The Group-theoretic Description of Musical Pitch Systems

Marcus Pearce

marcus.pearce@qmul.ac.uk

1 Introduction

Balzano (1980, 1982, 1986a,b) addresses the question of finding an appropriate level for describing the resources of a pitch system (e.g., the 12-fold or some microtonal division of the octave). His motivation for doing so is twofold:

> "On the one hand, I am interested as a psychologist who is not overly impressed with the progress we have made since Helmholtz in understanding music perception. On the other hand, I am interested as a computer musician who is trying to find ways of using our powerful computational tools to extend the musical [domain] of pitch ..." (Balzano, 1986b, p. 297)

Since the resources of a pitch system ultimately depend on the pairwise relations between pitches, the question is one of how to conceive of pitch intervals. In contrast to the prevailing approach which describes intervals in terms of frequency ratios, Balzano presents a description of pitch sets as mathematical groups and describes how the resources of any pitch system may be assessed using this description. Thus he is concerned with presenting an *algebraic* description of pitch systems as a competitive alternative to the existing *acoustic* description.

In these notes, I shall first give a brief description of the ratio based approach (§2) followed by an equally brief exposition of some necessary concepts from the theory of groups (§3). The following three sections concern the description of the 12-fold division of the octave as a group: §4 presents the nature of the group C12; §5 describes three perceptually relevant properties of pitch-sets in C12; and §6 describes three musically relevant isomorphic representations of C12. In §8 we review some psychological evidence for this way of characterising the resources of the 12-fold pitch system before proceeding to consider the extension of the approach to microtonal pitch systems in §7. Finally, we outline some related approaches to the representation of pitch in §9 and discuss the wider significance of the algebraic approach in §10.

2 Ratio Based Explanations

The description of pitch intervals as frequency ratios dates from Pythagoras, but was presented in a more complete and mechanistic manner by Helmholtz (1877). According to this theory, intervals are described as sets of ratios defined as powers of the prime numbers – mathematical objects of the form:

$$2^p . 3^q . 5^r \ldots$$

with p, q and r ranging over the positive and negative integers. In the twelve-fold system, the set of musical intervals relative to some fundamental frequency can be described in this way using only the prime factors 2, 3 and 5 (see Table 1).

A modern approach to describing pitch intervals is the cent system by which there are 1200 cents to the octave. Therefore, a semitone corresponds to 100 cents in the equal-tempered 12-fold system and

Interval	Ratio	$2^p.\ 3^q.\ 5^r$	Cents	Hz
I	1:1	2^0	0	440 (A)
m2	16:15	$2^4.\ 5^3$	111.73	469.33
M2	9:8	$2^{-3}.\ 3^3$	203.91	495
m3	6:5	$2^1.\ 3^1.\ 5^{-1}$	315.64	528
M3	5:4	$2^{-2}.\ 5^1$	386.31	550
p4	4:3	$2^2.\ 3^{-1}$	498.05	586.67
tt	45:32	$2^{-5}.\ 3^3.\ 5^1$	590.22	618.75
p5	3:2	$2^{-1}.\ 3^1$	701.96	660
m6	8:5	$2^3.\ 5^{-1}$	813.69	704
M6	5:3	$3^{-1}.\ 5^1$	884.36	733.33
m7	9:5	$3^2.\ 5^{-1}$	1017.60	792
M7	15:8	$2^{-3}.\ 3^1.\ 5^1$	1088.26	825
p8	2:1	2^1	1200	880

Table 1: A description of the 12-fold pitch system in terms of frequency ratios

to 60 in a 20-fold microtonal division of the octave. The number of cents in the interval between two frequencies f^1 and f^2 may be calculated as:

$$cents = 1200\ \frac{\log \frac{f^1}{f^2}}{\log 2}$$

The resources of a equal-tempered n-fold division of the octave are then a function of the goodness of fit between the equal log frequency intervals and the desired ratios as shown in Table 1. For example, in the equal tempered 12-tone system each semitone consists of 100 cents, leading to discrepancies of two cents on perfect fifths, 14 cents on major thirds and 16 cents on major sixths. The ear does not seem to mind these discrepancies, however, and equal temperament has the significant advantage that it avoids the wild tuning errors of other temperaments (e.g., Pythagorean, just or mean-tone) in certain keys. The ratio-based approach predicts that systems with *9*-, *31*- and *41*-fold divisions of the octave also allow the use of equal-tempered systems which closely approximate true tunings.

Helmholtz (1877) explained consonance and dissonance in terms of the coincidence and proximity of the overtones and differences tones that arise when simultaneously sounded notes excite non-linear physical resonators (e.g., the human ear). An interval is consonant to the degree that its most powerful secondary overtones exactly coincide, and is dissonant to the extent that any of its partials are separated by a frequency small enough to beat at a rate of around 33 Hz. Thus for a diatonic semitone (16 : 15) only very high, low-energy overtones coincide so it is weakly consonant, while the two fundamentals themselves produce beats (in the usual musical ranges) so it is strongly dissonant. For the perfect fifth (3 : 2), on the other hand, all its powerful partials coincide and only weak ones are close enough to beat; it is strongly consonant and weakly dissonant. Under the theory, the reason equal temperament works is because it distorts consonant intervals very little and the distortion to the already dissonant intervals is less noticeable.

Any study of pitch sets must account for the historical and cultural ubiquity of the diatonic and pentatonic scales. While there seems to be nothing special about these scales when viewed at the level of individual tones, they may be distinguished from other scales by the simplicity of the frequency ratios between their pitches and the close approximations they allow to important ratios such as 3 : 2, 4 : 3 and 5 : 4 (Helmholtz, 1877). At the level of ratios, therefore, these scales do appear to be somewhat special.

3 Basic Concepts in Group Theory

A group is one of the simplest algebraic structures consisting of a set G and an operation $*$. In order to constitute a group, the following conditions must be satisfied by $\langle G, * \rangle$:

1. G must include an identity element I such that: $\forall a \{a * I = a \mid a \in G\}$;
2. for each element e in G there must be an inverse element e^{-1} such that: $\{e^{-1} \in G \wedge e * e^{-1} = e^{-1} * e = I\}$;
3. the operation $*$ must be associative such that: $\forall a, b, c \{a * (b * c) = (a * b) * c \mid a, b, c \in G\}$;
4. G must be closed under the operation $*$ such that: $\forall a, b \{c \in G \mid a, b \in G \wedge c = a * b\}$.

If the operation $*$ is commutative:

$$\forall a, b \{a * b = b * a \mid a, b \in G\}$$

the group is said to be *Abelian*. A group G is said to be *finite* if the set G has a finite number of elements; otherwise G is an infinite group.[1] The number of elements in G is called the order of G denoted by $|G|$.

Let us consider a simple example. The set of positive integers is closed under the associative operation of addition but does not constitute a group without both an identity element, 0, and the negative integers which are the inverse elements of the positive integers. The integers, $\mathcal{Z}$ therefore form a group, C_∞, under integer addition. Since addition is commutative and $\mathcal{Z}$ has an infinite number of elements, C_∞ is an Abelian group of infinite order.

Groups that can be generated in their entirety by one element of the group are called *cyclic* groups. In order to precisely define a cyclic group, we shall use the following notation where a is an element in the group $\langle G, * \rangle$ and $n \in \mathcal{Z}^+$:

- $a^n = a * a * a \ldots * a$ (n times)
- $a^{-n} = a^{-1} * a^{-1} * a^{-1} \ldots * a^{-1}$ (n times)
- $a^0 = I$ (the identity element)

from which it follows (using the basic laws of exponents) that:

- $a^m * a^n = a^{m+n}$
- $(a^m)^n = a^{mn}$
- the inverse of a^n is a^{-n}.

$\langle G', * \rangle$ is said to be a *subgroup* of group $\langle G, * \rangle$ if the following conditions hold: G' is a subset of G; the identity element of G is also a member of G'; and $\langle G', * \rangle$ is a group under the operation $*$. The *cyclic subgroup* of G generated by a (a member of G) is then defined as:

$$\langle a \rangle = \{x \in G \mid x = a^n \wedge n \in \mathcal{Z}\}$$

The group G is a cyclic group if there exists an element a in G such that the cyclic subgroup generated by that element is equal to G:

$$\exists a \{G = \langle a \rangle \mid a \in G\}$$

[1] It is often convenient for the purposes of exposition to say that the set is the group, the operation being implied only.

The element a that satisfies this condition is known as a *generator* of G. In general, the inverse of any generator of a group (i.e., a^{-1} is also a generator of the group. Finally, it can be shown that all cyclic groups are Abelian groups. To take the example of C_∞, it can be seen that the element 1 (and its inverse -1) are group generators since that element and its (positive and negative powers) are sufficient to generate the entire group. Therefore, C_∞ is a cyclic group.

A further concept that we shall need to consider is that of group *homomorphisms* and *isomorphisms*. Let $\langle G,*\rangle$ and $\langle G',*'\rangle$ be groups and $\phi : G \to G'$ be a function. Then, ϕ is said to be a group homomorphism if:

- ϕ is an *onto* relation (it maps elements in G onto elements in G');
- any true statement that links elements in G is also true of the mapped images of those elements in G': $\forall a,b\{\phi(a*b) = phi(a) *' phi(b) \mid a,b \in G\}$.

If ϕ also constitutes a *one-to-one* mapping between elements in G and G' (i.e., it has an inverse) then we have a group isomorphism (a particular type of homomorphism) denoted by: $G \cong G'$. One of the most common examples of a group isomorphism is found in the theory of logarithms. Let G be the group of positive real numbers under multiplication and G' be the group of real numbers under addition. The logarithmic function, log, is a one-to-one relation since it has an inverse (the exponential function) and it can be seen that:

$$\forall a,b\{\log(a)+\log(b) = \log(a \times b) \mid a,b \in G\}$$

Therefore, log constitutes an isomorphism between the group of positive real numbers under multiplication and the group of real numbers under addition.

Finally, we shall need the concept of *direct product groups*. Again, let $\langle G,*\rangle$ and $\langle G',*'\rangle$ be groups. The set of all ordered pairs (a,a') such that $a \in G$ and $a' \in G'$ is called the direct product of G and G' denoted by $G \times G'$. The direct product $G \times G'$ is a group under the operation $\otimes$ such that:

$$\forall (a,a'),(b,b')\{(a,a') \otimes (b,b') = (a*b, a' *' b' \mid (a,a'),(b,b') \in G \times G'\}$$

To take an example, let G be the set $\{0,1,2\}$ under mod 3 addition and G' be the set $\{0,1\}$ under mod 2 addition. The direct product group $G \times G'$ contains six elements: $\{(0,0),(0,1),(0,2),(1,0),(1,1),(1,2)\}$. It can also be seen that for all elements $(a,a'),(b,b') \in G \times G'$:

$$(a,b) \otimes (a',b') = ([a+a']mod\ 2, [b+b']mod\ 3)$$

For example, $(1,1) \otimes (1,2) = (0,0)$.

Up to this point, we have given a *static* interpretation of groups where the members of the group are *elements* and the operation yields a *combination* of those elements. Every group also admits of a *dynamic* interpretation whereby the elements are viewed as *transformations* and the operation as *succession*. To take the example of C_∞, $3+4=7$ has a static interpretation as the addition of two integers (3 and 4) to yield 7 and a dynamic interpretation as the successive application of two transformations "Add 3" and "Add 4" to yield a third transformation "Add 7". It will be convenient to have a notation for the dynamic interpretation of groups and Balzano (1980) proceeds as follows. He defines an adjacency predicate $NEXT$ or N whence, for example, $2 = N(1)$ and $3 = N(2) = N(N(1)) = N^2(1)$. Furthermore, $1 = N^{-1}(2)$ and, for that matter, $1 = N(N^{-1}(1)) = N^0(1) = I(1)$. Under the dynamic interpretation, therefore, C_∞ consists of the set $\{\ldots,N^{-2},N^{-1},N^0,N^1,N^2,\ldots\}$.

4 The Group C12 and its Subgroups

Balzano (1980) begins his group-theoretic analysis of pitch by noting that any pitch system, or set of points on a (log) frequency continuum, is an example of C_∞. Under the static interpretation, each element is a pitch place in the system while under the dynamic interpretation, each element is a musical interval corresponding to a mapping between (sets of) pitch places. However, since C_∞ is common to all equal-tempered systems, this group is not very useful in describing the differences between such systems (Balzano, 1980, p. 68). The distinctive character of each system (of each different n-fold division of the octave) results from exploiting the redundancy of the system at every octave N^n. Balzano therefore introduces a homomorphic mapping of C_∞ to C_n for each n- fold pitch system. The mapping operates as an equivalence relation, partitioning the N^i of C_∞ into n equivalence classes:

$$\{ \quad \{\ldots,N^{-2n},N^{-n},I,N^n,N^{2n},\ldots\},$$
$$\{\ldots,N^{-2n+1},N^{-n+1},N^1,N^{n+1},N^{2n+1},\ldots\},$$
$$\{\ldots,N^{-2n+2},N^{-n+2},N^2,N^{n+2},N^{2n+2},\ldots\},$$
$$\ldots,$$
$$\{\ldots,N^{-2n+(n-1)},N^{-n+(n-1)},N^(n-1),N^{n+(n-1)},N^{2n+(n-1)},\ldots\} \quad \}$$

according to the function:

$$N^{jn+k} \to k(mod\ n)$$

where j ranges over the integers and k takes on the values $0,1,2,\ldots n-1$. When n is not clear from the embedding context, $k(mod\ n)$ is written as k_n.

The group C_{12} therefore consists of the set $\{0,1,2,3,4,5,6,7,8,9,10,11\}$ and the operation of modulo 12 addition. Unless explicitly stated otherwise, we shall use the symbol $*$ to denote modulo 12 addition. Each element in the set corresponds to a pitch class or octave equivalence class. The identity element, 0, has a dynamic interpretation as the set of pitch-class preserving transformations $\{\ldots,N^{-24},N^{-12},N^0,N^{12},N^{24},\ldots\}$, and a static interpretation as an element of potentially arbitrary origin to which the addition rules apply no less and no more than to the other elements. For any element, a in C_{12}, the inverse of that element is given by:

$$a^{-1} = (12-a)\ mod\ 12$$

Thus 1 and 11 are inverses, as are 2 and 10, and 6 is its own inverse. Modulo 12 addition is an associative operation and since the sum of any two integers is an integer, mod 12 reduction of the sum provides that it belongs to C_{12} and closure is thereby assured.

The complete group table (or *Cayley* table) for C_{12} is shown in Table 2. Although the static notation has been used for the sake of visibility, it is useful to conceive of the table as showing the combined effects of transformations. Thus $5*9=2$ represents the fact that $N^5(9)=2$: N^5 transforms pitch class 9 onto pitch class 2. We can also use this notation to express the effect of a transformation, N^i, on an entire pitch set S of size m:

$$N^i(\{S_0,S_1,S_2,\ldots,S_{m-1}\}) = \{S_0*i,S_1*i,S_2*i,\ldots,S_{m-1}*i\}$$

For example, $N^4(\{0,4,7\}) = \{4,8,11\}$.

In musical terms, the identity element, 0, may be defined to be any one of the twelve pitch classes in the chromatic set $\{C,C^\sharp,D,D^\sharp,E,F,F^\sharp,G,G^\sharp,A,A^\sharp,B\}$ (or their enharmonic equivalents) with the remaining elements corresponding to the remaining pitch classes in cyclic order. The dynamic interpretation of C_{12} as a set of intervals yields the musical interpretation given in Table 3. Balzano (1982) notes that while he adopts the convention of thinking of intervals as upward pitch transformations none of the results would be affected by assuming the opposite. He also draws attention to the congruence between

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

Table 2: The group table for C_{12} (adapted from Balzano, 1982)

0	1	2	3	4	5	6	7	8	9	10	11
p8/I	m2	M2	m3	M3	p4	tt	p5	m6	M6	m7	M7

Table 3: The elements of C_{12} as pitch intervals

mathematical and musical terminology in that the relation between, for example, the intervals m2 and M7 is called *inverse* by both.

Depending on the tuning system, these intervals translate approximately or exactly into log frequency differences. The interpretation of any interval N^i is independent of the particular temperament used so long as the system satisfies the criterion of *scalestep-semitone coherence* (see § 5.3). (Balzano, 1982, p. 68) points out, however, that when the system is equal-tempered, N^i has a natural interpretation in terms of equal log frequency intervals and that this property facilitates the interpretation of the group and its elements.

By taking subsets of C_{12} we may represent triads or chords and scales. For example, the pitch set $\{0,4,7\}$ mentioned above corresponds to a major triad. The transformation of a pitch set by N^i corresponds musically to a transposition of the pitch set up i semitones. For example, if we take 0 to correspond to C, then $\{0,4,7\}$ is a C major triad and $N^4(\{0,4,7\}) = \{4,8,11\}$ is an E major triad. If a pitch set S' may be obtained by the application of a transformation in C_{12} to pitch set S, the two sets are *transpositionally related* or are in the same *family* (Balzano, 1982, p. 324). For any pitch set, there are clearly 11 other pitch sets in the same family.

Mode	Pitch set
Ionian	$\{0,2,4,5,7,9,11\}$
Dorian	$\{2,4,5,7,9,11,0\}$
Phrygian	$\{4,5,7,9,11,0,2\}$
Lydian	$\{5,7,9,11,0,2,4\}$
Mixolydian	$\{7,9,11,0,2,4,5\}$
Aeolian	$\{9,11,0,2,4,5,7\}$
Locrian	$\{11,0,2,4,5,7,9\}$

Table 4: The church modes as rotationally equivalent members of the diatonic scale family

It is also important to make clear that the level at which pitch sets are described includes a notion of *rotational equivalence*. (Balzano, 1982, p. 326) stipulates that while the first element in a pitch set may be any element, the remaining elements must be written in ascending numerical order mod 12. All sets that correspond to rotations of such a set are considered to be in the same family. Taking the major diatonic scale, $\{0,2,4,5,7,9,11\}$, as an example, the pitch sets $\{2,4,5,7,9,11,0\}$, $\{4,5,7,9,11,0,2\}$ are considered to be in the same family. These sets correspond to the "church modes" with the tonic given by the first note in the pitch set (see Table 4). In each of the modes, the key is always given by the element, 0, and it can be seen that the Ionian mode corresponds to the major diatonic scale in a particular key, while the natural minor diatonic scale in the same key corresponds to the Aeolian mode transposed up a minor third:

$$N^3(\{9,11,0,2,4,5,7\}) = \{0,2,3,5,7,8,10\}$$

Therefore, the level of description employed by Balzano corresponds more closely to that of "diatonic scale family" than it does to "major scale". He justifies this by noting that "the fact that, of the seven possible tonics in this seven-note scale, fully six [all but the Locrian] had seen extensive musical use until about the 17th century, recommends that we leave open the possibility that a tonic may be determined mainly by temporal-contextual factors, and not by intrinsic set-structural properties" (Balzano, 1982, p. 325).

Subgroup	Pitch set	Distinct members	Musical Name
C_6	$\{0,2,4,6,8,10\}$	2	Whole-tone scale
C_4	$\{0,3,6,9\}$	3	Diminished-7th chord
C_3	$\{0,4,8\}$	4	Augmented triad
C_2	$\{0,6\}$	6	Tritone

Table 5: Subgroups of C_{12}

From the universal chromatic set of twelve pitch classes, one can select $2^{12} - 1 = 4095$ different pitch sets, including the full chromatic itself and the twelve single-note pitch-sets. We could say equivalently that the size of the power-set of C_{12} is 4095, not including the empty set. Partitioning this power-set according to a relation that counts members of the same family (i.e., rotationally and transpositionally equivalent sets) as equivalent results in a reduction to 351 distinct sets, 349 without the full chromatic and the single-note set. A number of these subsets of the full chromatic are special in that they correspond to subgroups of C_{12}. These subgroups are shown in Table 5, the third column of which contains the rotationally distinct members of the families of C_{12} to which these subsets belong. (Balzano, 1980, p. 70) notes that the subgroups of any C_n are simply all the C_k such that k divides n evenly. Furthermore, it can be seen that the subsets of C_{12} associated with each of these subgroups are all incomplete, since they consist of fewer than 12 distinct sets.

5 Distinguishing Properties of Pitch Sets

5.1 Overview

Balzano (1982) uses the analysis of C_{12} reviewed in §4 as a basis for demonstrating the uniqueness of the diatonic scale family in a manner that is entirely independent of frequency-ratios. In order to do so, he introduces three properties of pitch sets which he argues are of significant perceptual relevance to the listener. We shall review each of these properties in turn.

5.2 Uniqueness

Balzano (1982) suggests that the emergence of a tonic element in a pitch set, it must be possible to individuate the elements of the set by virtue of their relations with one another. Therefore, he defines a quality called *uniqueness*; a pitch set satisfies this property if each of its elements has a unique set of relations with the others and therefore has the potentiality for a unique musical role or "dynamic quality". The hypothesised relevance of this property to a listener is that "a melody based on a scale satisfying Uniqueness should be easier for a perceiver to deal with because the notes of the melody are individuated not only by their particular frequency locations, but also by their interrelations with one another" (Balzano, 1982, p. 326).

More formally, for each element of a pitch set $S = \{S_0, S_1, S_2, \ldots, S_{m-1}\}$ of cardinality m, Balzano defines a vector of relations, $V(S_i) = V_i = (v_{i_0}, v_{i_1}, \ldots, v_{i_{(m-1)}}$ such that:

$$v_{i_j} = [S_{[i+j] mod\ m} - S_i] mod\ 12$$

For example, in the diatonic scale $\{2,4,5,7,9,11,0\}$, $V(5) = V_2 = \{5-5, 7-5, 9-5, 11-5, 0-5, 2-5, 4-5\} = \{0,2,4,6,7,9,11\}$. A set satisfies Uniqueness if and only if the vector of relations associated with each element is distinct:

$$V_i = V_{i'} \iff i = i'$$

While members of the diatonic scale family exhibit this property, all the sets associated with subgroups of C_{12} (see Table 5) and sets such as $\{0,1,4,5,8,9\}$, $\{0,1,3,4,6,7,9,10\}$ and $\{0,2,3,5,6,8,9,11\}$ fail to satisfy Uniqueness. In all these cases, "it is the very symmetry and apparent elegance of the set that is its undoing with respect to Uniqueness" (Balzano, 1982, p. 326). Unfortunately, many more sets satisfy Uniqueness than fail it, so more than this property is required to allow thorough differentiation among potential scales.

5.3 Scalestep-Semitone Coherence

The vector of relations described in §5.2 is defined in terms of distance relationships based on the semitone unit of C_{12}. Any pitch set (or scale) short of the full chromatic also gives rise to a distance metric which (Balzano, 1982, p. 328) calls the *scalestep*. v_{i_j} describes the number of semitones contained in a distance of j scalesteps from a given scale member S_i. In the diatonic set $\{2,4,5,7,9,11,0\}$, for example, $S_2 = 5$ and $V_2 = \{0,2,4,6,7,9,11\}$; a distance of, for example, three scalesteps from scale element S_2 contains six semitones ($v_{i_j} = 6$).

The number of semitones contained in a distance of one scalestep may be different at different points in the scale. In the diatonic scales, for example, some scalesteps are two semitones wide while others are one semitone wide. In general, greater numbers of scalesteps from a particular scale degree correspond to greater numbers of semitones, but there is nothing that forces this to obtain everywhere in the scale. In the scale $\{0,2,4,9\}$, for example, a distance of two scalesteps from the element 0 (v_{0_2}) corresponds to four semitones while a distance of one scalestep from the element 4 (v_{4_1}) corresponds to five semitones; a larger number of scalesteps (v_{0_2} vs. v_{4_1}) is associated with a smaller number of semitones (four vs. five). For scales like these, the relation between scalesteps and semitones is *not coherent*.

Regarding the perceptual content of this property, Balzano (1982) notes that semitones are (approximately) a linear function of log frequency and that we do perceive musical pitch in terms of log frequency (e.g., perceptual invariance of melodies under constant log frequency shifts). If a scale is to satisfy Uniqueness, however, scalesteps cannot in turn be a linear function of semitones (see §5.2). Unless scalesteps are at least a monotonic, increasing function of semitones, perception mediated by log frequency will fail to yield consistent results in terms of scalesteps and perception would be "stalled" at the semitone level. To take an example, an interval of four semitones can be readily identified (in or out of context) as a major third as a scalestep-level property. An interval of six semitones, on the other hand, can often be identified only as the scale-neutral tritone, since it may function in a diatonic context as either an augmented fourth or a diminished fifth. If all intervals were like the tritone, scalestep-level perception could never occur outside the context of a specific piece of music and then only if the key didn't change too rapidly. Since learning to recognise intervals may require at least a temporary isolation

Cardinality	Family Size	Pitch sets	Musical Name
7	66	$\{0,2,4,5,7,9,11\}$	Diatonic Scale
5	66	$\{0,2,4,7,9\}$	Pentatonic Scale
		$\{0,2,4,6,9\}$	???
		$\{0,1,4,7,9\}$	???
		$\{0,1,4,6,9\}$	???
4	43	$\{0,4,7,11\}$	Dominant 7th Chord
		$\{0,3,6,10\}$	Leading 7th Chord
		$\{0,4,7,10\}$	Augmented 6th Chord
		$\{0,2,5,9\}$	???
3	19	All but C_3	???
2	6	All but C_2	???
1	1	$\{0\}$	???

Table 6: Pitch sets satisfying both Uniqueness and Coherence

of the interval from context, it is hard to see how learning to perceive scalestep-level qualities could occur unless scalestep-semitone coherence were satisfied.

Balzano (1982) defines the property of scalestep-semitone coherence formally as follows. A pitch set of cardinality m satisfies coherence if:

$$\forall S_i, S_{i'} \{ j < k \Rightarrow v_{i_j} < v_{i'_k} \mid j,k \in \{0,1,2,\ldots,m-1\}\}$$

Thus for scales satisfying coherence, larger numbers of scalesteps are always associated with a larger number of semitones; scalesteps are a monotone increasing function of semitones.[2] It can be shown, however, that any scale of odd cardinality containing a tritone will fail to satisfy coherence as defined above and that the failure will be localised to the tritone interval ($v_{i_j} = 6$) only. This can be remedied by relaxing the strict inequality on the right of the equation for that interval:

$$j < k \Rightarrow \begin{cases} v_{i_j} \leq v_{i'_k} & \text{if } v_{i_j} = 6 \\ v_{i_j} < v_{i'_k} & \text{otherwise.} \end{cases}$$

Coherence imposes powerful constraints on scales such that few pitch sets conform to it. Of the 66 distinct 5-note scale families, only four satisfy coherence, of which the familiar pentatonic scale, $\{0,2,4,7,9\}$ is one. Of the 80 distinct 6-note scale families, only two are coherent (one of which is the whole-tone scale) and both of which fail Uniqueness. Finally, of the 66 essentially different 7-note scale families, only the diatonic scale, satisfies Coherence. The pitch sets which satisfy both Uniqueness and Coherence are shown in Table 6. However, none of the 19 distinct 3-note pitch sets or the six distinct 2-note pitch sets fail to satisfy coherence, although C_3 and C_2 fail to satisfy Uniqueness (see §5.2).

5.4 Transpositional Simplicity

While Uniqueness and Coherence are based on features of a scale determined by relations between its elements, the third property described by Balzano (1982), *transpositional simplicity*, concerns relations between members of a scale's family. The basic idea is that it should be a simple task to move between different members of a scale family (transpose the scale) or, to put it another way, relationships between family members should be easy to define in terms of overlap of scale degrees.

[2]Note that the entailment applies in one direction only. The converse would require that all representatives of j scalesteps contain an identical number of semitones and this would lead to a failure of Uniqueness (see §5.2).

Balzano (1982) proceeds towards a formal definition of transpositional simplicity by defining predicates in terms of set operations. The first of these defines a *betweeness* relation among three scale family members as follows:

$$X/Y/Z \equiv [(X \cap Z) \subset (X \cap Y)] \wedge [(X \cap Z) \subset (Y \cap Z)]$$

where $X/Y/Z$ is to be read "Y is between X and Z". The character of betweeness can be more intuitively appreciated from an immediate consequence of the above definition:

$$X/Y/Z \Rightarrow (X \cap Z) \subset Y$$

For Y to be between X and Z, all elements shared by X and Z must also be contained in Y. From these definitions, it follows that:

1. no scale is between itself and any other scale: $\neg X/X/Z$;
2. betweeness is symmetrical with respect to its first and third arguments: $X/Y/Z \Rightarrow Z/Y/X$;
3. otherwise betweenness is asymmetrical: $X/Y/Z \Rightarrow \neg Y/X/Z \wedge \neg X/Z/Y$.

The next step is to define *besideness* in terms of betweenness. Two scales are beside one another if they enter into at least one betweenness relation and no other scale is between them:

$$Beside(X,Y) \equiv (\exists R : X/Y/R \vee Y/X/R) \wedge (\neg\exists S : X/S/Y)$$

Balzano (1982) suggests that the simplest scale families have a large number of betweenness relations and a small number of besideness relations. The simplest kind of besideness relations occur when each scale family member is beside exactly two others. This can arise in two different ways, exemplified by the harmonic minor scale family, $\{0,2,3,5,7,8,11\}$, and the diatonic scale family, $\{0,2,4,5,7,9,11\}$ (see Table 7). In the former, for any particular family member S, $N^0(S)$ is beside $N^3(S)$ which is beside $N^6(S)$ which is beside $N^9(S)$ which is beside $N^0(S)$, but there are no besideness relations connecting these four scales with the other eight family members. The scales associated with $\{N^1(S), N^4(S), N^7(S), N^{10}(S)\}$ and $\{N^2(S), N^5(S), N^8(S), N^{11}(S)\}$ are similar: "each set of four scales can be represented on the vertices of a square, but the relation between the three squares so obtained is indeterminate" (Balzano, 1982, p. 332).

Family Member	Harmonic minor	Diatonic
$N^0(S)$	$N^3(S), N^9(S)$	$N^5(S), N^7(S)$
$N^1(S)$	$N^4(S), N^10(S)$	$N^6(S), N^8(S)$
$N^2(S)$	$N^5(S), N^{11}(S)$	$N^7(S), N^9(S)$
$N^3(S)$	$N^0(S), N^6(S)$	$N^8(S), N^{10}(S)$
$N^4(S)$	$N^1(S), N^7(S)$	$N^9(S), N^{11}(S)$
$N^5(S)$	$N^2(S), N^8(S)$	$N^0(S), N^{10}(S)$
$N^6(S)$	$N^3(S), N^9(S)$	$N^1(S), N^{11}(S)$
$N^7(S)$	$N^4(S), N^{10}(S)$	$N^0(S), N^2(S)$
$N^8(S)$	$N^5(S), N^{11}(S)$	$N^1(S), N^3(S)$
$N^9(S)$	$N^0(S), N^6(S)$	$N^2(S), N^4(S)$
$N^{10}(S)$	$N^1(S), N^7(S)$	$N^3(S), N^5(S)$
$N^{11}(S)$	$N^2(S), N^8(S)$	$N^4(S), N^6(S)$

Table 7: Besideness relations for family members of harmonic minor and diatonic scale families (adapted from Balzano, 1982)

In the case of the diatonic scale, on the other hand, the family is fully connected by the besideness relation: $N^0(S)$ is beside $N^7(S)$ which is beside $N^2(S)$ which is beside $N^9(S)$ and so on back to $N^0(S)$. There is only one other 7-note scale family that leads to configuration of besideness relations as simple as this: $\{0,1,2,3,4,5,6\}$. For both these scales, besideness is essentially determined by a single group element $N^7(S)$ (and its inverse) or $N^1(S)$ (and its inverse). While there are scale families of other cardinalities that also have besideness relations determined by a single element, no other scale of any cardinality gives rise to a space of scale relations determined by any other element. We shall see why this is so in §6. For now it is enough to note that the only pitch sets with cardinality greater than three which satisfy uniqueness, coherence and transpositional simplicity are the diatonic and pentatonic scales.

5.5 Summary

Balzano (1982) sets out with the aim of demonstrating the significance of the diatonic scale family within his group-theoretic framework. He defines three properties of pitch sets (or scales) which he argues have psychological relevance to the listener:

1. *uniqueness*: there should be a unique *vector of relations* for each element in a pitch set so that the listener can orient herself unambiguously in relation to the other pitches.
2. *scalestep-semitone coherence*: the intervals of a scale proceed such that scalesteps form a monotonic, increasing function of semitones – this further aids the unambiguous identification of the position of intervals in the scale;
3. *transpositional simplicity*: there should be a simple and unambiguous relationship between members of scale family (i.e., transposed scales).

From the set of 325 pitch set families with more than three elements, the diatonic and pentatonic scales are shown to be unique in satisfying all three of these simple criteria: "Our enquiry into the the structure of scales and of the embedding group C_{12} has lead us to the two most nearly universal scales in the history of music." (Balzano, 1982, p. 335).

6 Isomorphisms of C12

6.1 Overview

Balzano (1980, 1982, 1986a,b) presents three isomorphic representations of C_{12}, based on different sets of generators. Although these representations are motivated almost entirely from mathematical concerns, Balzano demonstrates how they also address musical concerns, including the importance of the diatonic scales and major/minor triads and the fundamental attributes of melodic, harmonic and key relations in Western music.

Considering the subgroups of C_{12} (see Table 5), (Balzano, 1982, p. 333) notes that each subgroup contains generators which are of a certain period. For example, the group generator of C_6 ($\{N^0,N^2,N^4,N^6,N^8,N^{10}\}$) is N^2 (or its inverse, N^{10}) which is of period six since $(N^2)^6 = N^2$. Similarly, C_4 is generated by N^3 (or its inverse, N^9) which is of period four, C_3 is generated by N^4 (or its inverse, N^8) which is of period three and C_2 is generated by N^6 which is of period two. The only elements that do not appear in any of the subgroups of C_{12} are N^1 and N^7 (and their inverses). This is because these elements are of period 12 and generate the whole of C_{12}. If a subgroup of C_{12} did contain N^7 or N^1 it would also have to contain all the powers of those elements in order to satisfy closure and, thereby, would no longer be a proper subgroup.

Furthermore, Balzano (1982) notes that the representation of C_{12} generated by N^1 (see Figure 1) provides a map of besideness relations for the scale $\{0,1,2,3,4,5,6\}$, while the representation of C_{12} generated by N^7 (see Figure 2) provides a map of besideness relations for the diatonic scale $\{0,2,4,5,7,9,11\}$ (see §5.4). While no proper subgroup of C_{12} can contain the elements N^1 and N^7, it also follows that

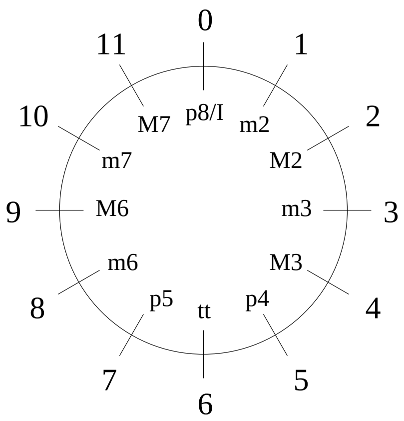

Figure 1: Semitone space: the first isomorphism of C_{12} generated by the semitone (m2, 1).

elements that are contained in a proper subgroup of C_{12} are incapable of generating the whole of C_{12}. Since N^1 and N^7 are the only elements not found in any of the subgroups of C_{12} it can perhaps be appreciated that these (and their inverses) are the only group generators of C_{12}. We consider the isomorphic representation of C_{12} generated by N^1 in §6.2 and the isomorphic representation of C_{12} generated by N^7 in §6.3 while in §6.4, we discuss a more complicated isomorphism of C_{12} based on a product group.

6.2 Semitone Space

The first representation of C_{12} is that generated by the (equal tempered) semitone or minor second interval: $(N^1)^{12}$ (see Figure 1). Note that $(N^1 1)^{12}$ is the result of running the cycle in the reverse order. Balzano calls this representation "semitone space" since closeness of pitch points in this space is a function of the number of semitones separating them. He notes that "when we refer to a melodic motion as 'small', I believe it is closely related to the sense in which distances in this space may also be 'small'." (Balzano, 1980, p. 69).

6.3 Fifths Space

The only other representation of C_{12} based on a single generator that is isomorphic to semitone space is that generated by the perfect fifth: $(N^7)^{12}$ (and its inverse $(N^5)^{12}$). Balzano calls this representation (shown in Figure 2) "fifths space". This is, of course, the familiar cycle of fifths which succinctly describes the relationships (e.g., "close" and "distant") between different key signatures (see Table 8). The isomorphic relationship between semitone space (A) and fifths space (B) is as follows:

$$A \longleftrightarrow B : i(mod 12) \longleftrightarrow 7i(mod 12)$$

The relation is one-to-one and structure preserving: although proximity relations among the elements have changed, any true statement about elements in system A is also true of their images in system B. To briefly illustrate, in the semitone group m2 $*$ M2 = m3 and the isomorphism ensures that in the fifths group p5 $*$ M2 = M6, which can be verified in Figure 2.

Balzano (1982) notes that while both of these isomorphisms are on an equal footing in the sense that neither is logically prior to the other, it could be argued that they are not on an equal *perceptual* footing: the closeness of two pitches corresponds more to the sense of closeness in semitone space than fifths space. The diatonic scale provides the additional constraint required to provide the proximity relations in fifths with perceptual information. Any fully connected region of seven points in fifths space corresponds to a diatonic scale and there are 12 such regions corresponding to the 12 transpositions of the scale (see Figure 3). Transforming the scale by a perfect fifth corresponds to a minimal rotation of the region in fifths space such that all but one element is common to any two adjacent scales. We have seen how

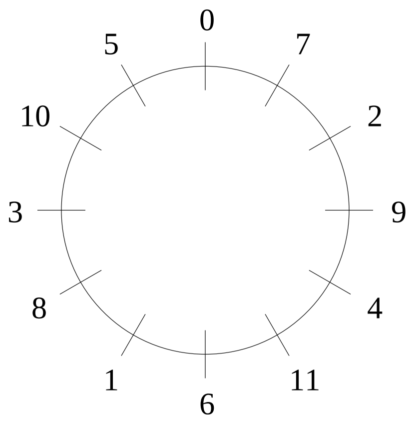

Figure 2: Fifths space: the second isomorphism of C_{12} generated by the perfect fifth (p5, 7).

unusual this property is (see §5.4) and it can perhaps be appreciated that only scales representable as regions in semitone or fifths space satisfy the property of transpositional simplicity.

Group element	No. of Sharps/Flats	Sharpened/Flattened Notes	Major Key	Relative Minor
0	No sharps/flats		C major	A minor
7	1 sharp	(F)	G major	E minor
2	2 sharps	(F C)	D major	B minor
9	3 sharps	(F C G)	A major	F♯ minor
4	4 sharps	(F C G D)	E major	C♯ minor
11	5 sharps	(F C G D A)	B major	G♯ minor
	7 flats	(B E A D G C F)	C♭ major	A♭ minor
6	6 sharps	(F C G D A E)	F♯ major	D♯ minor
	6 flats	(B E A D G C)	G♭ major	E♭ minor
1	7 sharps	(F C G D A E B)	C♯ major	A♯ minor
	5 flats	(B E A D G)	D♭ major	B♭ minor
8	4 flats	(B E A D)	A♭ major	F minor
3	3 flats	(B E A)	E♭ major	C minor
10	2 flats	(B E)	B♭ major	G minor
5	1 flat	(B)	F major	D minor

Table 8: Pattern of key signatures.

Balzano (1980, p. 71) uses the following features to differentiate the many scales that do form compact regions in semitone or fifths space:

1. scales with less than seven elements do not possess a sufficient number of intervals to contain representatives of all 12 intervals;

2. scales with more than seven elements show deteriorating overlap-distance relations: e.g., every pair of non-adjacent size-10 scales shares 8 of 10 elements while adjacent pairs share 9;

3. many of these scales, including all connected regions in semitone space and the connected region of size-6 in fifths space (Guido's hexachord), fail to satisfy coherence;

4. only the 5- and 7-note scales (i.e., the diatonic and pentatonic scales) lead to the following behaviour: transposing the scale by a p5 not only leads to a scale differing by just one element, but

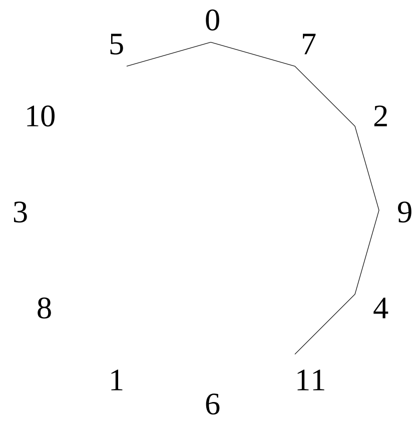

Figure 3: The diatonic scale as a fully connected region in fifths space.

also, since $(p5)^5 = (m2)^{-1}$ and $(p5)^7 = m2$, the changed element has undergone a minimal change in the sense given by semitone space and C_∞.

This latter property underlies the proximity relations between keys shown in Table 8. It also underlies the very possibility of key signatures which "were developed, it seems, with the diatonic scales specifically in mind" (Balzano, 1980, p. 70).

6.4 Thirds Space

Both of the previous isomorphisms have been one-dimensional in the sense that they were produced by a single generator. Balzano (1980, 1982) demonstrates that another isomorphism of C_{12} is generated by the direct product group of two of its subgroups C_3 and C_4. $C_4 \times C_3$ consist of the set $\{(0,0),(0,1),(0,2),(0,3),(1,0),(1,1),\ldots,(2,2),(2,3)\}$ and the operation:

$$(a,b) * (a',b') = ([a+a']mod\ 3, [b+b']mod\ 4)$$

The isomorphic relationship between the C_{12} structure shared by the first two representations and $C_4 \times C_3$ is as follows (see also Table 9):

$$C_4 \times C_3 \longleftrightarrow C_{12} : (a,b) \longleftrightarrow ([4a+3b]mod\ 12)$$

It can easily be verified that the 2-tuples of $C_4 \times C_3$ play analogous structural roles to their counterparts in C_{12}. For example, $(1,1)^3 = (0,3)$ and $7^3 = 9$ and 9 is the image of $(0,3)$.

$C_4 \times C_3$	(0,0)	(0,1)	(0,2)	(0,3)	(1,0)	(1,1)	(1,2)	(1,3)	(2,0)	(2,1)	(2,2)	(2,3)
C_{12}	0	3	6	9	4	7	10	1	8	11	2	5

Table 9: The isomorphism between $C_4 \times C_3$ and C_{12}.

Balzano (1980, 1982) gives the following interpretation of this isomorphism. In terms of musical intervals, there is one axis generated by major thirds (4_{12} generates C_3) and another generated by minor thirds (3_{12} generates C_4) and each interval is a point corresponding to the number of major and/or minor thirds contained in that interval. For example, a p5 may be broken down into a M3 and a m3 and it corresponds to $(1,1)$ in $C_4 \times C_3$. Balzano calls this space "thirds space" and notes that it should properly be represented on a torus which has been cut, unrolled and duplicated in Figure 4. To facilitate comparison with the other isomorphisms, the points of $C_4 \times C_3$ have been labelled with their C_{12} images rather than the 2-tuples.

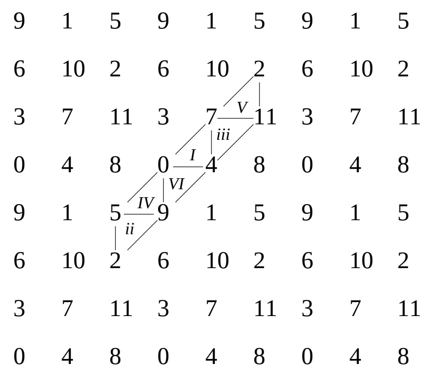

Figure 4: Thirds space: generated by the product group $C_4 \times C_3$. The horizontal axis is in major third units (4 *mod* 12) and the vertical axis is in minor third units (3 *mod* 12). The parallelogram represents the region of the diatonic scale.

In fifths space, the simplest or maximally compact shapes were intervals of the fifth itself and chains of such intervals. Correspondingly, in thirds space, we find chains of major and minor thirds. For example, chains of length three correspond to augmented triads ($\{0,4,8\}$) in the major third axis and diminished triads ($\{0,3,6\}$) in the minor thirds axis. But there also exist simple and compact 2-D shapes; the simplest of these is the unit triangle with a major and minor third constituting the perpendicular sides. If both sides are traced out in a positive going direction, there are two possible kinds of triangle representable as $\{0,4,7\}$ and $\{0,3,7\}$, the major and minor triad, respectively. Furthermore, we can chain these higher order structures together. For example, when we sweep out a region of thirds space by adjoining positive-going triangles together, we reach the point from which we began after obtaining six triangles interconnecting seven pitch places; the pattern of those seven pitch places (forming a convex, space-filling region of thirds space) is the diatonic scale.[3]

Other scales do not fare so well in attaining a simple structure in thirds space. The pentatonic scale, for example, does not form a compact structure in thirds space: it contains five notes but only two triads. The other connected-fifth scales ($m \neq 7$) also fail to form compact, space-filling shapes in thirds space. The semitone space diatonic analogue, $\{0,1,2,3,4,5,6\}$ can be found by altering the handedness of the space and constructing triangles with one positive-going and one negative-going side, which may be of the form $\{0,1,4\}$ or $\{0,3,4\}$. (Balzano, 1982, p. 337) notes that like the parent scale these pitch sets have seen virtually no musical usage. Finally, (Balzano, 1980, p. 74) notes that the family of lines corresponding to $y = x + c$ in thirds space are unwrapped cycles of fifths and the family of lines corresponding to $y = -x + c$ in thirds space are all cycles of semitones. These cycles require several circuits around the torus of thirds in order to close.

The concept of a tonic also has an interpretation in terms of thirds space (Balzano, 1980). While prior to the 17th Century there is little evidence that particular elements of the diatonic set were preferred as triads, with the development of harmony and the triadic basis of music, the diatonic set came to represent only two scales: the major and natural minor scales. The central notion was not so much one of a tonic as of a tonic triad. The diatonic set consists of three major and three minor triads (see Figure 4). The centrally located major and minor triads in the diatonic set are $\{0,4,7\}$ and $\{9,0,4\}$ which are none other than the tonic triads of the surviving major and natural minor modes of the diatonic set. Since neither of these triads, nor their roots, possess any privileged status in terms of fifths or semitone space, Balzano is:

> "tempted to speculate that only when music came to draw more fully upon the resources of the thirds representation did the related concepts of tonic and triad really come into full force." (Balzano, 1980, p. 74)

[3]The seventh triad is formed by the size-3 chain of minor thirds obtained by connecting the (11,2) side of the uppermost triangle with the (2,5) side of the lowermost triangle: it is a diminished triad belonging to the family $\{0,3,6\}$.

In any case, third-relatedness of triads in a given diatonic scale bears a striking resemblance to musical notions of harmonic closeness of chords. For example, it is sometimes said that a *ii* chord may be substituted for a *IV* chord in many contexts and these two triads are edge sharing sets in thirds space: $\{2,5,9\}$ and $\{5,9,0\}$.

$C_4 \times C_3$ is the only product group of subgroups of C_{12} that is isomorphic to C_{12}. Balzano (1982, p. 338) notes that the only other candidate, $C_6 \times C_2$, is not isomorphic to C_{12} for two reasons: first, there are no elements in the group having period 12 and therefore there are no group elements to correspond to semitones or fifths; and second, there are three elements of period two, so there are too many elements that potentially correspond to a tritone. Higher-order groups such as $C_2 \times C_2 \times C_3$ also fail to be isomorphic to C_{12}.

6.5 Summary

Balzano (1980, 1982) presents three isomorphic representations of C_{12}, based on the generators N^1 and N^7 and the product group $C_4 \times C_3$, which he calls semitone space, fifths space and thirds space respectively. These isomorphisms exhaust the structure of C_{12}: there are no other isomorphic representations based on single generators or product groups. Furthermore, each of these spaces has a natural interpretation in musical terms:

> "The space of semitones supplies a constraint that appears in even the simplest of music, guiding local note-to-note transitions and acting as the basic criterion for "smooth" or (more formally) "conjunct" melodic motion. Somewhat less locally, individual notes are constrained by membership in triads; the pitches of a melody may change although the underlying triad remains invariant. The triads themselves change more slowly in a piece of music, and here the space of thirds is the basis for triadic motion; in particular, major and minor triads that are third related (share an edge in thirds space) are often treated as substitutable for one another, a relation not shared by more distant pairs of triads. Diatonic scales serve as an even more global context of constraint for triads, which may change but still leave the underlying scale or key invariant. Analogously, scales may change in the course of a musical piece but will do so even more slowly than triads, and much more slowly than single notes. When they do, it is the space of fifths that provides the basis for "near" and "far" movement." (Balzano, 1986a, p. 222)

Underlying all these three levels of structure lies the constraints on pitch selection provided by the parent system C_{12}. We shall discuss the group-theoretic approach to microtonal pitch systems in §7.

Finally, it has been demonstrated that the diatonic scale family is strongly implicated in the structure of both fifths space and thirds space. Indeed it is unique in instantiating the higher-order relations of these spaces while still remaining coherent with respect to semitones.

7 Generalisation to *n*-fold Systems

Having described the important structural properties of C_{12}, Balzano (1980) considers the question of whether other C_ns also exhibit these properties. In particular, while all such systems have an analogous semitone space, he is concerned to find those which have two other groups isomorphic to this space, one of which is a single-generator cycle of keys and the other a product group of triads, such that a diatonic analogue emerges which is a connected subset in the cycle of keys, a connected structure in the product space and is coherent with respect to semitones.

It turns out that the only groups which exhibit these properties are of the following form:

$$C_n \cong C_k \times C_{k+1} \mid k \in Z$$

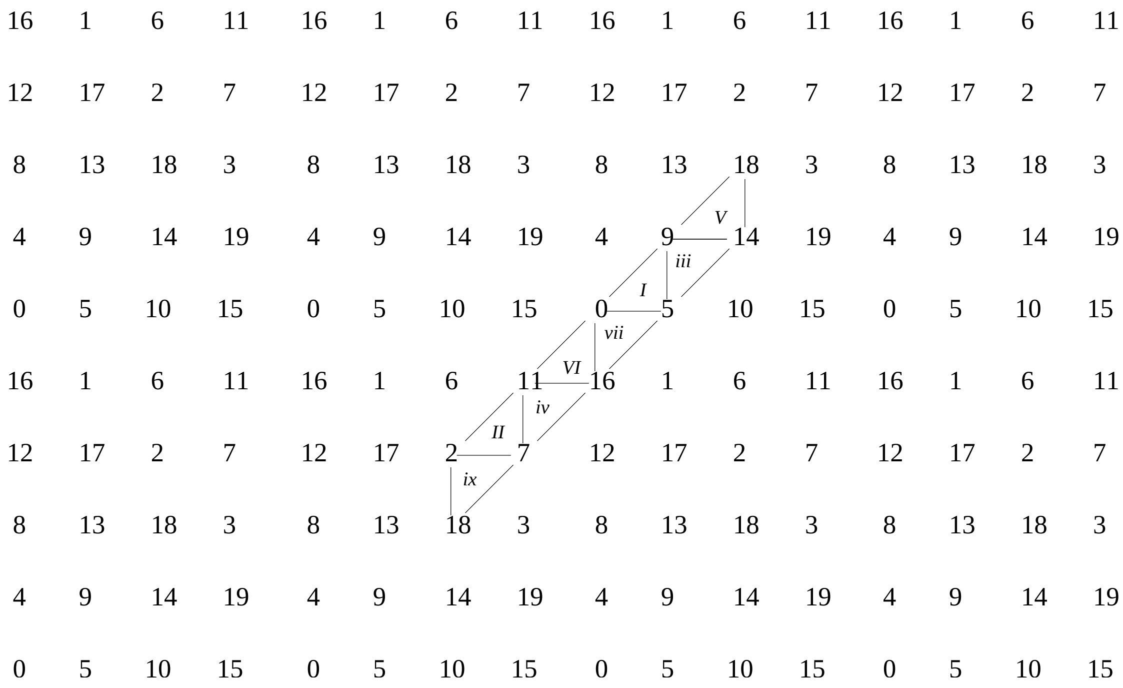

Figure 5: Thirds space for C_{20}.

and that this condition is both sufficient and necessary (Balzano, 1980, p. 74–75). Furthermore, Balzano demonstrates that C_{12} ($k = 3$) is the smallest group to satisfy the above conditions since with $k < 3$ there are not enough generators to produce a cycle of keys. When $k > 3$, however, there are increasingly many generators to choose from but in each case only one gives rise to a scale whose triadic structure is mirrored in the product space. That generator is the element $(2k+1)_n$, which behaves structurally like a p5, and the scale it generates has $2k+1$ scale degrees and shares the structural properties of the diatonic scale in C_{12}. Finally, the elements k_n and $(k+1)_n$ behave like minor and major thirds respectively and each diatonic scale analogue of size $2k+1$ have k major, k minor and 1 diminished triad.

Balzano (1980, p. 75–78) discusses C_{20} ($k = 4$) as an example of the above (see Figure 5). The analogues to fifths and semitone spaces can be read from the positive- and negative-going diagonal lines respectively. The diatonic scale analogue is generated by the element 9_{20}, consists of the set $\{0,2,5,7,9,11,14,16,18\}$ and consists of nine connected elements on 9_{20}. When the scale is transposed up a "p5" by the addition of 9 to each element, the result is a scale that differs minimally by just one semitone in one scale degree; 2 is replaced by 3 in the set $\{9,11,14,16,18,0,3,5,7\}$. It can be seen in Figure 5 that the diatonic scale analogue both a connected series of "fifths" and an overlapping series of "triads", of which four are major ($\{0,5,9\}$), four minor ($\{0,4,9\}$) and one diminished ($\{0,4,8\}$). Furthermore, the triads consist of "every other note" in the scale as expected.

The approach presented by Balzano (1980) predicts the existence of pitch systems with resources matching those of C_{12} with k corresponding to 30, 42, 56, 72, 90 and so on. Of these, he presents a brief discussion of C_{30} and C_{42}. While C_{20} contains the generators 1, 3, 7 and 9 (and their inverses), C_{30} not only contains several generators (1, 7, 11, 13 and their inverses) but there are three cyclic product groups of order 20 that are isomorphic to C_{30}: $C_5 \times C_6$, $C_2 \times C_{15}$ and $C_3 \times C_{10}$. Of these only $C_5 \times C_6$ in combination with the generator 11 mirrors the structural properties of C_{12}. Similar conclusions hold for C_{42}.

Balzano concludes that "in light of how scarce diatonic-scale supporting pitch systems are, it is truly remarkable that we have come up with these musical materials without group-theoretic thinking" (Balzano, 1980, p. 78). While the special nature of the diatonic scale in C_{12} is usually taken to be a function of the simple ratios it contains, the approach summarised here presents a case for thinking of the diatonic scale as an example of "survival of the fittest" independently of ratio concerns. Finally, Balzano notes that when we compare the two families of candidates for microtonal systems, one based on algebraic (e.g., $N = \{12, 20, 30, 42, 56, 72, 90\}$) and the other based on acoustic (e.g, $N = \{9, 12, 31, 41\}$) concerns, the 12-fold system emerges as the only system appearing in both families.

8 Empirical Evidence

8.1 Overview

The psychophysical and ratio-based approaches focus on the perception of such phenomena as beats, harmonics and combination tones. (Balzano, 1982, p. 338) argues that human sensitivity to these phenomena is rather low:

- many theoretically possible combination tones cannot be heard at all and those that can are only audible under restricted frequency-amplitude conditions;
- detecting harmonics requires a mode of listening anathemic to the perception of a running musical context and even under ideal conditions humans are insensitive to harmonics beyond the sixth or seventh;
- the presence of beats, besides being hard to distinguish from vibrato, is unrelated to the perception of "in-tuneness".

The experiments to be described, on the other hand, concern the perceptual sensitivity of human listeners to pitch set structure and the experiments are arranged to include: a preceding musical context; specifically musical intervals between the test tones; harmonically rich (and hence musical) tones; and the selection of listeners with some degree of musical background or ability (Shepard, 1982, p. 368).

8.2 Perception of Dynamic Qualities

Balzano (1982) first discusses the evidence that the perceptual judgements of human listeners demonstrate sensitivity to the pitch relations inherent in the structure of C_{12}. The first question to be asked concerns the perceived relation of scale degrees to the tonic. Under a strict psychoacoustic view, we might expect relatedness to be determined by tone height or, if we take in account octave-equivalence, tone chroma. Neither of these notions alone is sufficient.

Krumhansl & Shepard (1979) carried out an experiment in which subjects were played the first seven notes of an ascending or descending major scale, followed by a variable eighth tone which could be any one of the 13 pitches contained in octave between the first tone and the tone an octave above or below. The subjects were asked to judge how well the final note completed the sequence in each case. The results demonstrated that all subjects gave the highest rating to the tonic, its octave neighbour and the tones belonging to the scale and only the least musical subjects showed anything like an effect of pitch height. Multidimensional scaling of the data demonstrates that the results could be modelled by a combination of the distance from the tonic in semitone and fifths space (Shepard, 1982, p. 366).

In a related experiment, Krumhansl (1979) looked at judgements of perceived similarity among all possible pairs of pitch classes following a context-inducing diatonic scale or triad. The results for pitch pairs containing the tonic looked very similar to the results of Krumhansl & Shepard (1979). Multidimensional scaling of the entire matrix of rated similarities demonstrated that the most closely related pitch classes were those of the tonic triad ($\{0, 4, 7, 12\}$), followed by the remaining diatonic tones ($\{2, 5, 9, 11\}$), which were trailed by the remaining chromatic tones ($\{1, 3, 6, 8, 10\}$). Within each level of this configuration, perceived similarity appeared to be governed by proximity in semitone space. This research, therefore, demonstrates the perceptual importance of third relatedness (only hinted at in the data of Krumhansl & Shepard (1979).

Balzano (1982, p. 340) observes that it could be argued that the apparent relation between pitch classes in these experiments is a cognitive effect occurring during the process of forming a judgement and not a true perceptual effect. However, Balzano (1982) presents results which counter this hypothesis. Non-musical subjects taking a music course were presented with two octaves of an ascending major scale and were then presented with a tone belonging to one of the following subsets of the scale: $\{0,7\}$ or $\{0,11\}$. The task was to discriminate the tonic degree from the other degree and indicate the decision by pushing one of two buttons. Both correctness of response and latency were collected and the results demonstrated that the $\{0,11\}$ conditions were performed significantly faster ($p < 0.005$) and more accurately ($p < 0.001$) than the $\{0,7\}$ condition and these results were reliable over all 49 subjects and all five tones used.

Therefore, even on a speeded discrimination task, perceived similarity of scale degrees appears to be mediated by fifth relatedness rather than frequency separation. Furthermore, it cannot be argued that the result was due to "indoctrination" to music-theoretic beliefs since the same study repeated after three months of tuition revealed a reduction in the difference between the two conditions. Finally, regarding Uniqueness, the same set of students were given four melodies and asked to judge whether the melody was based on a pentatonic, diatonic major, harmonic minor, whole tone and chromatic scale. All but the last two of these scales satisfies uniqueness. The results revealed that the vast majority of identification errors involved whole-tone/chromatic or pentatonic/major/minor confusions. In general, scales satisfying uniqueness were hardly ever confused with scales failing to satisfy uniqueness.

In summary, these results suggest that listeners are sensitive to constraints on pitch relations defined by pitch-height, pitch class, fifth relatedness and third relatedness.

8.3 Scalestep-level Perception

Having argued that only coherent scales are conducive to learning scalestep-level properties of intervals (see §5.3), Balzano (1982) considers evidence that such learning does in fact occur. For example, **?** tested the recognition of intervals played at short durations by musicians. They found that most of the confusion errors involved intervals separated by a semitone, but of these, the overwhelming majority were to scalestep-equivalent intervals (e.g., m2 and M2). Killam *et al.* (1975) replicated and extended this finding to situation involving longer durations, sequential as well as harmonic intervals and non-expert subjects. These studies also demonstrated a significant trend towards confusing intervals with their inverses. Thus the p4 is confused with the p5 much more than either of these intervals is confused with the tritone and, similarly, seconds tend to be confused with sevenths, thirds with sixths and so on.

These findings were again replicated and extended by Balzano (1977a,b) who measure both latencies and errors in a slightly different experimental paradigm. Subjects were presented with a visual probe displaying the name of an semitone-level interval (i.e., m2 – p8) followed by a harmonic or melodic scalestep interval. The task was to decide if the probe and the stimulus were the same or different by pushing one of two buttons. The data demonstrated that, even when semitone differences between probe and stimulus were held constant, latencies were significantly longer and errors significantly more frequent when the probe and stimulus intervals were scalestep equivalent.

In a related experiment, Balzano (1977a) found that intervals may be recognised at the scalestep level *directly* without mediation through the semitone level. The experimental design was essentially the same except that a number of scalestep level visual probes (e.g., "third" or "seventh") were added. The scalestep level probes lead to responses that were significantly faster and more accurate than the semitone level probes. For a given interval, a major third for example, it was subjects were able to verify the interval faster and more accurately as the higher level category third than the lower level category major third.

Balzano & Liesch (1982) extended the experimental paradigm by using polyphonic music played by an orchestra played in a more natural listening environment. The results essentially replicated those described above: there were significantly more semitone-related confusions (particularly for melodic intervals) and scalestep-related confusions (particularly for harmonic intervals). However, there was no evidence that the p4 and p5 are perceived as more similar than the p4 or p5 and the tritone although intervals tended to be confused with their inverses. Furthermore, multidimensional scaling suggested that perceived interval space is best modelled by an interaction of scalestep equivalence and distance on a chroma circle than a linear scale of semitone distance.

Finally, addressing the question of whether scalestep-level properties are more perceptually salient than semitone-level properties, Dowling (1978) played subjects a standard melody followed by a comparison melody and asked whether they were the same melody or not. All melodies were based on diatonic scales and the comparison melodies fell into two groups: *same* melodies were exact transpositions of the standard melodies while *different* melodies were a scalestep preserving movement of the melody to a different point in the same scale. The results indicated that both musical and non-musical subjects had considerable difficulty in distinguishing the scalestep-equivalent non-transpositions from true transpositions.

In summary, scalestep level properties of pitch sets appear to be directly perceived by the listener in tandem with semitone level properties.

8.4 Key Relatedness

In §8.2, we reviewed evidence supporting the perceptual sensitivity of listeners to distances in fifths space and Balzano (1982) next considers the responses of listeners to pitch sets under transposition. In particular, given that a melody retains its identity under transposition he asks whether this relationship is a function of key relatedness. The evidence generally suggests a positive response.

Cuddy *et al.* (1979), for example, used a forced choice melody recognition task where a standard melody was followed by two comparison melodies, a target and a foil. While targets were exact transpositions of the standard, foils had a single note altered by a semitone. The transpositions were sometimes by fifth and sometimes by tritone and subjects were asked to choose the exactly transposed melody. The results demonstrated a small but consistent advantage for the fifth transpositions. In an extension of this work, Bartlett & Dowling (1980) employed a yes-no rather than a forced-choice paradigm and used foils which were melodies that had been both transposed and translated along the scale. These foils preserved scalestep-level relations but violated semitone-level relations. Both adults and children showed a reliable key-distance effect not on the targets but on the foils: foils were easier to recognise when they were transposed to a more distant key.

The next question asked by Balzano (1982) concerns the extent to which these results are a function of pitch set overlap. Using the same experimental paradigm as Cuddy *et al.* (1979) with either diatonic or non-tonal melodies (based on the scales $\{0,1,2,6,7,10\}$ and $\{0,2,4,6,8,10\}$), Cohen *et al.* (1977) found an interaction between the factors of pitch set and transposition. While the diatonic melodies produced an advantage for the p5 transformations, the non-tonal melodies showed a tritone advantage. Since both of the non-tonal scales used show greater pitch class overlap under N^6 than N^7, Balzano (1982) concludes that this evidence suggests that the ability to detect a correct transposition of a melody is a direct function of overlap (see §5.4).

The final set of experiments to be considered concern interval recognition by skilled musicians. Balzano (1977a) presented subjects with a base tone taken from the set E, A and C♯ ($\{0,4,7\}$) or the set G and C♯ ($\{0,6\}$). Since, for the former set, 2/3 of the intervals presented would be based on A and E which are close on the cycle of fifths, this would induce in the subjects a tonal context similar to that of A major. Therefore, recognition would be significantly worse for intervals based on C♯ which is more remote in fifths space. With the G-C♯ set, on the other hand, neither tone should show any perceptual advantage since the set exhibits a symmetrical relationship on the cycle of fifths. The results confirmed these expectations: for the E-A-C♯ basetones, intervals based on A and E were recognised significantly more rapidly and accurately than those based on C♯ while only one experiment out of three showed a significant advantage for intervals based on A over those based on E. There were no base-tone effects in the G-C♯ conditions.

In summary, these results indicate that perceptual distance between transposed melodies appear to be a function of pitch-set overlap. In the case of the diatonic scales this directly implicates the cycle of fifths which also seems to act as a default in the absence of appropriate pitch set constraints.

8.5 Group-Theoretic Constraints

Balzano (1986a, p. 218) notes that the group-theoretic approach predicts the sensitivity of listeners to two types of constraint on pitches: first, a *quantal* constraint which reduces the continuous frequency domain to a set of discrete (equally spaced) elements; and second, a *generative* constraint on the specific relationships among these values (there is a determinate generative relationship among the quantised values). He describes experiments which suggest that listeners are indeed sensitive to these constraints. Non-musical subjects were presented with computer generated pseudo-melodies and asked to give ratings of musicality or in-tuneness. The pseudo-melodies fell into three categories:

1. *low-constraint:* pitch values were continuously distributed throughout their range (no quantal or generative constraints);
2. *medium-constraint:* pitch values were quantised into an unequal (by up to .4 semitones) 12-fold division of the octave (quantal but not generative constraints);
3. *high-constraint:* pitch values were quantised into an equal log frequency division of the octave (quantal and generative constraints).

The results demonstrated a significant difference between the perceived musicality of highly constrained pseudo-melodies from medium-constraint pseudo-melodies but no significant difference between medium- and low-constraint pseudo-melodies. Thus quantal constraints appeared to be important while generative constraints did not contribute to the perceived musicality of the pseudo-melodies.

In a second experiment, involving diatonic scales, the following constraints were placed on the pseudo-melodies:

1. *low-constraint:* the pitches used were fixed deviations from two octave diatonic scale;
2. *medium-constraint:* the pitches used were diatonic but violated octave-equivalence by using one scale for the lower octave and another (transposed by a semitone) for the higher octave;
3. *high-constraint:* genuine two octave diatonic scales abiding by C_{12} constraints and octave equivalence.

The results showed the expected pattern with medium-constraint pseudo-melodies perceived as significantly more musical than low-constraint pseudo-melodies and high-constraint pseudo-melodies perceived as significantly more musical than medium constraint melodies. Balzano (1986a) argues that the difference between the low- and medium-constraint conditions demonstrates that even if octave equivalence is preserved, approximations to diatonic scales sound unmusical unless generative constraints are respected. The difference between the medium- and high-constraint conditions, on the other hand, demonstrates that octave equivalence (and its key defining property – the distance of a semitone is small in pitch height but large in fifths space) is also important (when generative constraints are assumed).

8.6 Summary

Balzano (1986a) takes a realist approach to understanding music perception: he is concerned with constraints present in the musical stimulus and how they are perceived by listeners rather than the contents of listeners representations of music. Under this view, music is viewed as a mode of expression subject to specific constraints on the global selection of pitches and music perception is viewed as the direct perception of those constraints (Balzano, 1986a, p. 217-218). His group-theoretic analysis suggests a number of such constraints and it seems that listeners are sensitive to these constraints. First, listeners

D♯	A♯	E♯	B♯	Fx	C x	G x
B	F♯	C♯	G♯	D♯	A♯	E♯
G	D	A	E	B	F♯	C♯
E♭	B♭	F	C	G	D	A
C♭	G♭	D♭	A♭	E♭	B♭	F
A♭♭	E♭♭	B♭♭	F♭	C♭	G♭	D♭

Figure 6: The space of pitch places (adapted from Longuet-Higgins, 1962a)

are sensitive to the similarities of pitches based on their differences in pitch class, fifth relatedness and third relatedness (see §8.2). Second, listeners appear to be sensitive to the constraints imposed by the scalestep-semitone coherence of the degrees of the diatonic scale (see §8.3). Third, listeners appear to be sensitive to the relatedness of keys based on distance in the cycle of fifths (see §8.4). Finally, listeners are sensitive to the basic constraints of quantisation of pitch, generative relations among pitch class elements and octave equivalence (see §8.5).

At a more general level, Balzano's account suggests the following constraints are placed on music:

- the 12-fold division of the octave (into pitch classes) is one of only a small number of such divisions which provide sufficient resources for the development of complex musical relations and is the only division that also satisfies psychoacoustic concerns;
- the diatonic scale is unique in satisfying Uniqueness and Coherence and underlies the properties of key-relatedness and harmonic distance in music.

Support for these constraints can be found in the cross-cultural prevalence of the 12-fold division and its presence in the most ancient tuning systems apparent in the archaeological records (Sloboda, 1985). Even when a 12-fold division of the octave is not used (as in the Indian system which postulates a theoretical 22-fold division of the octave):

- virtually all such scales are based on the octave whether or not it corresponds exactly to a 2:1 frequency ratio (Dowling, 1978; Sloboda, 1985);
- nearly all assign a central role to the perfect intervals: the fifth and its inversion (Dowling, 1978);
- most select a subset of either five or seven tones from each octave (Shepard, 1982);
- practice often diverges from theory in the direction of a 12-fold system (Sloboda, 1985).

These considerations provide support for the position that the constraints identified by Balzano are inherent in the nature of pitch systems rather than being imposed by the nature of human perception.

9 Related Approaches

9.1 Longuet-Higgins (1962a,b)

Longuet-Higgins (1962a,b) is concerned with the development of a formal identification of the nature of harmonic relations. His fundamental point is that a multidimensional space is required to describe such relations. Therefore, he describes a three dimensional space whose axes represent separation by fifths, major thirds and octaves respectively. Assuming octave equivalence, the space is suggests the pattern of pitch places shown in Figure 6. The space repeats itself in a South-Easterly direction and Longuet-Higgins (1962a) observes that the musician's notion of harmonic distance is very directly reflected by

Augmented Seventh	Augmented Fourth	Small Halftone	Augmented Fifth	Augmented Second	Augmented Sixth	Augmented Third
Imperfect Fifth	Minor Tone	Major Sixth	Major Third	Major Seventh	Tritone	Small Limma
Imperfect Third	Dominant Seventh	Perfect Fourth	Unison	Perfect Fifth	Major Second	Imperfect Sixth
False Octave	Minor Fifth	Minor Second	Minor Sixth	Minor Third	Minor Seventh	Imperfect Fourth
Diminshed Sixth	Diminished Third	Diminished Seventh	Diminished Fourth	Diminished Octave	Diminished Fifth	Great Limma

Figure 7: The space of intervals (adapted from Longuet-Higgins, 1962b)

a number of simple metrics in this space, of which summed "city-block" distance would be one and minimum spanning rectangle another. Thus, the diatonic scale forms a compact group such that a key may be defined as a region of this space. For example, the scale of C Major is boxed in Figure 6. Longuet-Higgins suggests that while listening to a piece of music we select a given region of the space thus orienting ourselves to a particular key. If the listener forces him to engage in large harmonic jumps in that region, it is abandoned in favour of another region in which the tones are more compactly represented, thus attributing a new key. Key relatedness is a function of overlap between the regions occupied by each key in the space.

Longuet-Higgins (1962b) applies his analysis to musical intervals by replacing the note names with their intervals from the tonic C (see Figure 7). Note that unlike the space of notes this space of intervals has no repeating pattern. This space of intervals has several interesting properties. First, to find the upward interval from a note X to a note Y in Figure 6, we simply have to superimpose the interval space over the note space such that "Unison" lies over note X. The square lying above the note Y then bears the name of the required interval. Second, the two intervals of a complementary pair (e.g., m3 and M6) are to be situated in diametrically opposite positions relative to the centre of the table ("Unison"). Third, the table can be used to calculate the sum of any two intervals by adding their respective displacements from the "Unison" square together. Fourth, the commonest and most primitive intervals lie near the centre of the space, whereas those near the edges are regarded as the most musically remote. Fifth, all intervals in common musical usage are to be found in the table.

Major key:	submediant subdominant	mediant tonic	leading note dominant	 supertonic
C major:	A F	E C	B G	 D
F major:	D B♭	A F	E C	 G

Table 10: The keys of C major and F major in the space

Longuet-Higgins (1962a) argues that it is wrong to equate enharmonic notes in the space and that the appearance of repetition in the space is illusory: it is just that we give different notes the same symbol. He illustrates the point by posing the question: How many notes are there in common between the keys of C major and F major? The answer he suggests is five and not six (see Table 10). The notes A, E, F, C and G are evidently common to both keys but the D in each key is different: it corresponds to a

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

Figure 8: The space of equal temperament (adapted from Steedman, 1994)

perfect fifth below A in F major and a perfect fifth above G in C major. Thus the supertonic of a key (D in C major) has a quite different musical definition from the submediant of the subdominant key (D in F major). Longuet-Higgins (1962a,b) backs up his argument with an analysis of several musical examples. In spite of these reservations, if we equate enharmonic intervals by assigning each note to one of the twelve equal-tempered pitch classes we obtain the representation shown in Figure 8. Note that the diagonals of this space yield semitone and minor third cycles – in fact, this space is equivalent to Balzano's thirds space rotated so that the minor third cycle forms takes the place of the p5 axis.

Steedman (1994) applies the theory of Longuet-Higgins (1962a,b) to a musical phenomenon that provides difficulties for the psychoacoustic approach. The distinctive characters of an augmented triad ($\{0,4,8\}$) or a diminished seventh chord ($\{0,3,6,9\}$), compared to major/minor triads, cannot be explained in terms of consonance or dissonance since all four chords are made up of equal-tempered minor and major thirds. In terms of Longuet-Higgins's theory, the listener when presented with an equal-tempered chord projects each ambiguous equal-tempered onto all possible interpretations onto a portion of the space defined by the traditional interval names. With the major and minor triads, one interpretation can be chosen that leads to the simplest (most compact) configuration where all intervals between pairs of notes in the triad are major/minor thirds or perfect fifths (or their inverses). Figure 9(a) demonstrates this for a major triad.

However, the augmented chord does not share these properties: first, all ways of selecting a single interpretation for all three notes forces one of the equally-tempered major thirds to be interpreted as a more remote augmented or diminished interval; and second, all interpretations are similarly compact – there is no way of choosing between them (see Figure 9(b)). Steedman (1994) observes that the ambiguity remains until we hear the next chord which provides resolution. For example, an augmented C chord might be interpreted as $\{C,E,G\sharp\}$ or $\{C,E,A\flat\}$ and if the following chord is an F major triad then the first interpretation is chosen. This resolution is strongly influenced by progressions of a semitone between the first and second chords such that the resolution in question is reinforced by the addition of a dominant seventh to the augmented chord, while the alternative resolution to a D♭ major triad is not

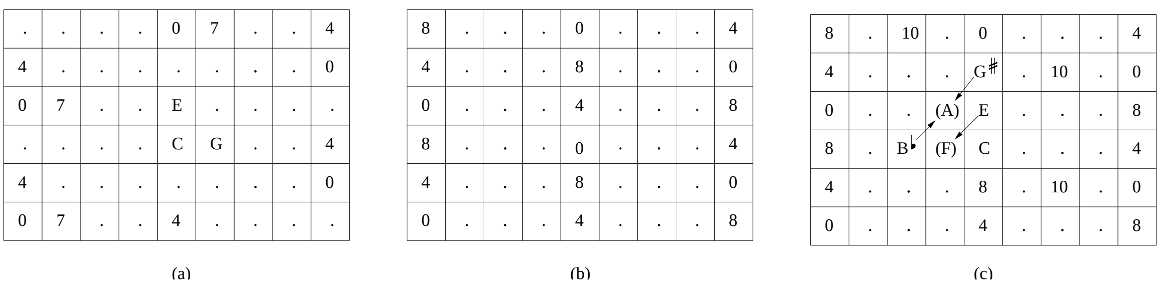

Figure 9: (a) The projection of an equally tempered chord of C Major; (b) The projection of an equally tempered augmented chord; (c) the projection of an equally tempered augmented seventh chord. (adapted from Steedman, 1994)

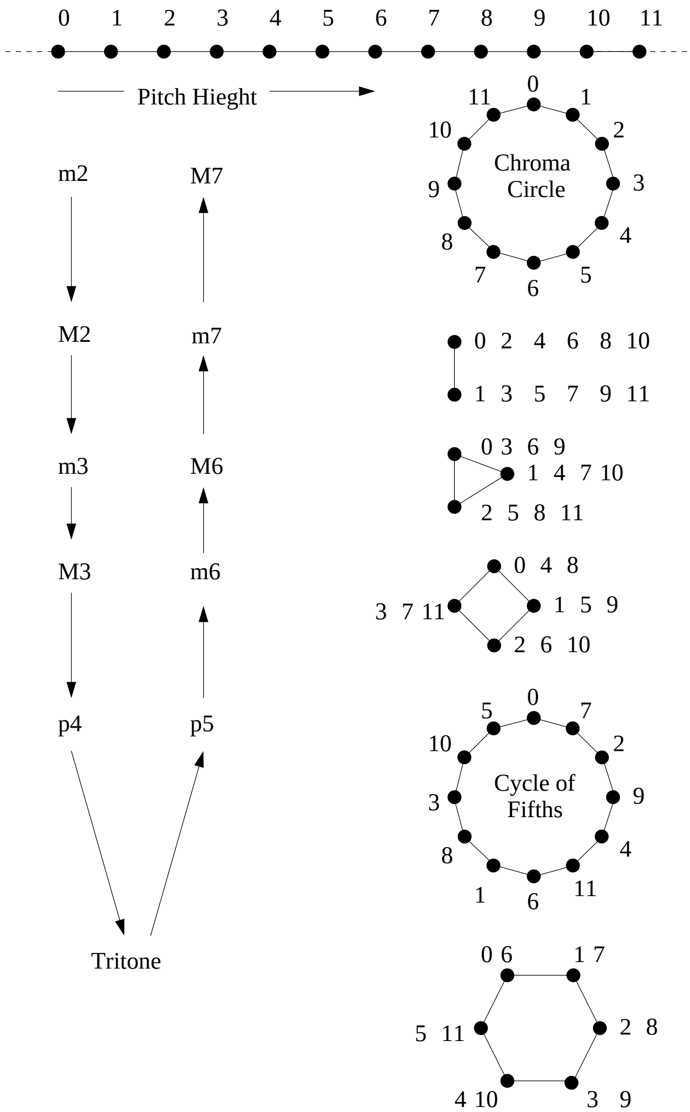

Figure 10: Seven structural components for the synthesis of helical structures for pitch (adapted from Shepard, 1982)

convincing (see Figure 9(c)). All of these findings also hold for the diminishes seventh chord.

9.2 Shepard (1982)

The model of pitch presented by Shepard (1982) is essentially geometrical in nature and uses the basic form of a helix, which has the attractive property that it can be continuously translated onto itself. It therefore facilitates the representation both of equivalence classes and of transposition of pitches.[4] Furthermore, complex helical structures embedded in multidimensional spaces may be composed out of simpler (circular and rectilinear) structural components. Shepard (1982, p. 359–361) suggests that for each musical interval that might characterise a perceptual similarity not adequately provided by pitch height, there should be a corresponding structural component in which the distance between all tones separated by that distance should be as small as possible relative to the other distances in that component (see Figure 10). Note that in order to maintain octave equivalence, the chroma circle and cycle of fifths are used to represent semitone and fifth relatedness (rather than equivalence).

[4]The use of such a geometrical representation implies certain assumptions (Shepard, 1982, p. 351–356).

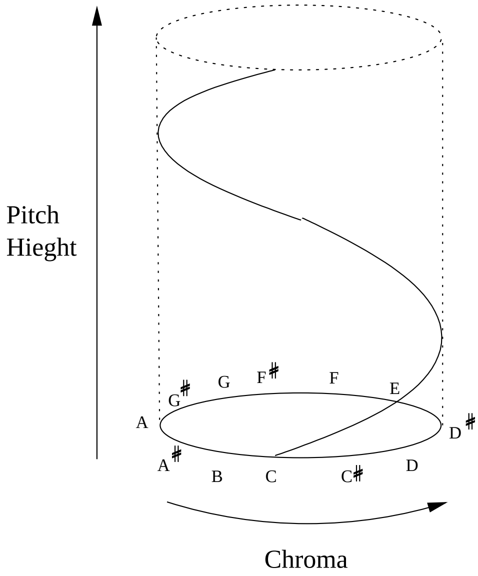

Figure 11: The simple helix of pitch class (adapted from Shepard, 1982)

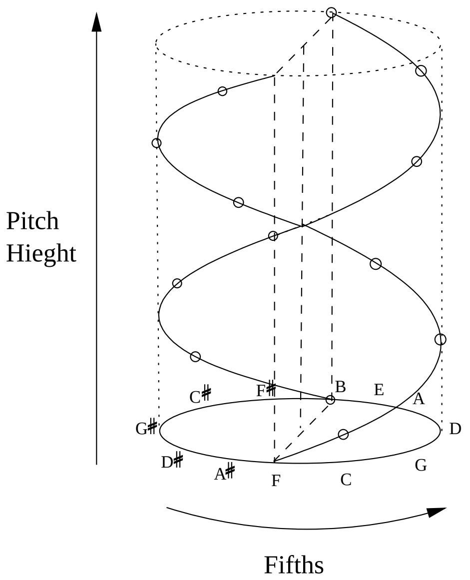

Figure 12: The three-dimensional double helix of musical pitch obtained by combining the cycle of fifths with pitch height (adapted from Shepard, 1982)

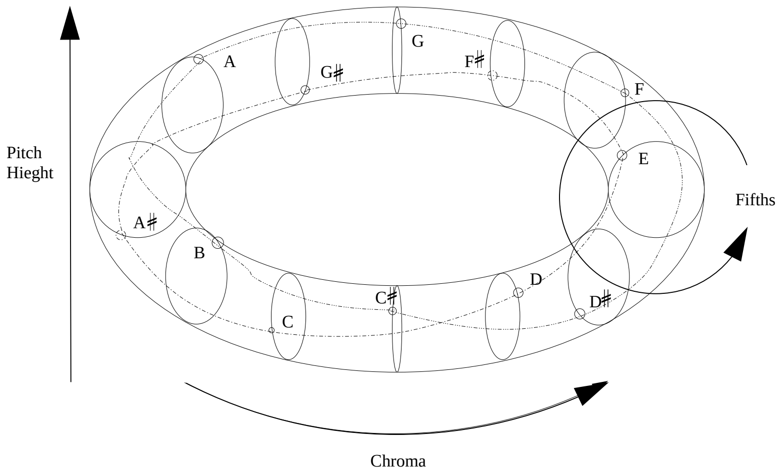

Figure 13: The double helix wrapped around a torus in four dimensions obtained by combining the cycle of fifths with pitch class (adapted from Shepard, 1982)

More complex helical structures may be formed through the weighted combination of these basic components. The combination of the chroma circle with pitch height, for example, results in the simple helical structure shown in Figure 11. Any pattern of intervals on the helix can be transposed by rotation of the spiral into itself, thus effecting a change of key. The combination of pitch height and the cycle of fifths, on the other hand, produces the double helical structure shown in Figure 12. There are several noteworthy features of this representation: first, the notes of any diatonic key can be divided from the notes not in that key by passing a plane through the central axis of the double helix (illustrated for the key of C major in Figure 12); and second, transposition into the most closely related keys is achieved by the smallest angles of rotation of the dividing plane about the central axis. The final structure considered by Shepard (1982) results from the combination of pitch height, the chroma circle and the cycle of fifths (see Figure 13). He argues that an appropriately weighted combination of these components can account for the perceived similarity of pitches which are close in pitch height and the heightened similarity of pitches at the octave and p5 (see § 8). He suggests, furthermore, that the heightened similarity of tones separates by major and minor thirds can be accounted for by forming appropriately weighted combinations of these elements with the others.

In Figures 12 and 13, the double helix is embedded in a two dimensional surface (the surface of a cylinder and a torus respectively) which can be cut and unfolded to give a two dimensional map (see Figure 14). The axes of this "melodic map" corresponds to intervals of a m2 and a M2 and it repeats itself indefinitely. Thus sequences in one dimension correspond to a chromatic scale while those in the other correspond to a whole tone scale. Figure 14 also demonstrates that the pitches in a particular diatonic key are to be found within a compact region of this space and the most closely related keys are obtained by the smallest horizontal shifts of this space. The characteristic zig-zag pattern of the scale results from shifts from one helix to the other (at m2 intervals). Furthermore, the relative minor and church modes are also represented by this pattern depending on which tone in the set is designated the tonic.

In accordance with Balzano (1982), Shepard notes that the irregularity associated with the diatonic scales enables each element to have a unique set of relations with the other elements thus allowing the listener to maintain orientation to a tonal centre. He calls the space a melodic map because the use of small intervals in sequential melodic sequences appears to be important for the perceptual integration of successive tones (cf Balzano's semitone space).

While tones differing by minor or major seconds often follow each other in melodic sequences, they are dissonant if sounded together. While melodies form compact representation in Figure 14, chords do not since they are based on intervals of major and minor thirds. However, because major and minor triads together form parallelograms in the melodic map that do not include any other tones (e.g., $\{C, E, G\}$ and

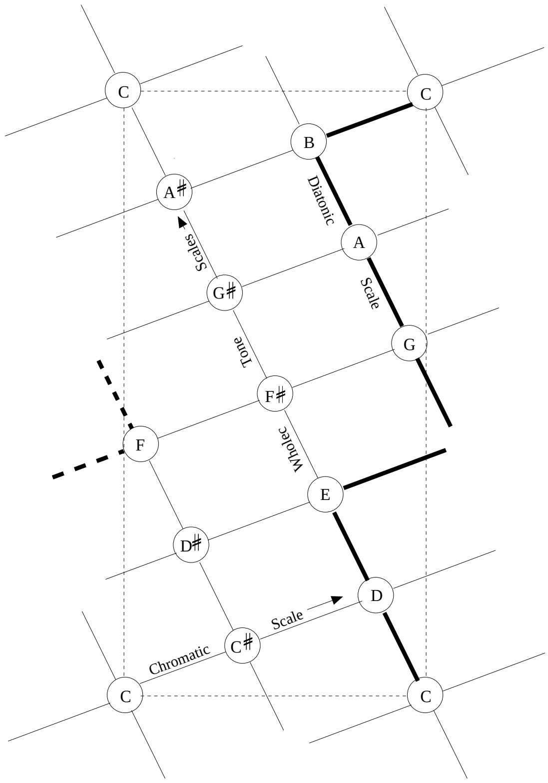

Figure 14: The two-dimensional melodic map obtained by unwrapping the double helical structure of Figures 12 and 13 (adapted from Shepard, 1982)

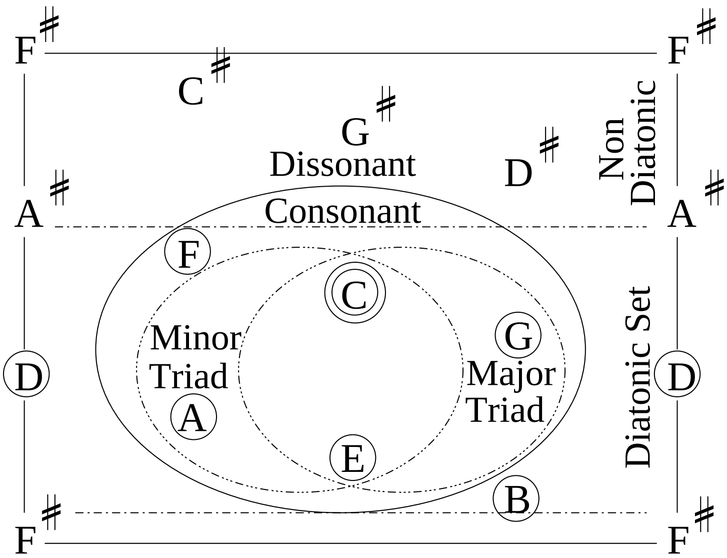

Figure 15: The two-dimensional harmonic map obtained by an affine transformation of the melodic map (adapted from Shepard, 1982)

{C, D♯/E♭, E} starting from the south-westerly C in Figure 14), this space can be transformed into one in which major and minor thirds do form compact groups. The transformation corresponds to a linear compression of the map along the C-G axis of the parallelogram and results in the so-called "harmonic map" shown in Figure 15. The set of tones consonant with a particular tonic do now form a compact group as do the common major, minor, diminished and augmented chords.

It can be appreciated from Figure 15 that the vertical axis represents intervals of a M3, while the South-Easterly going lines form sequences of fifths and the North-Easterly going lines form sequences of minor thirds. This space is, therefore, essentially equivalent to that proposed by Balzano (1980) and Longuet-Higgins (1962a,b). In an experiment related to that of Krumhansl & Shepard (1979), Krumhansl & Kessler (1982) extended the probe tone technique to the perception of chord proximity in different harmonic contexts. Multidimensional scaling of the data resulted in a configuration bearing a striking resemblance to this space. The cognitive-structural approach initiated by Shepard has been extensively refined and developed by Krumhansl and her colleagues (see Krumhansl, 1990) and has lead to one of the most influential research programmes in music cognition over the last 20 years (Cross, 1997).

10 Conclusion

By way of conclusion, let us examine the advantages and disadvantages offered by the group-theoretic approach in a more global context. The analysis presented by Balzano provides an integrated account of how melodic, harmonic and key relations as well as the diatonic scale are embedded in the structure of a 12-fold pitch system (among others) and it suggests reasons why these musical structures are of perceptual importance. These phenomena are difficult to interpret in an integrated and comprehensive manner at the acoustic level of description (Balzano, 1982, p. 348). However, the psychophysical and cognitive-structural approaches need not be mutually exclusive: a fundamental aspect of the perception of musical pitch concerns the relationship between the physical signal and its mental depiction (Cross, 1997). Just as the psychophysical approach captures what is physically different about, for example, different temperaments, so the algebraic approach accounts for what is similar about them at a higher level of description.

Equally there seem to be mutually compatible accounts that work at higher levels of description: such as that of highly overlearned structures based on pitch alphabets (Deutsch, 1982). To take one notable example, Lerdahl (1988) presents a representation of tonal pitch space which assumes the presence of five levels of structure (see Figure 16). The space is hierarchically organised with octave relationships at the highest level followed by the fifth, the diatonic triad, the diatonic scale and the chromatic scale. Noting that this space avoids the excessive symmetry in the models of Shepard and Balzano, he uses the space to express pitch-class relations, chord-proximity within and between tonal regions, and to regional proximity. In general, Lerdahl's model deals better with harmonic motion within and between keys than does Shepard's model (Dowling, 1991). Although Lerdahl's approach is primarily music-theoretic, there is experimental evidence that his hierarchical approach corresponds to the cognitive structures employed in music perception (see Krumhansl, 1990, p. 119–134).

Level *a*:	C												(C)
Level *b*:	C							G					(C)
Level *c*:	C				E			G					(C)
Level *D*:	C		D		E	F		G		A		B	(C)
Level *F*:	C	D♭	D	E♭	E	F	F♯	G	A♭	A	B♭	B	(C)

Figure 16: Tonal Pitch Space (adapted from Lerdahl, 1988)

At an even higher level of abstraction these theories of pitch space will constitute significant components in global accounts of the complexities of musical experience such as those of Lerdahl & Jackendoff (1983) and Narmour (1991).

Given that other approaches have resulted in descriptions of pitch systems that are essentially identical to those presented by Balzano, what is the importance of his analysis? In my opinion, the importance of the approach lies in the exposition of the intrinsic resources of a pitch system which are embedded in its simple algebraic structure. It therefore provides an integrated perceptual rationale for the basic structures assumed in higher level descriptions of pitch perception. Therefore, while it allows a complementary description of pitch at a higher level than the acoustic system, it also provides the foundations for examining music perception at higher levels of analysis. Many of the basic structures on which these higher level theories are based are grounded in the structure of C_n as a mathematical group. Balzano's work, therefore, has lead to new ways of thinking about pitch and he concludes that "it may even turn out that the most fascinating properties of human pitch perception still remain to be discovered." (Balzano, 1982, p. 348).

References

Balzano, G. J. (1977a). *Chronometric Studies of the Musical Interval Sense*. Doctoral dissertation, Stanford University.

Balzano, G. J. (1977b). On the bases of similarity of musical intervals. *Journal of the Acoustical Society of America*, *61*.

Balzano, G. J. (1980). The group–theoretic description of 12–fold and microtonal pitch systems. *Computer Music Journal*, *4*(4), 66–84.

Balzano, G. J. (1982). The pitch set as a level of description for studying musical pitch perception. In M. Clynes (Ed.), *Music, Mind and Brain* (pp. 321–351). New York: Plenum.

Balzano, G. J. (1986a). Music perception as detection of pitch–time constraints. In V. McCabe & G. J. Balzano (Eds.), *Event Cognition: An Ecological Perspective* (pp. 217–233). Hillsdale, NJ: Erlbaum.

Balzano, G. J. (1986b). What are musical pitch and timbre? *Music Perception*, *3*(3), 297–314.

Balzano, G. J. & Liesch, B. W. (1982). The role of chroma and scalestep in the recognition of musical intervals in and out of context. *Psychomusicology*, *2*(2), 3–31.

Bartlett, J. C. & Dowling, W. J. (1980). Recognition of transposed melodies: A key-distance effect in developmental perspective. *Journal of Experimental Psychology: Human Perception and Performance*, *6*, 501–515.

Cohen, A. J., Cuddy, L. L., & Mewhort, D. J. K. (1977). Recognition of transposed tone sequences. *Journal of the Acoustical Society of America*, *61*, 87–88.

Cross, I. (1997). Pitch schemata. In I. Deliège & J. Sloboda (Eds.), *Perception and Cognition of Music* (pp. 357–390). Hove, UK: Psychology Press.

Cuddy, L. L., Cohen, A. J., & Miller, J. (1979). Melody recognition: The experimental application of musical rules. *Canadian Journal of Psychology*, *33*, 148–157.

Deutsch, D. (1982). The processing of pitch combinations. In D. Deutsch (Ed.), *Psychology of Music* (pp. 271–316). New York: Academic Press.

Dowling, W. J. (1978). Scale and contour: Two components in a theory of memory for melodies. *Psychological Review*, *85*, 341–354.

Dowling, W. J. (1991). Pitch structure. In P. Howell, R. West, & I. Cross (Eds.), *Representing Musical Structure* (pp. 33–57). London: Academic Press.

Helmholtz, H. (1954/1877). *On the Sensations of Tone as a Physiological Basis for the Theory of Music* (Second ed.). New York: Dover. Translated by A. Ellis (originally published 1877).

Killam, R. N., Lorton, P. V., & Schubert, E. D. (1975). Interval recognition: Identification of harmonic and melodic intervals. *Journal of Music Theory*, *19*, 12–233.

Krumhansl, C. L. (1979). The psychological representation of musical pitch in a tonal context. *Cognitive Psychology*, *11*, 346–374.

Krumhansl, C. L. (1990). *Cognitive Foundations of Musical Pitch*. Oxford: Oxford University Press.

Krumhansl, C. L. & Kessler, E. J. (1982). Tracing the dynamic changes in perceived tonal organisation in a spatial representation of musical keys. *Psychological Review*, *89*(4), 334–368.

Krumhansl, C. L. & Shepard, R. N. (1979). Quantification of the hierarchy of tonal functions within a diatonic context. *Journal of Experimental Psychology: Human Perception and Performance*, *5*(4), 579–594.

Lerdahl, F. (1988). Tonal pitch space. *Music Perception*, *5*(3), 315–350.

Lerdahl, F. & Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. Cambridge, MA: MIT Press.

Longuet-Higgins, H. C. (1962a). Letter to a musical friend. *The Music Review*, *23*, 244–248.

Longuet-Higgins, H. C. (1962b). Second letter to a musical friend. *The Music Review*, *23*, 271–280.

Narmour, E. (1991). The top–down and bottom–up systems of musical implication: Building on Meyer's theory of emotional syntax. *Music Perception*, *9*(1), 1–26.

Shepard, R. N. (1982). Structural representations of musical pitch. In D. Deutsch (Ed.), *Psychology of Music* (pp. 343–390). New York: Academic Press.

Sloboda, J. (1985). *The Musical Mind: The Cognitive Psychology of Music*. Oxford: Oxford Science Press.

Steedman, M. (1994). The well–tempered computer. *Philosophical Transactions of the Royal Society, Series A, 349*, 115–131.