Math. Systems Theory 11, 353–372 (1978)

Mathematical Systems Theory

A Model for Pattern Perception with Musical Applications*
Part II: The Information Content of Pitch Structures

David Rothenberg

Department of Computer and Information Sciences, Speakman Hall, Temple University, Philadelphia, Pennsylvania 19122

Abstract. This is the second paper of a series which begins by treating the perception of pitch relations in musical contexts and the perception of timbre and speech. The first paper discusses in some detail those properties of musical scales required in order for them to function as "reference frames" which provide for the "measurement" of intervals such that ([1], p. 270), *Every melodic phrase, every chord, which can be executed at any pitch, can be also executed at any other pitch in such a way that we immediately perceive the characteristic marks of their similarity*. Here we continue this discussion by developing quantitative measures of the degree to which different scales possess the above properties. Then that property of musical scales which permits a listener to code the pitches of which it is constituted into "degrees" is examined and a corresponding quantitative measure developed. Musical scales are shown to be optimal choices with respect to both the former and latter measures, and a theory limiting those scales which are musically useful to a small fraction of possible sets of pitches is proposed. Existing scales which have been examined fall within the theory, which links the techniques of composition which may be used (i.e., those which produce perceptible relations between musical segments) to the above properties of the scale structures. This paper is not self-contained—reading of the previous paper in this series is required.

*This research was supported in part by grants and contracts AF-AFOSR 881-65, AF 49(638)-1738 and AF-AFOSR 68-1596.

0025-5661/78/0011-0353$4.00

9. Stability and Coverage

This section is pertinent to developing measures of (a) the tendency of scales to change into other related scales during musical use and (b) the degree to which a scale can accommodate the use of non-scale tones (see Part I, section 8) without losing its identity as a musical "reference frame." Initially we make the observation that the elimination of points from a proper set P (see Part 1, Sec. 3) may result in another set, P', which is also proper. Methods can be derived for finding such expendable points. Of greater importance is the fact that the elimination of points from a proper set, P, with many ambiguous pairs in $P \times P$ may result in a proper set, P', with few such ambiguous pairs. Since P is interpreted as a reference frame which is used for classification, the scarcity of these ambiguous pairs is significant. When P is proper and periodic (with period n and $\bar{n} = wn$ (see Part 1, Sec. 6) the proportion of ambiguous pairs in $P \times P$ is given by:[1]

$$W = 2\ \mathrm{card}((ij))|\alpha_{ij} = \left(\inf_k \alpha_{i+1,k}\right)/\bar{n}(\bar{n}-1), \qquad i = 1,\ldots,n-2;$$
$$j = 1,\ldots,\bar{n}$$

and the quantity $\bar{S} = 1 - W$ is called the *stability* of P. (In Part 1, Sec. 6, Ex. 1, the "major" scale, $\bar{S} = .9524$). Explanation: there are $\binom{7}{2} = 21$ two-element subsets of the seven element set $[C, D, \ldots, B]$, and of these only one $(\{F, B\})$ corresponds to an ambiguous interval; hence $\bar{S} = 20/21 = .9524$.

Various alterations can transform a P with low stability to one with higher stability, such as dropping points, adding points, and altering the positions of points. The extent of change resulting from each such alteration is difficult to evaluate. In most musical cultures, the cardinality of each of the scales used is, for the most part, fixed. In cultures using instruments with timbres in which harmonic partials predominate, certain intervals (e.g. the perfect fifth) often appear to resist alteration. However, it is convenient to conceive of a "gradient" (G) between two distinct proper scales (P_1 and P_2) as the difference of the squares of their stabilities $(\bar{S}_2^2 - \bar{S}_1^2)$ divided by the amount of alteration (to be discussed) necessary to transform one into the other. The stabilities are squared (as an approximation) because a scale with $\bar{S} = .25$ has less "tendency" to change to one with $\bar{S}_2 = .5$ than a scale with $\bar{S} = .75$ has to change to one with $\bar{S} = 1$, even though the amounts of alteration in both cases are the same. This is because it is difficult to use a proper musical scale with low stability *as a proper scale* (also to be discussed later). The amount of alteration in many cases may be taken as

[1] "card" indicates the number of elements in the set enclosed by "()". The number of α_{ij} equal to elements of preceding rows is the same as the number α_{ij} which are equal to elements of succeeding rows. Hence the factor of 2 in the above equation.

proportional to the fraction of elements of P in a period (n) whose position must be substantially changed (i.e. changed to a point not in any $\underline{R}$) (see Part 1, Sec. 4) or which must be dropped or added in order to effect the transformation. If A/n denotes this fraction, in such cases[2] let

$$\textbf{9.1.}\quad G(P_1,P_2)=\frac{\bar{S}_2^2-\bar{S}_1^2}{A/n}$$

Although crude, this notion of gradient indicates that two musical scales (P_1 and P_1') with equal stability will demonstrate different tendencies to decompose or be otherwise altered when there exists another scale (P_2) with respect to which one possesses a high gradient and when there exists none such with respect to the other. A few examples will indicate the significance of different values of G. The tendency to replace tones of the A melodic minor scale (P_1) by those of the A natural minor scale (P_2) (see Part 1, Sec. 6, examples 3 and 1) is indicated by $G(P_1, P_2) = 1.389$. If $P_1 \equiv (\mathrm{C, D, D^\#, E, F^\#, G, G^\#, A^\#, B})$, $(\psi(P_1) = (2,1,1,2,1,1,2,1,1))$ and $P_2 \equiv (\mathrm{C, D, E, F^\#, G^\#, A^\#})$, $(\psi(P_2)=(2,2,2,2,2,2))$, three points in P_1 are dropped and $G(P_1,P_2)=1.313$ which is of the same order of magnitude as that of the previous example. The same value of G results when P_2 is the twelve-tone scale $(\psi(P_1)=(1,1,1,1,1,1,1,1,1,1,1,1))$. An illustration of a high gradient may be constructed by adding a point between p_1 and p_2 of the Chinese pentatonic scale of Example 6 so that for the resulting scale, P_1, $\psi(P_1)=(1,2,2,3,2,2)$. Then, if P_2 is the Chinese pentatonic, $\bar{S}_2=1$, $\bar{S}_1=.4667$ and $G(P_1,P_2)=3.911$. Also, if a tone is added to P_1 between P_4 and P_5 to produce a major scale (Example 1), an only slightly lower gradient will result. Observe that, although P_2 and the major scale have been much used, it is difficult to find examples using P_1.

The concept of stability and gradient is pertinent to the tendency of a scale to "disintegrate". We next recall that, when musical scales are used together with many auxiliary (ornamental) tones or alternative tones (substitutions), the extent to which $\bar{R}$ (and $\underline{R}$) cover S is significant. When P is proper and periodic (with period n) and when S is finite within each cycle, the proportion of coverage may be defined as the cardinality of $\bar{R}$ contained within a cycle divided by the cardinality of S contained within a cycle. Let P be periodic, S be continuous and a metric on S be given. Let $\bar{R}_k^u$ denote the upper bound of $\bar{R}_k$

[2] Note, however, that if *all* the elements of P_1 are altered, $A=n$ and $G(P_1P_2)=(\bar{S}_2^2-\bar{S}_1^2)$, which, in some cases, can approach one as n increases (when in the periodic case some (p_ip_{i+1}) are ambiguous). It is possible to avoid this condition by making the denominator in (1) "$A/(n-A)$" and/or by adding constant multipliers in the expression. However, other distortions would be introduced. Since no cases will be encountered where an increase in $\bar{S}$ will result only when all elements of P are altered, (1) should suffice for all practical musical applications.

and $\bar{R}_k^l$ denote the lower bound. Then the *coverage* (C) of P is given by

$$C=\frac{\sum_{k=1}^{n}\left(\bar{R}_k^u-\bar{R}_k^l\right)}{\left(P_{n+1}-P_1\right)}$$

$$\left[\equiv\frac{\sum_{k=1}^{n}\left(\theta_k^{+}+\theta_k^{-}\right)}{K}\quad\text{when conditions in Part 1, Section 6 apply}\right]$$

(K and θ are discussed in Part 1, Sec. 6)

Notice that scales in all examples given thus far except example 5 (the harmonic minor scale), have a coverage of 1 (S is covered) when S is the 12-tone equal temperament system. However, if S is the reals, these coverages decrease differently in each example. (When S is continuous, coverage differences between scales tend to be of the same sign as their stability differences.) Most examples in this paper are drawn from the familiar 12-tone equal temperament system. Here a few proper scales with more than five elements and with high stability exist (all of which are tabulated in Figure 1, Section 13) and most familiar ones have $\bar{R}$'s which cover S. Those proper sets with fewer than five elements tend (to Western ears) to be heard as "chords", that is subsets of another P, and derive their stability from this latter proper set. Hence gradients connecting chords are not significant and, because of the small number of proper scales (in 12-tone equal temperament), gradient and coverage are of minor significance relative to stability when evaluating these musical scales. (An exception will be noted later.) However, these measures assume greater significance when scales using microtones are considered. In the application of this model to the generation of new musical materials (the last papers in this series) gradient and coverage are primary considerations.

Note that stability only applies to proper scales. Also, it does not really measure the degree to which a motif at a given pitch of a scale may be identified with (i.e., recognized as composed of the same intervals as) a "modal transposition" of that motif to another pitch in the scale (i.e., a sequence). To properly accomplish this, we would have to consider each subset of the scale paired with each of its modal transpositions and determine the fraction of these pairs, which have corresponding component intervals with the same diatonic distances (i.e., f).[3] Note that such a measure would apply to improper scales as well, in the sense of indicating the degree to which a motif may be followed by a modal transposition of itself such that corresponding intervals have identical diatonic distances. For reasons of computational economy we simplify the measure suggested above to obtain a "measure" which, in the case of proper scales, in

[3]To be thorough, it would be necessary to weight stepwise intervals more than skips when the former are more likely to occur in motifs. This would be required in the definition of stability as well. However, for simplicity, we here avoid such refinements.

practice is almost always ordered similarly to stability (assuming that the scales are periodic) and which also applies to improper scales:

If P is thought of as a ruler, $\bar{S}$ gives the proportion of ambiguous measurements under the assumption that all markers on the ruler will be used as edges (endpoints) when measuring distances (the intervals in $P \times P$). If we avoid using certain markers as endpoints (eliminate all such measurements from those points) all ambiguous and contradictory intervals may be avoided. Hence if the remaining set of markers are used as endpoints, all resulting measurements (all measurements from these endpoints) will be in an order consistent with the initial ordering. Here we are interested in all sets of such measurements from selected endpoints with such a consistent ordering.

For each element P there exists a set of intervals between that element and each of the others in P. Since $\delta_{ij} = p_{i+j} - p_j$ and $\|\alpha_{ij}\|$ preserves the order of $\|\delta_{ij}\|$, the j^{th} column of $\|\alpha_{ij}\|$ specifies the ordering of all the intervals (p_j, p_k), $p_k \in P$. The row position, i, of each of such intervals specifies its diatonic distance, i.e. $f(p_{i+j} - p_j) = i$. (Note that diatonic distance corresponds to the "measurements" referred to in the previous paragraph.) Hence the comparison of entries in two different rows of $\|\alpha_{ij}\|$ is in effect a comparison of the ordering of the measures of two intervals with their initial ordering (Part 1, Sec. 2). Let the set of diatonic distances corresponding to column j of $\|\alpha_{ij}\|$ be called the *endpoint set*, M_j. Then row by row comparison of the entries in two columns of $\|\alpha_{ij}\|$ is equivalent to the comparison of the ordering of the diatonic distances in two endpoint sets with the initial ordering of the corresponding intervals.

Here we wish to eliminate selected endpoint sets from comparison with the initial ordering. This is equivalent to deleting certain columns from $\|\alpha_{ij}\|$. When P is not strictly proper, certain such deletions will result in the elimination of ambiguous and contradictory intervals. In such a case the set of endpoint sets which have *not* been eliminated is called a *consistent set* and is specified by its component endpoint sets, $\{M_j, M_k, \ldots\}$. Thus $\{M_{k_1}, M_{k_2}, \ldots, M_{k_t}\}$ is a consistent set iff for all i and $j = k_1, k_2, \ldots, k_t$

$$\inf_j (\alpha_{i+1,j}) > \sup_j (\alpha_{ij}).$$

A consistent set which is a subset of no other consistent set is called a *maximum consistent set*. All subsets of a consistent set are clearly consistent.[4] (Examples: All the endpoint sets of a strictly proper set form a maximum consistent set. In Example 1, (Part 1, Sec. 6) maximum consistent sets are: $\{M_2, M_3, M_4, M_5, M_6, M_7\}$ and $\{M_1, M_2, M_3, M_5, M_6, M_7\}$.

For a given P let $c_k(P)$ equal the number of k-element consistent sets. Then the *consistency*, $\bar{C}$, of P is defined as

$$\bar{C} = \sum_{k=2}^{n} \left(c_k(P) \Big/ \binom{n}{k} \right) \Big/ (n-1)$$

Note that $\binom{n}{k}$ is the number of k-element consistent sets when P is strictly

[4]This is useful for computation—to be discussed in a later paper in this series.

proper. Hence $\bar{C}$, like $\bar{S}$, is a number between zero and one. (In Example 1, Sec. 6, $\bar{C}=.5556$.)

In this section we introduce an additional concept because of its interest to composers; i.e., the number of distinct values (i.e. α_{ij} or δ_{ij}) assumed by each collection of those intervals in $P\times P$ which all have the same diatonic distance. Hence we define the *variety*, V_i, of i as the number of distinct values assumed by $\alpha_{ij}, j=1,\ldots,n$ (i fixed), and we define the *mean variety* of P, $\bar{V}$, as

$$\bar{V}=\sum_{i=1}^{n-1} R_i/n-1.$$

(In Example 1, Sec. 6, $\bar{V}=2$)

10. Equivalence Classes of P

We now consider the question of when two musical scales are perceived as "mistunings" of a single scale as opposed to when they are perceived as distinct "different" scales. Of significance here is the fact that we do not assume a metric which imparts "absolute" perceived sizes to musical intervals. Hence we form equivalence classes of scales according to their initial ordering, which we *do* assume. Additional musical interpretation appears in Section 13.

Since an infinite number of $\|\delta_{ij}\|$ may map into a single $\|\alpha_{ij}\|$, $\|\alpha_{ij}\|$ will be referred to (by an abuse of language) as the *equivalence class* for all such $\|\delta_{ij}\|$. As we have seen (Part 1, Sec. 6) $\|\delta_{ij}\|$ may be specified by its first row, $\psi(P)=(\delta_{1,1},\delta_{1,2},\ldots,\delta_{1,n})\cdot\|\alpha_{ij}\|$ may be specified by $\bar{K}$ and the first $\dfrac{n(n-1)}{2}$ elements of the first $\left[\dfrac{n}{2}\right]$ rows.

Example 9. $\|\delta_{ij}(P_1)\|$ and $\|\delta_{ij}(P_2)\|$ are specified by their first row vectors, $\psi(P_1)$ and $\psi(P_2)$. The remainders of reduced matrices $\|\delta_{ij}(P_1)\|$ and $\|\delta_{ij}(P_2)\|$ are enclosed in brackets. Both belong to equivalence class $\|\alpha_{ij}\|$ which is specified by $\bar{K}$ and its first $\dfrac{n(n-1)}{2}$ elements (Part 1, Sec. 6).

$$\psi(P_1)=\begin{matrix}2 & 2 & 3 & 2 & 3\\ \end{matrix}\qquad\qquad \begin{matrix}3 & 3 & 4 & 3 & 4\end{matrix}$$

$$\psi(P_1)=\begin{pmatrix}4 & 5 & 5 & 5 & 5\\ 7 & 7 & 8 & 7 & 7\\ 9 & 10 & 10 & 9 & 10\end{pmatrix}\qquad \psi(P_2)=\begin{pmatrix}6 & 7 & 7 & 7 & 7\\ 10 & 10 & 11 & 10 & 10\\ 13 & 14 & 14 & 13 & 14\end{pmatrix}$$

$$\underset{\|\alpha_{ij}\|}{\text{equivalence class}}=\begin{bmatrix}1 & 1 & 2 & 1 & 2\\ 3 & 4 & 4 & 4 & 4\end{bmatrix}$$

$$\bar{K}=9$$

($\bar{K}=9$ tells us that there are no ambiguous intervals since $\alpha_{ij}<\bar{K}/2$ for all i,j shown above.)

Notice that the role of $\bar{K}$ in specifying $\|\alpha_{ij}\|$ may be replaced by specifying whether the interval corresponding to any max (α_{ij}) (where i and j range over the first $\dfrac{n(n-1)}{2}$ elements of the first $\left[\dfrac{n}{2}\right]$ rows) is ambiguous or not. If it is, since

$\alpha_{n-i,i+j}=\bar{K}-\alpha_{ij}$, $\bar{K}=2\max(\alpha_{ij})$; if it is not, $\bar{K}=2\max(\alpha_{ij})+1$. Henceforth, let an underlined $\max(\alpha_{ij})$ denote ambiguity. (In example 9 above, $\underline{\max}(\alpha_{ij})=\alpha_{22}$ (or α_{23} or α_{24} or α_{25}) $=4$ and $\bar{K}=2\alpha_{22}+1=9$).

Since the number of orderings of $\frac{n(n-1)}{2}$ terms is finite for any n, the number of equivalence classes is finite for any fixed n. These equivalence classes may be generated for any n and corresponding values of stability calculated; for stability is an invariant of this equivalence relation.

Since we assume that the ordering of all elements of $P\times P$ is known but that no metric is known, we may assume $\|\alpha_{ij}\|$ is known but not $\|\delta_{ij}\|$.

We now consider the number of scales in a particular discrete S (e.g. "temperament" system or subdivision of the octave into "cents") which belong to a particular equivalence class. Hence, when axioms (2.2) and (2.3) (Part 1, Sec. 1) hold,[5] we consider the mappings of P into the integers, such that the definitions of addition and subtraction (2.4) apply in their usual numerical interpretation. Then obviously

$$\delta_{ij}=\sum_{k=j}^{j+i-1}\delta_{ik} \quad \text{and} \quad K=\sum_{j=1}^{n}\delta_{1j} \tag{1}$$

In this application, $m=wK$ (Part 1, Sec. 6)[6] can be interpreted as an "equal temperament" system, and each $\|\delta_{ij}\|$ in equivalence class $\|\alpha_{ij}\|$ would then correspond to a "tuning" of a "scale" with reduced matrix $\|\alpha_{ij}\|$. Here a method will be shown which finds a canonical member $\|\delta_{ij}\|$ of $\|\alpha_{ij}\|$ such that K is minimal. This method can be generalized to find other $\|\delta_{ij}\|\in\|\alpha_{ij}\|$. Also, considerable economy results if an equivalence class, $\|\alpha_{ij}\|$ is represented by a member $\|\delta_{ij}\|$, which can be specified by its first row only (Part 1, Sec. 6). The solution of the following problem provides a mapping from a given $\|\alpha_{ij}\|$ to the first row of a $\|\delta_{ij}\|$ which will represent that $\|\alpha_{ij}\|$:

Find positive integers $\delta_{ij}, j=1,\ldots,n$, such that K is minimum where

$$K=\sum_{j=1}^{n}\delta_{1j} \tag{2}$$

subject to the constraints

$$\delta_{ij}R\delta_{kl}\ (\delta_{ij}\text{ is a positive integer}) \tag{3}$$

where R is a relation "$=$" or "$>$" and is determined by the ordering of α_{ij} and α_{kl}. One such constraint exists for each pair $i,j\neq k,l$ where i,j,k,l range over the first $\frac{n(n-1)}{2}$ terms of the first $\left[\frac{n}{2}\right]$ rows of $\|\alpha_{ij}\|$. An additional constraint exists if $\bar{K}=2\max(\alpha_{ij})$ (to be discussed).

---

[5]Techniques can also be developed for some cases where axioms (2.2) and (2.3) do not hold, but these are not treated here.

[6]m is the number of "units" in an octave. (In the 12-tone equal temperament system, $m=(12)$). K is the number of "units" in a cycle. (For the scale...C, D, D#, E, F#, G, G#, A#, B, C,..., $K=4$). w is the number of cycles per octave. (For that scale $w=3$.)

(3) may be rewritten:

$$\sum_{x=j}^{j+i-1} \delta_{1x} R \sum_{y=l}^{l+k-1} \delta_{1y} \tag{4}$$

Clearly, many of these constraints imply each other. All such redundant constraints are eliminated. Two constraints which imply a third will yield the third when added to each other (after cancellation of terms common to both sides).[7]

Let $\beta_1 = \min_j (\alpha_{ij})$ and in general

$$\beta_k = \min_j \{\delta_{1j} | \delta_{1j} > \beta_{k-1}\}$$

Let $\Delta_1 = \beta_1$, and in general

$$\Delta_k = \beta_k - \beta_{k-1}$$

Then, obviously

$$\beta_k = \sum_{y=1}^{k} \Delta_y. \tag{5}$$

Each $\delta_{1j} = \beta_k$ for some value of k; let $k \equiv f(j)$. Then

$$\delta_{1j} = \sum_{y=1}^{f(j)} \Delta_y. \tag{6}$$

Let C_k be the number of $\delta_{1j} = \beta_k$ and let $\bar{k}$ = the number of different magnitudes of δ_{1j}. Then (2) becomes

$$K = \sum_{k=1}^{\bar{k}} C_k \beta_k$$

and by (5)

$$K = C_1\Delta_1 + C_2(\Delta_1 + \Delta_2) + C_3(\Delta_1 + \Delta_2 + \Delta_3) + \cdots + C_{\bar{k}}(\Delta_1 + \cdots + \Delta_{\bar{k}}).$$

Hence

$$K = \sum_{k=1}^{\bar{k}} \left[\Delta_k \sum_{l=k}^{\bar{k}} C_l \right] \tag{7}$$

Applying (6), (4) may be rewritten

$$\sum_{x=j}^{j+i-1} \sum_{y=1}^{f(x)} \Delta_y R \sum_{y=l}^{l+k-1} \sum_{z=1}^{f(y)} \Delta_z. \tag{8}$$

[7]Other quicker but less simple elimination techniques exist. Also note that when addition of constraints produces contradictory constraints, the notation does not represent a valid equivalence class when axioms (2.2) and (2.3) hold.

We have an integer linear programming problem: find positive integers Δ_x, $x=1,\ldots,n$ such that K is minimum, subject to the constraints (8). If Δ_x need not be an integer but is a real $\geqslant 1$, a solution can certainly be found. Note that all terms on both sides of (8) are positive and that any Δ_k may be multiplied by a coefficient which is at most equal to i. Since $i \leqslant \left[\frac{n+1}{2}\right]$ (when $\bar{K}=2\max(\alpha_{ij})$, $i \leqslant \left[\frac{n}{2}\right]$), an upper bound on the least common denominator of the Δ_k can be computed and all the Δ_k will be rational. But if a rational solution exists, so does the integer solution obtained by multiplying by the least common denominator. Hence an upper bound on K when all Δ_x are integers exists and a finite solution to the problem can be obtained.

In most cases where P is proper and n is small, which are of interest here, a simple technique will supply a solution:

Example 10. Find the first row of the canonical $\|\delta_{ij}\| \in \|\alpha_{ij}\|$ where

$$\|\alpha_{ij}\| = \begin{bmatrix} 1 & 2 & 3 & 5 & 4 \\ 6 & 8 & 9 & \underline{10} & 7 \\ & & & & \\ \underline{10} & (\bar{K}=20) & & & \end{bmatrix}$$

(a) $\delta_{1,1}\delta_{1,2}<\delta_{1,3}<\delta_{1,5}<\delta_{1,4} \Rightarrow \delta_{2,1}<\delta_{2,2}<\delta_{2,3}<\delta_{2,4}$.

The remaining constraints are

(b) $\delta_{2,1}>\delta_{1,4}$; $\delta_{2,2}>\delta_{2,5}$; $\delta_{2,4}=\delta_{3,1}$;

all others are implied by (a) and (b). Expanding (b):

$$\text{(c)}\quad \delta_{1,1}+\delta_{1,2}>\delta_{1,4};\quad \delta_{1,2}+\delta_{1,3}>\delta_{1,5}+\delta_{1,1};$$
$$\delta_{1,4}+\delta_{1,5}=\delta_{1,1}+\delta_{1,2}+\delta_{1,3}$$
$$\delta_{1,1}=\Delta_1;\quad \delta_{1,2}=\Delta_1+\Delta_2;\quad \delta_{1,3}=\Delta_1+\Delta_2+\Delta_3;$$
$$\delta_{1,4}=\Delta_1+\Delta_2+\Delta_3+\Delta_4+\Delta_5;$$
$$\delta_{1,5}=\Delta_1+\Delta_2+\Delta_3+\Delta_4.$$

Substituting in (c)

$$2\Delta_1+\Delta_2>\Delta_1+\Delta_2+\Delta_3+\Delta_4+\Delta_5;$$
$$2(\Delta_1+\Delta_2)+\Delta_3>2\Delta_1+\Delta_2+\Delta_3+\Delta_4;$$
$$2(\Delta_1+\Delta_2+\Delta_3+\Delta_4)+\Delta_5=3\Delta_1+2\Delta_2+\Delta_3$$
$$\Delta_1>\Delta_3+\Delta_4+\Delta_5;\quad \Delta_2>\Delta_4;\quad \Delta_3+2\Delta_4+\Delta_5=\Delta_1.$$

Since all Δ_k are positive integers, Δ_1 is minimum when $\Delta_3=1$, $\Delta_4=1$, $\Delta_5=1$ and $\Delta_1=4$. So Δ_2 is minimum when $\Delta_2=2$. Then $\delta_{1,1}=4$, $\delta_{1,2}=6$, $\delta_{1,3}=7$, $\delta_{1,4}=9$, $\delta_{1,5}=8$,

$$\psi(P)=(4,6,7,9,8) \qquad \text{and} \qquad K=34.$$

Henceforth each equivalence class, $\|\alpha_{ij}\|$, will be represented by the first row, $\psi(P)$, of its member $\|\delta_{ij}\|$ as specified above. Ambiguous intervals, stability, etc. are identical for all members of a given $\|\alpha_{ij}\|$.

Now, analagously to our consideration of proper modifications, we consider all substitutions of $x \in S$ for $p_k \in P$ which will leave the equivalence class of P unaltered, and call this set the *equivalence range* (E-range or E_k) of p_k. Obviously, $E_k \subset R_k$, and by reasoning similar to that for (Part 1, 4.3 and 4.6), the determination of the extremes of x and y (where $E_k = [y, x]$) satisfying the following inequalities for all $i > 0$ is sufficient to determine E_k:

$$\begin{cases} (p_{k-1}x) < \min\{\delta_{lj} | (\delta_{lj} > \delta_{i,k-i}) \wedge (j \neq k-l)\} \\ (xp_{k+i+1}) > \max\{\delta_{lj} | (\delta_{lj} < \delta_{i+1,k}) \wedge (j \neq k)\} \\ (p_{k-i}x) < \min\{xp_{k+l} | \delta_{lk} > \delta_{i,k-i}\} \end{cases}$$

$$\begin{cases} (yp_{k+i}) < \min\{\delta_{lj} | (\delta_{lj} > \delta_{ik}) \wedge (j \neq k)\} \\ (p_{k-i-1}y) > \max\{\delta_{lj} | (\delta_{lj} < \delta_{i+1,k-i-1}) \wedge (j \neq k-l)\} \\ (yp_{k+i}) < \min\{p_{k-i-1}y | \delta_{i+1,k-i-1} > \delta_{ik}\} \end{cases}$$

The above inequalities, of course, need not be applied when

$$[(\delta_{lj} \sim \delta_{ik}) \wedge (j \neq k)] \vee [(\delta_{lj} \sim \delta_{i,k-i}) \wedge (j \neq k-l)] \tag{9}$$

because in such cases $x = y = p_k$.

In periodic cases where a metric is given by embedding P in the reals so that the definitions of addition and subtraction apply in their usual numerical interpretation (axioms (2.2) and (2.3) apply), the following formulae may be used: Let $E_k = [p_k - \varphi_k^-, p_k + \varphi_k^+]$. Then, after applying (9) wherever possible and making all column subscripts positive (mod n), we can set[8]

$$\varphi_k^+ = \min\begin{cases} [\min\{\delta_{lj} | (\delta_{lj} > \delta_{i,k-i}) \wedge (j \neq k-l)\}] - \delta_{i,k-i} \quad 0 < i \leqslant n, 0 < l \leqslant n \\ \delta_{i+1,k} - [\max\{\delta_{lj} | (\delta_{lj} < \delta_{i+1,k}) \wedge (j \neq k)\}], 0 < i \leqslant n, 0 < l \leqslant n \\ 1/2 \min_l \{(\delta_{lk} - \delta_{i,k-i}) | \delta_{lk} > \delta_{i,k-i}\}, 0 < l < n, 0 < i < n \end{cases}$$

$$\varphi_k^- = \min\begin{cases} [\min\{\delta_{lj} | (\delta_{lj} > \delta_{ik}) \wedge (j \neq k)\}] - \delta_{ik}, 0 < i \leqslant n, 0 < l \leqslant n \\ \delta_{i+1,k-i-1} - [\max\{\delta_{lj} | (\delta_{lj} < \delta_{i+1,k-i-1}) \wedge (j \neq k-l)\}], \\ 1/2 \min_l \{(\delta_{i+l,k-i-l} - \delta_{ik}) | \delta_{i+l,k-i-l} > \delta_{ik}\}, 0 < l < n, 0 < i < n \end{cases}$$

[8] Actually, the two extreme points of E_k as computed by these formulae should be eliminated from E_k.

Note that when $x > p_k$ and x replaces p_k, all δ_{ik} are reduced and all $\delta_{i,k-i}$ are increased. Since P is periodic, when δ_{ik} is decreased so is $\delta_{i,k-n}$ and when $\delta_{i,k-i}$ is increased so is $\delta_{i,k-i+n}$. But when $i=n$, $\delta_{i,k-i+n} \equiv \delta_{i,k} \sim \delta_{i,k-n} \equiv \delta_{i,k-i}$. Hence δ_{nk} is both decreased and increased and thus remains unaltered. For this reason in the last of the inequalities limiting both φ_k^+ and φ_k^-, i ranges only from 1 to $n-1$.[9]

Given any (not necessarily canonical) $\psi(P) \in \|\alpha_{ij}\|$, by computing φ_k^+ and φ_k^- for each k, if S is finite within a period, it is easy to see how all $\psi(P) \in \|\alpha_{ij}\|$ can be generated.

Just as previously the range was a union of proper modifications, the E-range is a union of modifications (E-modifications) which preserve equivalence class. The preceding formulae are easily adjusted to provide such E-modifications in a stepwise manner. Also, if we now consider the preservation of equivalence class when mapping from $S \times S$ to C, we obtain *SE-modifications*, which occupy the same relation to E-modifications that S-proper modifications had to proper modifications. While E-modifications apply to mistunings of single elements of a scale without alteration of its equivalence class membership, SE-modifications apply to such mistunings of more than one element (such as occur within a vocal performance). Again, half-modifications may be utilized as previously, and formulae for computation are easily obtained. Section 13 elaborates the musical applications.

11. Inverses, Descending Form

For any $\psi(P) = (\delta_{1,1}, \ldots, \delta_{1,n})$ and any $k \leqslant n$ we define

$$\psi^k(P) = (\delta_{1,(k+1)}, \ldots, \delta_{1,n}, \delta_{1,1}, \ldots, \delta_{1,k})$$

i.e. the cyclic permutation of R beginning with $\delta_{1,(k+1)}$.

We also define the *inverse* of $R(P)$ as

$$\hat{\psi}(P) = (\delta_{1,n}, \ldots, \delta_{1,1})$$

(The entries are in reverse order.)

If two P's are so related that their $\psi(P)$'s are cyclic permutations or inverses of each other, it is easy to see that their ambiguous intervals, stabilities and (if axiom 2.2 applies) their coverages are identical.

The *descending form* of a particular $\psi(P)$ is defined as the lexicographically latest permutation of it. (E.g. (2,2,2,1,2,2,1) is in descending form but (2,2,1,2,2,2,1) is not.)

---

[9] E_k is the set of all pitches which can replace p_k *jointly in all* octaves (or, more generally, periods) of P without changing its equivalence class. It is possible to define the ranges $\bar{R}_k$ analogously, i.e. as the set of all pitches q that can replace p_k *jointly in all periods* so that $P - \{p_k\} \cup \{q\}$ is still proper. For the computation of such $\bar{R}_k$, i must range only from 1 to $n-2$ in the expression $(\delta_{i+1,k} - \delta_{i,k-1})/2$ and its dual on 6.1 and 6.2 of Part 1. Also not all theorems about range necessarily hold for this modified notion if $n < 3$. Such a periodic definition of range is appropriately used only when a substitution for p_k occurs over *all* octaves and p_k does not immediately recur in any octave. It is not relevant to the use of added (auxiliary) tones.

For convenience, all elements p_i of P will hereafter be subscripted so that when $\psi(P)$ is in descending form,

$$\psi(P)\equiv(\delta_{1,1},\delta_{1,2},\ldots,\delta_{1,n})\equiv[(p_1p_2),(p_2p_3),\ldots,(p_np_1)] \tag{11.1}$$

Henceforth $\psi^0(P)$ and $\psi(P)$ (superscript omitted) will refer to the descending form. $\hat{\psi}^0(P)$ and $\hat{\psi}(P)$ will refer to the descending form of the inverse. $\psi(P)^*$ will refer to that member of an equivalence class which is used to represent it (K is minimum), $\hat{\psi}(P)^*$ to its (normalized) inverse, and both will be descending form.

For subsequent economy in the presentation of computed results, we note that certain $\psi(P)^*$ are their own inverses (that is $\hat{\psi}(P)^*=\psi(P)^*$) and that stability, coverage, etc. are the same for these. (E.g. $\psi(P)^*=(2,2,2,1,2,2,1)=\hat{\psi}(P)^*$. The inverse of $(2,2,2,1,2,2,1)$ is $(1,2,2,1,2,2,2)$ which in descending form is $(2,2,2,1,2,2,1)$ over again).

When axioms (2.2) and (2.3) do not hold, the definitions of *inverse* and *descending form* still apply. The descending form of $\|\alpha_{ij}\|$ ($\|\alpha_{ij}\|^*$) is defined as that cyclic permutation of the *columns* of $\|\alpha_{ij}\|$ such that the resulting first row $(\alpha_{1,1},\ldots,\alpha_{1,n})$ is latest in lexicographic order. The *inverse* of $\alpha_{ij}(\|\hat{\alpha}_{ij}\|)$ is defined as $\|\alpha_{ij}\|$ with the order of the columns reversed.

When different $\|\alpha_{ij}\|$ (or $\psi(P)$) are in descending form they can be mechanically compared for identity and can be listed in lexicographic order.

12. Keys, Modes, Tunings and Scales

To any equivalence class $\|\alpha_{ij}\|$ (or $\psi(P)^*$) there correspond many distinct sets, P. Consider the cyclic case where the period length (p_ip_{i+n}) is the same for all P. Suppose $P\in\|\alpha_{ij}\|$, $P'\in\|\alpha_{ij}\|$, $P\not\equiv P'$ and $x\in(P\cap P')$. It is possible that both $(x\in S)\equiv(p_i\in P)\equiv(p_j'\in P')$ and $i\neq j$. (Note that elements of P and P' have been indexed in accord with the descending form of $\psi(P)$—see 11.1.) That is, if S is sufficiently rich (as it is when axiom 2.2 holds), then for each element, x, of S and every value of i from 1 to n (the period) there exists a $P\in\|\alpha_{ij}\|$ such that x is the ith element of P (i.e., $x=p_i\wedge p_i\in P$). For any i and x, all $P\in\|\alpha_{ij}\|$ for which $x=p_i$ will be said to belong to *key* x of *mode* i of *scale* $\|\alpha_{ij}\|$. For any i and x there are at most as many members of a given key, mode and scale as there are $\|\delta_{ij}\|\in\|\alpha_{ij}\|$ (depending upon the richness and symmetry of S). That is, from a given S usually many distinct subsets, P, each with a different $\|\delta_{ij}\|$, may usually be selected, all of which have identical key, mode, and scale memberships. Each such $\|\delta_{ij}\|$ will be called a *tuning* of scale $\|\alpha_{ij}\|$. Thus any P is uniquely specified by i, x, $\|\delta_{ij}\|$ and $\|\alpha_{ij}\|$ (where, of course, $\|\delta_{ij}\|\in\|\alpha_{ij}\|$).

Given i and $\|\delta_{ij}\|$, any key can be simply generated by setting p_i equal to a particular element, x, of S and selecting remaining elements of P (provided such exist in S) which conform to $\|\delta_{ij}\|$. In the periodic case this can be done in clockwise order. According to our definition of addition, (2.4), (which is equivalent to algebraic addition when axioms (2.2) and (2.3) hold), if we set $p_j\equiv x$, then $p_{i+1} \equiv x + \delta_{1,j}$, $p_{i+2} \equiv x + \delta_{1,i} + \delta_{1,j+1}$, and in general,

$$p_{j+i} \equiv x + \sum_{k=j}^{j+i-1} \delta_{1,k} \equiv x + \delta_{i,j}.$$

Note, therefore, that if P and P' have identical $\|\delta_{ij}\|$, if P is key x of mode j and P' is key y of mode $j+i$, then if $y \equiv x + \delta_{ij}$, $P \equiv P'$. Thus any P can be uniquely specified by $\|\delta_{ij}\|$ and that point, $x \in S$, which corresponds to p_1 in P. (Indexing is in accord with the descending form of $\psi(P)^*$.) Accordingly any $P \in S$ will be denoted by $P_v(x)$ where $x \equiv p_1$ and v is an index which ranges over distinct $\|\delta_{ij}\|$. (E.g., $P_v(x)$ and $P_v(y)$ are different keys of the same tuning and $P_v(x)$ and $P_w(x)$ are the same key of different tunings and possibly different scales). The mode of P is not shown by such notation.

In the musical interpretation the terms "mode", "key", "tuning" and "scale" may be interpreted in terms of their customary musical usage (which, however, is rarely unambiguous).

When axioms (2.2) and (2.3) apply and there exists at least one $P \in \|\delta_{ij}\|$, the number of distinct keys of a given $\|\delta_{ij}\|$ (specified by its first row vector $\psi(P)$) is easily determined. Consider a $P \in \|\delta_{ij}\|$ with period n. (The *least* n for which there exists a K satisfying $p_{i+n} = p_i + K$.) Axiom (2.2) guarantees that if $x, y \in S$ and $x + \delta_{ij}$ is in S, so also is $y + \delta_{ij}$. Hence the number of distinct keys of $\|\delta_{ij}\|$ is equal to the number of points in a K-cycle (p_i, p_{i+n}). Thus if S is the integers $\bmod m$ (see Part 1, Section 6, $m = wK$ and $\bar{n} = wn$),[10] there are m points in an m cycle, and $K = \frac{m}{w} = \frac{mn}{\bar{n}}$ is the number of such keys.

13. Distinct Scales and Mistunings

That there is only a finite number of equivalence classes of any cardinality is musically significant in that only finitely many significantly differing musical scales may be constructed. Also note that, if a scale is conceived of as an ordering, the use of finer tunings and smaller intervals does not necessarily produce new scales. It is then reasonable to suppose that a listener learns and accepts some tuning of a particular equivalence class as correct, and perceives subsequent deviations from such tuning as "out of tune" (rather than as elements of new or different scales). Using twelve-tone equal temperament, no two proper scales of cardinality $\geqslant 5$ exist which fall into the same equivalence class. This may partially account for why, when producing non-diatonic music on a piano, very few mistakes sound "out of tune" (although such may be heard as dissonant or as "wrong notes" to a listener who knows the style or composition). However, when microtones are used (19, 22, 24, 31, 36, 48 and 53 equally spaced tones per octave have been used at various times in Western music),

[10]Because of our computation methods, it is convenient to choose S as the integers $\bmod m$ when axioms (2.2) and (2.3) apply.

many scales may appear which fall into the same equivalence class. It would be expected then, that the playing (several times) of some key of an unfamiliar scale on a microtonal instrument immediately followed by the playing of a different tuning (still within the same equivalence class, mode and key), would result in the latter being perceived as "out of tune". This should not occur when each of the tunings used is in a different equivalence class.[11]

In the early part of the century much experimentation in the use of microtones was conducted with the hope of finding new musical scales. For the most part this search was unsuccessful, and much resulted that is described by both naive listeners and musicians as familiar materials which are "out of tune". The above methods offer an explanation of such responses when examined in detail. The later papers in this series which deal with the development of new musical resources rely heavily upon the use of the techniques derived here.

We have suggested that the tunings of a proper scale are restricted so that propriety is preserved. Further restrictions are here imposed if equivalence class membership is to be retained. (Of course, adjustments are made to accommodate limitations on pitch discrimination—to be discussed in the next paper of this series). In the case of improper scales, such restrictions are the only bounds on tuning discussed thus far.[12] E-range clearly limits the mistuning of any single element of a scale. However, when simultaneous variations in tunings of all the elements of a scale are considered, SE-modifications are relevant. That is, any set of tunings of P which fall within such a modification may be simultaneously entertained without affecting equivalence class membership. Great demands for precision may result. It will be subsequently shown that the readiness (speed) with which a particular P (that is, a particular key of a scale) can be identified (it being assumed that the ordering of its pairs is already learned) depends upon such precision. However, once identification has been made (and the principal tones are determined—which in the case of improper scales rarely change) such precision is no longer needed. It is then usually sufficient that all contradictory and non-contradictory pairs in $P \times P$ retain their identities. In the case of proper scales, S-proper modifications satisfy this condition. When P is improper, the formulae for obtaining S-proper modifications can be trivially altered so that modifications which preserve the identities of contradictory and non-contradictory pairs are obtained.

It is interesting to note that according to B. Yserdraat (see Part 1, Section 8), Gamelon (orchestra) leaders in Sunda (West Java) periodically retune the fixed pitch instruments of the Gamelon to conform to (blend with) the changing tone colors (timbres) of the principal gongs. (As the gongs age, their timbres stabilize and this becomes less frequent. Note the suggestion here that the initial ordering depends upon timbre). At such times there is occassionally an awareness that a "new scale" has resulted. This theory suggests that this occurs when the retuning has resulted in a scale belonging to a new equivalence class.

---

[11]Such an experiment can be performed by altering frequencies on a harpsichord or moving frets on a guitar.

[12]Those familiar limitations on tuning which derive from the necessity of preserving characteristic harmonic properties of specific intervals (such as the coincidences between the harmonics of components of an interval or the relation of the beats and difference tones generated to such components) are not discussed here. These have been extensively treated in literature on the subject since Helmholtz.

Fig. 1. Table of all proper $\|\delta_{ij}\|$ when $m=12$.

All $\psi(P)$ are in descending form. Inverses are not shown; e.g., the major triad 5,4,3 is not shown because its inverse 5,3,4 is; the dominant seventh 4,3,3,2 is not shown because 4,2,3,3 is, etc. The column headed "Efficiency" will be explained later.

$\psi(P)$	Stability	Ambiguities[13]	T	Efficiency
615	.6667	6	0	.7778
624	.6667	6	0	.7778
633	.6667	6	0	1.0000
534	1.0000	—	2	.6667
552	1.0000	—	−2	.6667
444	1.0000	—	−4	.3333
5142	.6667	5,7	0	.6250
5151	1.0000	—	−1	.5833
5232	.5	5,7	0	.8333
4233	1.0000	—	1	.6250
4242	1.0000	—	2	.5833
4323	1.0000	—	1	.6667
4341	1.0000	—	1	.6667
4413	.5000	4,8	0	.6875
4422	.5000	4,8	0	.8125
3333	1.0000	—	3	.2500
41322	.6000	4,8,6	0	.5800
42222	.4000	4,8,6	0	1.0000
32322	1.0000	—	1	.8000
33132	.9000	6	0	.6000
33222	.9000	6	0	.6400
33312	.4000	3,6,9	0	.6400
312222	.7333	3,9,5,7	0	.6278
312312	.7333	3,9	0	.4556
313122	.6000	3,9,5,7	0	.5889
313131	1.0000	—	1	.4167
321222	.5333	3,9,5,7	0	.6333
321231	.5333	3,9,5,7	0	.6500
321312	.5333	3,9,5,7	0	.5611
322122	.4667	3,9,5,7	0	.7778
222222	1.0000	—	2	.1667
3121221	.4762	3,9,4,8,6	0	.6259
2221221	.9524	6	0	.7687
2222121	.7143	4,8,6	0	.6299
2222211	.2857	2,10,4,8,6	0	.6327
21212121	1.0000	—	1	.3250
22112121	.4643	2,10,4,8,5,7	0	.6616
22112211	.5714	2,10,4,8	0	.4964
22121121	.4286	2,10,4,8,5,7	0	.6750
211211211	.7500	2,10,6	0	.4683
212111211	.3611	2,10,3,9,5,7,6	0	.7262
2111121111	.4667	2,10,3,9,4,8	0	.6937
2111211111	.2667	2,10,3,9,4,8,5,7	0	.8543
21111111111	.1818	2,10,3,9,4,8,5,7,6	0	1.0000
111111111111	1.0000	—	1	.0833

[13] The entries in the table are δ_{ij} for all ambiguous $(p_i p_j)$. Hence each $(p_i p_j)$ such that δ_{ij} is an entry is ambiguous.

Figure 1 lists all $\|\delta_{ij}\|$ corresponding to proper sets when $m=12$ and S is the integers mod 12. Notice that all $\psi(P)$ with less than five terms and stability $>2/3$ (with the exception of $\Psi(P)=(5,1,5,1)$[14], appear as "chords" in music texts specifying standard "figured bass" notation. Chords with strong difference tones reinforcing one element of the set are those which music texts tell us have these elements as "roots". More will be said of this later.

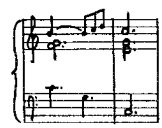
Since axiom (2.2) applies, instead of giving to T the values 0 and ± 1 only, we can measure it quantitatively by $T_i \equiv \underline{\delta}_{i+1} - \delta_i$, $T \equiv \min T_i$. On this basis, minimum tolerance would be maximal when $(p_i, p_{i+k}) = (p_j + p_{j+k})$ for all i, j, k (equal division of the octave). It might be expected that existing proper scales would have this property and/or be high in stability. The following discussion explains why this is not the case, and introduces other factors relevant to the evaluation of the proper sets in Figure 1.

14. Sufficient Sets; the Coding of P

We have hypothesized that when a listener is presented with a series of unfamiliar tonal stimuli, he must mentally construct a reference frame, P, to which all such stimuli are referred. Many proper P may satisfy this requirement. If the stimuli are sufficiently unfamiliar (as when one listens to music of an alien culture) many repeated hearings may be necessary during which a listener replaces a familiar P with one more appropriate for classifying the stimuli heard. The cardinality of the constructed P will depend upon the numbers of distinctions required by the particular musical language or, if the stimuli are not musical, upon the fineness of discrimination required by the recognition task to be performed. However, most listening to music involves the identification of a P whose corresponding $\|\delta_{ij}\|$ has already been learned. This would necessitate the identification of mode, key and tuning from a given subset of P (the stimulus) when $\|\alpha_{ij}\|$ is known. However, a listener who has learned a "scale" has learned not only an $\|\alpha_{ij}\|$ (an initial ordering) but a particular tuning $(\|\delta_{ij}\|)$ as well. That is, given any "interval" (pair) in P, he is able to recognize the possible positions its elements might occupy in P. In effect, given any pair $(p_i p_j)$ he can mentally supply (interpolate) a possible set of remaining elements of P which satisfy $\|\delta_{ij}\|$. Such an interpolation becomes unique after a sufficient number of elements in P are heard. This is equivalent to the identification of x (key) in a given $P_v(x)$. If we accept the indexing on elements of P given by the

[14]Scales with $\Psi(P)=(5,1,5,1)$ cannot be embedded in any major or minor key and are "node-minimal sets" for keys of the scale $(2,1,2,1,2,1,2,1,)$ which are not used in music of the classical period. Node-minimal sets and their applications will be subsequently discussed. $\Psi(P)=(5,1,4,2)$ is also sometimes omitted, but is occasionally used as an "altered seventh chord" and its inverse is also used. E.G.

(an arrangement of "Laura," a popular song



(a typical Mozart cadence)

descending form of $\psi(P)$ or by any of its cyclic permutations, the mode is, in effect, also so determined in the sense that the listener is able to identify the pitches heard as "degrees" of a particular key of a mode of $\|\delta_{ij}\|$. In this fashion the elements of P (within a period) may be coded into "scale degrees" as soon as a sufficient number of elements of P have been heard.

Accordingly we define a *sufficient set* for $P_v(x_1)$ as a subset, Q, of $P_v(x_1)$ such that if $x \neq x_1$, Q is *not* a subset of $P_v(x)$; i.e., as a set of p_i which is included in one and only one key of P. Thus the major triad C E G is not a sufficient subset of the C-major scale, for there are two other major scales (namely those on F and G) in which the set {C E G} is included. But there is no other major scale except that on C which contains the four notes G B D F, so the "dominant seventh chord" *is* a sufficient subset of the C-major scale. A *minimal* set is a sufficient set with no sufficient proper subset ('proper' of course in the sense of set inclusion!); G B D F is not a minimal set, but G B F is; for there is no major scale except that on C which contains G, B and F, while each of its proper subsets {G B}, {G F} and {B F} is included in some other major scale (e.g. G major, F major and F$^{\#}$ major respectively).

It is straightforward to verify that sufficient (and consequently minimal) sets are invariants of equivalence, i.e., they depend only on the α_{ij} and not on the δ_{ij}. More precisely: if $\{p_{i_j},\ldots,p_{i_k}\}$ is a sufficient (minimal) set for $P=\{p_j,\ldots,p_n\}$, and if $P'=\{p'_j,\ldots,p'_n\}$ is equivalent to P (the p_i and p'_i being arranged in corresponding orders, e.g., both in descending order), then $\{p'_{i_j},\ldots,p'_{i_k}\}$ is sufficient (resp. minimal) for P'.

15. Efficiency

Consider a language whose alphabet consists of $\bar{n}$ letters (phonemes). How many distinct n letter words can be formed using this alphabet? Of course, certain restrictions exist which limit the sequences of letters which can occur (e.g. no more than two consonants in a row). The more distinct words that can be formed whose length is less than or equal to some maximal n, the more efficient the alphabet may be said to be.

A similar situation applies when "words" are formed from sequences of intervals. Since interval sequences are formed from tone sequences (although not in a linear fashion) we consider sequences of the elements of some P. Also, since no new intervals are formed when an element is repeated, only non-repeating sequences will be considered. Since we are here concerned only with properties deriving from the structure of P, we will use the following criterion for the termination of a "word" (other criteria apply when "motifs", etc., are considered): When all remaining elements of P are determined by a sequence of some of its elements, the addition of elements will impart no further information of *this* type, and the "word" will be considered as terminated. That is, any sequence will be considered as complete as soon as a sufficient set occurs in that sequence, i.e., as soon as we know what key we are in. We now ask, given a particular $\|\alpha_{ij}\|$, how many distinct "words" can be formed using k elements where k varies from 1 to n. Again, we consider the "alphabet" formed by $\|\alpha_{ij}\|$ as more efficient when a greater number of words can be formed of length k

(averaged over all values of k). With this purpose the following definitions are made:

Consider all non-repeating sequences of $\bar{n}$ points[15] in P (there are $\bar{n}!$ such sequences). Let s_i be the number of elements in each such sequence which must appear before a sufficient set is encountered. Then $F(P)$ is defined as the average,

$$\left(\sum_{i=1}^{\bar{n}!} s_i\right) \Big/ \bar{n}!$$

$F(P)$ may be interpreted as the average number of elements in a non-repeating sequence of the $\bar{n}$ elements of $P_v(x)$ required to uniquely determine the key, x.

Efficiency, E, is defined as $F(P)/\bar{n}$ and *redundancy*, R, as $1-F(P)/\bar{n}$ (both numbers lie between zero and one).[16]

It should be noted that this kind of "efficiency" and "redundancy" differs intrinsically from the meanings these terms assume in information theory applications. The distinction is important and applies to alphabets in spoken natural languages as well as to musical "scales". The "redundancy" of information theory refers to a redundancy in the "message", not in the "code" (alphabet). In the discussion here, that property of the code which determines whether efficient (or redundant) messages *can* be constructed (if such are desired) is considered. This property is inherent in the code itself, and does not apply to the "message". Much confusion has resulted from the application of standard statistical "redundancy" measures to musical "messages" without considering the limitations introduced by the efficiency of the code being used.

Let us now classify scales according to their values of stability and efficiency. A crude classification would be:

Scale Type	Stability	Efficiency
(a) proper	high	high
(b) proper	high	low
(c) proper	low	high
(d) proper	low	low
(e) improper	—	high
(f) improper	—	low

Notice that in Figure 1 all scales in 12-tone equal temperament with which we are most familiar (the major, minor, Chinese Pentatonic) are relatively high in both stability and efficiency. (In fact, the major scale (of which the "natural minor" is a mode) has far higher stability and efficiency than any other

[15] $\bar{n}$ is the number of pitches per octave—e.g. $\bar{n}=7$ for the major scale.

[16] Methods exist for generating minimal sets directly and for computing efficiency without finding minimal sets. Generators for proper and strictly proper sets also exist. These will be set forth in the fourth paper in this series.

seven-tone scale shown).Next among seven-tone scales is the "melodic minor" (2,2,2,2,1,2,1). The "Chinese pentatonic" (3,2,3,2,2) stands out among scales of 5 and 6 tones. The use of any temperament system (other than 12-tone equal temperament—see the tables in the fourth paper in this series) which approximates the perfect fifth does not alter these results for the major and pentatonic scales. Here situation (a) above applies and its desirability is exemplified.

However, situation (b) applies to many scales with which we are familiar, such as the "whole-tone scale" or 12-tone scale". Note that when these are strictly proper scales, from the hearing of a sufficient set (any element) alone, it is not possible to code the elements of P into scale "degrees". That is, although $P \times P$ is coded by the proper mapping, there is no way to index elements of P except by arbitrary choice. *Thus, since in these cases intervals (pairs) are coded but tones are not, composition with these scales must involve relations which make use of motivic similarities rather than relations between scale degrees.* Hence the tone row basis of 12-tone music (which is essentially motivic in concept) is not surprising. An examination of Debussy's whole-tone piano prelude "Voiles" shows similar motivic dependency.

Now consider improper scales. $P \times P$ is not coded except by the employment of proper subsets or a fixed tonic (which, in effect, codes P into scale degrees). *Hence information is primarily communicated by the scale degrees.* Thus it is important that P be coded as quickly as possible, which is indicated by a high redundancy (low efficiency) as in case (f). It would be expected then that scales characterized by case (e) would be extremely difficult to use, except when the tonic is fixed by a drone or similar device and, in fact, we have not discovered such scales in any musical culture examined thus far. In general, the use of motivic sequences on different scale degrees of improper scales would *not* be expected (except within proper subsets of such scales). This is strongly supported by examination of Indian and other music using improper scales. However, it should be noted that the use of motivic sequences on different scale degrees of improper scales is possible when these sequences occur on mutually compatible degrees of the scale (i.e., as described in discussion of "consistent sets"—such sequences avoid the exposure of contradictory and ambiguous intervals).

We would also expect that proper scales characterized by low stability would tend to be used as improper scales, so that case (c) would resemble (e) and (d) resemble (f) and similar remarks apply.

Note that most musical "cadences" (in Western music) have the following characteristics: (a) a minimal set is contained in the cadence; (b) the tonic appears together with at least one other tone which forms a difference tone reinforcing the tonic; (c) each chord in the cadence progression is a stable $\left(\bar{S} > \frac{2}{3}\right)$ subset of the scale (or a subset of such a stable subset)[17] which is distinct from the stable subset (or subset of such a stable subset) formed by the succeeding chord. Condition (a) serves to uniquely determine key, (b) to fix the tonic, and (c) to present as many distinct stable subsets ("color changes") of the scale as possible.[18] The IV–V–I cadence and V_7–I cadence satisfies these

[17]Note that a subset of a stable set is not necessarily stable.

[18]Condition (c), although characteristic of all pre-twentieth-century cadences, is not a necessary condition for a cadence. However, it also applies to nearly all twentieth-century "cadences".

conditions. (The use of a dominant seventh chord in the V_7–I progression provides a tone without which condition (a) would not be satisfied). The IV–I (plagal) and the V–I (not V_7–I) cadence fail to satisfy condition (a) and are used only when key has been previously established. Note that the "resolution" of an ambiguous interval to an unambiguous interval emphasizes the cadential effect. This is satisfied by the V_7–I progression $((F,B)\rightarrow(E,C))$. Similar conditions are satisfied by non-diatonic cadences (to be discussed later).

References

1. H. L. F. Helmholtz, *On the Sensations of Tone as a Physiological Basis for the Theory of Music*, (translated by Alexander J. Ellis, 1885) Peter Smith, New York, 1948.
2. P. Boomsliter and W. Creel, The long pattern hypothesis in harmony and hearing, *J. Music Theory*, 1961 **5**, 2–31.
3. J. C. R. Licklider, Three Auditory Theoreies, In: S. Koch (Ed.) *Psychology: a study of a science*, 41–144, McGraw-Hill, New York, 1959.
4. J. Kunst, *Music in Java*, Martinus Nijhoff, The Hague, 1949.
5. Mantle Hood, *The Nuclear Theme as a Determinant of Patet in Javanese Music*. J. B. Wolters—Groningen, Djakarta, 1954.
6. Mantle Hood, Slendro and Pelog Redefined, *Selected Reports*, Institute of Ethnomusicology, U.C.L.A. 1966.
7. D. Rothenberg, *A mathematical model for the perception of redundancy and stability in musical scales*. Paper read at Acoustical Society of America, New York, May 1963. Also: Technical Reports to Air Force Office of Scientific Research 1963–1969 (Grants and Contracts AF-AFOSR 881-65, AF 49(638)-1738 and AF-AFOSR 68-1596).
8. G. M. Stratton, Vision without inversion of the retinal image, *Psych. Rev.* **4** (1897), 341–60 and 463–81.
9. R. H. Thouless, Phenonmenal regression to the real object, *Brit. Jour. Psychol.* **21** (1931), 339–59.
10. M. von Senden, *Raum-und Gestaltauffassung beioperierten Blindgeborenen vor und nach der Operation*. Barth, Leipzig, 1932.
11. Carrol C. Pratt, Comparison of tonal distance, also: Bisection of tonal intervals larger than an octave, *Jour. of Experimental Psychol.* **11** (1928), 77–87 and 17–36.
12. H. Munsterburg, Vergleichen der Tondistanzen, *Beitrage zur Experimentelle Psychology*, **4** (1892), 147–177.
13. Poincaré, *The Value of Science*, Dover Publications, New York, 1954.
14. P. Hindemith, *The Craft of Musical Composition*, Associated Music Publishers, New York, 1941.
15. Olivier Messiaen, *Technique de mon Langage Musical*, Alphonse Leduc, Paris, 1944.
16. J. F. Schouten, R. J. Ritsma, and B. Lopez, Cardozo, Pitch of the Residue, *Jour. of the Acoustical Society of America*, **34** (1962), 1418–1424.
17. J. E. Evetts, The Subjective Pitch of a Complex Inharmonic Residue, unpublished report, Pembroke College, England, 1958.

Received March 1969 and in revised form June 1976; final version August 29, 1977.