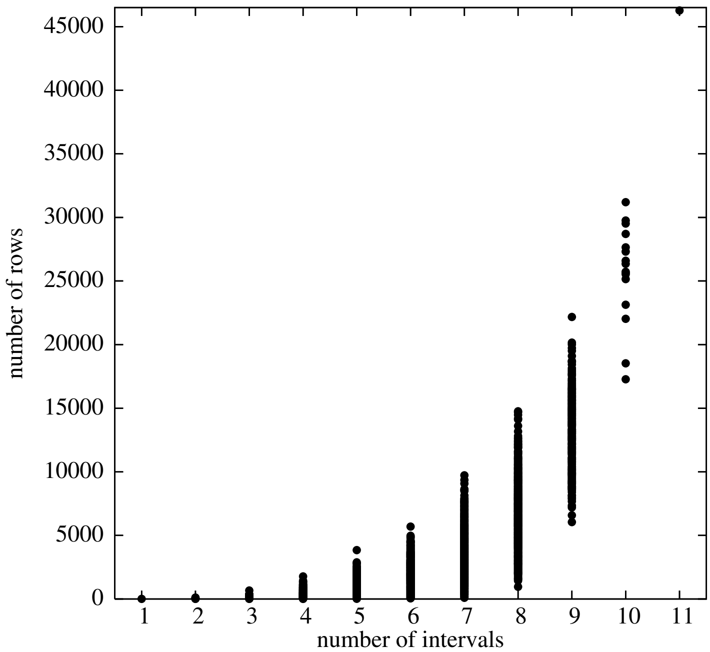


Figure 12.6: The distribution of unordered INT contents of rows. The horizontal axis denotes the number of distinct ordered intervals in unordered INT contents and the vertical axis denotes the number of rows with such contents.

minimum value is 0, denoting maximal similarity. The maximum value is 60, and is obtained by means of several permutations – one of which retrogrades the succession of intervals.

As noted in Section 12.3, there are 301666 different unordered INT contents of twelve-tone rows. Consequently, INTERVAL DISPLACEMENT is applicable only within each of these 301666 disjoint sets of rows. Figure 12.6 shows the distribution of the unordered INT contents of rows. Each dot in the figure represents one or several unordered INT contents with the indicated number of distinct intervals and the indicated number of rows with that such contents. Note that the "popularity" of an interval content correlates to some degree with the number of distinct intervals, or the "evenness" of the distribution. At the low end are the rows with maximally uneven distributions – rows with only one interval – which are either the chromatic scale or the cycle of fifths. Each of the four possible unordered INT contents with only a single interval contains only twelve rows. At the high end are the rows with maximally even distribution of intervals, which are the 46272 all-interval rows in which each of the eleven non-zero intervals is present exactly once.[10]

INTERVAL DISPLACEMENT is different from all the other row measures discussed in this work in that it defines a value for only some row pairs. Hence, it is difficult to compare its values to those of other similarity measures. The following example illustrates how the similarity of intervals does not necessarily correlate with other similarity measures, however. Let us consider rows $P = 5409728136\mathrm{AB}$ and $I_9P = 4590271836\mathrm{BA}$. Since only six adjacent dyads have been flipped these

[10]Curiously enough, while an all-interval row is a special case, it is not a rarity but an example of the most representative interval contents. Hence, if we choose a random twelve-tone row, with respect to its interval contents it is most likely to be an all-interval row.

two rows are relatively similar according to most similarity measures based on order relations. They are maximally dissimilar with respect to INTERVAL DISPLACEMENT, however, since the INT is retrograded between them.

Extending INTERVAL DISPLACEMENT to segments

According to the permutational approach, the primitive elements are the permutations of the twelve pitch classes. We could, however, define the successions of ordered pitch-class intervals as primitive elements. (In this case, naturally, we would ignore the transpositional level.) If we allow the application of permutations to the successions of intervals, the resulting successions of pitch classes may contain duplications.

For example, the INT of row $P = 5409728136\mathrm{AB}$ is $\langle 11 \quad 8 \quad 9 \quad 10 \quad 7 \quad 6 \quad 5 \quad 2 \quad 3 \quad 4 \quad 1 \rangle$. If we exchange the two first intervals the resulting succession of intervals is $\langle 8 \quad 11 \quad 9 \quad 10 \quad 7 \quad 6 \quad 5 \quad 2 \quad 3 \quad 4 \quad 1 \rangle$. Now, any pitch-class segment with this succession of intervals contains a duplicated pitch class. Therefore, this succession of ordered pitch-class intervals is not the INT of any twelve-tone row. For example, if we set the first pitch class at 0 we obtain the succession of pitch classes $08742938\mathrm{A}156$, which is not a twelve-tone row since pitch class 8 is duplicated and pitch class 11 does not appear at all. INTERVAL DISPLACEMENT could thus be extended to evaluate the similarity of any two sequences of pitch classes that have identical unordered INT contents.

CHAPTER 13

Partial orders, graph theory, and similarity

This chapter introduces a method of using partial orders and graph theory to evaluate the similarity of twelve-tone rows. In particular, I will use graph theory to relate two similarity measures discussed in Chapter 9: ORDER INVERSIONS and BADNESS OF SERIAL FIT. Appendix C provides a concise introduction to the theory of partially ordered sets, and the pertinent concepts of graph theory are introduced in Appendix D.

I will begin in Section 13.1 with a relatively technical discussion about the construction of a graph of all partial orders on a given set. On the basis of this discussion, in Section 13.2 I will analyze two similarity measures based on the conception of a row as a set of ordered pairs: ORDER INVERSIONS and BADNESS OF SERIAL FIT. Finally, in Section 13.3 I will consider how the ideas developed in the previous sections could be applied to the analysis of similarity measures based on other conceptions of twelve-tone rows.

13.1 The graph of partial orders

This section combines the theories of partial orders and of graphs. The focus is on the graph as formed by all partial orders on a finite set.

The vertices of the graph of partial orders are all possible partial orders on a given set, and as there is only a finite number of possible partial orders on a finite set, the graph is finite. The edges are defined by the inclusion relation: two vertices are adjacent, that is, connected by an edge, only if one is a subset of the other and there is no "intermediate" subset.[1]

Let us postpone the formal definition of the graph of partial orders pending consideration of a concrete example. Let

$$A = \{(a,a),(b,b),(c,c),(a,c)\} \text{ and } C = \{(a,a),(b,b),(c,c),(a,b),(b,c),(a,c)\}$$

[1] The adjacency of partial orders in the graph should not be confused with the adjacency of elements in a partial order. For the latter, see Definition C.3 in Appendix C.

be two partial orders on the set $\{a, b, c\}$.[2] Clearly, $A \subset C$ since C contains all the four pairs that A contains.[3] Partial order A contains one non-reflexive pair and partial order C contains three non-reflexive pairs. However, there is an *intermediate* partial order between partial orders A and C, namely

$$B = \{(a, a), (b, b), (c, c), (b, c), (a, c)\}.$$

Clearly, $A \subset B \subset C$. Therefore, because there is an intermediate partial order, A and C are not adjacent in the graph. The definition below formalizes these observations.

DEFINITION 13.1 The vertices of graph G are the partial orders on a given finite set. Two vertices A and B are adjacent if and only if the two following conditions hold: (i) $A \subset B$ (or $B \subset A$) and (ii) there does *not* exist a partial order X such that $A \subset X \subset B$ (or $B \subset X \subset A$).

The first requirement of the definition is that one of the vertices must be a proper subset of the other, and the second is that there must not be an intermediate partial order.

Admittedly, the definition is somewhat complex. The rest of this section is devoted to a discussion on its implications, and in particular on what kind of partial orders are adjacent in the graph.

Let $\#A$ denote the number of ordered pairs in partial order A. Since we are restricted to finite partial orders, it follows immediately that if partial order A is a proper subset of partial order B, then of necessity $\#A < \#B$.

The following lemma provides a simple criterion for the adjacency of vertices. It turns out that if two vertices A and B are adjacent, then of necessity $|\#A - \#B| = 1$. It should be noted that as both A and B are partial orders they are both reflexive and contain all pairs of the type (x, x). Thus, the pair missing from the smaller of the adjacent vertices must be a non-reflexive pair, that is, a pair of type (x, y) in which $x \neq y$.

LEMMA 13.1 The two vertices A and B are adjacent in the graph of partial orders if and only if $A \subset B$ (or $B \subset A$) and $|\#A - \#B| = 1$.

Proof. The lemma carries two implications. The one from right to left is straightforward to prove. Let A be a subset of B. Because by assumption the cardinalities of A and B differ by one, A contains all the pairs in B except exactly one. Thus, A is a proper subset of B. In addition, there cannot exist an intermediate partial order because any subset of B that is larger than A must contain the missing pair and is thus equal to B. Therefore, vertices A and B are adjacent.

The implication from left to right is most easily proved by indirect means. Let us assume that the implication does not hold and show that this leads to a contradiction.

If A is not a subset of B, and nor is B a subset of A, they cannot be adjacent by definition. Let us therefore assume that $A \subset B$ and $|\#A - \#B| > 1$. Let Z denote the set of pairs in B that are

[2]For the sake of brevity, I occasionally omit the reflexive pairs. If I write, for instance, that $A = \{(a, c)\}$ is a partial order on the set $\{a, b, c\}$, the interpretation is that $A = \{(a, a), (b, b), (c, c), (a, c)\}$.

[3]I remind the reader that in this study the sign $\subset$ denotes a *proper subset*.

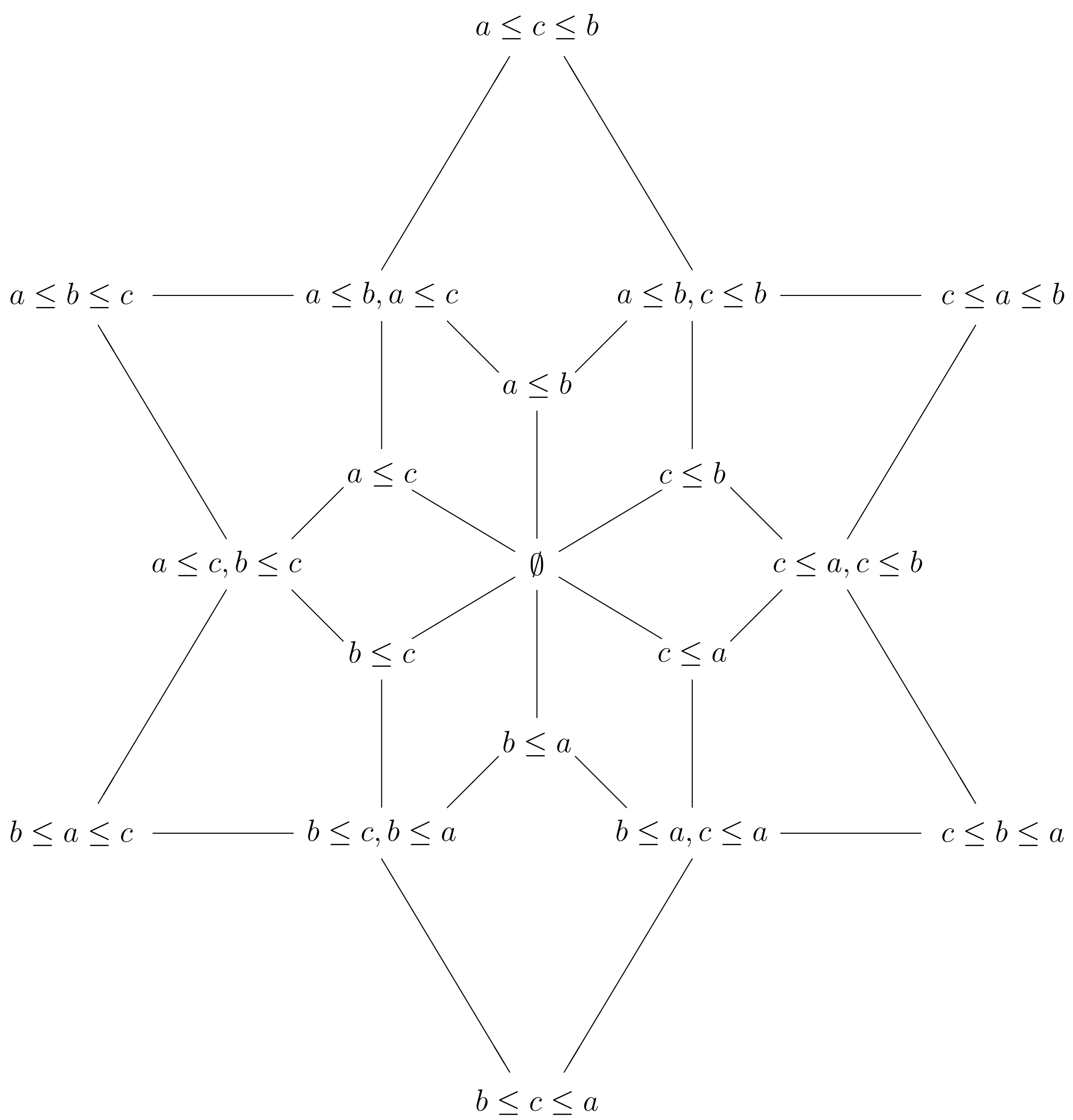


Figure 13.1: The graph formed by all partial orders on the set $\{a, b, c\}$.

not in A: in other words, $Z = B \setminus A$ and $\#Z = |\#A - \#B| > 1$. I will show that this assumption leads to a contradiction.

According to Lemma C.3 in Appendix C, set Z as a non-empty subset of a finite partial order contains at least one minimal element. Let one of these minimal elements be denoted by x. As Z does not contain reflexive pairs of the type (a, a) there is an element y such that pair (x, y) is in Z. It can be assumed that x and y are adjacent in Z.

Now, elements x and y are either adjacent in B or they are not. Both cases must be checked. If they are adjacent in B we can form the partial order $C = B \setminus \{(x, y)\}$ which according to Lemma C.2 in Appendix C is a partial order. Clearly, then, $A \subset C \subset B$. Therefore A and B are not adjacent so the antithesis is disproved and $|\#A - \#B| = 1$.

The second case is that x and y are not adjacent in B, but there are one or several elements z such that $x < z < y$. Therefore B contains the chain

$$x = z_0 < z_1 < \ldots < z_{n-1} < z_n = y$$

in which all elements are adjacent. Now, at least one of the pairs (z_i, z_{i+1}), $0 \leq i < n$, must be in Z, let it be denoted by (z_k, z_{k+1}). If this were not the case, then the whole chain would be in A and then, because of the transitivity partial order A, $x \leq y$ would also be in A and then $x \leq y$ would not be in Z. Then we could form the partial order $D = B \setminus \{(z_k, z_{k+1})\}$, which according to Lemma C.2 in Appendix C is a partial order. Clearly then, $A \subset D \subset B$. Then again, A and B are not adjacent so the antithesis is disproved and $|\#A - \#B| = 1$. Both cases have been checked and this proves the lemma. □

Figure 13.1 illustrates the lemma and depicts a graph of all the partial orders on the set $\{a, b, c\}$.[4] In order to avoid extra clutter on the graph the reflexive pairs are not drawn. Note that all adjacent partial orders differ only by a single ordered pair, as proved in Lemma 13.1. The empty set $\emptyset$ in the middle of the graph is the trivial partial order that contains only the reflexive pairs.

The remaining theoretical task is to discuss the distances between partial orders in the graph defined above. Distance in a graph is naturally defined as the length of the shortest path between two vertices, that is, the number of edges crossed while traveling from one vertex to the other. Naturally, there may be multiple shortest paths. Lemma C.1 in Appendix C proves that the intersection of two partial orders is also a partial order. It turns out that of the shortest paths between two partial orders, at least one runs via their intersection. However, finding one such path is sufficient for the present purposes.

My claim is that the distance between two partial orders A and B is given by the formula

$$|\#A - \#(A \cap B)| + |\#B - \#(A \cap B)|.$$

[4] As the number of vertices grows exponentially it is not feasible to draw complete graphs of partial orders on sets containing more than three elements. The graph of partial orders on twelve elements would contain as many as 414864951055853499 nodes (Erné and Stege 1991). Incidentally, since $12! \cdot 12! < 414864951055853499$, this also shows that not all partial orders can be expressed in terms of an intersection of two linear orders.

In other words, the distance is the number of pairs that exist in one partial order but do not exist in the other. The intersection contains the ordered pairs shared by both partial orders.[5]

The proof of the claim is not trivial. Two issues must be proved, the first that there exists a path running via the intersection the length of which is defined by the above formula, and the second that there does not exist any shorter path. Let us start by proving that there exists a path from an arbitrary partial order to any of its subsets, the length of which is the difference of their cardinalities.

LEMMA 13.2 Let $A \subset B$. If $|\#B - \#A| = N > 1$ then there exists a partial order C such that $A \subset C \subset B$ and $|\#B - \#C| = 1$ and $|\#A - \#C| = N - 1$.

Proof. Let $Z = B \setminus A$. We must prove that Z contains a pair (a, b) such that $B \setminus \{(a, b)\}$ is a partial order. According to Lemma C.2 in Appendix C it is sufficient to prove that Z contains a pair that is adjacent in B.

The strategy for proving this lemma is very similar to that used in proving Lemma 13.1. According to Lemma C.3 in Appendix C, set Z contains at least one minimal element. Let one of these minimal elements be denoted by x. Again there is an element y such that (x, y) is in Z, and we can assume that pair (x, y) is an adjacent pair in Z. Now, elements x and y are either adjacent in B or they are not. Both cases must be checked. If x and y are adjacent in B we can form the partial order $C = B \setminus \{(x, y)\}$, which according to Lemma C.2 is a partial order. Clearly, then, $A \subset C \subset B$, $|\#B - \#C| = 1$ and $|\#B - \#C| = N - 1$.

The second case is that x and y are not adjacent in B, but that there is one or several elements z such that $x < z < y$. Therefore there exists a chain

$$x = z_0 < z_1 < \ldots < z_{n-1} < z_n = y$$

in B in which all elements are adjacent. Now, at least one of the pairs (z_i, z_{i+1}), $0 \leq i < n$, must be in Z: let it be denoted by (z_k, z_{k+1}). If this were not the case, then the whole chain would be in A and then, because of the transitivity of the partial order A, $x \leq y$ would also be in A and then $x \leq y$ would not be in Z. We could then form the partial order $C = B \setminus \{(z_k, z_{k+1})\}$, which according to Lemma C.2 is a partial order. Clearly, then, $A \subset C \subset B$ and $|\#B - \#C| = 1$ and $|\#B - \#C| = N - 1$. Both cases have been checked and this proves the lemma. □

LEMMA 13.3 The distance between partial order B and its subset $A \subset B$ is given by the formula $|\#B - \#A|$.

Proof. Let $A \subset B$ and $N = |\#B - \#A|$. If $N = 1$, the partial orders are adjacent. If $N > 1$, then according to Lemma 13.2 there exists a partial order C such that $A \subset C \subset B$, where $|\#B - \#C| = 1$ and $|\#C - \#A| = N - 1$. Using Lemma 13.2 iteratively gives the result that there exists a path

[5] My claim is slightly more general than would be necessary in the analysis of ORDER INVERSIONS and BADNESS OF SERIAL FIT. Namely, I will be considering twelve-tone rows as partial orders, and they contain exactly 66 non-reflexive pairs. Hence, in the case of twelve-tone rows we know that the equation $|\#A - \#(A \cap B)| = |\#B - \#(A \cap B)|$ holds.

with cardinality $|\#B - \#(A \cap B)|$. Moreover, there does not exist a shorter path because at every edge exactly one pair is added or removed. Therefore the distance is exactly $|\#B - \#A|$. □

Having now completed all the difficult stages we obtain as a corollary the result proposed earlier concerning the distance between two partial orders.

COROLLARY 13.1 The distance between two partial orders A and B is given by the formula

$$|\#A - \#(A \cap B)| + |\#B - \#(A \cap B)|.$$

Proof. As $A \cap B$ is a subset of A, Lemma 13.1 proves that there is a path from A to $A \cap B$ of length $|\#A - \#(A \cap B)|$. Correspondingly, there is a path from B to $A \cap B$ of length $|\#B - \#(A \cap B)|$. Thus there exists a path from A to B of length $|\#A - \#(A \cap B)| + |\#B - \#(A \cap B)|$, and there does not exist a shorter path because at every edge exactly one pair is added or removed. On the way from A to B $|\#A - \#(A \cap B)|$ pairs must be removed and $|\#B - \#(A \cap B)|$ pairs added. Therefore the distance between two total orders is $|\#A - \#(A \cap B)| + |\#B - \#(A \cap B)|$. □

To conclude this section, let us briefly discuss the concept of *linear extension*. The linear extension of a given partial order is a total order that contains the given partial order as a subset. The number of different linear extensions describes how much the partial order has left "undecided." The set of linear extensions is what Daniel Starr labels the *total order class*: the set of rows that satisfy a protocol (Starr 1984, 188).

Counting the linear extensions of a given partial order is not a trivial task. In fact, Graham Brightwell and Peter Winkler have proved that it is a $\#P$ complete problem (Brightwell and Winkler 1991). In complexity analysis it is believed that there is no polynomial-time algorithm for solving $\#P$ complete problems. Thus, it is difficult to say from looking at a partial order what is the exact number of its linear extensions. (This translates directly into the fact that it is difficult to say from looking at two twelve-tone rows what the BADNESS OF SERIAL FIT value is.) However, Gara Pruesse and Frank Ruskey have developed – within the limits of possibility – an efficient algorithm for generating linear extensions; the running time of the algorithm depends on the number of linear extensions to be generated (Pruesse and Ruskey 1997, 273).[6] In technical terms the algorithm is $O(N)$, where N is the number of objects generated. However, as the size of the set increases, the maximum number of linear extensions of a partial order grows exponentially and thus the calculation time required grows exponentially.

13.2 Representing similarity measures using partial orders

13.2.1 Background

The graph of partial orders on the set $\{0, 1, \ldots, 11\}$ can be used to analyze the relationships and similarity of twelve-tone rows. According to the conception of a row as a set of ordered pairs, a

[6] An implementation of the algorithm written in C by Kenny Wong and Frank Ruskey is available under the GNU General Public License.

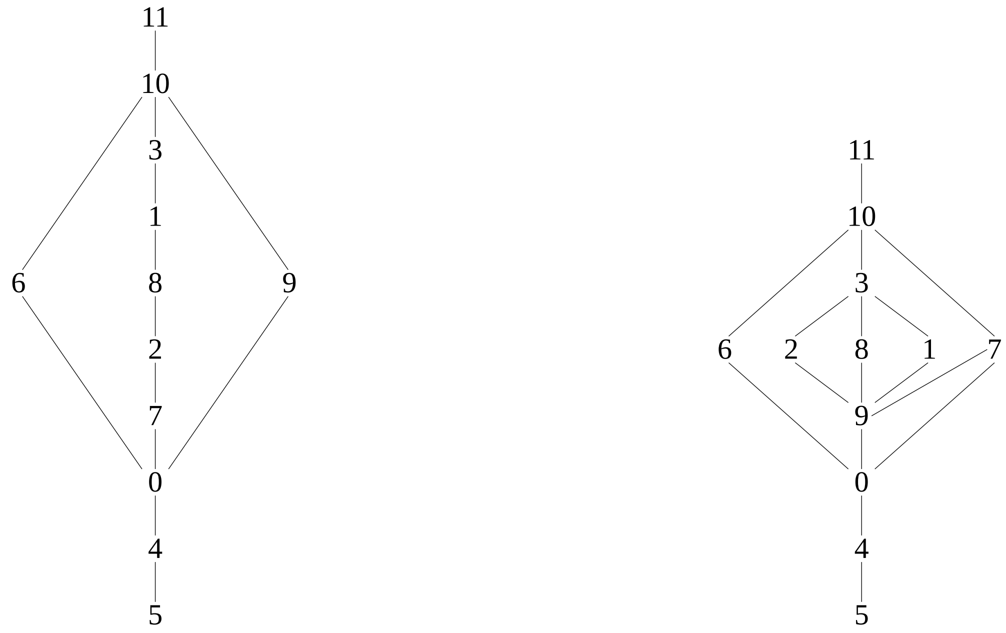


Figure 13.2: Two Hasse diagrams. The diagram on the left depicts the partial order defined by rows $P = 5409728136\text{AB}$ and $Q = 5406728139\text{AB}$ and the diagram on the right depicts the partial order defined by rows $P = 5409728136\text{AB}$ and $S = 5406918237\text{AB}$.

twelve-tone row is a total order, and therefore the intersection of any two twelve-tone rows is a partial order. This observation leads to two different ways of evaluating the similarity of two rows using the graph of partial orders. First, a natural method for defining distance between partially ordered sets is to define the distance as the length of the shortest path from one partial order to another in the graph presented above. The second method is to count the linear extensions of the intersections of the rows. These two methods correspond to the similarity measures ORDER INVERSIONS and BADNESS OF SERIAL FIT.

A list of ordered pairs is not a very illustrative representation of partial orders, and *Hasse diagrams* are more helpful (Stanton and White 1986, 27).[7] By convention, the largest elements are drawn at the top and the smallest elements at the bottom of the diagram. Figure 13.2 shows two partial orders as Hasse diagrams: one defined by rows P and Q in Figure 7.3 and the other defined by rows P and S.

Only cover relations are drawn in a Hasse diagram: the partial order depicted is the transitive closure of the cover relations. Therefore, in both diagrams in Figure 13.2 a line goes up from 5 to 4 denoting the relation $5 \leq 4$, and another line goes up from 4 to 0 denoting the relation $4 \leq 0$. Nevertheless, due to the transitivity of partial orders it is implied that the relation $5 \leq 0$ is also included. Starr refers to the function that reduces a partial order to a set of cover relations as Prn (shorthand for "pruning"), and to the function that extends a set of cover relations to a proper partial order as Ext (shorthand for "extension") (Starr 1984, 190–192).

[7]Both Lewin (1976) and Starr (1984) provide visual means of illustration that are similar to Hasse diagrams.

13.2.2 ORDER INVERSIONS and partial orders

Let us now discuss twelve-tone rows in terms of partial orders. Let us use rows P and Q in Figure 7.3 as our example. A twelve-tone row is a total order and therefore it can be written as a single chain. Row P can be written as a single chain as follows:

$$5 \leq 4 \leq 0 \leq 9 \leq 7 \leq 2 \leq 8 \leq 1 \leq 3 \leq 6 \leq 10 \leq 11.$$

Correspondingly, row Q can be written as a single chain as follows:

$$5 \leq 4 \leq 0 \leq 6 \leq 7 \leq 2 \leq 8 \leq 1 \leq 3 \leq 9 \leq 10 \leq 11.$$

The theory of partial orders can now be used to analyze the relationship between these two rows. Below is a complete list of pairs that are in row P but not in row Q (all of them are naturally non-reflexive):

$$9 \leq 7,\ 9 \leq 2,\ 9 \leq 8,\ 9 \leq 1,\ 9 \leq 3,\ 9 \leq 6,\ 3 \leq 6,\ 1 \leq 6,\ 8 \leq 6,\ 2 \leq 6,\ 7 \leq 6.$$

The list of pairs that are in row Q but not in row P would contain precisely the inversions of these pairs. Incidentally, the above list is also the set $P \setminus (P \cap Q)$. The inverse elements of that set form the set $Q \setminus (P \cap Q)$, which has the same number of elements as set $P \setminus (P \cap Q)$. The number of order inversions is the length of the path from row P or row Q to the intersection $P \cap Q$. Thus we obtain the identity

$$\begin{aligned} &2 \cdot \text{number of order inversions between rows } P \text{ and } Q \\ &= \text{the length of the path from row } P \text{ to row } Q. \end{aligned}$$

Removing the pairs from the partial order P one at a time in exactly the order in which they are enumerated in the above list gives a path from P to $P \cap Q$. It is left as an exercise for the reader to verify that it consists of adjacent partial orders.

I have shown in Section 9.2 that ORDER INVERSIONS is a similarity measure that defines a metric. The structure of the graph makes this rather obvious. The distances are symmetric: the distance between two total orders A and B is the length of the shortest path between them and it is, of course, equal to the distance between B and A. The distance from a total order to itself is zero, and two non-identical total orders have a positive distance. That triangle inequality holds is also rather obvious: if the length of the path is denoted by d, the length of the composition of the paths from A to B and from B to C is the sum of the lengths of the two paths and thus $d(A, C) \leq d(A, B) + d(B, C)$.

The idea of measuring similarity according to the distances in a graph is not new. John Rahn writes as follows.

> . . . any path through a graph, or network, can be viewed as a similarity relation quanti-

fied by "nearness" in the path, or more precisely, by the number of intervening element-nodes. (Rahn 1979–80, 496)

This idea is reiterated in Morris (1987, 103–105).

13.2.3 BADNESS OF SERIAL FIT and partial orders

Let us now state the BADNESS OF SERIAL FIT similarity measure in terms of partial orders. It is based on the notion of protocol pairs introduced by Milton Babbitt (1962). David Lewin analyzed the properties of protocol pairs further, and invoked the notion of partial orders (Lewin 1976). He also introduced the idea of measuring the similarity of twelve-tone rows based on how badly they "fit together." However, he did not take the step to formulate the similarity measure fully in terms of partial orders. The necessary observation here is that a twelve-tone row is also a partial order: a special type partial order in which the order of every element is defined, but a partial order nevertheless. In addition, the intersection of two partial orders is a partial order (see Lemma C.1 in Appendix C). Hence, the protocol defined by two rows – the set of common pairs – is a partial order that is simply the intersection of the rows. The method for creating the protocol Lewin gives in the appendix of his article (Lewin 1976, 257) is much more cumbersome than calculating the intersection of two partial orders.

According to the theory of partial orders, BADNESS OF SERIAL FIT is equivalent to counting the number of linear extensions of the intersection of two rows. For example, it was established in Section 9.3 that the BADNESS OF SERIAL FIT value for rows P and Q was 42. In terms of partial orders this means that the partial order defined by the intersection of rows P and Q has 42 linear extensions.

An attractive property in this approach is that it is not restricted to the comparison of two rows. All of the above definitions and theorems can be generalized by induction to any number of partial orders (or rows). The intersection of any number of twelve-tone rows is well defined, and defines the area that fits all given rows.

For example, the protocol defined by rows A, B, and C is $A \cap B \cap C$, which could equivalently be defined by two protocols of two rows: $A \cap B \cap C = (A \cap B) \cap (B \cap C)$. Note that the set of linear extensions of protocol $A \cap B \cap C$ is usually a proper superset of the union of the sets of linear extensions of $A \cap B$ and $B \cap C$.

13.2.4 Conclusions

Lewin notes that the two measures ORDER INVERSIONS and BADNESS OF SERIAL FIT are related (Lewin 1976, 256). The theory of partial orders explains why and how: they are two different aspects of the graph formed by partial orders. Further explorations of this theory might give more insight into the relationships between twelve-tone rows. While the whole graph is far too large to be fully explored, some parts of it, such as in the neighborhood of total orders, are within reach.

While ORDER INVERSIONS and BADNESS OF SERIAL FIT are both based on ordered dyads, they are two different approaches to the measurement of the similarity of twelve-tone rows. Figures 6.5 and 6.6 illustrate these approaches. ORDER INVERSIONS is a traditional measurement of distance, while BADNESS OF SERIAL FIT could be seen as describing the distance between rows in terms of the sector required to cover both. If the rows are close to each other, a small sector is enough, and if they are far away, a large sector is required.

BADNESS OF SERIAL FIT is related to but not the same as relative similarity or percentiles (see Section 6.4). In order to illustrate the difference, let us consider row $P = 5409728136\mathrm{AB}$ and row $P' = 4509728136\mathrm{AB}$, which is obtained by exchanging the two first pitch classes of row P. Obviously, there is exactly one order inversion between the two rows. The BADNESS OF SERIAL FIT value for these two rows is 2; the intersection of the two rows $P \cap P'$ has only two linear extensions – rows P and P'. If we then consider the relative similarity, we need to find all rows that are at most at a distance of 1 from row P, in other words that have at most one order inversion with P. There are a total of 11 such rows, for example row $P'' = 5049728136\mathrm{AB}$. Now, row P'' has the ordered pair $(0, 4)$ that neither row P nor row P' has. Hence, the approach that BADNESS OF SERIAL FIT represents cannot be reduced to relative similarity.

We are now better able to understand the different approaches to the measurement of similarity discussed in Section 6.4. ORDER INVERSIONS represents the approach of traditional distance: the similarity of rows is directly related to the number of different ordered pairs in them and other twelve-tone rows play no role. In the relative approach we still first measure the number of different ordered pairs in the rows, but then continue the evaluation by asking how common it is to have such a value. However, only BADNESS OF SERIAL FIT is able to provide information that goes beyond the mere quantitative measurement of differences: it gives information about the specific configurations of differences that amount to the distance measured by means of ORDER INVERSIONS.

In order to clarify the difference and to obtain a deeper understanding of the unique features of BADNESS OF SERIAL FIT, let us consider one more example involving three twelve-tone rows: $A = 0123456789\mathrm{AB}$, $B = 2103456789\mathrm{AB}$, and $C = 1032546789\mathrm{AB}$. In particular, the idea is to compare rows A and B and rows A and C. In both pairs of rows the number of order inversions is three: $OI(A, B) = OI(A, C) = 3$. Yet, we notice immediately that the order inversions in rows A and B involve three pitch classes while those in rows A and C involve six. Hence, rows A and B are similar in a different way from rows A and C even if the distances are the same. BADNESS OF SERIAL FIT is, nevertheless, able to distinguish these cases: $BSF(A, B) = 6$ and $BSF(A, C) = 8$. The moral of this comparison is that according to some similarity measure two rows may be equally distant from a third row, but it is built up from different components.

13.3 Applications to other conceptions of rows

In the preceding sections graph theory was applied to the analysis of two similarity measures – ORDER INVERSIONS and BADNESS OF SERIAL FIT – both of which based on the conception of twelve-tone rows as a set of ordered pairs. We could apply a similar approach to the analysis of similarity measures based on other conceptions of twelve-tone rows. The following sections give a brief outline of graphs based on the conceptions of twelve-tone rows as a vector and as a set of subsegments.

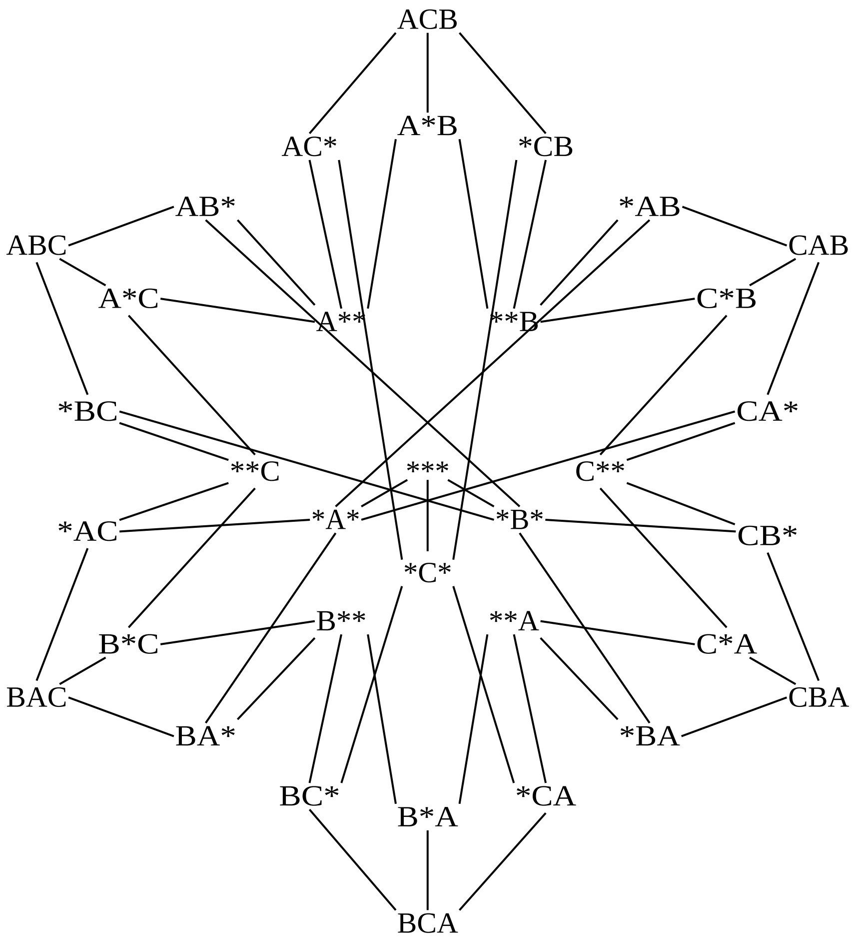


Figure 13.3: A graph of partially defined vectors of three elements.

13.3.1 Graph theory and the vector approach to twelve-tone rows

According to the vector approach, a twelve-tone row comprises twelve order positions filled with twelve pitch classes (or, correspondingly, twelve positions filled with order numbers), and the graph would contain as its vertices vectors in which the entries in some of the order positions may be undefined. An order position with undefined entry is denoted in the following with an asterisk. For example, in vector $5409728136\mathrm{AB}$ the entries in all twelve order positions are defined, while in vector $5{*}0{*}7{*}8{*}3{*}\mathrm{A}{*}$ six of them are defined and six are undefined.

Figure 13.3 depicts a graph of vectors of length 3. I will refer to this construct as the *graph of partially defined vectors*. The vertices are partially defined vectors with some entries defined (denoted by letters) and some entries undefined (denoted by an asterisk). The edges of the graph are based on the inclusion relation: two vertices are adjacent if and only if one is a "subset" of the other and there does not exist an intermediate vector. For example, entries ACB and $\mathrm{A{*}B}$ are adjacent since the latter is a "subset" of the former (the entries that are defined in the latter are defined similarly in the former) and there does not exist an intermediate vector (a vector that would be a proper superset of $\mathrm{A{*}B}$ and a proper subset of ACB).

The graph of partial orders gave rise to two similarity measures: ORDER INVERSIONS and

BADNESS OF SERIAL FIT. Inspired by the correspondence between the graph of partially defined vectors and that of partial orders, we might similarly define two similarity measures based on the former. The similarity measure corresponding to ORDER INVERSIONS would be DERANGEMENT: in the graph of vectors of length 12 the distance between two twelve-tone rows is exactly twice the DERANGEMENT value.

The similarity measure corresponding to BADNESS OF SERIAL FIT might be tentatively be labeled BADNESS OF VECTOR FIT. In order to calculate the BADNESS OF VECTOR FIT value for two twelve-tone rows we would first look for pitch classes that the two rows have in the same order positions (the number of such pitch classes is twelve minus the DERANGEMENT value). The value would then be the number of rows that have those pitch classes in the same order positions as the two rows. The value of BADNESS OF VECTOR FIT, unlike that of BADNESS OF SERIAL FIT, is easily computed. Namely, if two rows contain n pitch classes in the same order positions then the number of rows that have those n pitch classes in those order positions is $(12-n)!$, which is simply the number of permutations of the remaining $12-n$ pitch classes.

13.3.2 Graph theory and the subsegment approach to twelve-tone rows

In the subsegment approach a twelve-tone row is conceived of as a set of subsegments. Chapter 10 described several approaches to the measurement of the similarity of twelve-tone rows based on subsegments. Basically, we could concentrate on subsegments of some fixed length or consider them all, or we could examine the longest shared subsegment.

Graphs based on subsegments of length 2 (ordered pairs) were discussed extensively above. It turns out that the graph of ordered pairs is also essentially the graph of all the subsegments, in other words the ordered pairs of a row define all its subsegments of any length. Therefore, we would not obtain much new information by constructing a graph using the set of subsegments as a basis.

However, we can construct a graph based on the idea of the longest shared subsegment in two rows. Figure 13.4 depicts a graph of segments of at most three elements. I will refer to this construct as the *graph of subsegments*. The vertices of the graph are subsegments. The edges are based on the inclusion relation: two vertices are adjacent if and only if one is a subsegment of the other and there does not exist an intermediate subsegment. Consequently, the lengths of the subsegments in adjacent vertices differ by exactly one. For example, vertices ACB and AB are adjacent since the latter is a subsegment of the former and there does not exist an intermediate subsegment (a segment that would be a proper supersegment of AB and a proper subsegment of ACB).

Again we can create two similarity measures based on the graph. The measure corresponding to ORDER INVERSIONS would be ULAM'S DISTANCE: in a graph of subsegments of at most twelve elements the distance between two twelve-tone rows is exactly twice the ULAM'S DISTANCE value.

The similarity measure corresponding to BADNESS OF SERIAL FIT might be tentatively be labeled BADNESS OF SUBSEGMENT FIT. In order to calculate the BADNESS OF SUBSEGMENT FIT value for two twelve-tone rows we would first look for the longest shared subsegments (the length of that subsegment is twelve minus the ULAM'S DISTANCE value). The value would then be the number of rows that contain that subsegment. The value of BADNESS OF SUBSEGMENT FIT is

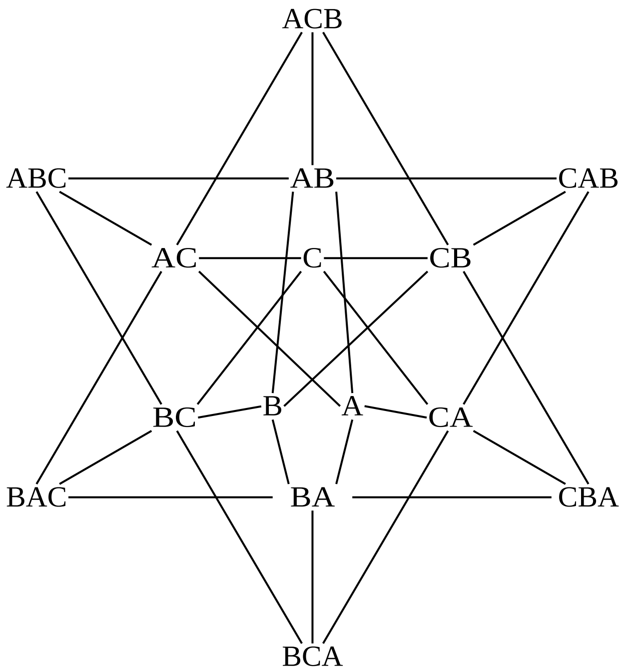


Figure 13.4: A graph of subsegments of at most three elements.

easily computed, unlike that of BADNESS OF SERIAL FIT. Namely, if two rows contain a shared subsegment of length n, the number of rows that contain that subsegment is $12!/n!$ – there are $n!$ possible permutations of the n elements and only one of those is the order in the shared subsegment.

As noted in Section 10.4, focusing on the longest shared subsegment fails to take into account their multiplicity. The same caveat applies to the graph of subsegments. As long as we focus only on a single longest subsegment we must ignore the issue of multiplicity. If we were to take into account all shared subsegments between two rows, we would end up with the similarity measures ORDER INVERSIONS and BADNESS OF SERIAL FIT: these two measures precisely take into account all shared subsegments of rows, even if they are expressed in terms of shared ordered dyads.

CHAPTER 14

Conclusions

In the previous chapters I introduced a number of similarity measures and analyzed some of their properties. I also considered how we could use graph theory in discussing the relations of twelve-tone rows.

I will take a broader view in this chapter in an attempt to evaluate the similarity measures *in toto*. I will begin by comparing them from two different angles: first, in Section 14.1 I will discuss how the usual row operations relate to them and then in Section 14.2 I will examine whether the similarity measures support my claim that there are multiple different approaches to the similarity of twelve-tone rows. I will suggest some potential analytical applications in Section 14.3, and offer some topics for further research in the final Section 14.4.

14.1 Row operations and similarity

As noted in the preceding chapters, according to some similarity measures two rows belonging to the same row class may be maximally dissimilar. In other words, a row operation may transform one row into another that is maximally dissimilar. In this section I will examine the relation between classic row operations and similarity measures in more general terms. The basis of this examination is the division of row operations into pitch-class operations and order-number operations.

14.1.1 Pitch-class operations and similarity

The majority of the similarity measures discussed in this work measure similarity in the order-number dimension. Given the permutational nature of row operations, applying a pitch-class operation to a row may change the order positions of the pitch classes. Consequently, a pitch-class operation also moves pitch classes in the order-number dimension, but it depends on the structure of the row *how* it does so. For example, if a row begins with pitch class 0 and we transpose it by some non-trivial pitch-class interval, it depends on the structure of the row how the transposition moves pitch class 0 in the order-number dimension, in other words at which order position pitch class 0 will be in the transposed row. Therefore, we cannot generalize about the similarity of the rows related by a pitch-class operation in the order-number dimension.

Let us consider transposition T_6. Rows $A = 061728394\text{A}5\text{B}$ and $T_6A = 60718293\text{A}4\text{B}5$ are related by transposition T_6. Since $T_6A = \boldsymbol{\alpha_1}A$, these two rows are also very similar in the order-number dimension – using virtually any similarity measure based on the GIS of order number-rows except DERANGEMENT.[1] Rows $P = 5409728136\text{AB}$ and $T_6P = \text{BA}6318279045$ are also related by transposition T_6. Since $T_6R = RP$, these two rows are very dissimilar in the order-number dimension – using virtually any similarity measure based on the GIS of order-number rows. Hence, in the order-number dimension the similarity of rows related by a pitch-class operation depends to a very high degree on the structures of the rows.

In the pitch-class dimension, however, we can predict the similarity of rows related by a pitch-class operation since the similarity of the rows is defined in terms of the complexity of the pitch-class operation mapping one row into the other.

With respect to the INT of rows, note that transpositions keep them invariant. Inversions, however, alter the ordered pitch-class intervals, and whether two rows related by an inversion are similar or dissimilar with respect to their unordered INT contents depends on their interval structure. For example, for rows $A = 061728394\text{A}5\text{B}$ and $I_1A = 1706\text{B}5\text{A}49382$, $ID(A, I_1A) = 10$, which is a relatively low value. However, for rows $B = 0369147\text{A}258\text{B}$ and $I_1B = 1\text{A}740963\text{B}852$, $ID(B, I_1B) = 62$, which is close to the maximum value.

The degree to which a pitch-class operation changes the subset contents of rows depends on the structure of the rows. For example, the transpositionally related rows $C = 012345\text{BA}9876$ and $T_6C = 6789\text{AB}543210$ have identical subset contents. This happens because rows related by retrograde have identical subset contents and row C is one of those that are invariant under operation RT_6. In contrast, the transpositionally related rows $D = 012346579\text{B}8\text{A}$ and $T_6D = 6789\text{A}0\text{B}13524$ do not share a single non-trivial subset.

In sum, how a pitch-class operation changes a row depends on the structure of the row given any conception of rows except the vector approach in the pitch-class dimension.

14.1.2 Order-number operations and similarity

What applies to pitch-class operations and similarity in the pitch-class/order-number dimensions also applies *mutatis mutandis* to order-number operations and similarity in the order-number/pitch-class dimensions. In the order-number dimension we can predict the similarity of rows related by an order-number because it is defined in terms of the complexity of the order-number operation that transforms one row into the other.

In the pitch-class dimension the degree of similarity of rows related by an order-number operation depends on their structure. For example, the two retrograde-related rows $E = 013579\text{A}8642\text{B}$ and $RE = \text{B}2468\text{A}975310$ are relatively similar in the pitch-class dimension since, according to

[1] Operation $\boldsymbol{\alpha_1}$ refers here to an order-number operation that belongs to a family of row operations known as alpha-operations (Morris 1982; Ilomäki 2005). Operation $\boldsymbol{\alpha_1}$ is the order-number operation $\mathbf{1032547698BA}$ that is equivalent to exchanging adjacent pitch classes at order positions $\mathbf{0}$ and $\mathbf{1}$, $\mathbf{2}$ and $\mathbf{3}$, $\mathbf{4}$ and $\mathbf{5}$, etc.

PITCH-CLASS DISPLACEMENT, $PCDIS(E, RE) = 12$ (since in the pitch-class dimension retrograding a row displaces every pitch class by at least one semitone, the minimum PITCH-CLASS DISPLACEMENT value for retrograde-related rows is $12 \cdot 1 = 12$). On the other hand, the retrograde-related rows $F = 012345\text{BA}9876$ and $RF = 6789\text{AB}543210$ are maximally dissimilar in the pitch-class dimension since, according to PITCH-CLASS DISPLACEMENT, $PCDIS(F, RF) = 72$.

When a row is retrograded, its succession of ordered pitch-class intervals is retrograded and inverted. Consequently, the degree to which the successions of ordered pitch-class intervals of rows related by retrograde are similar depends on the structure of the rows. For example, the retrograde-related rows $F = 012345\text{BA}9876$ and $RF = 6789\text{AB}543210$ have identical INTs and, hence, they are maximally similar. On the other hand, the retrograde-related rows $G = 012387\text{A}56\text{B}49$ and $RG = 94\text{B}65\text{A}783210$ have maximally dissimilar successions of ordered pitch-class intervals.

As noted in Chapter 11, the two similarity measures based on the conception of rows as a set of subsets – NESTINGS and SCATTERING – do not distinguish between rows related by retrograde. However, other order-number operations, such as rotations, do alter the subset contents of rows.

14.1.3 Similarity measures modulo a group of row operations

Similarity measures NESTINGS, SCATTERING and INTERVALLIC DISTANCE have the special property that they are invariant under a nontrivial group of row operations: NESTINGS and SCATTERING are invariant under operations $\{T_0, R\}$ and INTERVALLIC DISTANCE is invariant under operations $\{T_0, T_1, \ldots, T_{11}\}$. It is worthwhile to consider the implications of these invariances to the pitch-class domain and the order-number domain: if the two domains were isomorphic or structurally identical (compare Section 2.1.2), we should find corresponding invariances in both domains. Let us examine whether that is the case.

The group of order-number operations that corresponds to the group of pitch-class operations $\{T_0, T_1, \ldots, T_{11}\}$ is the group of rotations $\{r_0, r_1, \ldots, r_{11}\}$. These two groups are isomorphic (both are cyclic groups of cardinality 12). The INTERVALLIC DISTANCE similarity measure is invariant "modulo transposition." A similarity measure would be invariant "modulo rotation" only if rows were treated as cyclic entities and only the relations of adjacent order numbers were taken into account. However, as already mentioned in Section 2.1.2, a row is not usually considered a cyclic structure, and as discussed in Section 12.1, no convincing base was found for similarity measures based on order-number intervals. Hence, none of the similarity measures discussed in this work are defined "modulo rotation."

The group of order-number operations $\{T_0, R\}$ is illustrative since there does not exist a corresponding group of pitch-class operations and thus it reveals a conceptual difference between the pitch-class domain and the order-number domain. Retrograde is the inversion $\boldsymbol{I_{11}}$ applied to the order numbers. Now, while there are twelve groups of pitch-class operations $\{T_0, I_k\}$ there is no justification for selecting one of them as the one that corresponds to the retrograde operation (compare Section Definitions and conventions). Hence, it would unfounded even to attempt to find a corresponding group of pitch-class operations. Consequently, there are no similarity measures that are defined modulo inversion.

14.2 The similarity measures compared

My working hypothesis was that there are different aspects of similarity, and that the degree of similarity of two rows depends on which aspect is in focus. I will now consider whether that premise can be justified in terms of the similarity measures, in other words whether there is correlation between values based on different conceptions. I will also assess coherency of a conception by examining the correlation of the similarity measure values based on it.

I introduced five different conceptions of twelve-tone rows in my discussion of similarity measures: vector, ordered pairs, subsegments, subsets, and interval contents. In addition, the vector approach divides into the pitch-class domain and the order-number domain. In the following, I will compare some similarity measures belonging to these categories. I will not examine all of the possible combinations in detail: my aim is to find some general trends.

In addition to the five conceptions I also introduced three transformational procedures: adjacent exchanges, exchanges, and moves (see Sections 9.4.1, 9.4.2, and 10.8). These three procedures are related to the conception of a row as a set of ordered pairs, as a vector, and as a set of subsegments, respectively. Naturally, adjacent exchanges and exchanges are related. I will also examine whether there is more correlation between the conceptions of rows these two procedures suggest than between exchanges and moves or between adjacent exchanges and moves.

14.2.1 PITCH-CLASS DISPLACEMENT versus DISPLACEMENT

Even if the similarity measures PITCH-CLASS DISPLACEMENT and DISPLACEMENT are based on the same idea of a vector space, and both have the same structure and range, the fact that they measure similarity in different dimensions means that their values do not necessarily coincide. A pair of rows may have a small PITCH-CLASS DISPLACEMENT value and a large DISPLACEMENT value, and vice versa. As an example, let us consider row $P = 5409728136\text{AB}$ in Figure 7.3 and row $X = 63\text{BA}81724590$. Figure 14.1 depicts a comparison of these two rows using both DISPLACEMENT and PITCH-CLASS DISPLACEMENT, and also provides a method for visually illustrating the difference between these two similarity measures. The former measures how the pitch classes move in the horizontal dimension (the order-number space), whereas the latter measures how they move in the vertical dimension (the pitch-class space). The DISPLACEMENT value for rows P and X is 72, and the PITCH-CLASS DISPLACEMENT value for rows P and X is 12. Hence, the values of these two similarity measures do not coincide.

Rows P and X represent an extreme case in which the two rows are maximally dissimilar with respect to DISPLACEMENT but relatively similar with respect to PITCH-CLASS DISPLACEMENT. Similarly, we can construct two rows $Y = 061728394\text{A}5\text{B}$ and $T_6Y = 60718293\text{A}4\text{B}5$ that are maximally dissimilar with respect to PITCH-CLASS DISPLACEMENT but relatively similar with respect to DISPLACEMENT. The DISPLACEMENT value for rows Y and T_6Y is 12, and the PITCH-CLASS DISPLACEMENT value is 72.

Figure 14.2 depicts a comparison of the DISPLACEMENT and PITCH-CLASS DISPLACEMENT values. The horizontal axis denotes the DISPLACEMENT values and the vertical axis the PITCH-

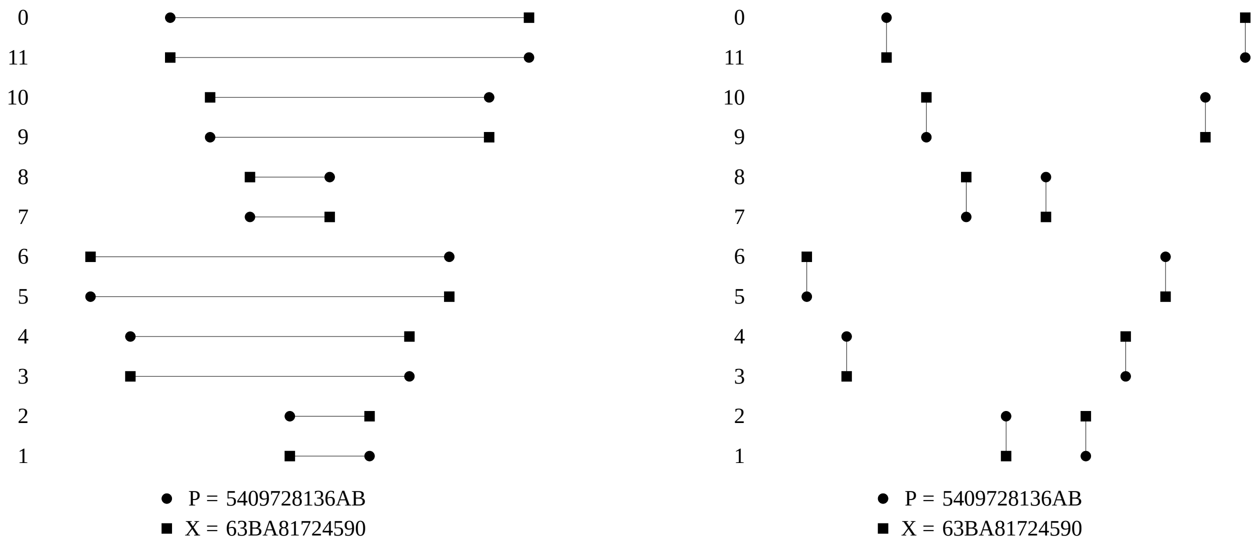


Figure 14.1: A comparison of thef DISPLACEMENT and PITCH-CLASS DISPLACEMENT values. In both pictures rows $P = 5409728136\text{AB}$ and $X = 63\text{BA}81724590$ are depicted in a two-dimensional space, in which the horizontal dimension represents the order-number space and the vertical dimension represents the pitch-class space. The figure on the left illustrates the measurement of similarity with respect to DISPLACEMENT between two rows and the figure on the right illustrates the measurement of similarity with respect to PITCH-CLASS DISPLACEMENT.

CLASS DISPLACEMENT values. The correlation is negligible: the DISPLACEMENT value provides little help in predicting the PITCH-CLASS DISPLACEMENT value, and vice versa. Consequently, the division between similarity in the pitch-class domain and in the order-number domain is sound.

14.2.2 DERANGEMENT versus DISPLACEMENT and PITCH-CLASS DISPLACEMENT

DERANGEMENT is arguably the most straightforward similarity measure based on the vector approach. Its values also provide a relatively good estimate for the two other similarity measures based on this approach: DISPLACEMENT and PITCH-CLASS DISPLACEMENT. Figure 14.3 depicts the correlations between DERANGEMENT and DISPLACEMENT and between DERANGEMENT and PITCH-CLASS DISPLACEMENT in graphical format.

Let us derive the following inequalities for the DERANGEMENT and DISPLACEMENT values:

$$2 \cdot \lceil D(X,Y)/2 \rceil \leq DIS(X,Y) \leq \sum_{0<k\leq D(X,Y)/2} 2 \cdot (11 - 2 \cdot (k-1)).$$

The lower limit stems from the facts that the minimum (non-zero) displacement of a pitch class in the order-number dimension is one order position and that only an even number of pitch classes can all be simultaneously displaced by one order position. Hence, if an even number of pitch classes are displaced then the minimum DISPLACEMENT value is obtained if all pitch classes are displaced by one order position. Similarly, if an odd number of pitch classes are displaced then the minimum DISPLACEMENT value is obtained if one pitch class is displaced by two order positions and the other pitch classes are displaced by one order position.

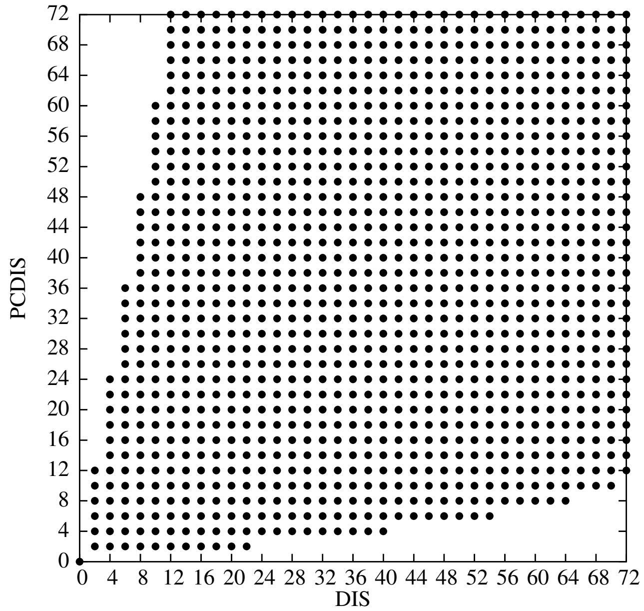


Figure 14.2: The correlation of the DISPLACEMENT and PITCH-CLASS DISPLACEMENT values.

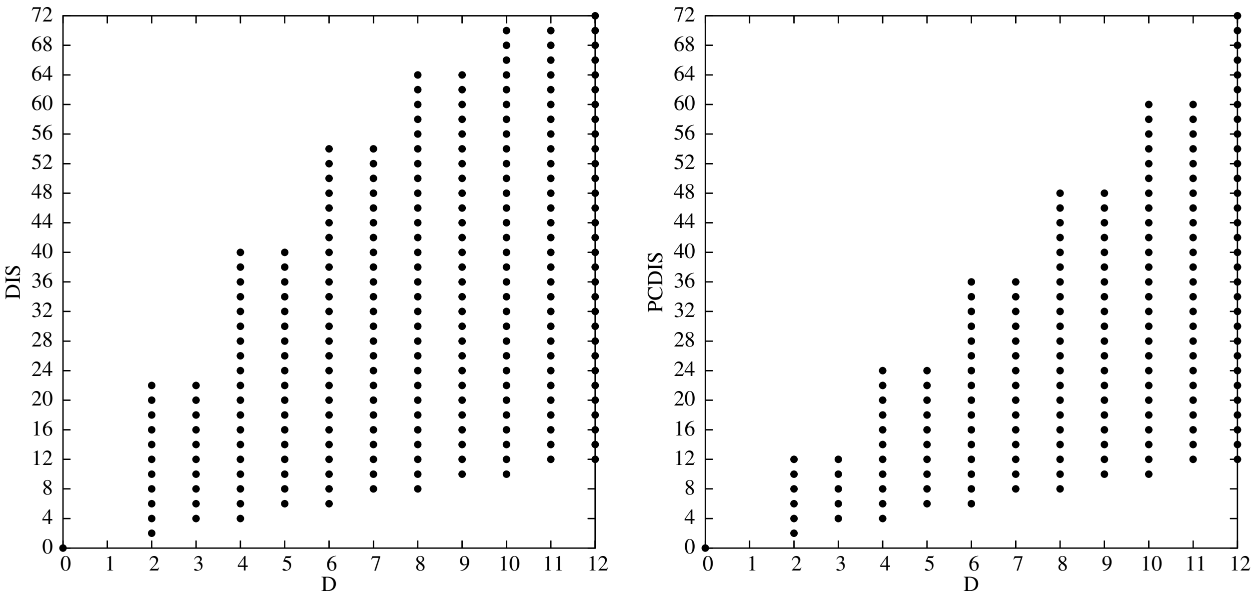


Figure 14.3: The correlation of the similarity measures DERANGEMENT and DISPLACEMENT on the left, and of DERANGEMENT and PITCH-CLASS DISPLACEMENT on the right.

The formula for the upper limit is slightly awkward. Basically, if $D(X, Y) = 0$ then the sum is empty and by convention the value of an empty sum is 0. If $D(X, Y) = 2$ or $D(X, Y) = 3$ then the sum is $2 \cdot (11 - 2 \cdot 0) = 2 \cdot 11 = 22$. Similarly, if $D(X, Y) = 4$ or $D(X, Y) = 5$ then the sum is $(2 \cdot (11 - 2 \cdot 0)) + (2 \cdot (11 - 2 \cdot 1)) = (2 \cdot 11) + (2 \cdot 9) = 40$. The formula stems from the observation that the maximum DISPLACEMENT value is obtained when pitch classes are exchanged pairwise as many order positions as possible. Hence, if two pitch classes are displaced, then the maximum DISPLACEMENT value is obtained if the first and last pitch classes of a row are exchanged, in which case both pitch classes are displaced by eleven order positions. If four pitch classes are displaced then the maximum DISPLACEMENT value is obtained if the first and last (both are displaced by eleven order positions) and the second and next to last (both are displaced by nine order positions) are exchanged.

Similarly, we can derive the following inequalities for the DERANGEMENT and PITCH-CLASS DISPLACEMENT values:

$$2 \cdot \lceil D(X,Y)/2 \rceil \leq PCDIS(X,Y) \leq 12 \cdot \lfloor D(X,Y)/2 \rfloor.$$

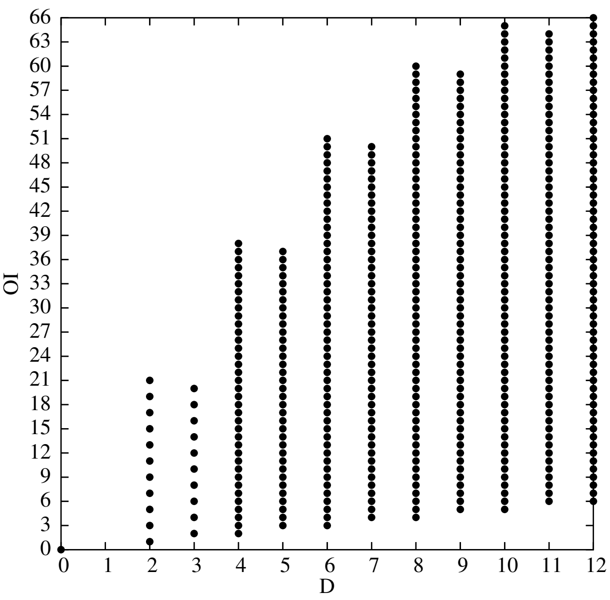
These inequalities result from a few simple observations. The lower limit stems from the facts that the minimum displacement of a pitch class in the pitch-class dimension is a semitone and that only an even number of pitch classes can all be simultaneously displaced by a semitone. Hence, if an even number $2k$ of pitch classes are displaced then the minimum PITCH-CLASS DISPLACEMENT value is obtained if all $2k$ pitch classes are displaced by one semitone. Similarly, if an odd number $2k + 1$ of pitch classes are displaced then the minimum PITCH-CLASS DISPLACEMENT value is obtained if one pitch class is displaced by two semitones and the remaining $2k$ pitch classes are displaced by one semitone.

The upper limit stems from the facts that the maximum displacement of a pitch class in the pitch-class dimension is a tritone and that only an even number of pitch classes can all be simultaneously displaced by a tritone. Hence, if an even number of pitch classes are displaced, then the maximum PITCH-CLASS DISPLACEMENT value is obtained if all pitch classes are displaced by a tritone. Similarly, if an odd number of pitch classes are displaced then the maximum PITCH-CLASS DISPLACEMENT value is obtained if three pitch classes are displaced by four semitones and the other pitch classes are displaced by a tritone.

14.2.3 DERANGEMENT versus ORDER INVERSIONS and CAYLEY DISTANCE

As discussed in Sections 9.2 and 9.5, the similarity measures ORDER INVERSIONS and CAYLEY DISTANCE are associated with the number of adjacent exchanges and the number of exchanges in the transformation that transforms one row into the other. A comparison of these two measures with DERANGEMENT gives some fundamental information about the number of distinct pitch classes that are displaced when a given number of exchanges is applied to a row.

Figure 14.4 depicts the correlation of the DERANGEMENT and ORDER INVERSIONS values and of the DERANGEMENT and CAYLEY DISTANCE values in graphical format. Let us derive the



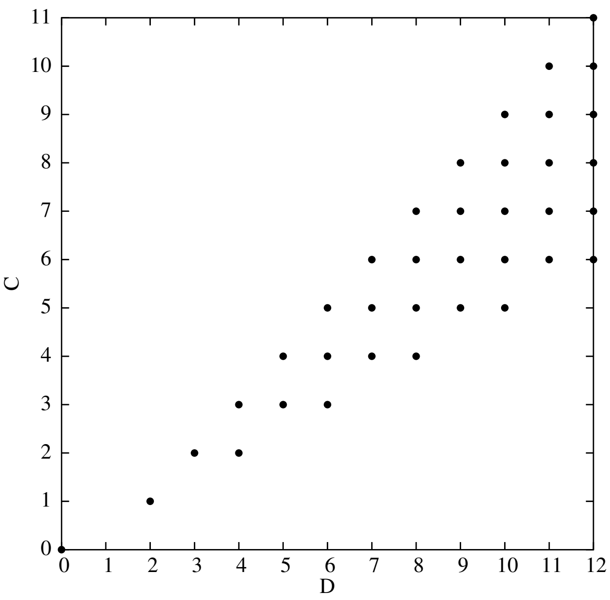


Figure 14.4: The correlation of the DERANGEMENT and ORDER INVERSIONS values on the left, and of the DERANGEMENT and CAYLEY DISTANCE values on the right.

following inequality for the DERANGEMENT and ORDER INVERSIONS values:

$$\lceil D(X,Y)/2 \rceil \leq OI(X,Y).$$

This inequality stems from the fact that exchanging two adjacent pitch classes results in one order inversion. Hence, for an even number of displaced pitch classes we obtain the smallest number of order inversions when we exchange pairs of adjacent pitch classes. For an odd number of displaced pitch classes (which must be at least three) we obtain the smallest number of order inversions when we rearrange one triplet of adjacent pitch classes $x_1x_2x_3$ as $x_2x_3x_1$, and for the remaining pitch classes we again exchange pairs of adjacent pitch classes.

The largest ORDER INVERSIONS value with respect to a given DERANGEMENT value is obtained when we exchange pairwise pitch classes that are far from each other. For example, exchanging the first and last pitch classes of a row results in $11 + 10 = 21$ order inversions even if only two pitch classes are displaced. Similarly, exchanging the first and last, and the second and the next to last results in $11 + 10 + 9 + 8 = 38$ order inversions even if only four pitch classes are displaced. Curiously, the maximum ORDER INVERSIONS value when an odd number $2k + 1$ of pitch classes are displaced is smaller than when only $2k$ pitch classes are displaced. This stems from the fact that if an odd number of pitch classes are displaced we cannot use the scheme of exchanging pitch classes pairwise.

The DERANGEMENT and CAYLEY DISTANCE values are strongly correlated. Let us derive the following inequalities for their values:

$$\lceil D(X,Y)/2 \rceil \leq C(X,Y) \leq D(X,Y) - 1.$$

sequence of exchanges	row	DIS	OI
P	5409728136AB	0	0
$F_{0,1}P$	4509728136AB	2	1
$F_{1,2}F_{0,1}P$	4059728136AB	4	2
$F_{0,1}F_{1,2}F_{0,1}P$	0459728136AB	4	3

Figure 14.5: A sequence of exchanges of adjacent pitch classes (operation $F_{n,m}$ denotes exchanging pitch classes at order positions n and m) applied to the Lyric Suite row $P = 5409728136\text{AB}$, and the DISPLACEMENT and ORDER INVERSIONS values of the resulting rows compared to row P.

The lower bound stems from the fact that exchanging two pitch classes displaces two pitch classes. Hence, if $2k$ or $2k-1$ pitch classes have been displaced, then at least k exchanges are involved. The upper bound stems from the fact that n pitch classes are displaced in the cycle $(x_1 \quad x_2 \quad \dots \quad x_k)$, and this cycle can be decomposed into $k-1$ exchanges.

14.2.4 ORDER INVERSIONS versus DISPLACEMENT

ORDER INVERSIONS and DISPLACEMENT turn out to be closely related similarity measures – even if they are based on different conceptions. Both measure in their own ways how the pitch classes have strayed in the order-number dimension from their original positions. Even if the scale of DISPLACEMENT (from 0 to 72) is slightly larger than that of order inversions (from 0 to 66), the resolution of ORDER INVERSIONS is better than that of DISPLACEMENT: the latter gives 37 while the former gives 67 distinct values.

I will illustrate some of the differences between these two measures by considering their values in consecutive exchanges of adjacent pitch classes. Let us use here the following notation: $F_{i,j}$ denotes an exchange (see Definition 9.4) that exchanges the pitch classes at order positions i and j. Given the focus here on exchanges of adjacent pitch classes, it will always hold that $j = i + 1$. Figure 14.5 shows a sequence of exchanges applied to row $P = 5409728136\text{AB}$ in Figure 7.3 and the DISPLACEMENT and ORDER INVERSIONS values of the resulting rows compared to the original row. The ORDER INVERSIONS value increases at every exchange, but the DISPLACEMENT value is the same for rows P and $F_{1,2}F_{0,1}P$ and rows P and $F_{0,1}F_{1,2}F_{0,1}P$. ORDER INVERSIONS has a better resolution than DISPLACEMENT: Figure 14.5 shows a case in which it is able to distinguish a minor difference whereas DISPLACEMENT is not.

I will now give a more detailed analysis of the relation between the ORDER INVERSIONS, DISPLACEMENT and CAYLEY DISTANCE values. Diaconis and Graham (1977, 264) proved an inequality that I express below in terms of the similarity measures discussed in this work:

$$OI(X,Y) + C(X,Y) \leq DIS(X,Y) \leq 2 \cdot OI(X,Y). \tag{14.1}$$

This inequality imposes restrictions on the values of the three similarity measures ORDER INVERSIONS, DISPLACEMENT and CAYLEY DISTANCE. If we know the values of two of the measures, we can make a relatively good estimate of the value of the third one.

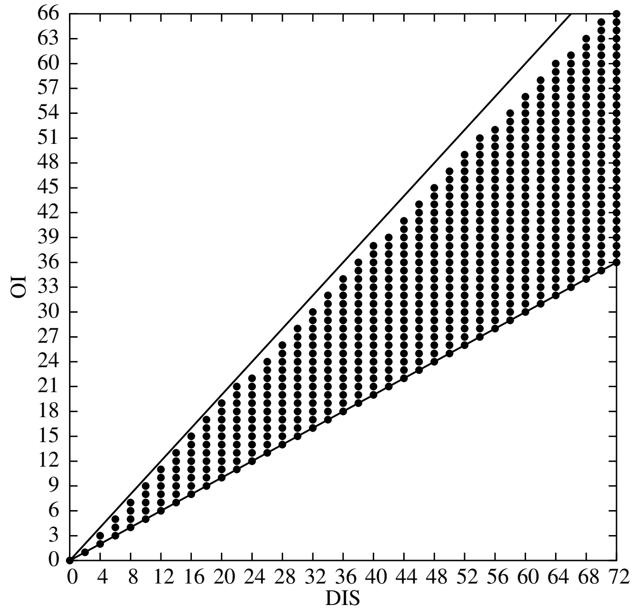


Figure 14.6: The correlation of the ORDER INVERSIONS and DISPLACEMENT values.

Figure 14.6 shows a comparison of the ORDER INVERSIONS and DISPLACEMENT values. The horizontal axis denotes the DISPLACEMENT and the vertical axis the ORDER INVERSIONS values. We can see inequality 14.1 at work here: all the values lie between two diagonal lines. First, the lower line denotes the equation $DIS(X,Y) = 2 \cdot OI(X,Y)$: all pairs of values being on or above that line; hence, $2 \cdot OI(X,Y) \geq DIS(X,Y)$ for all rows X and Y. Secondly, the upper line denotes the equation $OI(X,Y) = DIS(X,Y)$: all pairs of values lying on or below that line; hence, $OI(X,Y) \leq DIS(X,Y)$ for all rows X and Y. Thirdly, the narrow blank strip between the upper line and the dots is defined by CAYLEY DISTANCE. We can derive the following inequality from inequality 14.1:

$$C(X,Y) \leq DIS(X,Y) - OI(X,Y).$$

Therefore the difference between the DISPLACEMENT and ORDER INVERSIONS values for any two rows must be at least the CAYLEY DISTANCE value for those rows.

14.2.5 DERANGEMENT versus ULAM'S DISTANCE and DIVISIONS

As discussed in Sections 10.4 and 10.6, ULAM'S DISTANCE and DIVISIONS are based on the concept of a row as a set of subsegments, while DERANGEMENT is based on the concept of a row as a vector. By comparing these we obtain some fundamental information on the relation of the two approaches.

Figure 14.7 depicts the correlations of DERANGEMENT and ULAM'S DISTANCE values and of DERANGEMENT and DIVISIONS values in graphical format. There seems to be little correlation between them and the difference seems to stem precisely from the different conceptions. We can thus derive the inequality

$$U(X,Y) \leq D(X,Y)$$

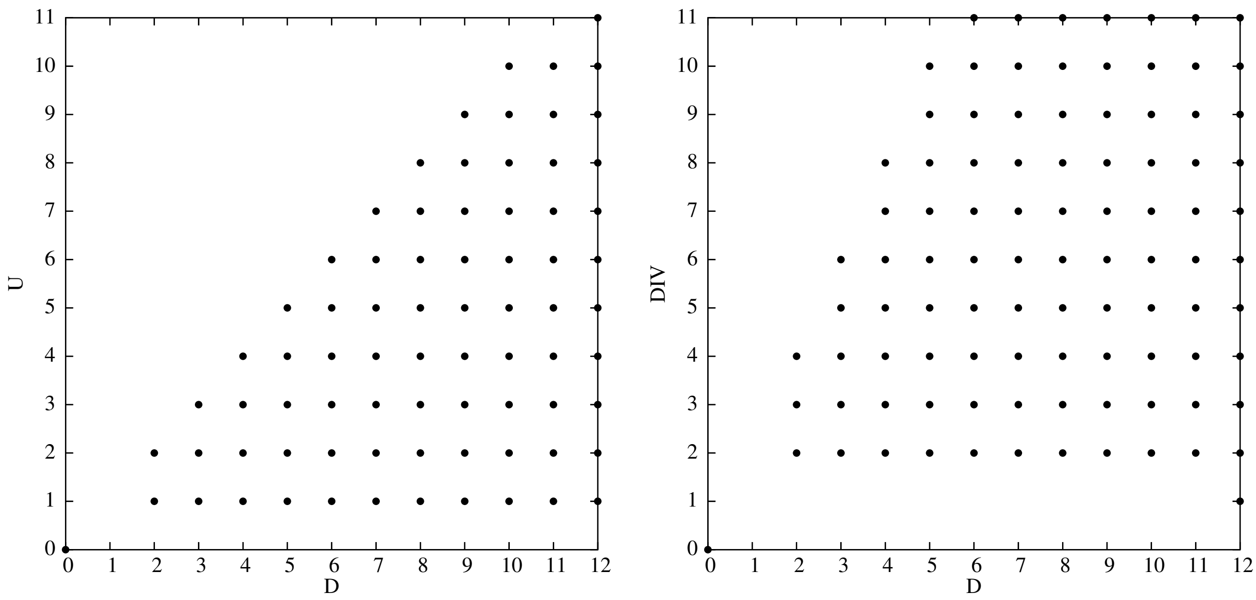


Figure 14.7: The correlation of DERANGEMENT and ULAM'S DISTANCE values on the left, and of DERANGEMENT and DIVISIONS values on the right.

for the DERANGEMENT and ULAM'S DISTANCE values. This inequality is, however, almost trivial since DERANGEMENT denotes the number of pitch classes that are displaced and ULAM'S DISTANCE denotes the number of pitch classes that do not belong to the longest shared subsegment. Therefore, the pitch classes that are *not* displaced can be used to form a shared subsegment that gives an upper limit for the ULAM'S DISTANCE value. Hence, a low DERANGEMENT value implies a low ULAM'S DISTANCE value. In contrast, a low ULAM'S DISTANCE value does not imply a low DERANGEMENT value. For example, for any row X, $U(X, r_1X) = 1$, whereas $D(X, r_1X) = 12$.

We can derive the inequality

$$DIV(X,Y) \leq 2 \cdot D(X,Y)$$

for the DERANGEMENT and DIVISIONS values. This inequality stems from the fact that by displacing k pitch classes we can introduce $2k$ divisions. For example, by exchanging the two pitch classes x_2 and x_4 in the segment $x_1x_2x_3x_4x_5$ we obtain segment $x_1x_4x_3x_2x_5$, which does not share a single nontrivial contiguous subsegment with the original segment. Furthermore, a low DIVISIONS value does not imply a low DERANGEMENT value since by adding a single division we can displace all twelve pitch classes. For example, for any row X, $DIV(X, r_6X) = 1$, whereas $D(X, r_6X) = 12$.

In sum, there is little correlation between DERANGEMENT and ULAM'S DISTANCE or between DERANGEMENT and DIVISIONS. In addition, the differences in the measures stem precisely from the different conceptions behind them. A comparison of the DISPLACEMENT values with those of ULAM'S DISTANCE and DIVISIONS would yield similar results.

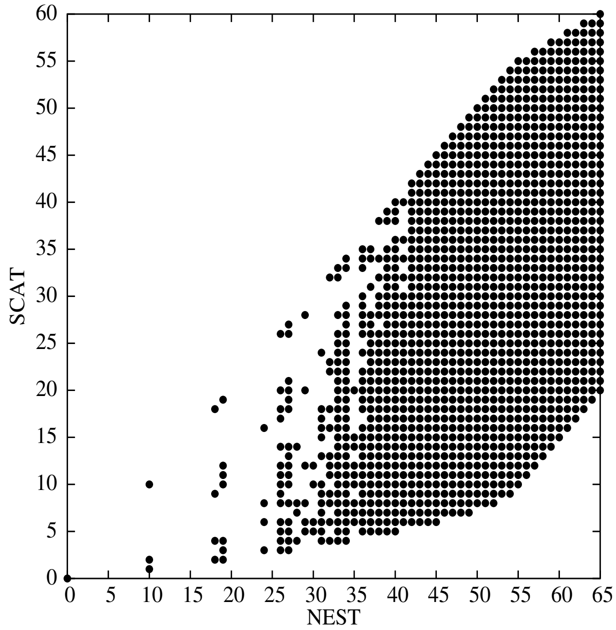


Figure 14.8: The correlation of NESTINGS and SCATTERING similarity measures.

14.2.6 NESTINGS versus SCATTERING

The NESTINGS and SCATTERING similarity measures are both based on the conception of a row as a set of subsets. This relatedness is reflected in the fact that both measures deem rows related by retrograde maximally similar.

Figure 14.8 depicts the correlation of the NESTINGS and SCATTERING values. Let us derive the inequality

$$SCAT(X, Y) \leq NEST(X, Y)$$

for these values. The basis for this inequality becomes clear if we consider what happens when we move one pitch class. Let us begin with row $X = x_0 x_1 \ldots x_{11}$. If we move the first pitch class x_0 between pitch classes x_1 and x_2 the resulting row will no longer contain the ten (contiguous) subsets $\{x_1, x_2\}$, $\{x_1, x_2, x_3\}$, ..., and $\{x_1, x_2, \ldots, x_{11}\}$ of the original row. Yet the SCATTERING value for these two rows is only 1. If instead we move the first pitch class to the last pitch class the resulting row will no longer contain the ten subsets $\{x_0, x_1\}$, $\{x_0, x_1, x_2\}$, $\ldots$, and $\{x_0, x_1, \ldots, x_{10}\}$ of the original row. In this case the SCATTERING value for these two rows is 10. The maximum number of subsets "destroyed" by moving one pitch class is 10, which is also the maximum SCATTERING value that can be introduced by moving one pitch class. Hence, this gives some idea of why the NESTINGS value is an upper bound for the SCATTERING value.

On the other hand, in order to see why the SCATTERING value for two rows may be relatively low even if the corresponding NESTINGS value is high, let us consider the two order-number operations **60718293A4B5** and **13579B02468A**. The NESTINGS value for rows related by either one of these operations is the maximum value 65, while in the former most of the adjacent pitch classes are not moved far. For example, only the pitch class at order position **0** is inserted between the

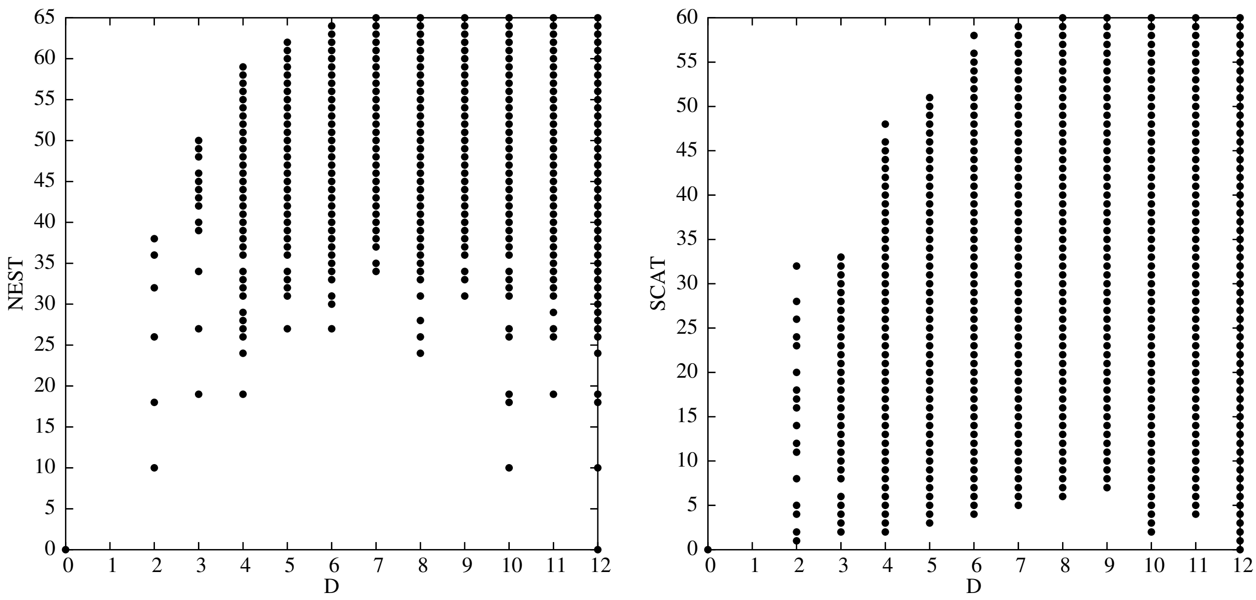


Figure 14.9: The correlation of the DERANGEMENT and NESTINGS values on the left, and of DERANGEMENT and SCATTERING values on the right.

pitch classes at order positions $\mathbf{6}$ and $\mathbf{7}$, and only the pitch class at order position $\mathbf{7}$ is inserted between the pitch classes at order positions $\mathbf{0}$ and $\mathbf{1}$, etc. In contrast, in order-number operation $\mathbf{13579B02468A}$ all adjacent pitch classes are moved five or six order positions apart, which results in a large SCATTERING value.

The correlation of the NESTINGS and SCATTERING values is not unexpected. Namely, a large SCATTERING value is obtained when an order-number transformation moves pitch classes that were originally adjacent far from each other. Therefore, since the subset contents of a row is defined precisely by pitch classes that are adjacent, breaking the adjacencies breaks the subset contents as well. Hence, the correlation of these two measures stems from the fact that they are both based on the same conception of a row.

14.2.7 DERANGEMENT versus NESTINGS and SCATTERING

Let us now compare the DERANGEMENT similarity measure to the two measures based on the subset contents of rows: SCATTERING and NESTINGS. Figure 14.9 depicts the correlations of DERANGEMENT and NESTINGS values and of DERANGEMENT and SCATTERING values in graphical format.

The only correlation between the values of DERANGEMENT and the NESTINGS and SCATTERING values seems to be that a low value in the former implies a low value in the latter two. A low NESTINGS or SCATTERING value does not imply a low DERANGEMENT value, however.

We need to bear in mind the fact that both NESTINGS and SCATTERING are invariant under retrograde, in other words any two rows related by retrograde have identical subset contents. Since such rows are maximally dissimilar according to DERANGEMENT, it is no surprise that there is little correlation between the DERANGEMENT values and the NESTINGS and SCATTERING values.

This lack of correlation can be illustrated by the following two examples. First, the DERANGEMENT value for rows related by order-number operation **1520496B83A7** is 7, but the NESTINGS value is the maximum 65. Secondly, the DERANGEMENT value for rows related by order-number operation **0127456B89A3** is 3, but the NESTINGS value is 50. Hence, only three pitch classes are displaced but the rows share only 15 nontrivial subsets.

14.2.8 A summary of the relations between the similarity measures

We can draw three conclusions based on the above discussion. First, the pitch-class dimension and the order-number dimension are truly two different dimensions. As shown in Section 14.2.1, similarity in one dimension does not necessarily imply similarity in the other. There are a few exceptions to this principle, however. These are the similarity measures in which the two dimensions coincide: DERANGEMENT and CAYLEY DISTANCE. (These are also the measures that are transformationally coherent under the exchange operation.)

Secondly, the division of similarity measures into categories based on the conceptions they imply seems justified. There was higher correlation between measures based on the same conception of rows than between those based on a different conception. NESTINGS and SCATTERING provide a prime example since they deem rows related by retrograde to be maximally similar whereas most of the other similarity measures deem them maximally dissimilar.

Thirdly, some of the conceptions are closer than others. The close relation between the vector approach and the approach based on ordered pairs was discussed in Section 14.2.4. Furthermore, in both approaches the transformational procedures based on exchanges is the common denominator.

Finally, recall that even if the pitch-class dimension seems related to the approach based on intervals, as noted in Section 12.2, the similarity of the INTs of rows and that of rows in the pitch-class space do not necessarily coincide.

The above considerations show that the working hypothesis was correct: there are different aspects of similarity of twelve-tone rows. Many of the different aspects emerge from the fact that twelve-tone rows are ordered entities. Indeed, four of the conceptions – vector, ordered pairs, subsegments and subsets – could be applied to permutations of twelve apples and only in the fifth conception – row as a sequence of ordered pitch-class intervals – it is essential that the entities permuted are pitch classes. Hence, as opposed to the similarity of set classes, precisely because of the ordering the aspects of similarity of twelve-tone rows are not "deeply and inextricably interrelated" (compare Section 3.6.3).

14.3 Some examples of analytical uses

This section gives some examples of analytical applications of the similarity measures discussed, drawn from Alban Berg's *Lyric Suite*. These examples are not intended as a complete analysis of the piece, however, they rather illustrate the ways in which similarity measures could be used as one analytical tool among others. They also show the importance of taking the distribution of values into account.

I will focus on two issues. In Section 14.3.1 I will discuss how similarity measures can be used

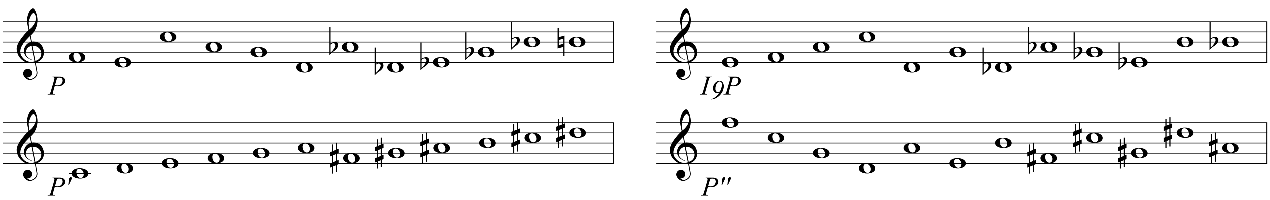


Figure 14.10: Four row forms that are used in the first movement of the *Lyric Suite*: prime form P, inversion I_9P, and two non-standard transformations of row P, labeled P' and P''.

to decipher row forms in twelve-tone compositions, and in Section 14.3.2 I will consider the analysis of transformational processes.

14.3.1 Deciphering row forms

Sometimes deciphering the row forms in a composition may be tricky: composers do not always confine themselves to the strict rules of twelve-tone composition, and reorder the pitch classes in rows to suit their musical demands. In order to explain how the composer has altered the ordering of the pitch classes we must first discover the row form that has been altered. Indeed, in some cases it is far from unequivocal which are the rows that are used in a piece. While labeling the row forms is not important per se, the question of which row has been altered and why cannot be answered unless we can identify them.

The following example, however, is a case in which the similarity measures *cannot* decipher the row form. This sheds light on the peculiarity of the row forms and alterations in question.

The first movement of Alban Berg's *Lyric Suite* introduces some novel uses of twelve-tone rows. One of these is that the pitch classes of the rows are not only presented in their normal order, but are also reordered. Figure 14.10 depicts four row forms that are used in this first movement. The two upper rows P and I_9P are related by inversion. The hexachords of row P are reordered into a scalar form and a cycle of fifths in the two bottom rows P' and P''.

Both rows P' and P'' constitute a thorough reorganization of the hexachords of row P: in row P' the hexachords form two ascending diatonic hexachords and row P'' is a cycle of fifths. Indeed, the reorganizations are so thorough that we cannot even decipher unequivocally what their sources are: namely, both rows $P = 5409728136\text{AB}$ and $I_9P = 4590271863\text{BA}$ contain the very same hexachords and the very same dyads. Even our copious repository of similarity measures does not help since, curiously enough, both P' and P'' are precisely as far from P as they are from I_9P using any of the similarity measures DERANGEMENT, DISPLACEMENT, PITCH-CLASS DISPLACEMENT, ORDER INVERSIONS, BADNESS OF SERIAL FIT, CAYLEY DISTANCE, SUBSET CONTENT DIFFERENCE, ULAM'S DISTANCE, ORDER-INTERVAL INVARIANT N-TUPLES, DIVISIONS, FRAGMENTATION, NESTINGS, SCATTERING, and INTERVALLIC DISTANCE. Only the CORRELATION COEFFICIENT indicates a small difference in the similarity of these rows, but since all other measures do not indicate that difference (and, perhaps, also due to the peculiarity of CORRELATION COEFFICIENT), it would not be wise to make a judgement one way or the other based on similarity.

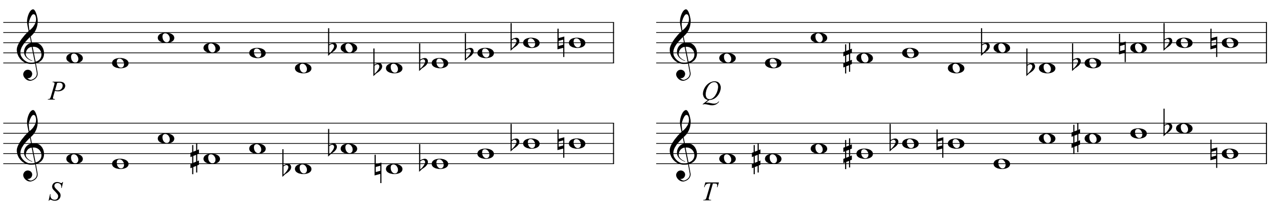


Figure 14.11: Four rows from Alban Berg's *Lyric Suite*, labeled P, Q, S, and T.

Rows P' and P'' illustrate the utmost peculiarity of Berg's row. A row class with two distinct row forms that are equally distant from both a scalar ordering P' and a cycle-of-fifths ordering P'' is rather exceptional (even if not unique: row $0123456789\mathrm{AB}$, for example, has similar properties).

Joel Lester emphasized the familiarity of the diatonic hexachord (Lester 1989, 147). The original row P as well as the modified rows P' and P'' employ two diatonic hexachords, and this familiarity allows us to recognize the connection between rows P' and P'' and the original row P: if the underlying hexachord was a less familiar one, the recognition of this connection would be more difficult.

In sum, the association between rows P' and P'' and row P relies on establishing hexachord areas. Thus the first row of the *Lyric Suite* allows for a compositional practice, which is reminiscent of Josef Hauer's tropes, that affords plenty of local variance while maintaining overall coherence through the limited and systematic use of hexachord areas.

14.3.2 Deciphering transformational processes

Berg uses rows belonging to four different row classes in the *Lyric Suite*; these rows are depicted in Figure 14.11, repeated from Figure 7.3. We can trace the sequence of row transformations during the piece. Row P is the starting point, and rows Q, S, and T stray further and further away from it. First, we obtain row Q by exchanging pitch classes 9 and 6 in row P, then we obtain row S by exchanging pitch classes 9 and 7 and pitch classes 1 and 2 in row Q, and finally we obtain row T by partitioning and concatenating the pitch classes of row S at order positions $\mathbf{0346AB}$ and $\mathbf{125789}$ (Headlam 1996, 248).[2]

The concern in this section is whether the similarity measures provide some insight into Berg's row derivations and use of these rows. Figure 14.12 shows the ORDER INVERSIONS, DISPLACEMENT, and PITCH-CLASS DISPLACEMENT values for rows P, Q, S, and T. There are several observations to be made about the rows based on these values. First, the values confirm my initial observation that the further we go in the sequence of transformations, the more dissimilar to the starting point P the rows become: in all three similarity measures the values increase as we go further in the sequence of transformations. Nevertheless, the first three rows are all mutually relatively similar, the last one being the odd one out. This is not surprising since rows Q and S are obtained

[2]Berg illustrates these successive row derivations in his notes and sketches: see, for example, his notes to Schoenberg, reproduced in number 2 of The International Alban Berg Society Newsletter, or his *Neun Blätter zur "Lyrischen Suite für Strichquartett"*, notes to Kolish Quartert, reproduced in Rauchhaupt (1971). Berg also found the retrograde symmetry of row P a "disadvantage," and so his alterations destroy this property.

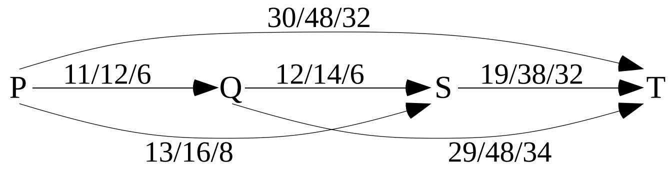


Figure 14.12: Distances of rows P, Q, S, and T. The numbers denote similarity measures ORDER INVERSIONS (scale 0 to 66), DISPLACEMENT (scale 0 to 72) and PITCH-CLASS DISPLACEMENT (scale 0 to 72).

by exchanging one or two pairs of pitch classes, but row T is obtained through a considerably more complex transformation.

Secondly, the first three rows are significantly more dissimilar measured in the order-number space than in the pitch-class space.[3] This disparity stems from the fact that the exchanged pitch classes are relatively close in the pitch-class space in the transformations that derive the rows (such as pitch classes 6 and 9 in transforming row P into row Q and pitch classes 7 and 9 in transforming row Q into row S), but they are not that close in the order-number space (such as order numbers **3** and **9** in transforming row P into Q and order numbers **4** and **9** in transforming row Q into row S).

Thirdly, the ORDER INVERSIONS and DISPLACEMENT values are approximately the same between rows P and Q and between rows Q and S, but not between rows S and T. This is because row T is composed of two segments of row S – hence there are no order inversions within the two segments even if the pitch classes are moved: all order inversions occur between pitch classes of the two segments. This observation is backed up if we look at the limits on the relations between the ORDER INVERSIONS and DISPLACEMENT values. Recall inequality 14.1

$$OI(X,Y) + C(X,Y) \leq DIS(X,Y) \leq 2 \cdot OI(X,Y)$$

from Section 14.2.4. Between rows S and T, $OI(S,T) = 19$ and $DIS(S,T) = 38 = 2 \cdot 19$. Hence, rows S and T are a pair of rows with a maximal displacement of pitch classes given the number of order inversions.

The fact that row Q is obtained by exchanging two pitch classes in row P suggests that we can use CAYLEY DISTANCE to analyze the magnitudes of the transformations in the process of transforming row P into T via rows Q and S (CAYLEY DISTANCE denotes the minimum number of exchanges in a decomposition of a permutation into exchanges, see Section 9.5). In terms of transformations, we can depict the process as follows:

$$P \xrightarrow{QP^{-1}} Q \xrightarrow{SQ^{-1}} S \xrightarrow{TS^{-1}} T.$$

[3]They are so also if we compare the relative values. By comparing row pairs P and Q, Q and S, and P and S, and using relative values scaled to the interval $[0, 1]$, we obtain the following values: 0.00024944, 0.00048724, and 0.00090164 for ORDER INVERSIONS; 0.00001593, 0.00005189, and 0.00015041 for DISPLACEMENT; and 0.00000022, 0.00000022, and 0.00000128 for PITCH-CLASS DISPLACEMENT. Hence, even if PITCH-CLASS DISPLACEMENT tends to give smaller absolute values than its counterpart DISPLACEMENT, the relative values confirm my initial observation about the rows being more similar in the pitch-class space than in the order-number space.

	OI	BSF	DIS	$PCDIS$	C	U
absolute value	19	241	38	32	9	5
maximum value	66	479001600	72	72	11	12
scaled value	0.287879	0.0000005	0.52778	0.44444	0.81818	0.45455
relative value	0.022368	0.002583	0.12949	0.21920	0.34994	0.00503

Table 14.1: A comparison of the values of six similarity measures on rows S and T, and three ways of interpreting them. The absolute value is the value that the similarity measure returns. The maximum value of each measure is given for reference. The scaled value is the absolute value divided by the maximum value of the similarity measure; here 0 denotes maximum similarity and 1 maximum dissimilarity. The relative value takes into account the distribution of the values: it is the number of rows that are closer to S than T is, divided by the number of all rows.

The next task is then to analyze transformations QP^{-1}, SQ^{-1}, and TS^{-1} in terms of exchanges of pitch classes.[4] These three transformations can be written as products of exchanges as follows:[5]

$$\begin{aligned} QP^{-1} &= (6\ \ 9) \\ SQ^{-1} &= (1\ \ 2)(7\ \ 9) \\ TS^{-1} &= (0\ \ 2)(0\ \ 7)(0\ \ 11)(0\ \ 1)(0\ \ 3)(0\ \ 10)(0\ \ 9)(4\ \ 8)(4\ \ 6). \end{aligned}$$

Transformation TS^{-1} is significantly more complex than transformations QP^{-1} and SQ^{-1}: while the latter two can plausibly be described in terms of exchanges, the same cannot be said about TS^{-1} due to the sheer number of exchanges involved. Hence, we could actually "reverse engineer" something about the derivation of the rows: it is plausible to explain the derivation of the two first transformations in terms of exchanges, but for the last one we need another explanation. We might find this if we compare the values of various similarity measures.

Rows S and T provide an excellent example of a case in which the similarity measures do not agree. Table 14.1 enumerates the values, scaled values, and relative values of six similarity measures for rows S and T: ORDER INVERSIONS, BADNESS OF SERIAL FIT, DISPLACEMENT, PITCH-CLASS DISPLACEMENT, CAYLEY DISTANCE, and ULAM'S DISTANCE. Incidentally, the values in the table also demonstrate the importance of taking into account the distribution of the values of the similarity measures. For example, in the case of CAYLEY DISTANCE both the value and the scaled value suggest dissimilarity, but the relative value rather suggests similarity, and in the case of ULAM'S DISTANCE both the value and the scaled value suggest neither similarity nor dissimilarity, but the relative value suggests rather strong similarity. Thus, ignoring the distribution would give us a distorted picture of the similarity of the rows.

Three similarity measures – ORDER INVERSIONS, BADNESS OF SERIAL FIT, and ULAM'S DISTANCE – give significantly lower (relative) values to rows S and T than the others. From the very low relative value of ULAM'S DISTANCE we can deduce that rows S and T contain a relatively

[4] These transformations could be discussed in terms of exchanges of pitch classes and in terms of exchanges of order numbers. I chose the former simply because we are more used to discussing rows in terms of pitch classes.

[5] Recall from Section 9.5 that while the decomposition of a permutation into a product of exchanges is not unique, the minimum number of exchanges in such a decomposition is unequivocal.

long shared subsegment. By definition, there are no order inversions between the pitch classes of the longest shared subsegment. In addition, the relatively low values of ORDER INVERSIONS and BADNESS OF SERIAL FIT indicate that there are not many order inversions between the pitch classes that are not in the longest shared subsegment, or between those that are in the longest shared subsegment and those that are not.

Naturally, we know from the writings and sketches of Berg how he *actually* derived the rows. Nevertheless, the above discussion serves simply to demonstrate how we could use the similarity measures to find out information about the derivation of the rows if we did not have prior knowledge.

14.4 Future research

The present work gives an introduction to the similarity measures for twelve-tone rows. While I believe that this reflects the status of current research, a number of issues would merit further consideration, and could be topics for future research.

Firstly, I have made some observations on the correlation between various similarity measures. The focus was mostly on finding some boundary conditions for the values of one similarity measure when we know the values of another. It would be useful to devise some statistical methods for furthering understanding of the correlations between the values. For example, even if we are unable to create an inequality expressing a boundary condition between the values of two similarity measures, it may well be that, in general, the values correlate, but there are some isolated "discrepancies" that break the inequality.

Secondly, it was noted that some row classes, such as symmetric row classes, tend to have fewer row classes at small distances than others. This case is anything but closed, however. There is still a need to discover how the structure of rows, for example their interval contents, affects the relation between one row class and the others.

Thirdly, different row classes feature different sets of distances to the other rows in that class. In some row classes all rows might be neither particularly similar nor particularly dissimilar, whereas in others they might group into sets of mutually similar rows. Row classes of the latter kind could be compositionally particularly suggestive. I developed this idea only in terms of the degree of symmetry of rows, but I did not cover all the distance relations within a row class. The question remains whether there is some structural property of rows (for example interval contents) that would allow us to predict how close the various rows in the row class are. In addition, we might enquire whether the special properties of a row (multiple order-function rows, self-deriving rows, supersaturated rows, etc.) affect its distances to the other rows in its class.

Fourthly, the rows used by the twelve-tone composers provide an interesting corpus that could be analyzed using the vocabulary and methodology developed in this work. For example, we know that the three members of the second Viennese school preferred different interval contents of rows, and that Webern favored more symmetric rows than Schoenberg and Berg. Are there any other aspects of rows that distinguish the composers? Should we decipher the conceptions they used based on the rows, or on the way in which they were used? Are the rows of one composer closer to each other than to the rows of the other composers? Are the rows used by members of the second Viennese school different from those used by later composers?

Fifthly, the different conceptions of rows suggest different compositional practices and vice versa. The dynamics of the triangle formed by the properties and conceptions of rows, and by compositional practices, sets us a challenge that has not been fully addressed. The philosophical aspects of the transformational approach and how it relates to these conceptions is also uncharted territory.

Finally, I have only scratched the surface of the potential analytical applications of similarity measures. The existing literature on the similarity of twelve-tone rows is biased more towards the theoretical aspects than the analytical application. My hope that this work will serve as a tool to help music analysts to balance that bias.

Appendices

APPENDIX A

Permutations

This appendix gives some basic definitions and results of permutations. Section A.1 examines two definitions, then in Section A.2 we will walk through how they are composed, and their decomposition into cycles is discussed in Section A.3.

A.1 Two definitions of permutations

A permutation is the formalization of the idea that a set of objects can be arranged in different orders. A well-known result of combinatorics states that the number of different permutations on a set of n elements is $n!$ (read "n factorial"). As the target here is to apply the theory of permutations to twelve-tone rows, in the following I will use permutations of the twelve integers modulo 12 as my example. Applying the formula gives

$$12! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6 \cdot 7 \cdot 8 \cdot 9 \cdot 10 \cdot 11 \cdot 12 = 479001600$$

distinct twelve-tone rows.

The mathematical literature provides two definitions for a permutation: *linear order* and *function*.[1] I will employ both definitions in the following.

Let us first consider the notion of a permutation as a linear order. For easy reference, the formal definition is given below.

DEFINITION A.1 A permutation on set S is a linear ordering of the elements of S.

A linear ordering means that the elements are enumerated in some order in such a way that each element appears exactly once.

[1] See, for example, Bóna (2004, 1, 73).

The notion of a permutation as a linear ordering is applicable to pitch-class rows and order-number rows. Let us use one-line notation for the linear orderings:

(A.1)
$$5\ 4\ 0\ 9\ 7\ 2\ 8\ 1\ 3\ 6\ 10\ 11,$$

or in shorthand $5409728136\mathrm{AB}$. The twelve elements are put into order and are read from left to right.

The notion of a permutation as a linear ordering describes a static finished product. The second notion, a permutation as a function, is more dynamic.

DEFINITION A.2 A permutation on set S is a bijective mapping $S \to S$.

A bijective (one-to-one and onto) mapping $S \to S$ maps every element of S into some element of S. In addition, no two distinct elements are mapped into the same element, and for each one, some element is mapped into it.

The notion of a permutation as a mapping gives rise to the following two-line notation. Let us enumerate between the parentheses how each element of a set is mapped. For example, in the permutation

(A.2)
$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 5 & 4 & 0 & 9 & 7 & 2 & 8 & 1 & 3 & 6 & 10 & 11 \end{pmatrix},$$

element 0 is mapped into element 5, element 1 is mapped into element 4, element 2 is mapped into element 0, etc. It is customary (although not necessary) to enumerate the upper row in some readily accessible order. Note that no two distinct elements are mapped into the same element (no element appears twice in the lower row), and for each element, some element is mapped into it (each element appears in the lower row).

The two above definitions are associated. We could consider the permutations of formulas A.1 and A.2 interchangeable depending on the context since the lower row in the latter is the same as the single row in the former. In the present context it is imperative that the entries in the upper row in formula A.1 are enumerated from 0 to 11. Conceptually, they are different types of entities, however. A linear ordering is a linear ordering and a mapping is a mapping. This is analogous to the conceptual difference between objects and operations, and twelve-tone rows and row operations (see Section 1.2.2). Twelve-tone rows are linear orderings and row operations are mappings. Nevertheless, the action of the group S_{12} (permutations in the sense of mappings) on the set of all linear orderings (permutations in the sense of linear orderings) given in Section 2.2.3 was based on this association.

A.2 The binary operation of permutations

If we consider permutations as mappings, we could justifiably introduce the notion of composing permutations: a product of two permutations of n elements results in a new permutation of n elements. Let us use left orthography – hence the permutations are composed from right to left:

$$\overbrace{\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 0 & 1 \end{pmatrix}}^{T} \cdot \overbrace{\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 5 & 4 & 0 & 9 & 7 & 2 & 8 & 1 & 3 & 6 & 10 & 11 \end{pmatrix}}^{P}$$
$$= \overbrace{\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 7 & 6 & 2 & 11 & 9 & 4 & 10 & 3 & 5 & 8 & 0 & 1 \end{pmatrix}}^{TP}.$$

The above example depicts the product of two permutations, tentatively labeled P and T, and their product TP. In calculating the product let us begin with the rightmost permutation and trace the mapping of each element. For example, in permutation P element 0 is mapped into 5 and in permutation T element 5 is mapped into 7; hence, in the product TP element 0 is mapped into 7. The following diagram illustrates this process:

$$0 \xrightarrow{P} 5 \xrightarrow{T} 7.$$

Similarly, in permutation P element 1 is mapped into 4 and in permutation T element 4 is mapped into 6; hence, in the product TP element 1 is mapped into 6. Tracing each element gives the resulting permutation TP.

Label P suggests a twelve-tone row, label T suggests a transposition (in the musical sense of the word), and label TP suggests a transposed twelve-tone row. I should stress once more that, conceptually, P and T represent different domains: P represents a permutation in the sense of a linear ordering whereas T represents a permutation in the sense of a mapping. Nevertheless, we can define the multiplication of a linear order by a mapping just as if the linear order were also a mapping.

To conclude this section, let us consider the notion of inverse permutation. The permutations define a group: one of the requirement of a group structure is that every element has an inverse element such that the product of the element and its inverse result in the identity element. Consider the following product of permutations:

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 2 & 7 & 5 & 8 & 1 & 0 & 9 & 4 & 6 & 3 & 10 & 11 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 5 & 4 & 0 & 9 & 7 & 2 & 8 & 1 & 3 & 6 & 10 & 11 \end{pmatrix}$$
$$= \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \end{pmatrix}.$$

This example depicts again the product of two permutations. In calculating the product we begin with the rightmost permutation and trace the mapping of each element. For example, in the leftmost permutation element 0 is mapped into 5 and in the rightmost permutation element 5 is mapped back into 0; hence, in the product element 0 is mapped into 0. Correspondingly, if we take any element and trace its mapping, we notice it is mapped into itself in the product. Consequently, all elements are mapped into themselves. The result of the product is the identity permutation that is the identity element of the *symmetric group* S_{12}, that is, the group of permutations of twelve elements.

Finally, note that the notion of the inverse of a mapping is well defined, whereas that of the inverse of a linear ordering is not. Similarly, row operations have inverse operations but twelve-tone rows do not. Nevertheless, when I define the GIS of twelve-tone rows in Section 5.4 in terms of permutations I utilize their dual nature: in order to discover the row operation that transforms a row into another row I use permutations denoting twelve-tone rows as if they were mappings. This is done only in order to simplify the formalism, and it does not compromise the conceptual difference between linear orderings and mappings.

A.3 Cycles

The *cycle structures* of the 48 twelve-tone operations or TTOs[2] have been widely studied (Rothgeb 1967; Morris 1977; Starr and Morris 1977; Starr 1978; Lord 1978; Morris 1987; Morris 1991; Buchler 1997; Morris 2001). I show in Section 2.2.5 that pitch-class operations and order-number operations are permutations acting on pitch-class rows and order-number rows, respectively. Here I explore the cycle structure of permutations, paying special attention to their decomposition into products of *cycles*. Since cycles are an essential concept in this work, I provide a brief introduction to the pertinent concepts and notation in the following.

Cycles are a notation for permutations (in the sense of a mapping). Let us consider the permutation

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 & 0 \end{pmatrix} \tag{A.3}$$

and how each element is mapped in it. Element 0 is mapped into 1, element 1 is mapped into 2, element 2 is mapped into 3, element 3 is mapped into 4, element 4 is mapped into 5, and element 5 is mapped into 0, thus closing the circle. We can express the same information by writing the cycle

$$(0\ 1\ 2\ 3\ 4\ 5)\,. \tag{A.4}$$

The parentheses indicate the denotation of a permutation in the sense of a mapping: they are not used in the one-line notation for a permutation in the sense of a linear ordering.

Figure A.1 illustrates how this notation should be read: each element is mapped into the element that is on its right side except the last one, which is mapped into the first.

Formulas A.3 and A.4 denote the very same permutation. For example, in both cases 0 is mapped into 1. This mapping was derived in formula A.3 because 0 of the upper row is mapped into 1 of the lower row, and in formula A.4 because 1 is the next to the right of 0.

It should be noted that, due to its cyclic nature, it is immaterial which element is placed first in

[2]See, for example, Chapter 3 in Morris (1987) for a discussion of the TTOs.

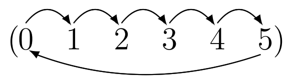


Figure A.1: An interpretation of the cycle $(0\ \ 1\ \ 2\ \ 3\ \ 4\ \ 5)$.

a cycle. Hence, cycles $(0\ \ 1\ \ 2\ \ 3\ \ 4\ \ 5)$ and $(3\ \ 4\ \ 5\ \ 0\ \ 1\ \ 2)$ denote the same permutation: in both element 0 is mapped into 1, element 1 is mapped into 2, element 2 is mapped into 3, element 3 is mapped into 4, element 4 is mapped into 5, and element 5 is mapped into 0.

A permutation cannot always be written as a single cycle, however. For example, consider the permutation

$$\begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 0 & 4 & 5 & 3 \end{pmatrix}.$$

Here, 0 is mapped into 1, 1 is mapped into 2, and 2 is mapped into 0. In addition, 3 is mapped into 4, 4 is mapped into 5, and 5 is mapped into 3. Elements 0, 1, and 2 form cycle $(0\ \ 1\ \ 2)$, and we cannot describe how the three other elements are mapped by continuing it. Therefore, this permutation cannot be written as a single cycle, but we can write it as a product of two cycles as follows:

$$(0\ \ 1\ \ 2)\,(3\ \ 4\ \ 5) = (3\ \ 4\ \ 5)\,(0\ \ 1\ \ 2)\,.$$

If two cycles are disjoint, in other words they do not contain common elements, their order of the cycles in a product is immaterial. Hence, as cycles $(0\ \ 1\ \ 2)$ and $(3\ \ 4\ \ 5)$ are disjoint, the two products of these cycles above denote the same permutation.[3]

[3] A well-known result in the theory of permutations is that every finite permutation can be represented as a product of disjoint cycles (see, for example, Theorem 5 in Section 1.4 in Nicholson (1999)). This representation is unequivocal except for the order of the cycles.

APPENDIX B

Relations

This appendix introduces some basic definitions of sets and relations, which provide the basis for discussion of a special type of relation known as partial order in Appendix C.

DEFINITION B.1 Let S and T be sets. The Cartesian product $S \times T$ is the set of all possible pairs (s, t) where $s \in S$ and $t \in T$.

A Cartesian product of two sets contains ordered pairs in which the first element is a member of the first set and the second element is a member of the second set. For example, if $S = \{0, 1\}$ and $T = \{2, 3, 4\}$, the Cartesian product $S \times T$ is $\{(0, 2), (0, 3), (0, 4), (1, 2), (1, 3), (1, 4)\}$. Of course, we may have a Cartesian product between a set and itself. For example, if $S = \{0, 1\}$, the Cartesian product $S \times S$ is $\{(0, 0), (0, 1), (1, 0), (1, 1)\}$.

We can now define a binary relation using the notion of the Cartesian product.

DEFINITION B.2 Let S be a set. A (binary) relation on S is a subset of the Cartesian product $S \times S$.

A relation on a set declares that some of its elements are related to some other elements. In this case, set S could be, for example, the set of all pitch-class sets or the set of all twelve-tone rows.

Let us use the notation $(s, t) \in R$ to denote that s is related to t. Note that the order of the elements matters: if $(s, t) \in R$ then s is related to t, but t is not necessarily related to s.

Any subset of $S \times S$ is a relation — it does not have to "make sense" to a human observer. However, we are often interested in relations with special properties. Four such properties are defined in the following: reflexivity, symmetry, transitivity and antisymmetry.

DEFINITION B.3 Let S be a set and R be a binary relation on S. R is reflexive if $(s, s) \in R$ for all $s \in S$. Relation R is symmetric if $(s, t) \in R$ implies $(t, s) \in R$ for all $s, t \in S$. Relation R is transitive if $(s, t) \in R$ and $(t, u) \in R$ imply $(s, u) \in R$ for all $s, t, u \in S$. Relation R is antisymmetric if $(s, t) \in R$ and $(t, s) \in R$ imply $s = t$ for all $s, t \in S$.

The definition of reflexivity is straightforward: it means that every element of a set is related to itself. The definitions of symmetry, transitivity, and antisymmetry are slightly more complicated as they involve implication. In symmetry, always when the pair (s, t) is in the relation then so must be the pair (t, s). In other words, if the element s is related to the element t, then t must also be related to s. In transitivity, always when the pairs (s, t) and (t, u) are in the relation then so must be the pair (t, u). Antisymmetry is not the exact opposite of symmetry – we can construct a relation that is both symmetric and antisymmetric. Antisymmetry means that there are no symmetric pairs of distinct elements in the relation – in other words, for any two distinct elements s and t, if the pair (s, t) is in the relation then the pair (t, s) cannot be in the relation.

Note that symmetry, transitivity and antisymmetry do not require that the relation is nonempty. The formulation of these properties is that *if* there are some pairs in the relation, *then* they must satisfy the pertinent criteria. Therefore, somewhat counterintuitively, an empty relation – a relation in which nothing is related – is always symmetric, antisymmetric and transitive.

Let us next formally define an equivalence relation.

DEFINITION B.4 An equivalence relation is a relation that is reflexive, symmetric, and transitive.

The three conditions are intuitive. The first requirement, reflexivity, states that every element is equivalent to itself. The second condition, symmetry, states that equivalence is bilateral: it is not possible that one element is equivalent to another but not vice versa. The third condition, transitivity, states that equivalence is "chained:" if X is equivalent to Y and Y is equivalent to Z then X must be equivalent to Z.

An equivalence relation partitions a set into equivalence classes. Elements in an equivalence class are considered equivalent. Every element belongs to exactly one equivalence class.

APPENDIX C

Partially ordered sets

This appendix concerns the concept of *partially ordered set* or *partial order*. Partial orders have been discussed in the literature on music theory by Daniel Starr (1984) and Robert Morris (1987, 2001), for example. Both use somewhat proprietary notation, but I will stick here to standard mathematical notation. The introductory chapters of Fishburn (1985) and Trotter (1992) and Chapter 3 in Stanley (1997) provide a comprehensive introduction to partial orders.

Partial orders are relations that are reflexive, transitive and antisymmetric. They are often notated using the symbol $\leq$. Thus, if P is a partial order, notations $(a, b) \in P$ and $a \leq b$ are equivalent. The shorthand $a < b$ denotes that $a \leq b$ and $a \neq b$, and is used when we wish to stress that the pertinent pair is not a reflexive pair of type (a, a).

The formal definition of the partial order is given below.

DEFINITION C.1 Partial order $\leq$ on set S is a binary relation that is reflexive, transitive and antisymmetric. In other words, for every x, y and z in S the three following properties hold: (i) $x \leq x$ (reflexivity), (ii) $x \leq y$ and $y \leq z$ implies $x \leq z$ (transitivity), and (iii) $x \leq y$ and $y \leq x$ implies $x = y$ (antisymmetry).

Let us consider a concrete example. The relation

(C.1) $$P = \{(a, a), (b, b), (c, c), (d, d), (b, c), (c, d), (b, d), (a, d)\}$$

is a partial order on set $S = \{a, b, c, d\}$. First, it contains all the "reflexive pairs" (a, a), (b, b), (c, c) and (d, d). Secondly, it is transitive because the pair (b, d), which contains the "endpoints" of the two pairs (b, c) and (c, d), is in the relation. Thirdly, it is antisymmetric since it contains no two symmetric pairs such as (a, b) and (b, a). It should also be noted that not all elements are necessarily *comparable* (which is why it is called a *partial* order). The order of a and d is defined by the pair (a, d), but the order of a and b, for example, is undefined and they are not comparable.

A *total order* or *linear order* (compare Appendix A) is a special case of a partial order (sic!) in which the order of all pairs of elements is defined. Given two elements of a set, their order is defined by the total order. A *trivial partial order* is the opposite: it contains only the reflexive pairs and the order of no distinct two elements is defined. The formal definitions are given below.

DEFINITION C.2 A partial order on set S is a *total order* if it satisfies the following comparability requirement: for all elements $a, b \in S$ either $a \leq b$ or $b \leq a$. A partial order on set S is a *trivial order* if $a \leq b$ implies $a = b$.

The following lemma proves that the intersection of two partial orders on a given set is also a partial order.

LEMMA C.1 The intersection of two partial orders on a given set S is a partial order.

Proof. Let P_1 and P_2 be two arbitrary partial orders on a set S. We must verify that the intersection $P_1 \cap P_2$ meets the three requirements of Definition C.1. First, the intersection is reflexive, because for all elements $x \in S$, $(x, x) \in P_1$ and $(x, x) \in P_2$ and therefore $(x, x) \in P_1 \cap P_2$. Secondly, if the two pairs (x, y) and (y, z) are in the intersection, then they are also both in P_1 and P_2. But then, because both P_1 and P_2 are transitive relations, $(x, z) \in P_1$ and $(x, z) \in P_2$, which implies $(x, z) \in P_1 \cap P_2$, the intersection is transitive. Thirdly, antisymmetry is satisfied, as no two pairs violating the requirement of antisymmetry exist in P_1 or in P_2, and therefore no such two pairs can exist in the intersection. As the intersection of two partial orders is reflexive, transitive and antisymmetric, it is a partial order. $\square$

Let us again look at an example. The relations $P_1 = \{(a, a), (b, b), (c, c), (a, b), (b, c), (a, c)\}$ and $P_2 = \{(a, a), (b, b), (c, c), (b, a), (a, c), (b, c)\}$ are total orders on the set $\{a, b, c\}$. The intersection $P_1 \cap P_2 = \{(a, a), (b, b), (c, c), (a, c), (b, c)\}$ contains all pairs that P_1 and P_2 have in common and it is – as claimed in the previous lemma – a partial order.

It behooves us to examine elements that are *adjacent* in a partial order. These are called *adjacent pairs* or *cover relations*. The following definition provides an exact formulation of this concept.

DEFINITION C.3 The pair (a, b) in a partial order is called an *adjacent pair*, if $a < b$ and there exists no element x such that $a < x < b$. We could also say that b *covers* a.

For example, in the partial order P_1 above the adjacent pairs or the cover relations are $a < b$ and $b < c$. Pair $a < c$ is not a cover relation since b is "between" a and c. Note that the use of the symbol $<$ in Definition C.3 implies that the elements in a cover relation are distinct elements: a reflexive pair cannot be a cover relation.

We can express a partial order by writing only the cover relations (and implying that we are dealing with a partial order); the remaining elements can be deduced from the cover relations as their *transitive closure*. For example, from the two cover relations $a < b$ and $b < c$ we can deduce that the relation also contains $a < c$. In addition, the reflexive pairs can be omitted (but implied) as they are part of every partial order and enumerating them adds no information.

Long lists of pairs may be difficult to grasp. The concept of the *chain* allows for a more intuitive notation.

DEFINITION C.4 Let P be a partially ordered set. A chain in P is a set of pairwise comparable elements (i.e., a totally ordered subset).

Under the chain notation, the chain that the relation P_1 above forms is conveniently written as $a \leq b \leq c$. Thus, only the cover relations ($a \leq b$ and $b \leq c$) are written out.

The following lemma proves that removing any cover relation from a partial order results in another partial order.

LEMMA C.2 If P is a partial order then $P \setminus \{(a, b)\}$ is a partial order if and only if (a, b) is an adjacent pair in P.

Proof. Let us first prove the implication from right to left. Let $P' = P \setminus \{(a, b)\}$. Clearly P' is a subset of P that is antisymmetric and reflexive. Therefore of the three criteria of partial orders only transitivity needs to be verified. Let us use an indirect strategy and assume that transitivity is violated in P'. Because P is a partial order and is thus transitive, and the only difference between P and P' is the pair (a, b), the violation of transitivity in P' must be that there is some element x such that $a < x$ and $x < b$ are in P', but $a < b$ is not. This means that $a < x < b$ in P, which is against the assumption of (a, b) being an adjacent pair. Therefore the antithesis is disproved and the implication from right to left is proved.

To prove the implication from left to right, let us again use an indirect strategy. Let us assume that P' is a partial order but (a, b) is not an adjacent pair in P. There are two possibilities. Either (a, b) is a reflexive pair, i.e., $a = b$, or there exists an intermediate element x such that $a < x < b$. But if $a = b$ then P' does not contain all reflexive pairs and it cannot be a partial order. Moreover, if there exists an intermediate element x such that $a < x < b$, then P' contains (a, x) and (x, b) and, because as a partial order it is transitive, also (a, b). Both possibilities lead to a contradiction, so the antithesis is disproved and the implication from left to right is proved. $\square$

Figure C.1 shows the previous lemma in action: the removal of any of the three adjacent pairs of the partial order $a \leq b \leq c \leq d$ results in another partial order.

Let us next define the *minimal element* of a set. This definition is general enough to apply to any relation on a set that is a partial order, or at least a subset of a partial order.

DEFINITION C.5 Let P be a relation on set S that is a partial order (on the set S) or a subset of a partial order. Element $x \in S$ is minimal if P contains a pair (x, y) and $y \leq x$ implies $x = y$.

The definition of a minimal element is indirect. The interpretation is that an element x is minimal if, when it would appear that there is a "smaller" element, it always turns out that the smaller element is identical to it. In other words, there is no element that is smaller than x.

We can now prove that every nonempty finite subset of a partial order has at least one minimal element, and that every nontrivial finite partial order always has at least one adjacent pair. Here it is

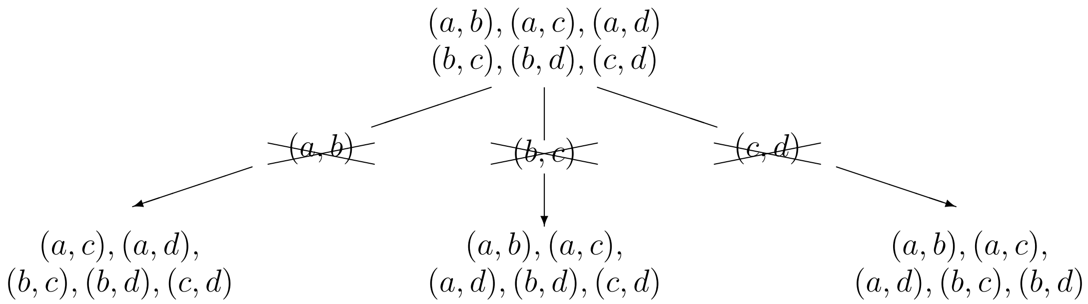


Figure C.1: The topmost partial order is $a \leq b \leq c \leq d$. The three partial orders at the bottom are obtained by removing an adjacent pair from the topmost partial order. To avoid extra clutter the reflexive pairs have been omitted.

essential that we are dealing with finite sets and partial orders: the following two lemmas hold only for finite partial orders and finite sets.

LEMMA C.3 Let P be a nonempty finite relation on a set S that is a partial order (on the set S) or a subset of a partial order. Then P contains a minimal element.

Proof. The subset is nonempty so it contains at least one pair, let us denote such a pair by (a, b). If there exists no element x such that $x < a$, then a is a minimal element and the lemma is proved. Otherwise a finite sequence

$$x_k, x_{k-1}, \ldots, x_1, x_0, a$$

can be constructed by iterating through every element in the set and adding each element to the beginning of the sequence if it is smaller than the currently first element. The subset is finite and the pairs are a subset of a partial order and therefore there cannot be duplicates in the sequence. The iteration thus ends in a finite number of steps. Therefore in the sequence $x_i < x_{i-1}$, and when the sequence has been constructed, there exists no element x such that $x < x_k$. Element x_k is then a minimal element. □

LEMMA C.4 A nontrivial finite partial order always has at least one adjacent pair.

Proof. Let P be a nontrivial finite partial order and let P' be a set that contains all pairs of P except the reflexive pairs. P' is nonempty and therefore Lemma C.3 states that P' contains a minimal element. Let us denote the minimal element by m. There is a finite number of pairs with m as the left element, so we can choose a pair (m, x) that has no intermediate element. Thus (m, x) is an adjacent pair. □

To conclude this appendix, I will prove a well-known property of the subset relation $\subseteq$, namely that the subset relation induces a partial order on a set of sets.

LEMMA C.5 If S is a set of sets, then the subset relation $\subseteq$ induces a partial order on S where for any two sets $A, B \in S$, $A \leq B$ if and only if $A \subseteq B$.

Proof. We need to show that the subset relation $\subseteq$ is reflexive, transitive, and antisymmetric. All these properties follow directly from the properties of the subset relation. First, for any A set, $A \subseteq A$; hence, the relation is reflexive. Secondly, the subset relation is transitive since if $A \subseteq B$ and $B \subseteq C$ then $A \subseteq C$. Finally, the relation is antisymmetric since $A \subseteq B$ and $B \subseteq A$ implies $A = B$. Thus, the subset relation $\subseteq$ induces a partial order. $\square$

To illustrate the partial order that the subset relation defines, let us take a concrete example. Let us first define four sets: $A = \{0\}$, $B = \{1\}$, $C = \{1, 2\}$ and $D = \{0, 1, 2\}$. The space S is then the set containing these four sets: $S = \{A, B, C, D\}$. Since any set is always a subset of itself, we have the following subset relations: $A \subseteq A$, $B \subseteq B$, $C \subseteq C$, and $D \subseteq D$. These are the reflexive pairs in the partial order. In addition, we have the following four subset relations: $A \subseteq D$, $B \subseteq C$, $B \subseteq D$, and $C \subseteq D$. The relation is antisymmetric since it does not contain any symmetric pairs. It is also transitive because the pair $B \subseteq D$, which contains the "endpoints" of the two pairs $B \subseteq C$ and $C \subseteq D$, is in the relation. Hence, the relation is a partial order. In fact, if we change the labels of the sets to lowercase letters, it is the same partial order as defined above in Formula C.1.

In the partial order induced by the subset relation for two sets A and B, $A \leq B$ if and only if $A \subseteq B$. Symbol $\subseteq$ allows the two sets to be equal. If we wish to stress the fact that the two pairs are not identical we use the symbol $\subset$.

Once more let us stress that not all pairs of elements in a partial order are necessarily comparable: in the above example of a partial order induced by the subset relation the sets $A = \{0\}$ and $B = \{1\}$ are not comparable since neither $A \subseteq B$ nor $B \subseteq A$.

APPENDIX D

Graphs

This appendix introduces the basic concepts of (finite) graphs. For a more comprehensive and formal treatment, see, for example, Liu (1968), Diestel (2000) or West (2001). Let us start by first defining the graph.

DEFINITION D.1 A *graph* G consists of a set $V(G)$ of objects called *vertices* and set $E(G)$ of unordered pairs of elements of $V(G)$ called *edges*.

Figure D.1 depicts a graph with four vertices represented by dots and four edges represented by lines connecting them. Two vertices are *adjacent* if they are joined by an edge. For example, vertices A and B are adjacent in the graph in Figure D.1, but vertices A and D are not.

If we can "travel" from one vertex to another by following the edges, we say that there is a path from one vertex to the other. If there is a path between any two vertices we say that the graph is connected. For example, there is a path between any two vertices of the graph in Figure D.1, hence it is connected.

DEFINITION D.2 The distance between two connected vertices is the length of the shortest path between them.

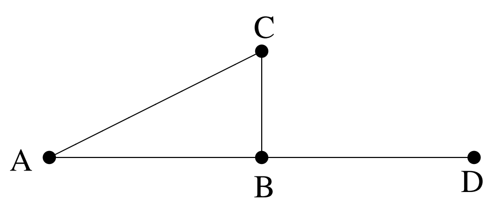


Figure D.1: A graph with four vertices and four edges.

For example, in Figure D.1 there are multiple paths from vertex A to vertex D. The shortest path is A-B-D and since it consists of two edges, the distance between vertices A and D is 2.

DEFINITION D.3 Two graphs X and Y are isomorphic if there is a bijection f from the vertices of X to the vertices of Y such that vertices A and B in graph X are connected by n edges if and only if vertices $f(A)$ and $f(B)$ are connected by n edges.

Isomorphic graphs are structurally identical. Hence, if graphs X and Y are isomorphic, we can turn graph X into Y by simply relabeling the vertices and the edges.

APPENDIX E

Similarity measures for twelve-tone rows

Table E.1 enumerates the 17 similarity measures discussed in this study. The measures are grouped based on the conception of rows they suggest. For each measure the six columns of the table denote the conception on which it is based, its name, its definition (or a short description when the value cannot be expressed as a brief formula), whether it defines a metric, whether it is transformational, and its range.

The definitions give the value of the measures for pitch-class rows X and Y or order-number rows $\boldsymbol{X}$ and $\boldsymbol{Y}$. Transformation $int(X, Y)$ is the transformation in the GIS of pitch-class rows that transforms row X into row Y, and $int(\boldsymbol{X}, \boldsymbol{Y})$ is the transformation in the GIS of order-number rows that transforms row $\boldsymbol{X}$ into row $\boldsymbol{Y}$.

conception	measure	definition	metric	transf.	range
vector approach	DERANGEMENT	$\#\{n \mid x_n \neq y_n\}$	×	×	$[0, 12]$
	DISPLACEMENT	$\sum_{n=0}^{11} \|oint(\boldsymbol{x_n}, \boldsymbol{y_n})\|$	×	×	$[0, 72]$
	PITCH-CLASS DISPLACEMENT	$\sum_{n=0}^{11} \|ic(x_n, y_n)\|)\|$	×	×	$[0, 72]$
ordered pairs	ORDER INVERSIONS	$\#\{(i,j) \| (i,j) \in X$ and $(i,j) \notin Y\}$	×	×	$[0, 66]$
	BADNESS OF SERIAL FIT	number of rows satisfying the protocol that X and Y define		×	$[0, 479001600]$
	CAYLEY DISTANCE	smallest number of exchanges in decompositions of $int(X, Y)$	×	×	$[0, 11]$
	CORRELATION COEFFICIENT	correlation coefficient of $int(\boldsymbol{X}, \boldsymbol{Y})$ and id		×	$[-1, 1]$
sub-segments	GENERALIZED ORDER INVERSIONS	$\#\{(a_1, a_2, \ldots, a_n) \mid (a_1, a_2, \ldots, a_n) \in X \setminus Y\}$	×	×	$[0, \binom{12}{n}]$
	SUBSEGMENT CONTENT DIFFERENCE	$\sum_{n=0}^{12} OI_n(X, Y)$	×	×	$[0, 4083]$
	ULAM'S DISTANCE	12 minus the length of the longest shared subsegment	×	×	$[0, 11]$
	ORDER-INTERVAL INVARIANT N-TUPLES	$\#\{(S_n, OINT(S_n)) \mid (S_n, OINT(S_n)) \in X \setminus Y\}$	×	×	$[0, 4083]$
	DIVISIONS	number of adjacent pairs in $int(\boldsymbol{X}, \boldsymbol{Y})$ that are not consecutive ascending pairs	×	×	$[0, 11]$
	FRAGMENTATION	$66 - \sum \frac{n_i(n_i-1)}{2}$ where n_i are the lengths of the shared maximal disjoint subsegments of X and Y	×	×	$[0, 66]$
subsets	NESTINGS	78 minus the number of nodes in the nesting that X and Y define	×	×	$[0, 65]$
	SCATTERING	$\sum_{n=1}^{11} \|g_n - g_{n-1}\| - 11$ where g_n is the nth element of $int(\boldsymbol{X}, \boldsymbol{Y})$		×	$[0, 60]$
intervals	INTERVALLIC DISTANCE	$\sum_{n=0}^{10} ic(int(x_n, x_{n+1}), int(y_n, y_{n+1}))$	pseudo		$[0, 66]$
	INTERVAL DISPLACEMENT	$\min\{INTDIS(G) \mid G(INT(X)) = INT(Y)\}$ where $INTDIS(G) = \sum_{n=1}^{11} \|g_n - n\|$			$[0, 66]$

Table E.1: The definitions and main properties of the 17 similarity measures for twelve-tone rows.

References

Agmon, E. (2002). The Multiplicative Norm and Its Implications for Set-Class Theory. *Perspectives of New Music 40*(1), 216–234.

Alegant, B. (1993). *The Seventy-Seven Partitions of the Aggregate: Analytical and Theoretical Implications*. Ph.D. dissertation, University of Rochester, Eastman School of Music.

Babbitt, M. (1946/1992). *The Function of Set Structure in the Twelve-Tone System*. Ph.D. dissertation, Princeton University.

Babbitt, M. (1955). Some Aspects of Twelve-Tone Composition. *The Score and I.M.A. Magazine 12*, 53–61.

Babbitt, M. (1960). Twelve-Tone Invariants as Compositional Determinants. *The Musical Quarterly 46*(2), 246–259.

Babbitt, M. (1961a). Past and Present Concepts of the Nature and Limits of Music. In J. Larue (Ed.), *Report of the Eighth Congress*, Volume 1, New York, pp. 398–403. International Musicological Society.

Babbitt, M. (1961b). Set Structure as a Compositional Determinant. *Journal of Music Theory 5*(1), 72–94.

Babbitt, M. (1962). Twelve-Tone Rhythmic Structure and the Electronic Medium. *Perspectives of New Music 1*(1), 49–79.

Babbitt, M. (1965). The Structure and Function of Musical Theory: I. *College Music Symposium 5*, 49–60.

Babbitt, M. (1974). Since Schoenberg. *Perspectives of New Music 12*(1–2), 3–28.

Batstone, P. (1972a). Multiple Order Functions in Twelve-Tone Music. *Perspectives of New Music 10*(2), 60–71.

Batstone, P. (1972b). Multiple Order Functions in Twelve-Tone Music (Part II). *Perspectives of New Music 11*(1), 92–111.

Bauer-Mengelberg, S. and M. Ferentz (1965). On Eleven-Interval Twelve-Tone Rows. *Perspectives of New Music 3*(2), 93–103.

Block, S. and J. Douthett (1994). Vector Products and Intervallic Weighting. *Journal of Music Theory 38*(1), 21–41.

Bóna, M. (2004). *Combinatorics of permutations*. Boca Raton: Chapman & Hall/CRC.

Brand, J., C. Hailey, and D. Harris (Eds.) (1987). *The Berg-Schoenberg correspondence: selected letters*. Basingstoke: Macmillan.

Brightwell, G. and P. Winkler (1991). Counting Linear Extensions. *Order 8*(3), 225–242.

Buchler, M. (1997). *Relative Saturation of Subsets and Interval Cycles as a Means for Determining Set-Class Similarity*. Ph.D. dissertation, Eastman School of Music.

Callender, C. (2005). Some thoughts on measuring voice-leading distance. Paper presented at the Society for Music Theory Conference, Boston, Massachusetts.

Cameron, P. (1999). *Permutation Groups*. Cambridge: Cambridge University Press.

Castrén, M. (1994). *RECREL: A Similarity Measure for Set-Classes*. Ph.D. dissertation, Sibelius Academy.

Chailley, J. (1961). *40000 ans de musique*. Paris: Librairie Plon.

Chomsky, N. (1957). *Syntactic structures*. The Hague: Mouton.

Covach, J. (2000). *Schoenberg's "Poetics of Music," the Twelve-Tone Method, and the Musical Idea*, pp. 309–346. In Cross and Berman (2000).

Critchlow, D. E. (1985). *Metric Methods for Analyzing Partially Ranked Data*. Lecture Notes in Statistics. Berlin: Springer-Verlag.

Cross, C. and R. Berman (Eds.) (2000). *Schoenberg and Words: The Modernist Years*. New York: Garland Publishing.

David, F. N. and D. Barton (1962). *Combinatorial chance*. London: Griffin.

David, F. N., M. G. Kendall, and D. Barton (1966). *Symmetric function and allied tables*. Cambridge: Cambridge University Press.

de Lannoy, C. (1972). Detection and Discrimination of Dodecaphonic Series. *Interface 1*(1), 13–27.

Demske, T. (1995a). Relating Sets: On Considering a Computational Model of Similarity Analysis. *Music Theory Online 1*(2).

Demske, T. (1995b). Reply to Richard Hermann. *Music Theory Online 1*(4).

Diaconis, P. (1988). *Group Representations in Probability and Statistics*. Institute of Mathematical Statistics Lecture Notes – Monograph Series, 11. Hayward: Institute of Mathematical Statistics.

Diaconis, P. and R. Graham (1977). Spearman's Footrule as a Measure of Disarray. *Journal of the Royal Statistical Society, Series B 39*(2), 262–268.

Diestel, R. (2000). *Graph Theory* (2 ed.). New York: Springer.

Dowling, W. (1972). Recognition of melodic transformations: Inversion, retrograde, and retrograde inversion. *Perception & Psychophysics 12*(5), 417–421.

Erdös, P. and G. Szekeres (1935). A Combinational Problem in Geometry. *Compositio Mathematica 2*, 463–470.

Erné, M. and K. Stege (1991). Counting finite posets and topologies. *Order 8*(3), 247–265.

Fishburn, P. C. (1985). *Interval orders and interval graphs*. Wiley-Interscience Series in Discrete Mathematics. New York: Wiley.

Forte, A. (1964). A Theory of Set-Complexes for Music. *Journal of Music Theory 8*(2), 136–183.

Forte, A. (1973a). The Basic Interval Patterns. *Journal of Music Theory 17*(2), 234–272.

Forte, A. (1973b). *The Structure of Atonal Theory*. New Haven: Yale University Press.

Francés, R. (1958). *La Perception de la Musique*. Paris: Librairie Philosophique J. Vrin.

French, P. A. and H. K. Wettstein (Eds.) (2003). *Meaning in the Arts*, Volume 27 of *Midwest studies in Philosophy*. Boston, MA: Blackwell Publishing.

Friedmann, M. (1985). A Methodology for the Discussion of Contour: Its Application to Schoenberg's Music. *Journal of Music Theory 29*(2), 223–248.

Fripertinger, H. (1992). Enumeration in Musical Theory. *Beiträge zur Elektronischen Musik 1*.

Goldstone, R. L. (1994). The role of similarity in categorization: providing a groundwork. *Cognition 52*, 125–157.

Goldstone, R. L. (1999). *Similarity*, pp. 763–765. In Wilson and Keil (1999).

Goodman, N. (Ed.) (1972a). *Problems and Projects*. Indianapolis: The Bobbs-Merrill Company.

Goodman, N. (1972b). *Seven Strictures on Similarity*, pp. 437–446. In Goodman (1972a).

Graham, R. L., D. E. Knuth, and O. Patashnik (1994). *Concrete Mathematics: a Foundation for Computer Science* (Second ed.). Reading, MA: Addison-Wesley.

Haimo, E. and P. Johnson (1984). Isomorphic Partitioning and Schoenberg's Fourth String Quartet. *Journal of Music Theory 28*(1), 47–72.

Hamming, R. (1950). Error-detecting and error-correcting codes. *Bell System Technical Journal 29*(2), 147–160.

Hauer, J. M. (1925). *Vom Melos zur Pauke*. Wien: Universal Edition.

Hauer, J. M. (1926). *Zwölftontechnik*. Wien: Universal Edition.

Headlam, D. (1985). The Derivation of Rows in Lulu. *Perspectives of New Music 24*(1), 198–233.

Headlam, D. (1992). Fritz Heinrich Klein's 'Die Grenze der Halbtonwelt' and Die Maschine. *Theoria 6*, 55–96.

Headlam, D. (1996). *The Music of Alban Berg*. New Haven and London: Yale University Press.

Headlam, D. (2006). Walter O'Connell's "Tone Spaces": An Early View of the Order Position/Pitch-Class Exchange. *Studies in Music from The University of Western Ontario 19–20*, 113–159.

Hermann, R. (1994). *A General Measurement for Similarity Relations: A Heuristic for Constructing or Evaluating Aspects of Possible Musical Grammars*. Ph.D. dissertation, Eastman School of Music, University of Rochester.

Hermann, R. (1995). Towards a New Analytic Method for Post-Tonal Music: A Response to Thomas R. Demske. *Music Theory Online 1*(3).

Hunter, D. and P. von Hippel (2003). How Rare Is Symmetry in Musical 12-Tone Rows? *American Mathematical Monthly 110*(2), 124–132.

Ilomäki, T. (2003). Similarity of Twelve-Tone Row Classes: A Metric Point of View. In G. Lasker, J. Lily, and J. Rhodes (Eds.), *Systems Research in the Arts, Vol V: Music, Environmental Design and the Choreography of Space*, pp. 29–34. The International Institute for Advanced Studies in Systems Research and Cybernetics.

Ilomäki, T. (2005). Group Structure and Equivalence Classes in Extended Twelve-Tone Operations. In *Proceedings of the 2005 International Computer Music Conference*, pp. 487–490.

Ilomäki, T. and Y. Kortesniemi (2004). Stochastic Estimation of Bsf. In *Proceedings of the 2004 International Computer Music Conference*, pp. 145–148.

Isaacson, E. (1990). Similarity of Interval-Class Content between Pitch-Class Sets: The IcVSIM Relation. *Journal of Music Theory 34*(1), 1–28.

Isaacson, E. (1992). *Similarity of Interval-Class Content Between Pitch-Class Sets: The IcVSIM Relation and Its Application*. Ph.D. dissertation, Indiana University.

Isaacson, E. (1996). Issues in the Study of Similarity in Atonal Music. *Music Theory Online 2*(7).

Jedrzejewski, F. (2006). *Mathematical theory of music*. Paris: Editions Delatour.

Klumpenhouwer, H. J. (1991). *A generalized model of voice-leading for atonal music*. Ph.D. dissertation, Harvard University.

Knuth, D. (1998). *The Art of Computer Programming* (Second ed.), Volume 3. Reading, Massachusetts: Addison-Wesley.

Koivisto, T. (1996). *The moment in the flow: understanding continuity and coherence in selected atonal compositions*. Ph.D. dissertation, University of Michigan.

Kolman, O. (2004). Transfer Principles for Generalized Interval Systems. *Perspectives of New Music 42*(1), 150–192.

Kowalski, D. (1987). The Construction and Use of Self-Deriving Arrays. *Perspectives of New Music 25*(1–2), 286–361.

Krumhansl, C., G. Sandell, and D. Sergeant (1987). The Perception of Tone Hierarchies and Mirror Forms in Twelve-Tone Serial Music. *Music Perception 5*(1), 31–78.

Kuusi, T. (2001). *Set-Class and Chord: Examining Connection between Theoretical Resemblance and Perceived Closeness*. Ph.D. dissertation, Sibelius Academy.

Lester, J. (1989). *Analytic Approaches to Twentieth-Century Music*. New York: Norton.

Lewin, D. (1962). A Theory of Segmental Association in Twelve-Tone Music. *Perspectives of New Music 1*(1), 89–116.

Lewin, D. (1966). On Certain Techniques of Re-Ordering in Serial Music. *Journal of Music Theory 10*(2), 276–287.

Lewin, D. (1976). On Partial Ordering. *Perspectives of New Music 14(2)/15(1)*, 252–257.

Lewin, D. (1977). A Label-Free Development for 12-Pitch-Class Systems. *Journal of Music Theory 21*(1), 29–48.

Lewin, D. (1979–80). A Response to a Response: On PC Set Relatedness. *Perspectives of New Music 18*(1–2), 498–502.

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

Liu, C. (1968). *Introduction to Combinatorial Mathematics*. New York: McGraw-Hill.

Lord, C. (1978). *An explication of some recent mathematical approaches to music analysis*. Ph.D. dissertation, Indiana University.

Lord, C. (1981). Intervallic Similarity Relations in Atonal Set Analysis. *Journal of Music Theory 25*(1), 91–111.

Mannila, H. (1985). Measures of Presortedness and Optimal Sorting Algorithms. *IEEE Transactions on Computers 34*(4), 318–325.

Martino, D. (1961). The Source Set and its Aggregate Formations. *Journal of Music Theory 5*(2), 224–273.

Marvin, E. and P. Laprade (1987). Relating Musical Contours: Extensions of a Theory for Contour. *Journal of Music Theory 31*(2), 225–267.

Mead, A. (1988). Some Implications of the Pitch-Class/Order-Number Isomorphism Inherent in the Twelve-Tone System: Part One. *Perspectives of New Music 26*(2), 96–163.

Mead, A. (1989). Some Implications of the Pitch-Class/Order-Number Isomorphism Inherent in the Twelve-Tone System: Part Two: The Mallalieu Complex: Its Extensions and Related Rows. *Perspectives of New Music 27*(1), 180–233.

Mehlhorn, K. (1979). Sorting Presorted Files. In *Proceedings of the 4th GI-Conference on Theoretical Computer Science*, London, pp. 199–212. Springer-Verlag.

Morris, R. (1977). On the Generation of Multiple Order-Function Twelve-Tone Rows. *Journal of Music Theory 21*(2), 238–262.

Morris, R. (1979–80). A Similarity Index for Pitch-Class Sets. *Perspectives of New Music 18*(1–2), 445–460.

Morris, R. (1982). Set Groups, Complementation, and Mappings Among Pitch-Class Sets. *Journal of Music Theory 26*(1), 101–144.

Morris, R. (1983–84). Set-Type Saturation among Twelve-Tone Rows. *Perspectives of New Music 22*(1–2), 187–217.

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

Morris, R. (1991). *Class Notes for Atonal Music Theory*. Lebanon, NH: Frog Peak Music.

Morris, R. (1993). New Directions in the Theory and Analysis of Musical Contour. *Music Theory Spectrum 15*(2), 205–228.

Morris, R. (2001). *Class Notes for Advanced Atonal Music Theory*. Lebanon, NH: Frog Peak Music.

Morris, R. (2003a). Compositional Theory, Musical Spaces, and Compositional Designs. Paper presented at the conference Resonances 2003, Rencontres Scientifiques, IRCAM, Paris.

Morris, R. (2003b). Precursors of Transformational Theory. Paper presented at the Mannes Institute for Advanced Music Theory.

Morris, R. and D. Starr (1974). The Structure of All-Interval Series. *Journal of Music Theory 18*(2), 364–389.

Nicholson, W. (1999). *Introduction to Abstract Algebra* (2nd ed.). New York: Wiley.

O'Connell, W. (1962). Tone Spaces. *Die Reihe 8*, 35–67.

Orpen, K. and D. Huron (1991). The Measurement of Similarity in Music: A Quantitative Approach for Non-parametric Representations. *Computers in Music Research 4*, 1–44.

Pedersen, P. (1975). The Perception of Octave Equivalence in Twelve-Tone Rows. *Psychology of Music 3*(2), 3–8.

Perle, G. (1991). *Serial Composition and Atonality: An Introduction to the Music of Schoenberg, Berg, and Webern* (6th ed.). Berkeley, CA: University of California Press.

Polansky, L. (1996). Morphological Metrics. *Journal of New Music Research 25*(4), 289–368.

Polansky, L. and R. Bassein (1992). Possible and Impossible Melody: Some Formal Aspects of Contour. *Journal of Music Theory 36*(2), 259–284.

Pruesse, G. and F. Ruskey (1997). Generating Linear Extensions Fast. *SIAM Journal on Computing 23*(2), 373–386.

Quinn, I. (2001). Listening to Similarity Relations. *Perspectives of New Music 39*(2), 108–158.

Raffman, D. (2003). *Is Twelve-Tone Music Artistically Defective?*, pp. 69–87. Volume 27 of French and Wettstein (2003).

Rahn, J. (1979–80). Relating Sets. *Perspectives of New Music 18*(1–2), 483–498.

Rahn, J. (1980). *Basic Atonal Theory*. New York: Longman.

Rahn, J. (1989). Toward a Theory of Chord Progression. *In Theory Only 11*(1–2), 1–10.

Rauchhaupt, U. v. (1971). *Die Streichquartette der Wiener Schule; Schoenberg, Berg, Webern. Eine Dokumentation.* München: H. Ellerman.

Read, S. (1997). Combinatorial problems in the theory of music. *Discrete Mathematics 167–168*, 543–551.

Reiner, D. (1985). Enumeration in Music Theory. *American Mathematical Monthly 92*, 51–54.

Roeder, J. (1987). A Geometric Representation of Pitch-Class Series. *Perspectives of New Music 25*(1–2), 362–409.

Rogers, D. (1999). A Geometric Approach to Pcset Similarity. *Perspectives of New Music 37*(1), 77–90.

Rothgeb, J. (1967). Some Ordering Relationships in the Twelve-Tone System. *Journal of Music Theory 11*(2), 176–197.

Rudin, W. (1987). *Real and Complex Analysis* (3 ed.). New York: McGraw-Hill.

Rufer, J. (1966). *Die Komposition mit zwölf Tönen* (2 ed.). Kassel: Bärenreiter.

Schoenberg, A. (1975). *Style and Idea: Selected Writings of Arnold Schoenberg*. Transl. Leo Black. Berkeley and Los Angeles: University of California Press.

Scott, D. and E. Isaacson (1998). The Interval Angle: A Similarity Measure for Pitch-Class Sets. *Perspectives of New Music 36*(2), 107–142.

Sidorenko, A. (1992). Inequalities for the Number of Linear Extensions. *Order 8*(4), 331–340.

Solomon, L. (1973). New Symmetric Transformations. *Perspectives of New Music 11*(1), 257–264.

Spearman, C. (1906). Footrule for measuring correlation. *British Journal of Psychology 2*, 89–108.

Stanfield, M. (1984). Some Exchange Operations in Twelve-Tone Theory: Part One. *Perspectives of New Music 23*(1), 258–277.

Stanfield, M. (1985). Some Exchange Operations in Twelve-Tone Theory: Part Two. *Perspectives of New Music 24*(1), 72–95.

Stanley, R. (1997). *Enumerative Combinatorics*, Volume I. Cambridge: Cambridge University Press.

Stanton, D. and D. White (1986). *Constructive Combinatorics*. Undergraduate Texts in Mathematics. New York: Springer-Verlag.

Starr, D. (1978). Sets, Invariance, and Partitions. *Journal of Music Theory 22*(1), 1–42.

Starr, D. (1984). Derivation and Polyphony. *Perspectives of New Music 23*(1), 180–257.

Starr, D. and R. Morris (1977). A General Theory of Combinatoriality and the Aggregate (Part 1). *Perspectives of New Music 16*(1), 3–35.

Starr, D. and R. Morris (1978). A General Theory of Combinatoriality and the Aggregate (Part 2). *Perspectives of New Music 16*(2), 50–84.

Steen, L. and J. Seebach (1978). *Counterexamples in Topology* (2nd ed.). New York: Springer-Verlag.

Teitelbaum, R. (1965). Intervallic Relations in Atonal Music. *Journal of Music Theory 9*(1), 72–127.

Toiviainen, P. (1996). Optimizing auditory images and distance metrics for self-organizing timbre maps. *Journal of New Music Research 25*(1), 1–30.

Trotter, W. T. (1992). *Combinatorics and partially ordered sets*. Johns Hopkins Series in the Mathematical Sciences. Baltimore, MD: Johns Hopkins University Press.

van Lint, J. and R. Wilson (2001). *A Course in Combinatorics* (2nd ed.). Cambridge: Cambridge University Press.

Ward, J. (1992). *Theories of similarity among ordered pitch class sets*. Ph.D. dissertation, The Catholic University of America.

Webern, A. (1960). *Der Weg zur Neuen Musik*. Wien: Universal Edition.

Webern, A. (1963). *The path to the new music*. Bryn Mawr, Pennsylvania: Theodore Presser company.

West, D. (2001). *Introduction to Graph Theory* (2nd ed.). Upper Saddle River: Prentice Hall.

Wilson, R. and F. Keil (Eds.) (1999). *The MIT encyclopedia of the cognitive sciences*. Cambridge, Mass.: MIT Press.