Math. Systems Theory 11, 199-234 (1978)

Mathematical
Systems Theory

A Model for Pattern Perception with Musical Applications*
Part I: Pitch structures as order-preserving maps

David Rothenberg

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

Abstract. The initial five papers of this series on pattern perception treat first, the perception of pitch in musical contexts and then, the perception of timbre and speech. Each sound is considered to be embedded in a "context" consisting of those sounds which surround it or coincide with it. The apprehension of a musical pattern depends upon the perceptibility of certain relations between, and properties of, its parts (e.g. "motif *A* is similar to motif *B*" or "G is the tonic"). It is hypothesized that, because of the limitations of short term memory, the perception of specific relations and properties requires that certain "mental reference frames" be extracted from the various contexts. However, a reference frame which supports the perception of any specified relation may be extracted from only very few of all possible contexts. The choices of musical materials in both Western and non-Western music are shown to avoid precisely such difficulties. When they are not avoided, distortions of perception are predicted and methods for experimental verification are suggested. This theory is then applied to suggest new materials for the composition of both "microtonal" and "tone-color" music. This is done in a manner which exposes the correspondence between each choice of musical materials and those musical properties and relations whose perception is (or is not) thereby supported.

This first paper discusses the relation between the ability to perceive relative sizes of musical intervals and the choice of reference frame from a given musical context.

Introduction

Since the publication of Helmholtz's Tonempfindungen [1] in 1862 many attempts have been made to explain or predict the perception of complex tones.[1] These have derived from speculations about difference and combination tones,

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

[1]Complex tones are tones containing harmonics of the fundamental frequency

0025-5661/78/0011-199 $07.20

beats and coincidences between harmonics, possible similarities between neural pathways and electronic circuits and the accidents of cultural conditioning [2]. As yet, none of these has produced a satisfactory model for comprehending complex auditory phenomena [3]. Those theories susceptible to application to the perception of complex tones in their musical context have treated musical "scales" in an essentially similar manner: Certain musical "intervals" are chosen for specific acoustical properties (usually their frequencies form a simple ratio) and musical scales are formed by their superimposition.

It is now known that highly developed musical cultures exist which use scales with great precision and consistency which contain no intervals derivable from any of the above theories (except that of accidents of cultural conditioning).[2] Also, to most listeners, identical pairs of complex tones are not recognizable as the same when embedded in different sequences of tones. Such acoustical "illusions" deriving from context differences remain unexplained and, in general, unpredictable,[3] except in certain cases involving classical Western harmony.

Helmholtz, however, appears to have anticipated an alternative to the above approaches, which is the basis of the work to be presented here ([1] pp. 252, 270):

> As we have seen, then, melody has to express a motion, in such a manner that the hearer may easily, clearly, and certainly appreciate the character of that motion by *immediate perception*. This is only possible when the steps of this motion, their rapidity and their amount, are also exactly measurable by immediate sensible perception.[4]
>
> It is also necessary that the alteration of pitch should proceed by intervals, because motion is not measurable by immediate perception unless the amount of space to be measured is divided off into degrees. Even in scientific investigations we are unable to measure the velocity of continuous motion except by comparing the space described with the standard measure, as we compare time with the seconds pendulum. The individual parts of a melody reach the ear in succession. We cannot perceive them all at once.
>
> We cannot observe backwards and forwards at pleasure. Hence for a clear and sure measurement of the change of pitch, no means was left but progression by determinate degrees. This series of degrees is laid down in the musical scale. When the wind howls and its pitch rises or falls in insensible gradations without any break, we have nothing to measure the variations of pitch, nothing by which we can compare the latter with the earlier sounds, and comprehend the extent of the change.[5]... *The musical*

---

[2]Java, for example [4], [5], [6]. (Square brackets indicate references.)

[3]Some explanations and predictions were generated by a primitive version of the model described here [7].

[4]The italics are Helmholtz's in [1].

[5]The reader can verify that it is virtually impossible to differentiate between two moderately slow glissandi between the same two pitches when one follows a linear and the other, say, an exponential curve.

> *scale is as it were the divided rod, by which we measure progression in pitch, as rhythm measures progression in time.*[6]
>
> Upon this resposes also the characteristic resemblance between the relations of the musical scale and of space, a resemblance which appears to me of vital importance for the peculiar effects of music. *It is an essential character of space that at every position within it like bodies can be placed, and like motions can occur.*[6] Everything that is possible to happen in one part of space is equally possible in every other part of space and is perceived by us in precisely the same way. This is the case also with the musical scale. *Every melodic phrase, every chord, which can be executed at any pitch, can also be executed at any other pitch in such a way that we immediately perceive the characteristic marks of their similarity.*[6]

In the third of the above quotations, a scale is a *divided rod.* However, for example, the major scale is *unequally divided.* The fourth quotation imposes in effect a conditions of "*like motions being possible at every position.*" One of the main motivations of this present work is to make precise the nature of the constraint which this condition imposes upon the (unequal) division of the rod.

Helmholtz's attention was centered upon European music rather than, say, Javanese music,, which uses musical intervals comprised of tones with irrational frequency ratios not approximating those to be found low in the overtone series and which employs instruments which produce inharmonic partials. This focus upon harmonic music blurred the distinction between his belief that the entire musical scale was "the divided rod by which we measure progression in pitch" and his conviction that "scale degrees" were chosen so that musical intervals which correspond to simple frequency ratios would result. It is unclear whether it is the intervals or the scale which provides the measurements and, although he states the latter, his explanations rely throughout on the former view ([1], pp. 253, 370):

> Let us inquire, then, what motive there can be for selecting one tone rather than another in its neighborhood for the step succeeding any given tone. We remember that in sounding two tones together such a relation was observed. We found that under such circumstances certain particular intervals, namely the consonances, were distinguished from all other intervals which were nearly the same, by the absence of beats. Now some of these intervals, the Octave, Fifth, and Fourth, are found in all the musical scales known.[7]
>
> Moreover, it is by no means a merely external indifferent regularity which the employment of diatonic scales, founded on the relationship of compound tones, has introduced into the tonal material of music, as, for instance, rhythm introduced some such external arrangement into the words of poetry. I have shown, on the contrary, in Chapter XIV, that this construction of the scale furnished a means of measuring the intervals of their tones, so that the equality of two intervals lying in different sections

---

[6] The italics are mine.

[7] False, of course (even if the fifths and fourths are defined by reasonable approximations to simple ratios) e.g., Thailand and Indonesia.

of the scale would be recognised by immediate sensation. Thus the melodic step of a Fifth is always characterised by having the second partial tone of the second note identical with the third of the first. This produces a definiteness and certainty in the measurement of intervals for our sensations, such as might be looked for in vain in the system of colours, otherwise so similar, or in the estimation of mere differences of intensity in our various sensual perceptions.

Here we investigate the possible function of the "musical scale" (from our point of view "chords" are a special case of scales) as a "reference frame" or "context" which both provides measurements of musical intervals and serves to identify the "degrees" of its member tones. The treatment is independent of the rationality of the frequency ratios which correspond to the musical intervals between the degrees of the scale. The "rules of musical composition" employed by different musical cultures (e.g., the "rules of voice leading") are interpreted as restrictions which preserve such properties of the musical scales employed as are required to measure its intervals and/or permit identification of its degrees. (e.g., the prohibition of approaching a "tritone" by "skip" in early European music will be so analyzed.) Different techniques of musical composition, both in different musical cultures and within the same culture, are considered to be structural (i.e., pattern-producing) devices which are here studied only in terms of the way in which their use is restricted by such of the above properties as characterize each scale and thereby determine the recognizability of patterns produced by the application of each technique to the musical materials inherent in that scale. (e.g., the relative absence of melodic ("modal") sequences in compositions using the Javanese "Pelog" scale or the "Hungarian Minor" scale (in the West) will be interpreted in this manner.)

We have said above that a fundamental theme of this work is the investigation of that property of musical scales which assures that [1] "every phrase, every chord, which can be executed at any pitch, can also be executed at any other pitch in such a way that we immediately perceive the characteristic marks of this similarity." This property (which is obviously pertinent to the recognizability of melodic ("modal") and harmonic sequences) determines a classification of musical scales into "proper" and "improper" ones. This notion will be introduced in the first of this series of papers, which deals solely with the ways in which scales may provide measurement of intervals (independently of the frequency ratios between the degrees of the scales).

The second paper develops quantitative measures of the degree to which a scale provides such measurement and also of the degree to which rapid identification of the scale elements corresponding to its different "degrees" is facilitated by the structure of each scale. This property has an explanatory function with regard to the *inequality* of intervals in most musical scales and is obviously relevant to "tonality" and the differing importance (i.e., "function") of differing scale degrees. These two quantitative measures are then related to the perceptibility of musical structural relations resulting from the use of corresponding techniques of musical composition.

The third paper extends the theory to apply to the perceptions of a listener (such as a contemporary Western musician) who is familiar with many different musical scales and their corresponding styles.

The fourth paper describes relevant techniques of computation and contains samples from an available table of the computed values of relevant properties (those above and other related ones) of both existing and possible scales which may be extracted from thirty-one tone equal temperament (twelve was done earlier). This table may be examined to verify that (close approximations to) the known scales of existing musical cultures are optimal choices (given the timbres of the instruments in those cultures) and for the purpose of constructing listener experiments to test the theory. *The primary purpose of the whole work is to suggest new materials for musical composition in a manner which exposes their relevant musical properties*, and it is suggested that the table be utilized for this purpose.

The fifth of the initial series of papers is a generalization setting forth a theory of properties preserved when a continuous space is represented by a discrete space and is intended for application to the perception of phonemes of a spoken language, to certain aspects of visual perception and *primarily, to the perception of musical timbres in the context of a given set of such timbres*. Again, the primary purpose is the development of new musical materials for musical composition (of "tone-color" music).

A later paper in this series will relate musical *syntax* to the above properties. The perception of rhythmic patterns will also be treated.

The author owes a special debt of gratitude to John Myhill who, unasked, assumed the formidable task of rewriting a major portion of the material which follows. Whatever clarity it now possesses was his contribution to the opaque presentation of a novice. The original version was produced in 1965, and apologies must be given for a few out-of-date remarks and for the omission of references to a good deal of pertinent recent work.

1. The Coding of Relevant Parameters

Clearly our perception of music is, in large part, learned. That is, (1) a listener must extract from the raw sensory input (i.e., focus his attention on) relevant properties of and relations between portions of the data (in effect, such properties must also be determined) and (2) he must code (classify and label) the relevant parameters of these properties and relations in a manner suitable for interrogation by means of feedback (checking whether his choice of relevant parameters was correct) and suitable for further processing (e.g., combining "intervals" to form "motifs", "motifs to form "phrases", etc.). Thus by the "coding" of a sensory input is meant the partitioning of the stimuli from that input into classes each of whose elements is equivalent in musical function. Such coding is, of course, different for each musical culture (as it is for each linguistic culture; i.e., different sounds form equivalent phonemes in different spoken languages). Since coding places stimuli in equivalence classes, ambiguity is avoided. Also, to the extent that music resembles a linguistic system, the efficiency with which the constructed codes can carry relevant information is most important.

2. The Initial Ordering

Although the perception of pitch and timbre (musical "tone color") are interdependent[8], the model presented in this paper deals only with the perception of pitch when variation in timbre is sufficiently restricted so that discrete tones can be simply ordered (to be discussed later). The simultaneous perception of pitch and timbre when both may vary freely requires the use of the general model to be described in the fifth paper in this series.

First the mapping from sensory inputs to a "code" will be considered. In effect, a possibly continuous space (over which physical stimuli range) is mapped into a space of discrete points (the classifications of such stimuli).

We will assume that a subject can determine which of two tones with similar quality (tone color) is higher (in frequency) or that they are the same. Such determinations are, of course, within some tolerance. That is, two tones, which differ by less than some listener-dependent tolerance, may be considered the same.[9] Thus a simple ordering is possible of a series of discrete tones of similar quality. Since such stimuli carry information in most music, it is from the relations between these that a code must be selected. (It is not necessary at this point to state precisely what such "information" is. It is sufficient that in most music such phenomena as "wrong notes" and "being out of tune" exist).

The question which now must be considered is "What is being coded?" Since few people have absolute pitch, we may conclude that the isolated frequencies of the tones are not directly mapped into a code. However, we do know that, given that all tones have similar quality, each pair, (x,y) forms a "musical interval"[10] which can be compared with all other such "intervals" for equivalence. This, together with the condition that all tones can be simply ordered by pitch, permits comparison of all pairs sharing an endpoint and finally of all other pairs as well. That is, to compare (z,w) with (x,y) add u on the same side of z as w so that (x,y) is perceived as "similar (in size) to" (z,u) (denoted by "$(x,y) \sim (z,u)$"). Then determine whether u is internal to (z,w) and, if so, define (x,y) as "smaller than" (z,w) (denoted by "$(x,y) < (z,w)$"). In this manner a preordering (i.e., transitive, connected, and reflexive) of all musical intervals between tones in a stimulus may be obtained.[11] Such a preordering will be referred to as an *initial ordering*.[12]

However, it is by no means clear that, on the simplest morphological level, "interval" size is the parameter that is coded. We also know that a series of stimuli forms a "reference frame" or "context". Such a reference frame is often

---

[8] It may be worth noting here that pilot experiments have been performed in which subjects have selected an interval of frequency ratio 7:5 (between tones) as equal to a previously heard interval of ratio 3:2 when appropriate changes in timbre were made between the initial hearing and the selection.

[9] We temporarily assume that such pairs do not occur in sequences long enough to produce intransitivities in the equality relation.

[10] Note that a "musical interval" is a pair, not an "interval" in its usual mathematical meaning.

[11] No metric assigning distance values to each interval is assumed. This is consistent with present neurological evidence. [8], [9], [10], [11], [12].

[12] It may be possible to assign timbres to each tone such that there exist four intervals, i_1, i_2, i_3, i_4 where, apparently $i_1 < i_2 \sim i_3 < i_4$ and $i_1 \sim i_4$. Such cases (if they can indeed be constructed) are properly treated by the general model to be presented later.

called a "scale" or "chord". (Motifs and other morphological units on a higher level are not relevant at this stage). It may be that the degrees (successive elements) of such a scale are coded (do, re, mi, fa, etc.). However, since absolute pitch is rare, such degrees cannot be determined unless the "scale" is determined, for which it is either necessary to discriminate significant "interval" sizes or to sharply emphasize one particular tone (the tonic) so as to fix its position. (Otherwise one could not determine which scale degree (element) was the first, the second, etc.)

Let us state our assumptions precisely. S (the set of all tones or pitches) is simply ordered (by the property of being perceived as "higher" or "lower" in pitch). A relation $\leqslant$ (the initial ordering) is defined on $S \times S$ which is a preorder, i.e., *transitive*, *connected* and *reflexive*. Define $(xy) \sim (zw)$ to mean $(xy) \leqslant (zw) \wedge (zw) \leqslant (xy)$ and $(xy) < (zw)$ to mean $(xy) \leqslant (zw) \wedge \neg((zw) \leqslant (xy))$. Require that always $(xy) \sim (yx)$, and moreover

$$x < y < z \Rightarrow (xy), (yz) < (xz),^{13} \tag{2.1}$$

i.e., any proper subinterval of an interval (xz) is smaller than (xz). This completes the list of our assumptions about the orderings on S and on $S \times S$. When musical timbre is such that these assumptions are *not* satisfied, this model does *not* apply, although the general model to be presented subsequently applies to a large number of such cases.

Let us consider further reasonable requirements, which are satisfied in most musical applications, *but will not be assumed in the following* except where explicitly noted: In case the relations $<$ on S and *equality* ("$\sim$") of intervals are taken as primitive, the initial ordering can be constructed out of them by the method suggested a couple of paragraphs back, i.e., we can define

$$(xy) < (zw) \Leftrightarrow (\exists u)((xy) \sim (zu) \wedge (z < u < w \text{ or } w < u < z)) \tag{2.2}$$

and this definition will guarantee that 2.1 holds. However, 2.1 holds also in some models in which 2.2 does not, and it is conceivable that such models may find application to situations in which changes in timbre with pitch lead to violations of 2.2. For an example of a model in which 2.1 is satisfied but 2.2 is not, let $S = \{A, B, C\}$, $A < B < C$ (these are the letter names for musical tones, not variables), and $(BC) < (AB) < (AC)$; then there is no u between A and B for which $(BC) \sim (Au)$, contradicting 2.2. (Intuitively we would explain this by a 'missing note' between A and B.) Even in models which do satisfy 2.2. the requirement of *additivity* may not be satisfied, i.e. we may not have

$$x < y < z \wedge u < v < w \wedge (xy) \sim (uv) \wedge (yz) \sim (vw) \Rightarrow (xz) \sim (uw). \tag{2.3}$$

For example, let S_1 be the set of all points on a "V"-shaped figure with an obtuse angle, let $a < b$ mean that a lies to the left of b and let $(ab) < (cd)$ mean that the distance from a to b (measured along a straight line) is less than that from c to d; then 2.1–2.2 are satisfied but 2.3 is not.

[13]Note that "$<$" denotes relative pitch height when applied to elements of S, and denotes relative interval size when applied to elements of $S \times S$.

For convenience, when 2.3. holds, we define addition and subtraction by

$$x<y<z\to[(xz)\sim(xy)+(yz)]$$
$$\wedge[(xy)\sim(xz)-(yz)]\wedge[(yz)\sim(xz)-(xy)]. \quad (2.4)$$

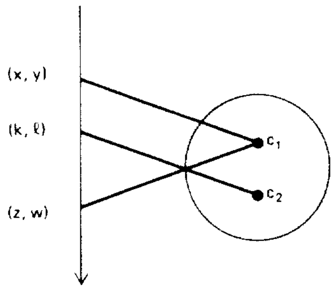
3. Proper Mappings

We will now consider all possible ways of defining a function which will map $S\times S$ (all pairs of tones forming musical intervals) into a set of discrete points, C (the code), by some function F: $S\times S\to C$. These mappings will be classified according to whether the following conditions are met for all $x,y,z,w,k,l\in S$ where $(xy)\leqslant(zw)$:[14]

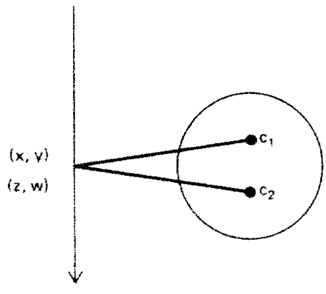
$$[F(xy)=F(zw)]\wedge[(xy)<(kl)<(zw)]\Rightarrow F(kl)=F(xy) \quad (3.1)$$

$$(xy)\sim(zw)\Rightarrow F(xy)=F(zw) \quad (3.2)$$

A pictorial representation of a violation of condition (1) and (2) is convenient. (In this diagram, if $(xy)\sim(zw)$, both will be represented as a single point on line $S\times S$).



(1) is violated



(2) is violated

Thus condition (1), restated, says that all points within an interval in $S\times S$ whose endpoints have the same image, will also have that same image. Hence, in this case, the capability for discriminating differences between musical "intervals" (pairs) interior to such intervals in $S\times S$ is not required. Since the discrimination of pitches differing by less than some tolerance, t, is not possible, the ordering of musical "intervals " which differ by less than such tolerance is also impossible. Thus it would appear that condition (1) is essential where (xy) and (zw) differ by less than t. It will be later shown that t is dependent upon F when F is defined by the configuration of points in a "reference frame".

[14]In the general model to be discussed subsequently $(xy)>(zw)\to F(x,y)>F(z,w)$ is used in place of 3.1. Later we will see, that given our assumptions and the definition of F to be used, these are equivalent.

Henceforth mappings satisfying condition (1) as well as their domains will be called *proper*, those satisfying (1) and (2) will be called *strictly proper* and all others *improper*. Elements (xy), (zw) of $S \times S$ not satisfying condition (2) will be called *ambiguous*, and each element (kl) of $S \times S$ not satisfying condition (1) will be called *contradictory*.

A mapping is proper iff the inverse image of any point of C is a (connected) *piece* of $S \times S$, i.e. a set $A \subset S \times S$ for which

$$(xy)<(kl)<(zw)\wedge(xy),(zw)\in A \Rightarrow (kl)\in A.$$

For such a mapping, all ambiguous tones are at the endpoints of the connected pieces. It is strictly proper iff there are no such tones.

Since a particular set of tones usually comprises the material of a given musical composition or section thereof, and since at least some of these tones form a local "reference frame", usually called a "musical scale" or "chord", a set $P \subset S$, the reference frame, will replace S in the preliminary discussion. Till further notice P will be assumed discrete, countable and infinite in both directions, and indexed according to the ordering on S (see 2.1), i.e.

$$P \equiv \{\ldots,p_{-2},p_{-1},p_0,p_1,p_2,\ldots\}.$$

The initial ordering of $S \times S$ induces an ordering of $P \times P$; the function $f: P \times P \to Z$ defined by

$$f(p_i,p_j) \equiv |i-j|,$$

which we can call the *diatonic distance* of p_i and p_j, suggests itself[15] as a natural candidate for a coding of the intervals (p_i,p_j) into the natural numbers. (Intuitively, this means we measure the distance between two notes of the scale P by counting the number of scale notes between them.) We easily prove the

Proposition. (a) *$f: P \times P \to$ positive integer is proper iff*

$$|i-j|<|k-l| \Rightarrow (p_i,p_j) \leqslant (p_k,p_l)$$

(b) *f is strictly proper iff*

$$|i-j|<|k-l| \Rightarrow (p_i,p_j) < (p_k,p_l)$$

Note that 3.1 *follows trivially from*

$$\text{(c)} \quad (xy)<(zw) \to F(xy) \leqslant F(zw)$$

regardless of the definition of F. However (c) *does not follow from* (3.1) *unless $F \equiv f$ as defined above. When $F \equiv f$* (2.2) *is* not *required to prove that* (c) *follows from* (3.1) (*and hence the equivalence of* (c) *and* (3.1):

[15]Justification for the choice of f is given in the description of the general model.

Proof. Suppose (c) is false when $F \equiv f$, i.e. $p_ip_j < p_kp_l$ and $f(p_i,p_j) > f(p_k,p_l)$. By the definition of f, since P is simply ordered $\exists q(f(p_i,p_q) = f(p_k,p_l) \wedge [(p_i \subset p_q \subset p_j) \vee (p_j \subset p_q \subset p_i)])$, i.e. $p_q \in P$. But by (2.1), $(p_i,p_q) < (p_i,p_j)$. Hence (3.1) is violated, i.e., $f(p_i,p_q) = f(p_k,p_l) \wedge (p_i,p_q) < (p_i,p_j) < (p_k,p_l)$ and $f(p_i,p_j) > f(p_k,p_l)$. □

Intuitively, f is proper if knowing the number of scale steps between two scale notes is equivalent to knowing a certain range between which the musical interval of those notes must lie—i.e., if it makes sense to speak of (diatonic) "seconds", "thirds", "fourths", etc. *relative to the scale P*; condition (a) is simply that no second should be bigger than a third, no third bigger than a fourth, etc. f is *strictly* proper if knowing the musical interval between two scale notes determines the number of intervening scale steps, i.e. all (diatonic) seconds are strictly *less* than all thirds, all thirds less than all fourths, etc. (not merely less than or equal to). A scale which is proper but not strictly proper contains ambiguous intervals, i.e. equal intervals which contain different numbers of scale notes between the endpoints. However, the amount of ambiguity is limited: an ambiguous interval can have at most *two* sizes, measured by diatonic distance.

To familiarize the reader with the notions of propriety and strict propriety, which are fundamental to all that follows, consider a few scales. The ordinary *major scale* can be represented as

...C		D		E		F		G		A		B		C		D		E...
	2		2		1		2		2		2		1		2		2	

where the numbers written underneath represent the distance in semitones between adjacent notes. We tabulate; in the left column are listed diatonic distances, in the right the possible corresponding intervals (measured in semitones). Asterisks mark the ambiguous intervals.

DIATONIC DISTANCE		INTERVAL	
"Second";	$f=1$	1 (E-F, B-C);	2 (C-D etc.)
"Third";	$f=2$	3 (E-G etc.);	4 (C-E etc.)
"Fourth";	$f=3$	5 (C-F etc.);	6* (F-B)
"Fifth";	$f=4$	6*(B-F);	7 (C-G etc.)
"Sixth";	$f=5$	8 (E-C etc.);	9 (C-A etc.)
"Seventh";	$f=6$	10 (D-C etc.);	11 (F-E, C-B)
"Eighth";	$f=7$	12	

Obviously the mapping f which assigns to any scale-interval its diatonic distance is proper (no second is larger than any third, etc.) but not strictly proper; the half-octave is ambiguous since it can represent a diatonic distance of either a fourth or a fifth. Notice that the diatonic distance in the left-hand column is given in the usual musical terminology: this is greater by one than the value of $f(p_i,p_j) \equiv |i-j|$. This convention will be used with other scales below: i.e. to say two scale-notes have a diatonic distance of a "second" is to say there are no scale notes between them, to say they have a distance of a "third" is to say there is one scale-note between, and so on. Names of diatonic distances will be written in quotes when this usage conflicts with standard musical terminology. Instead of saying that the mapping f is proper but not strictly proper we shall sometimes say that the major scale itself is proper but not strictly proper; likewise for the other scales about to be discussed.

For an example of a strictly proper scale, consider Chinese pentatonic, i.e.,

...	C		D		E		G		A		C		D	...
		2		2		3		2		3		2		

The table is

DIATONIC DISTANCE	INTERVAL
"Second"	2 (C-D etc.) ; 3 (E-G, A-C)
"Third"	4 (C-E) ; 5 (D-G etc.)
"Fourth"	7 (C-G etc.) ; 8 (E-C)
"Fifth"	9 (C-A, G-E); 10 (D-C etc.)
"Sixth"	12

Finally here is an example of a scale which is not proper at all—the "Japanese Pentatonic":

...	A		B		C		E		F		A		B	...
		2		1		4		1		4		2		

with the table

DIATONIC DISTANCE	INTERVAL
"Second"	1 (B-C, E-F); 2 (A-B); 4 (C-E, F-A)
"Third"	3 (A-C) ; 5 (B-E); 6 (F-B)
"Fourth"	6 (B-F) ; 7 (A-E etc); 9 (C-A)
"Fifth"	8 (A-F, E-C); 10 (B-A); 11 (C-B, F-E)
"Sixth"	12

This scale is improper because e.g. the "second": C-E is bigger than the "third" A-C.

We define (when P is finite)[16]

$$\delta_{ij} \equiv (p_{i+j}, p_j)$$

$$\underline{\delta}_i \equiv \min_j \delta_{ij}, \qquad \bar{\delta}_i \equiv \max_j \delta_{ij}$$

and *row tolerance* T_i and *minimum tolerance* T as follows:

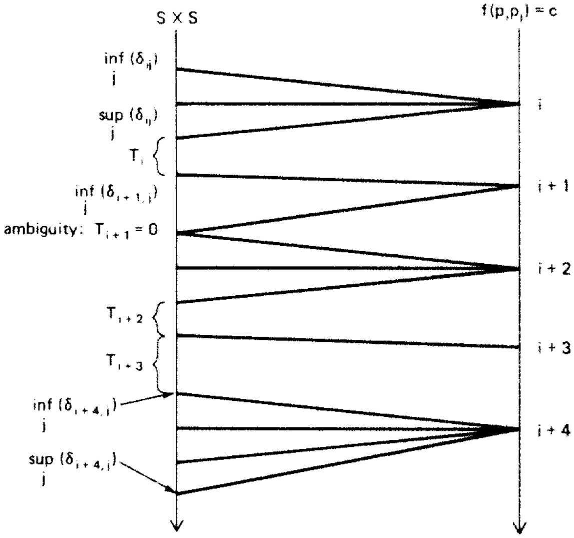
$$T_i \equiv \begin{cases} 1 & \text{if } \underline{\delta}_{i+j} > \bar{\delta}_i \\ 0 & \text{if } \underline{\delta}_{i+1} \sim \bar{\delta}_i \\ -1 & \text{if } \underline{\delta}_{i+1} < \bar{\delta}_i \end{cases}$$

[16]Note that δ_{ij} denotes that interval which has lower endpoint p_j and diatonic distance i.

By Propositions (a) and (b):

$$T \equiv \min T_i = \begin{cases} 1 & \text{for a strictly proper scale} \\ & (\delta_{i+1,j} > \delta_{i,k},\quad j,k,i=1,2,3,\ldots) \\ 0 & \text{for a proper, not strictly proper one} \\ & (\delta_{i+1,j} \geqslant \delta_{i,k},\quad j,k,i=1,2,3,\ldots) \\ -1 & \text{for an improper one} \end{cases}$$

If 2.2 applies and $(p,g) \sim \underline{\delta}_{i+1}$ and $(r,s) \sim \bar{\delta}_i$, there is an $x \in S$ with $(g,x) \sim T_i$. If 2.3 applies, we define $T_i \equiv \underline{\delta}_{i+1} - \bar{\delta}_i$ instead of as above ($\bar{\delta}_0 \equiv 0$ by definition) and again $T = \min T_i$. The following diagram illustrates a proper mapping:



Note that f, when considered as a "distance function" provides a pseudometric on P; i.e., the following conditions are satisfied:

(a) $f(p_i,p_j) = f(p_j,p_i)$

(b) $f(p_i,p_i) = 0$

(c) $f(p_i,p_j) + f(p_j,p_k) \geqslant f(p_i,p_k)$

"$\geqslant$" in (c) is really "$=$", of course.

In applications to the perception of pitch, distortions of perception of musical "intervals" corresponding to ambiguous pairs (in $S \times S$) may be expected. However, these "intervals" may be perceived in terms of other musical "intervals" which are not ambiguous but which have an endpoint in common with the ambiguous "interval" in question (more will be said of this later). In fact, music is written so that this is possible. (E.g. in music using the "major scale" in which the "tritone" is ambiguous, skips from either of the tritone's endpoints are consistently avoided so that it may be perceived as a "perfect fourth" or "perfect fifth" according to the tone(s) adjacent to its endpoint(s) (i.e., if the tritone is FB and the B moves to a C or an A or if the F moves to an E or a G, it is a "fourth"; if the B moves to an D^{b} or a B^{b} or if the F moves to $D^{\#}$ or an $F^{\#}$ it is a "fifth").

4. Mapping from $P \times S$ into C

We have described in the previous section a subset P of the "stimulus-space" S; P is to be thought of as a "scale" used in measuring distances between points of S. If the distance between two points $p_i, p_j \in P$ is to be measured, we simply count the number of intervening points of P, i.e. $f(p_i, p_j) \equiv |i-j|$. Now we want to extend f so as to be able to "measure" the distance between any two points of S, or at least of as many pairs as possible.

We shall approach this problem in stages. First consider the case where we only want to measure the distances from points of P to points of S-P. The natural way to do this is to divide S (or as large a subset of S as possible) into neighborhoods $R_1, R_2, R_3, \ldots$ with each $p_i \in R_i$. The distance of $x \in R_j$ from p_i is $|i-j|$; i.e. to measure the distance from a "scale" tone to a "chromatic" tone we measure its distance from the "scale tone" of which the latter is felt to be an "alteration".

Formally we proceed as follows: A *modification* of P is defined as an assignment to each p_i of a "neighborhood", i.e. a set $R_i \subset S$ containing p_i. The *induced distance-function* g is defined by $g(p_i, x) \equiv |i-k|$ when $x \in R_k$, and the modification $\{R_i\}$ is *proper* if g is. (The definition of g only makes sense if the R_i are disjoint; this is so in all interesting cases. Cases where the R_i are not disjoint will be treated later.) We do not in general require $\cup_i R_i = S$, but we seek to make $\cup_i R_i$ as large as possible.

We also seek to make $\{R_i\}$ proper. Clearly this can be done only if P is; so assume P proper and let $p_i \in P$. The *range* R_i of p_i is the set of all $x \in S - P$ for which $P' = (P - \{p_i\}) \cup \{x\}$ is still proper (where, of course, f is applied to the elements of p_i as indexed according to the ordering on S). We prefer that the range of each p_i is an interval $\subset [p_{i-1}, p_{i+1}]$; the elements of $R_i - [p_{i-1}, p_{i+1}]$ are called *wild points*.

Example worked out. Let S be *all* pitches (the whole real line) and let P be the major scale

	-1	$-\gamma$	0	δ	2	4	5	7	9	11	12	
...												...
	B		C		D	E	F	G	A	B	C	

Here the seconds	range from	1–2 semitones
Here the thirds	range from	3–4 semitones
Here the fourths	range from	5–6 semitones
Here the fifths	range from	6–7 semitones
Here the sixths	range from	8–9 semitones
Here the sevenths	range from	10–11 semitones

What is the range of C[17], i.e. how far can we change the pitch of C and still have a proper scale? If we change its pitch *downwards*, say by γ semitones, then:

The seconds	will range from	$1-\gamma$ to	$2+\gamma$ semitones
The thirds	will range from	$3-\gamma$ to	$4+\gamma$ semitones
The fourths	will range from	$5-\gamma$ to	$\max(6,5+\gamma)$ semitones
The fifths	will range from	$\min(6,7-\gamma)$ to	$7+\gamma$ semitones
The sixths	will range from	$8-\gamma$ to	$9+\gamma$ semitones
The sevenths	will range from	$10-\gamma$ to	$11+\gamma$ semitones

For propriety it is necessary and sufficient that $2+\gamma \leqslant 3-\gamma$, $4+\gamma \leqslant 5-\gamma$, etc. i.e. $\gamma \leqslant 1/2$. If on the other hand we change the pitch of C upwards, say to $+\delta$, then:

The seconds will range from $\min(1,2-\delta)$ to $\max(2,1+\delta)$

The thirds will range from $\min(3,4-\delta)$ to $\max(4,3+\delta)$

The fourths will range from $5-\delta$ to $\max(6,5+\delta)$ etc.

For propriety, $\max(4,3+\delta) \leqslant 5-\delta$ i.e. $\delta \leqslant 1$ is necessary: if we tabulate the possible sizes of fifths, sixths, and sevenths in the scale $(P-\{C\}) \cup \{\delta\}$ we see that $\delta = 1$ is sufficient, and so *the range of* C is $[-1/2, +1]$. Likewise we compute:

The range of D is $\left[1\frac{1}{2}, 2\frac{1}{2}\right]$

The range of E is $\left[3, 4\frac{1}{2}\right]$

The range of F is $[5,6]$

The range of G is $\left[6\frac{1}{2}, 8\right]$

The range of A is $\left[8, 9\frac{1}{2}\right]$

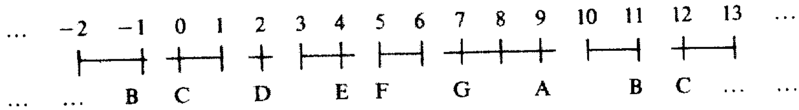
The range of B is $[10,11]$.

[Note: if D is pushed up to $D^{\#}$, the "second" C-$D^{\#}$ is larger than the "third" $D^{\#}$-F; if it is pushed down to $D^{\flat}$, the "second" $D^{\flat}$-E is larger than the "third" B-$D^{\flat}$: either way the resulting scale is improper. Note also that if F is lowered *at*

[17]Henceforth "C" will denote the musical pitch by that name, rather than the "code" previously mentioned, unless explicitly noted to the contrary.

all, or B raised at all, the "fourth" F-B becomes larger than the "fifth" B-F; hence F can only be modified upwards, and B only downwards].

In this case (major scale) *the range of each note is an interval containing that note*:



which is what we want; however, there are gaps not covered by this partition (i.e. $\cup \bar{R}_i \neq S$).

It is unfortunately not in general true however that the partition into ranges has this neat appearance. In particular consider the following *non-periodic* scale (i.e. it does not repeat at the octave).

...	C	C$^\#$	D	E	F$^\#$	G$^\#$	B^b	...	(*)
...	0	1	2	4	6	8	10	...	

formed by adjoining *one only* C$^\#$ to a whole-tone scale. This scale is proper, for

The seconds range from 1–2 semitones
The thirds range from 2–4 semitones
The fourths range from 4–6 semitones
The fifths range from 6–8 semitones
The sixths range from 8–10 semitones

and in general the $n$ths ($n \geqslant 3$) range from $2n-4$ to $2n-2$ semitones. Evidently replacement of C$^\#$ by any note in the scale would not disturb propriety. Thus the range of C$^\#$ is not an interval, but includes the "wild notes" D$^\#$, F, G, A, B, all of which are separated from C by one or more scale notes. The above scale has a plethora of ambiguous intervals: more specifically it has as many as possible, i.e. for $n=2,3,4,5$ there are diatonic n^{ths} which are equal (in the sense of $\sim$) to diatonic $n+1^{\text{sts}}$. This motivates two definitions and a theorem.

Definition. *$q \in S - P$ is called wild if for some $i, q \in R_i$ but $q \notin [p_{i-1}, p_{i+1}]$.*

Definition. *P is called highly ambiguous if for $i = 1,2,3,\ldots$ there exist j,k with $(p_{j-1}, p_{j+1}) \sim (p_k, p_{k+i})$.*

Theorem. *A necessary condition for the existence of wild notes is that P be highly ambiguous.*

Proof. Let q be wild, say $q \in \bar{R}_t - [p_{t-1}, p_{t+1}]$. Without loss of generality we can assume $q > p_{t+i}$. Let p_s (where $s \leqslant t+i$) be the highest note of P which is below q. To show that P is highly ambiguous, it suffices to prove that

$$(p_{t-1}, p_{t+i}) \sim (p_s, p_{s+i}) \tag{1}$$

for each i > 0. Since the former interval has diatonic distance $i+1$ (in P) and the latter has diatonic distance i and P is proper

$$(p_{t-1},p_{t+1}) \geqslant (p_s,p_{s+i}). \tag{2}$$

Let $P' = P - \{p_t\} \cup \{q\}$; P' is proper since $q \in \bar{R}_t$. The diatonic distance from p_{t-1} to p_{t+1} in P' is i, because p_t has been left out; and that from p_s to p_{s+i} is $i+1$, because q has been put in. By propriety

$$(p_{t-1},p_{t+1}) \leqslant (p_s,p_{s+i})$$

which combined with (2) yields (1) and hence the theorem. □

Note that if $T_i > 0$ for *any* i, no wild notes exist.
The following facts are just as easily proved:
If P is proper and without wild notes, then

I. The range of each scale-note (i.e. element of P) is on interval about that scale note;
II. The intersection of the ranges of two consecutive scale notes is either empty or consists of a single point;
III. If *every* two consecutive ranges intersect, then

$$\cup_i \bar{R}_i = S.$$

Note that the union of the ranges may not be the whole of S.

The work of computing ranges can be shortened and systematized as follows: Let P be proper and without wild notes, and let it be required to compute the range $\bar{R}_k$ of p_k. Define for $i = 1,2,3,\ldots$, $\underline{\delta}_i^+ =$ the smallest (according to the initial order on $S \times S$) "$i+1^{st}$" (i.e. (p_l,p_m) such that $|l-m| = i$) except possibly the one whose right endpoint is p_k.

$\underline{\delta}_i^- =$ the smallest "$i+1^{st}$" except the one whose left endpoint is p_k,

$\bar{\delta}_i^+ =$ the largest "$i+1^{st}$" except the one whose left endpoint is p_k.

$\bar{\delta}_i^- =$ the largest "$i+1^{st}$" except the one whose right endpoint is p_k.

We have easily for $x \in \bar{R}_k$

$$\left.\begin{aligned} (p_{k-i},x) &\leqslant \underline{\delta}_{i+1}^+ \quad (3)\\ (x,p_{k+i}) &\leqslant \underline{\delta}_{i+1}^- \quad (4)\end{aligned}\right\} \quad i = 0,1,2,\ldots$$

$$\left.\begin{aligned} \bar{\delta}_i^+ &\leqslant (x,p_{k+i+1}) \quad (5)\\ \bar{\delta}_i^- &\leqslant (p_{k-i-1},x) \quad (6)\end{aligned}\right\} \quad i = 1,2,3,\ldots$$

Consider for example (6); the interval (p_{k-i-1},x) is a diatonic $i+2^{\text{nd}}$ in the scale $P'=P-\{p_k\}\cup\{x\}$; by propriety it cannot be less than any diatonic $i+1^{\text{st}}$ in that scale, and in particular not less than any diatonic $i+1^{\text{st}}$ in P except possibly one of those with endpoint p_k, which do not exist in P'. However neither can it be less than (p_k,p_{k+1}); for if it were either (Case I) $x<p_k$; but then the diatonic $i+1^{\text{st}}$ (x,p_{k+i}) in P' would be greater than (p_k,p_{k+i}) which would in turn be greater than the $i+2^{\text{nd}}$ (p_{k-i-1},x), so that P' would be improper; or else (Case II) $x>p_k$ and $(p_k,p_{k+i})\geqslant(p_{k-i-1},x)>(p_{k-i-1},p_k)\geqslant(p_k,p_{k+i})$ (by propriety of P), contradiction. Thus the only diatonic $i+1^{\text{st}}$ in P which (p_{k-i-1},x) can be less than is (p_{k-i},p_k); but that is exactly what (6) says. (3)–(5) are proved similarly.

It would help greatly in the solution of our problem if for proper P without wild notes we could prove that the modification $\{\overline{R}_i\}$ was proper. In fact this is not so, in general, even if P is strictly proper. *Counterexample*: Let P be the four-note scale CEGB, thus

−5	−1	0	$4-\gamma$	4	$4+\delta$	7	11	12	16	19	23	...
G	B	C		E		G	B	C	E	G	B	...

We leave it to the reader to verify strict propriety and the absence of wild notes.

The "seconds" range from 1–4 semitones
The "thirds" range from 5–7 semitones
The "fourths" range from 8–11 semitones
The "fifths" are all 12 semitones

We inquire as to the range of the note E. If we change its pitch *downwards*, say by γ semitones, then:

The "seconds" range from $\min(1,4-\gamma)$ to $\max(4,3+\gamma)$ semitones
The "thirds" range from $5-\gamma$ to $7+\gamma$ semitones
The "fourths" range from $\min(8,9-\gamma)$ to 11 semitones

For propriety we must have $4\leqslant 5-\gamma, 7+\gamma\leqslant 8, 7+\gamma\leqslant 9-\gamma$, i.e., $\gamma\leqslant 1$. If we *raise* E, say by δ semitones, then:

The "seconds" range from $\min(1,3-\delta)$ to $4+\delta$ semitones
The "thirds" range from $\min(5,7-\delta)$ to $\max(7,5+\delta)$ semitones
The "fourths" range from $8-\delta$ to $\max(11,9+\delta)$ semitones

and we need $4+\delta\leqslant 5$, $4+\delta\leqslant 7-\delta$, $7\leqslant 8-\delta$, $5+\delta\leqslant 8-\delta$, $9+\delta\leqslant 12$, i.e. $\delta\leqslant 1$. Hence the range of E is $[\text{E}^\flat,\text{F}]=[3,5]$.

Now consider the modification $\{\overline{R}_i\}$. It induces the distance-function g, defined by

$$g(p_i,x)=|i-k| \qquad \text{for } x\in\overline{R}_k.$$

In particular $g(\mathrm{C},\mathrm{F})=1$ while $g(\mathrm{B},\mathrm{E}^\flat)=2$. But $(\mathrm{C},\mathrm{F})>(\mathrm{B},\mathrm{E}^\flat)$, so g, and hence $\{\bar{R}_i\}$, is improper. So our hope, that we could make proper modifications, and hence extend f: $P\times P\to \mathrm{C}$ to g: $P\times S\to \mathrm{C}$ by simply taking the ranges as our neighborhoods, is not supported. However, a variation of the range idea will now be shown to work.

Suppose a scale P is given, proper and without wild notes, and suppose we wish to determine a maximum proper modification of P. This means putting an interval R_k round each $p_k\in P$ which is as large as possible consistent with propriety. Consider for example the "scale"

$$\begin{array}{cccccccc} \cdots & p_{-1} & p_0 & p_1 & p_2 & p_3 & p_4 & \cdots \\ \cdots & \mathrm{B} & \mathrm{C} & \mathrm{E} & \mathrm{G} & \mathrm{B} & \mathrm{C} & \cdots \end{array}$$

just discussed. We must pick, for each of the four notes CEGB, both an upper and a lower limit of its R_i, and define R_i as the closed interval bounded by these two limits. There are thus eight points to be chosen altogether, two for each note of the "scale". There are 8! orders in which these eight decisions can be made and these will lead to 8! maximal proper modifications of P not all of which are distinct.

Our chosen order might be for example

Upper limit of R_{E}

Lower limit of R_{E}

Upper limit of R_{C}

Lower limit of R_{B}

Upper limit of R_{G}

Lower limit of R_{C}

Lower limit of R_{G}

Upper limit of R_{B}

and at each step we use inequalities analogous to (3) and (5) above to determine an upper limit, and inequalities analogous to (4) and (6) to determine a lower one.

To find the upper limit of R_{E} (henceforth we set $R_i\equiv[p_i^-,p_i^+]$) we must satisfy

$$(p_{1-i},\mathrm{E}^+)\leqslant \underline{\delta}_{i+1} \qquad (i=0,1,2,\ldots) \tag{3'}$$

and

$$\bar{\delta}_i\leqslant(\mathrm{E}^+,p_{i+2}) \qquad (i=1,2,3,\ldots) \tag{5'}$$

analogous to (3) and (5) (here $\underline{\delta}_i$ and $\bar{\delta}_i$ are the smallest and largest $i+1^{\mathrm{sts}}$, respectively, in P. Also we must have

$$(p_{1-i},\mathrm{E}^+)\leqslant(\mathrm{E}^+,p_{i+2}) \qquad (i=1,2,3,\ldots)$$

Joint satisfaction of these three inequalities is equivalent to propriety of the modification

$$\{C\}, [E, E^+], \{G\}, \{B\}.$$

It is easily seen that $E^+ = F$ is the highest E^+ satisfying these conditions.

Now we want to choose an E^- so that $\{C\}, [E^-, E^+]\ \{G\}, \{B\}$ is proper. The corresponding inequalities are

$$(E^-, p_{i+1}) \leqslant \underline{\delta}_{i+1} \qquad (i=0,1,2,\dots) \tag{4'}$$

$$\bar{\delta}_i \leqslant (p_{-i}, E^-) \qquad (i=1,2,3,\dots) \tag{6'}$$

$$(E^-, p_{i+1}) \leqslant (p_{-i}, E^-) \qquad (i=1,2,3,\dots)$$

which gives $E^\flat$ as a lower bound on E^-. This is not strict however, for we saw above that

$$\{C\}, [E^\flat, F], \{G\}, \{B\} \tag{**}$$

is not a proper modification of C, E, G, B. This is because the above modification (**) calls the interval $(B, E^\flat)$ a "third" and the larger interval (C, F) a "second". Hence E must not only satisfy the given inequalities but also

$$(p_{1-i}, F) \leqslant (p_{-i}, E^-) \qquad (i=1,2,3,\dots)$$

which gives $E^- = E$ as the only possible value *compatible with* $E^+ = F$.

Proceeding in this way we obtain successive modifications

$$R^{(3)} = [C, C^+], [E, F], \{G\}, \{B\}$$

$$R^{(4)} = [C, C^+], [E, F], \{G\}, [B^-, B]$$

$$\vdots$$

$$R^{(8)} = [C^-, C^+], [E, F], [G^-, G^+], [B^-, B^+]$$

of which the last is evidently maximal.

General algorithm. Let $P = \{p_1, \dots, p_n\}$ be proper and without wild notes. Let $\alpha = \{(q_i, \sigma_i)\}$ be a sequence of $2n$ terms such that each p_j occurs exactly twice amongst the q_i, once with the corresponding $\sigma = +$ and once with the corresponding $\sigma = -$. In the example just treated, α is

$$((1,+), (1,-), (0,+), (3,-), (2,+), (0,-), (2,-), (3,+)).$$

Then the maximal proper modification of P induced by α (notation R^α) is obtained recursively as follows:[18]

$$R^\alpha = [p_1^-, p_1^+], \dots, [p_n^-, p_n^+]$$

[18]We do not claim that every maximal proper modification is induced by such an α.

where $p_{a_k}^{\sigma_k}$ (with k indexing the p_k in the sequence, α and a_k indexing the p_i in P) is the greatest x satisfying

$$(p_{a_k-i},x) \leqslant \underline{\delta}_{i+1}^k \qquad (i=0,1,2,\ldots)$$

$$\bar{\delta}_i^k < (x,p_{a_k+i+1}) \qquad (i=1,2,3,\ldots)$$

and

$$(p_{a_k-i},x) < (x,p_{a_k+i+1}) \qquad (i=1,2,3,\ldots)$$

if σ_k is $+$, or the least x satisfying

$$(x,p_{a_k+i}) \leqslant \underline{\delta}_{i+1}^k \qquad (i=0,1,2,\ldots)$$

$$\bar{\delta}_i^k < (p_{a_k-i-1},x) \qquad (i=1,2,3,\ldots)$$

and

$$(x,p_{a_k+i}) < (p_{a_k-i-1},x) \qquad (i=1,2,3,\ldots)$$

if σ_k is $-$; and where $\bar{\delta}_i^1 = \bar{\delta}_i, \underline{\delta}_i^1 = \underline{\delta}_i$,

$$\bar{\delta}_i^{M+1} = \begin{cases} \max\left(\bar{\delta}_i^M, (p_{a_k-i}, P_{a_k}^+)\right) & \text{if } \sigma_M = + \\ \max\left(\bar{\delta}_i^M, (p_{a_k}^-, p_{a_k+i})\right) & \text{if } \sigma_M = - \end{cases}$$

and

$$\underline{\delta}_i^{M+1} = \begin{cases} \min\left(\underline{\delta}_i^M, (p_{a_k}^+, p_{a_k+i})\right) & \text{if } \sigma_M = + \\ \min\left(\underline{\delta}_i^M, (p_{a_k-i}, p_{a_k}^-)\right) & \text{if } \sigma_M = -. \end{cases}$$

If R^α is defined as above and $R_k^\alpha \equiv [p_k^-, p_k^+]$ then evidently

$$\bar{R}_k = \bigcup_\alpha R_k^\alpha; \qquad \bigcup_k \bar{R}_k = \bigcup_\alpha R^\alpha$$

We now define $\bar{R} = \cup_k \bar{R}_k$, $\underline{R}_k = \cap_\alpha R_k^\alpha$, and $\underline{R} = \cup_k \underline{R}_k$, and call $\underline{R}_k$ the *blur* of p_k. We have evidently, for all α, $\underline{R} \subset R^\alpha \subset \bar{R}$. For any x,y with $p_{k-1} < y < p_k < x < p_{k+1}$ define

$$\bar{\delta}_i^*(x,y) \equiv \max\Big(\bar{\delta}_i, \max_k \{(p_{k-i},x) | (p_{k-i},x) < (x,p_{k+i+1})\},$$
$$\max_k \{(y,p_{k+i}) | (y,p_{k+i}) < (p_{k-i-1},y)\}\Big)$$

$$\underline{\delta}_i^*(x,y) \equiv \min\Big(\underline{\delta}_i, \min_k \{(x,p_{k+i}) | (p_{k-i+1},x) < (x,p_{k+i})\},$$
$$\min_k \{(p_{k-i},y) | (y,p_{k+i-1}) < (p_{k-i},y)\}\Big)$$

Then the endpoints a,b of $\underline{R}_k$ satisfy

$$(p_{k-i},a)\leqslant\underline{\delta}^*{}_{i+1}(a,b)\qquad(i=0,1,2,\ldots)\tag{3''}$$

$$(b,p_{k+i})\leqslant\underline{\delta}^*{}_{i+1}(a,b)\qquad(i=0,1,2,\ldots)\tag{4''}$$

$$\bar{\delta}_i{}^*(a,b)\leqslant(a,p_{k+i+1})\qquad(i=1,2,3,\ldots)\tag{5''}$$

$$\bar{\delta}_i{}^*(a,b)\leqslant(p_{k-i-1},b)\qquad(i=1,2,3,\ldots)\tag{6''}$$

For each k let that portion of the range, $\overline{R}_k$ which is $\geqslant p_k$ be denoted by $\overline{R}_k^+$ and let the portion which is $\leqslant p_k$ be denoted by $\overline{R}_k^-$. These will be called *upper* and *lower ranges* respectively. We similarly define upper and lower blur ($\underline{R}_k^+$ and $\underline{R}_k^-$) and R_k^+ and R_k^- (where R is a proper modification). For any proper modification R, $R^+\equiv\{R_k^+\}$ and $R^-\equiv\{R_k^-\}$ will be called the *upper* and *lower half-modifications* respectively.

Since $\overline{R}_k\equiv\cup_\alpha\overline{R}_k^\alpha$ and $\overline{R}_k$ is an interval containing p_k, each $\overline{R}_k^+$ or $\overline{R}_k^-$ is a subset of R_k in some proper modification. Then, since for any i, $\underline{R}_i$ is common to all proper modifications $\{R^\alpha\}$, the set of blurs, $\{\underline{R}_i\}$ where for any k, $\underline{R}_k$ is replaced by $(\underline{R}_k\cup\overline{R}_k^+)$ or $(\underline{R}_k\cup\overline{R}_k^-)$, is a proper modification (but rarely a maximal proper modification). This is useful because it is far less laborious to determine $\overline{R}$ and $\underline{R}$ than R^α for all α.

Since axioms (2.2) and (2.3) hold for all familiar examples in the application to the perception of pitch, these will apply to all examples shown. The reader's attention will be drawn to some of the consequences of adding these axioms to our treatment, which may be easily verified (see preceding discussion).

$$\exists k(\delta_{2,k-1}\sim\delta_{1,j},\ \text{all}\ j\neq k-1,k)\Rightarrow\overline{R}_k\equiv S\tag{7}$$

When no wild points exist:

$$(\forall_j)(\delta_{1,j}\sim\delta_{1,j+2})\Rightarrow(\overline{R}\equiv S)\tag{8}$$

$$(\forall_j)(\delta_{1,j}\sim\delta_{1,j+1})\Rightarrow(\underline{R}\equiv\overline{R}\equiv S)\tag{9}$$

(7) states that when P is an equally spaced grid with one point added the range of that point is all of S (c.f. the scale (*) above). (8) assures us that $\overline{R}$ covers S when "intervals" between adjacent points in P form a repeating sequence of two alternating magnitudes only. (9) assures that $\underline{R}$ covers S when all "intervals" between adjacent points in P are equivalent.[19]

5. Assignment of "Distances" to Pairs of Non-Scale Tones (Mapping from (part of) $S\times S$ into C)

Let us recapitulate the discussion of the last section. We have supposed that a "scale" P was given, $P=\{\ldots,p_{-2},p_{-1},p_0,p_1,p_2,\ldots\}$ with distance-function $f(p_i,p_j)\equiv|i-j|$. We have shown how, under suitable conditions on P (propriety

[19]Note that when S is the real numbers, the implications in (8) and (9) may be reversed.

and absence of wild notes) f can be extended to a map g: $P \times R \rightarrow C$ which is still proper and indeed maximal amongst such proper mappings. This is done by assigning to each p_i a "neighborhood" R_i and then defining $g(p_i,x) \equiv |i-j|$ for $x \in R_j$. This enables us to measure the distance between any "scale tone" p_i and "non-scale tone" x. We now wish to construct a G which will measure the distance between two non-scale tones. This motivates a definition and a theorem.

Definition. *The modification* $\{R_i\}$ *is called S-proper if the mapping* G: $\cup_i R_i \rightarrow C$ *defined by*

$$G(x,y) \equiv f(p_i,p_j) = |i-j| \qquad (x \in R_i, y \in R_j)$$

is proper.

Theorem. *For any proper modification* $\{R_i\}$, *the modifications* $\{R_i^+\}$ *and* $\{R_i^-\}$ *are S-proper.*

Proof. Let $\{R_i\}$ be proper; then the mapping g defined by

$$g(p_i,u) = |i-j| \qquad (u \in R_j)$$

is proper. For each i let $R_i^+ = [p_i, x_i]$. We need to show that the function G defined by

$$G(x,y) = |i-j| \qquad (x \in R_i^+, y \in R_j^+)$$

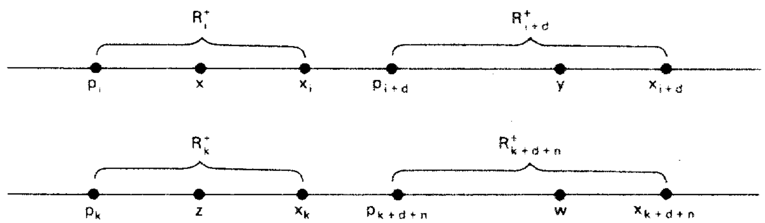
is proper, i.e. that

$$G(x,y) < G(z,w) \Rightarrow (x,y) \leqslant (z,w)$$

for all $x,y,z,w \in \cup R_k^+$. Let $x \in R_i^+$, $y \in R_{i+d}^+$, $z \in R_k^+$, $w \in R_{k+d+n}^+$; then the assumption $G(x,y) < G(z,w)$ means $n > 0$. If now, contrary to hypothesis, $(x,y) > (z,w)$, we would have

$$(p_i, x_{i+d}) \geqslant (x,y) > (z,w) \geqslant (x_k, p_{k+d+n})$$

(A glance at the diagram will make all this clear).



But $g(p_i, x_{i+d}) = d$ and $g(p_{k+d+n}, x_k) = d+n$; hence $(p_i, x_{i+d}) < (p_{k+d+n}, x_k)$ by propriety of g; hence the assumption that G is improper leads to a contradiction, q.e.d. Likewise for R^-. □

It will not in general be the case, however, that if R is a *maximal* proper modification then $R \equiv R^+$ is a *maximal* S-proper modification. However, it can be corrected into one by expanding each of the $R_i^+ = [p_i, x_i]$ downwards to $R_i = [y_i, x_i]$, the y_i being so chosen as to satisfy the simultaneous inequalities

$$(y_k, x_{k+j}) \leqslant (x_i, y_{i+j+1})$$

for all i, k and all $j \geqslant 0$,

$$(x_{k-j}, y_k) \geqslant (y_i, x_{i+j-1})$$

for all i, k and all $j > 0$. The successive y's are adjusted in an arbitrary order $y_{\alpha(1)}, \ldots, y_{\alpha(n)}$, where P has n notes; thus for each maximal proper modification R, R^+ induces $n!$ (not necessarily distinct) maximal S-proper modifications

$$R^\alpha \equiv \{[y_{\alpha(1)}, x_{\alpha(1)}], \ldots, [y_{\alpha(n)}, x_{\alpha(n)}]\};$$

so, dually, does R^-.

From the above it is clear that the union of all S-proper modifications is $\{\bar{R}_k\}$.

When axioms (2.2) and (2.3) hold, if c_1 and c_2 are constants, for all $x \in R^+$ and $y \in R^-$ let $x' = x - c_1$ and $y' = y + c_2$ and let $\overset{<+}{R}$ denote the set of all x' and $\overset{>-}{R}$ denote the set of all y'. When $P \subset \overset{<+}{R}$ (or $P \subset \overset{>-}{R}$) and when $0 \leqslant c_1$, $c_2 < \min_j \delta_{1j}$, $\overset{<+}{R}$ (or $\overset{>-}{R}$) will be called a *shifted half-modification*. Note that only strictly proper P have shifted half-modifications. All shifted half-modifications are S-proper modifications.

For realistic application all formulae should be adjusted so that a listener's inability to distinguish two pitches which differ by less than some listener-dependent tolerance, ϵ, is considered. Because of the resulting mathematical complication, this will be deferred till the end of the paper.

6. The Periodic Case, the Matrix $\|\alpha_{ij}\|$

P has period n if for all $i,j,(p_i p_k) \sim (p_{i+n} p_{k+n})$ and if n is the least positive integer satisfying the condition. That is, for all $i,j, \delta_{ij} \sim \delta_{i,j+n}$. When P is periodic the numbers of inequivalent pairs less than any given pair, $(p_i p_j)$ is obviously finite.

This guarantees that the positive integers suffice to index the rank order of all pairs, δ_{ij}, according to the preordering (initial ordering) on $P \times P$. Thus when P is periodic we define the matrix $\|\alpha_{ij}(P)\|$ (the argument P will usually be omitted) whose elements, α_{ij}, are the rank of each δ_{ij} in the preordering on $P \times P$ (i.e. the images under the mapping onto the positive integers). Also, when $i = 0$ we define $\alpha_{ij} = 0$. Since $\delta_{ij} \sim \delta_{i,j+n}$, all entries in the matrix $\|\alpha_{ij}\|$ are determined by columns $1, \ldots, n$.

Note that all formulae and propositions can be restated with an "α" substituted for "δ" with no loss of validity (since the mapping from $\|\delta_{ij}\|$ onto $\|\alpha_{ij}\|$ is order preserving). Also when a metric assigning values to each pair in $P \times P$ is known, the matrix $\|\delta_{ij}\|$, which contains such values, is also defined.

When axioms (2.2) and (2.3) apply, all entries in matrix $\|\alpha_{ij}\|$ are determined by the first $n-1$ rows of columns $1 \ldots n$:

$$\begin{matrix} \alpha_{1,1}, & \cdots & \alpha_{1,n} \\ \vdots & & \vdots \\ \alpha_{(n-1),1} & \cdots & \alpha_{(n-1),n} \end{matrix}$$

which we call the *reduced matrix* of P. Define $\bar{K}$ as the first α_{ij} which is independent of j (i.e. $\bar{K} = \alpha_{n,j}$ for any j). Then the reduced matrix is determined by $\bar{K}$ and the first $n(n-1)/2$ elements of the first $[n/2]$ rows[20] of $\|\alpha_{ij}\|$, since

$$\alpha_{n-i,i+j} = \bar{K} - \alpha_{ij} \qquad (i = 1, \ldots, n)$$

(proof is left to the reader—note that axioms 2.2 and 2.3 need not apply).

It is clear that in the periodic case no wild notes exist. For example in the ordinary major scale which repeats itself after 7 notes, all octaves are equal; in general, in any periodic scale that repeats itself after n notes, all $n+1^{st}$'s are equal —i.e., the scale is not highly ambiguous, which we say in Section 4 was a necessary condition for wildness.

In those periodic cases where a metric is given by embedding P in the reals so that the definitions of addition and subtraction (of musical intervals) apply in their usual numerical interpretation (i.e., axiom 2.3 holds and definition 2.4 applies), we have

$$(p_j p_{i+j}) = (p_i p_{i+1}) + (p_{i+1} p_{i+2}) + \ldots + (p_{i+j-1} p_{i+j}),$$

$$\delta_{ij} = \sum_{k=j}^{j+i-1} \delta_{1,k}$$

and so $\|\delta_{ij}\|$ may be specified by its first row

$$\psi(P) = (\delta_{1,i}, \delta_{1,2} \ldots \delta_{1,n}).$$

Also, in such cases, the formulae for the computation of $\bar{R}$ and $\underline{R}$ can be simplified and these are easily computed. Then $\bar{R}_k = [P_k - \theta_k^-, p_k + \theta_k^+]$, where, if all column subscripts are made positive (mod$n$; i.e., $j + n \equiv j$) and row $n+1$ is added to reduced matrix $\|\delta_{ij}\|$, we set

$$\theta_k^+ = \min_{q=1\ldots n} \left[(\delta^{\,+}_{-q+1} - \delta_{q,k}), (\delta_{q+1,k} - \delta_{q,k-q})/2 \right] \tag{1}$$

$$\theta_k^- = \min_{q=1\ldots n} \left[(\delta^{\,-}_{-q+1} - \delta_{q,k}), (\delta_{q+1,k-q-1} - \delta_{q,k})/2 \right] \tag{2}$$

The reader can also verify that $\underline{R}_k$ can be computed from the following simple formulae, where σ_k^+ and σ_k^- correspond to upper and lower bounds on

[20] In this discussion the use of square brackets "[ ]" indicates that the fractional portion of the enclosed quantity is to be dropped.

$(\underline{R}_k - p_k)$:

$$\left.\begin{array}{l}\sigma_k^+ = \min\limits_{i=1} (\underline{\delta}'_{i+1} - \delta_{i,k-i}) \\ \sigma_k^- = \min\limits_{i=1} (\underline{\delta}'_{i+1} - \delta_{i,k})\end{array}\right\} \qquad \underline{\delta}'_{i+1} = \min_k (\delta_{i+1,k} - \max(\theta_k^+, \theta_{k+i+1}^-))$$

All other inequalities in this section can easily be replaced by similar formulae (when the stated necessary conditions hold). These methods apply to all examples that will be shown.

In our application we must consider that most "musical scales" are duplicated at the octave, although many have a period which is a fractional portion of the octave. Since n is the period, $\exists K \forall_j (K \sim \delta_{n,j})$ and this K is the interval which comprises a cycle. However, it is convenient to represent intervals as fractions of the octave cycle (when such exists). Accordingly we define $\bar{n} = wn$ where w is the number of n-cycles in an octave[21] and we let $\psi(P)$ contain $\bar{n}$ terms:

$$\psi(P) = (\delta_{1,1}\delta_{1,2}\ldots\delta_{1,\bar{n}})$$

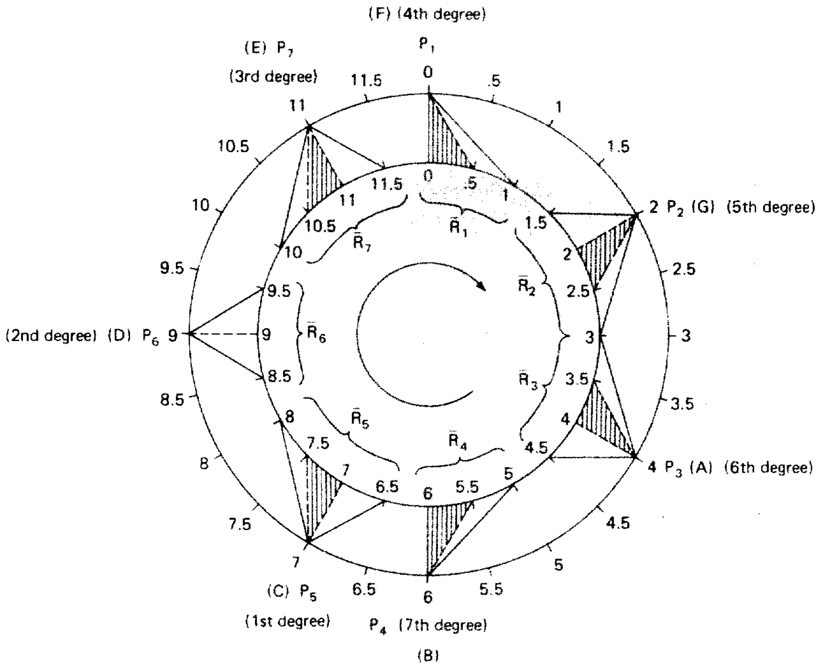
Then let $m = wK$ and m will be the number of units (e.g. cents or semitones) in an octave cycle. When m is an integer, it may be interpreted as an "equal temperament" system. Those $\psi(P)$ where $w > 1$ will be called *symmetric*.

Now some examples will be shown and the results will be interpreted:

Example 1. This is the familiar "major scale" of Western music, in its tuning in the 12-tone equal temperament system. S is assumed to be the reals, axioms (2.2) and (2.3) apply, and P is periodic ($\bar{K} = 12$)—hence a clockwise circular diagram is used where each pair is an arc on the circle. To avoid using a spiral representation, $(p_i p_j)$ and $f(p_i p_j)$ will be measured clockwise from p_i to p_j. Thus, instead of writing "$(p_i p_{j+n})$" when $j < i$, we write only "$(p_i p_j)$" which distinguishes it from $(p_j p_i)$, and all our arithmetic is mod $\bar{K} = 12$. Also, $f(p_i p_j) + f(p_j p_i) = n = 7$ (the number of elements of P in a period). In this example numerical values will be assigned to each $(p_i p_j)$ according to the number of semitones in each such musical "interval" (in 12-tone equal temperament). Note that in this example, for all $i, j, \delta_{ij} = \alpha_{ij}$.[22] The diagram shows the mapping from $\bar{R}$ (on the inner circle) onto P (on the outer circle). The numbers on the circles are elements of S. Elements of P are written on the outside of the outer circle. Solid lines show the boundaries of each $\bar{R}_i$ and dotted lines those of the $\underline{R}_i$ (the $\underline{R}_i$ are shaded in the diagram). P is indexed so that p_i is the "fourth degree" of the major scale (F in C major). (The reason for using the Lydian mode will be explained in the next article in this series.) For convenience, the musical names of each p_i when P is the C major scale are shown enclosed in brackets on the outer rim of the outer circle, as are the "degree" names of each p_i in conventional musical language:

[21] An octave need not always be used. It is possible to cycle at frequency ratios other than 2 : 1, although musical examples are extremely rare.

[22] $\|\alpha_{ij}\|$ is obtained from $\|\delta_{ij}\|$ by replacing the smallest element by 1, the next smallest by 2 and so forth. $\|\alpha_{ij}\| = \|\delta_{ij}\|$ because all chromatic distances (in the 12-tone system) occur as diatonic distances in the major scale.

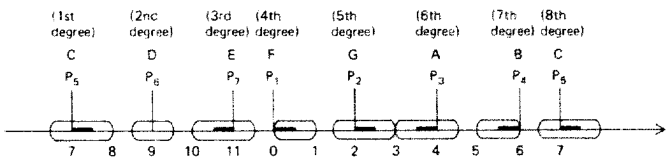


$$\|\alpha_{ij}\| = \|\delta_{ij}\| = \begin{bmatrix} 2 & 2 & 2 & 1 & 2 & 2 & 1 \\ 4 & 4 & 3 & 3 & 4 & 3 & 3 \\ \textcircled{6} & 5 & 5 & 5 & 5 & 5 & 5 \\ 7 & 7 & 7 & \textcircled{6} & 7 & 7 & 7 \\ 9 & 9 & 8 & 8 & 9 & 9 & 8 \\ 11 & 10 & 10 & 10 & 11 & 10 & 10 \\ 12 & 12 & 12 & 12 & 12 & 12 & 12 \end{bmatrix} \begin{matrix} T_0=1 \\ T_1=1 \\ T_2=1 \\ T_3=0 \\ T_4=1 \\ T_5=1 \\ T_6=1 \end{matrix}$$

{all entries in row zero = 0, see page 9}

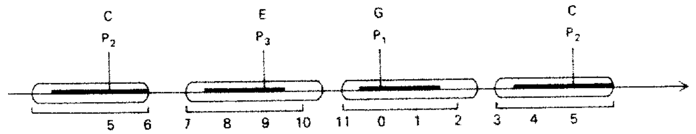
$T=0$, P is proper, (p_1,p_4) and (p_4,p_1) are ambiguous.

For convenience, the above diagram will be drawn in linear fashion showing only one cycle. Each $\bar{R}_i$ will be shown by brackets enclosed on top and bottom and each $\underline{R}_i$ by a darkened rod. This technique will also be used for ensuing examples:



Example 2. This is the familiar "major triad" in its tuning in the 12-tone equal temperament system (discussion of previous example applies). If P is a C-major triad, p_1 is the "fifth" of that triad (G). The proper modification obtained from

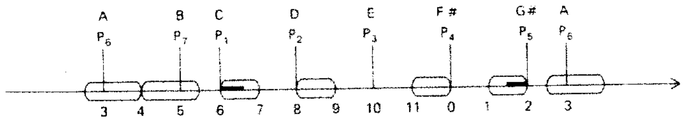
the stepwise maximization of $R_2^-, R_2^+, R_1^-, R_3^-, R_1^+, R_3^+$ is shown by "└──┘". (The same result will be produced by any other sequence where all R_k^+ and R_k^- whose corresponding extremes of $\bar{R}_k^+$ or $\bar{R}_k^-$ are quarter tones are computed last).



$$\|\delta_{ij}\| = \begin{bmatrix} 5 & 4 & 3 \\ 9 & 7 & 8 \\ 12 & 12 & 12 \end{bmatrix} \qquad \|\alpha_{ij}\| = \begin{bmatrix} 3 & 2 & 1 \\ 6 & 4 & 5 \\ 7 & 7 & 7 \end{bmatrix}$$

P is strictly proper (there are no ambiguous intervals). $\|\alpha_{ij}\|$ will be discussed in the next paper in this series.

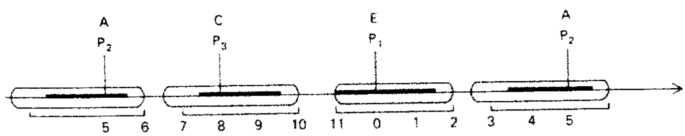
Example 3. This is the "melodic minor scale" in its ascending form in the 12-tone equal temperament system. Again, in this case $\|\alpha_{ij}\| = \|\delta_{ij}\|$. If P is an A melodic minor scale, p_1 is the "third degree" of that scale:



$$\|\alpha_{ij}\| = \|\delta_{ij}\| = \begin{bmatrix} 2 & 2 & 2 & 2 & 1 & 2 & 1 \\ \textcircled{4} & \textcircled{4} & \textcircled{4} & 3 & 3 & 3 & 3 \\ \textcircled{6} & \textcircled{6} & 5 & 5 & \textcircled{4} & 5 & 5 \\ \textcircled{8} & 7 & 7 & \textcircled{6} & \textcircled{6} & 7 & 7 \\ 9 & 9 & \textcircled{8} & \textcircled{8} & \textcircled{8} & 9 & 9 \\ 11 & 10 & 10 & 10 & 10 & 11 & 10 \\ 12 & 12 & 12 & 12 & 12 & 12 & 12 \end{bmatrix}$$

P is proper, all encircled α_{ij} correspond to ambiguous intervals.

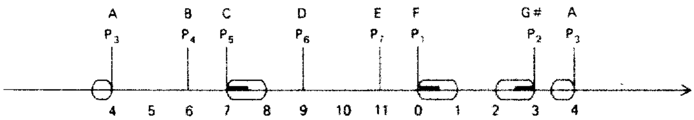
Example 4. This is the familiar "minor triad". If P is an A minor triad, p_i is the "fifth" of that triad (E). The proper modification obtained from the stepwise maximization of $R_1^-, R_1^+, R_2^+, R_3^+, R_2^-, R_3^-$ is shown by "└──┘". (Any other sequence described in Example 2 will produce the same result).



$$\|\delta_{ij}\| = \begin{bmatrix} 5 & 3 & 4 \\ 8 & 7 & 9 \\ 12 & 12 & 12 \end{bmatrix} \qquad \|\alpha_{ij}\| = \begin{bmatrix} 3 & 1 & 2 \\ 6 & 4 & 5 \\ 7 & 7 & 7 \end{bmatrix}$$

P is strictly proper, ($\|\alpha_{ij}\|$ will be discussed in the next paper in this series).

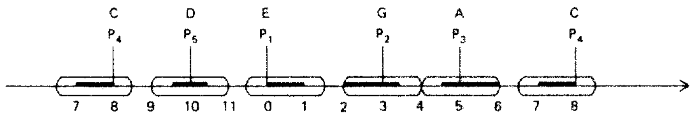
Example 5. This is the "harmonic minor scale" in 12-tone equal temperament. If P is an A harmonic minor scale, p_1 is the "sixth degree" of that scale.



$$\|\alpha_{ij}\| = \|\delta_{ij}\| = \begin{bmatrix} \textcircled{3} & 1 & 2 & 1 & 2 & 2 & 1 \\ \textcircled{4} & \textcircled{3} & \textcircled{3} & \textcircled{3} & \textcircled{4} & \textcircled{3} & \textcircled{4} \\ \textcircled{6} & \textcircled{4} & 5 & 5 & 5 & \textcircled{6} & 5 \\ 7 & \textcircled{6} & 7 & \textcircled{6} & \textcircled{8} & 7 & 7 \\ \textcircled{9} & \textcircled{8} & \textcircled{8} & \textcircled{9} & \textcircled{9} & \textcircled{9} & \textcircled{8} \\ 11 & \textcircled{9} & 11 & 10 & 11 & 10 & 10 \\ 12 & 12 & 12 & 12 & 12 & 12 & 12 \end{bmatrix}$$

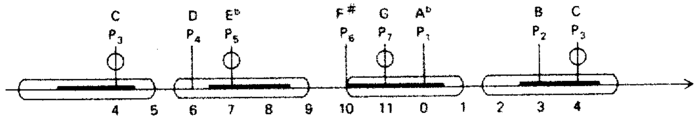
P is proper, but all encircled δ_{ij} are ambiguous.

Example 6. This is a "Chinese Pentatonic" scale tuned accoriding to 12-tone equal temperament. If $P \equiv \{C, D, E, G, A\}$, $p_1 \equiv E$.



$$\|\delta_{ij}\| = \begin{bmatrix} 3 & 2 & 3 & 2 & 2 \\ 5 & 5 & 5 & 4 & 5 \\ 8 & 7 & 7 & 7 & 7 \\ 10 & 9 & 10 & 9 & 10 \\ 12 & 12 & 12 & 12 & 12 \end{bmatrix} \qquad \|\alpha_{ij}\| = \begin{bmatrix} 2 & 1 & 2 & 1 & 1 \\ 4 & 4 & 4 & 3 & 4 \\ 6 & 5 & 5 & 5 & 5 \\ 8 & 7 & 8 & 7 & 8 \\ 9 & 9 & 9 & 9 & 9 \end{bmatrix}$$

Example 7. This is a "Hungarian minor" scale in 12-tone equal temperament. If $P \equiv \{C, D, E^{b}, F^{\#}, G, A^{b}, B\}$, $p_1 = A^{b}$ (the "sixth degree" of the scale). P is improper. Hence the ranges and blurs shown in the diagram are those of the minor triad on the "first degree" of the scale, to be discussed in Section 8. That is, this triad $\equiv \{p_3, p_5, p_7\}$ (or $\{C, E^{b}, G\}$—these are marked by circles on the diagram).



$$\|\alpha_{ij}\| = \|\delta_{ij}\| = \begin{bmatrix} [3] & 1 & \textcircled{2} & 1 & [3] & 1 & 1 \\ \textcircled{4} & \textcircled{3} & \textcircled{3} & \textcircled{4} & \textcircled{4} & [2] & \textcircled{4} \\ \textcircled{6} & \textcircled{4} & \textcircled{6} & 5 & 5 & 5 & 5 \\ 7 & 7 & 7 & \textcircled{6} & \textcircled{8} & \textcircled{6} & 7 \\ [10] & \textcircled{8} & \textcircled{8} & \textcircled{9} & \textcircled{9} & \textcircled{8} & 8 \\ 11 & [9] & 11 & \textcircled{10} & 11 & [9] & 11 \\ 12 & 12 & 12 & 12 & 12 & 12 & 12 \end{bmatrix}$$

P is improper, all bracketed δ_{ij} are contradictory. All encircled δ_{ij} are ambiguous.

7. Musical "Scales" and "Intervals"

In this application S will correspond to the set of pitches (with certain restricted timbres—see Section 2) which may occur as stimuli. P is a subset of S which is used to classify (measure) "musical intervals" (pairs) in $P \times S$. Note that, in this application (except when using S-proper modifications) we do not measure intervals in $S \times S$ with those $P \subset S$ containing no endpoint of the interval (see last paragraph of Section 3). When we judge a musical interval to be "sharp" or "flat", we mean with respect to one of the constituent tones, usually the lower. Similarly, when measuring a physical distance with a ruler, one of the markers of the ruler is placed at an endpoint of this distance. This analogy may be useful in interpreting much of the discussion which follows.

P may be intuitively interpreted as a musical "scale", but in the sense of "ruler", *not* in that musical usage which implies tonality (root) or mode (in French "echelle" not "gamme"). It may also sometimes be similarly interpreted as a "chord". In psychological terminology P corresponds to a "reference frame". How such reference frames are extracted from sequences of stimuli is treated later.

Since P provides measures (classifications), distortions of perception of ambiguous intervals may be expected. Alternatively, these intervals may be perceived in terms of (i.e., measured as one greater or one less than) other intervals which are not ambiguous but which have an endpoint in common with the ambiguous interval in question (e.g. if P is the major scale, the tritone ($\delta_{3,1}=6$) may be perceived in terms of a perfect fifth ($\delta_{4,1}=7$) or perfect fourth ($\delta_{3,4}=5$) in the same column. Thus (except when a chord is being outlined) it would be expected that skips from an endpoint of an ambiguous interval would be avoided (as is actually the case with the tritone in the "major scale"). Thus an ambiguous interval is one whose tones lack independence of adjacent tones in the scale (see Example 8, which follows).

Again, if P is considered a scale (in the sense of "ruler"), then "tonality" (in its common musical usage) corresponds to that marker of the ruler from which the measurements, $f(p_j x)$, are performed. But each column of $\|\alpha_{ij}\|$ shows all distances in $P \times P$ from such a common endpoint, p_j. Then if tonality (the tonic) never changes, even momentarily (as in modal sequences), we need consider the array $\|\alpha_{ij}\|$ as containing only one column, j (p_j is the tonic), and obviously no ambiguous distances can exist. Such is the case when an ostinato is used or when (as in some Indian music) a drone is employed.[23] However, when first one scale element, then another is emphasized so that each temporarily appears to be a "tonic" or "root" (as in modal sequences in the major scale), ambiguous intervals can occur. So that such cases can be conveniently described, the term "temporary tonic" will here refer to an element which temporarily assumes prominence, so that (for that time period) it becomes, in effect, an endpoint of all intervals measured. These may occur when motivically similar passages beginning on different scale "degrees" (elements) succeed each other, when harmony changes, or when different scale elements are rhythmically emphasized in succession. Since, in musical usage, the "tonic" of a key of a scale is usually considered fixed, the term "temporary tonic" is more nearly approximated by the musical term "root" (although "root" usually refers only to principal tones in "chords"). However, the composer Paul Hindemith's well-known notion of "degree progression" (both melodic and harmonic) [14] corresponds rather closely to that of "sequence of temporary tonics". When scales composed primarily of intervals forming irrational frequency ratios are considered (these occur frequently in the Orient), Hindemith's notion of "step progression" resembles that of "sequence of temporary tonics".

If the tonic of an improper scale is fixed, clearly neither ambiguous nor contradictory intervals are evident. This reasoning can be extended: Again, to all pairs appearing in column j of $\|\alpha_{ij}\|$, p_j is a common endpoint. Suppose columns are eliminated from $\|\alpha_{ij}\|$ so that $T \geqslant 0$ for the remaining matrix. If the endpoints corresponding to eliminated columns are not used as temporary tonics in conjunction with the endpoints corresponding to the remaining columns, contradictory intervals in the scale will not be exposed. (That is, measurements from certain markers of the ruler are avoided so that contradictory measurements do not occur.) Similarly, ambiguous intervals may be avoided by the elimination of columns of $\|\alpha_{ij}\|$. Hence the exposure of contradictory or ambiguous intervals can be avoided by either fixing the tonic of an improper scale or by sharply restricting the use of temporary tonics. (Expected consequences of these disturbances in perception are described in the section dealing with listener experiments in a later paper in this series.)

Note that in music where harmony is of importance, a tone that forms intervals with other tones in a scale of such a kind that it is reinforced by the resulting difference tones and overtones is more easily used as a tonic or root than tones for which such is not the case. Provided enough is known about the timbre of the tones being considered to determine those intervals which produce

[23] This can account for the phenomenon familiar to musicians that almost any pitches may be played above an ostinato.

strong difference tones reinforcing their roots, an examination of each column, j, of any $\|\delta_{ij}\|$ for the values of such intervals will show which p_j's are easily used as tonics or roots.[24] In such cases it is often convenient to think of the K corresponding to such $\|\delta_{ij}\|$ as an equal temperament system (although such need *not* be the case).

Example 8. The following may be helpful in obtaining an intuitive impression of the difference between ambiguous and unambiguous intervals; play the following tones on a piano repeatedly with the tonic emphasized and/or on bottom of the sequence until the metal impression of the scale is firmly fixed: E♭, F, G, A♭, B♭, C, D, E♭; E♭ is the tonic (although any tonic will do, E♭ makes it the familiar major mode). Now play in sequence and listen carefully to: A♭, C. This interval is unambiguous in this scale. Now repeat the entire procedure using the scale: E, F#, G#, A, B, C, D, E; E is the tonic (E is supported as the tonic by difference tones with G# and B and serves to expose the interval, E, G#). Now again play in sequence: G#, C (same as A♭, C). (G# is momentarily a "tonic" or "root" (in the sense of "being measured from") when G#, C is played). This interval is ambiguous in this scale (and many listeners will have the impression that it is larger than when played after the first scale). Any attempt to explain the effect as resulting from the augmented triad: E, G#, C will be obviated by repeating the procedure with scale: C, D, E, F#, G#, A#, C, which also contains the triad but in which the interval: G#, C is not ambiguous.

8. Alternative and "Auxiliary" Tones, Improper "Scales"

Many attempts have been made to account for the predominance of the familiar modes on the first degree ("major") and the sixth degree ("minor") of the "major" scale (Example 1). All of these arguments rely upon harmonic considerations. So also do arguments that justify the "seventh degree leading to the tonic", "the supertonic (second degree) leading to the mediant (third degree)", "the subdominant (fourth degree) leading to the mediant" and "the submediant (sixth degree) leading to the dominant (fifth degree)". Less successful have been attempts to account for the variations on the minor mode (Examples 3 and 5). Here we shall offer additional explanations consistent with the others, but which do not so heavily rely upon harmonic considerations.

For a considerable period of time, many modes of the major scale (Example 1) were in use. Gradually notes were "altered" and added at special places in the music so that many pitches not in the scale were occasionally present. These pitches naturally were referred to certain pitches in the scale ("degrees" of the scale), since the "reference frame" remained esentially the same. These added pitches were in the ranges (R_k) of those already in the scale (see Example 1) and

[24]In many cases such an investigation of difference tones and overtones will show overwhelming dominance of one tone. This tone will usually be a tonic and modes of the scale will rarely be used when harmony is involved.

each could substitute for some degree of the scale (as determined by $\bar{R}_k$). Some such substitutions would change the key (scale) itself and cause a modulation. Others would not. Also, pitches could be added without a substitution, in which cases they functioned as "auxiliary tones" or "ornamental tones". An examination of the $\bar{R}_k$ for different k shows that if such tones are added at random (double sharps and double flats included) the first, third, fifth and the sixth degrees of the scale (p_5, p_7, p_2, p_3) would be most emphasized thereby.[25] (Range is appropriately used since these auxiliary tones were generally added one at a time. The union of P and any upper or lower range is a proper modification—see Section 4). Note that, of these, the first and third degrees (together) and the first and fifth degrees (together) generate difference tones reinforcing the first degree. The sixth and third degrees generate a difference tone reinforcing the sixth. No other such reinforcements occur. Thus it is not surprising that the first degree would receive most emphasis and the sixth next most (the major and minor mode).

Now consider the major mode. The first, third and fifth degrees, as has been seen, reinforce the tonic. These degrees also form a strictly proper subset of the major scale (Example 2—the major triad), and hence may function as an independent "reference frame". Note that all the elements of the major scale are in the ranges ($\bar{R}_k$) of elements of this major triad. However, more than one element of the major scale which is foreign to this major triad may be present at one time (e.g. a tonic ninth chord). For this reason proper modifications (rather than ranges) are appropriately used. Reference to Example 2 will show that all tones of the major scale are also contained in a proper modification, R. Hence if the tones of the major triad are considered as principal tones and the remaining tones as auxiliary to those tones, p_k, in whose R_k they are contained, the traditional rules for the "resolution" of one degree of the major scale to another are derived. That is, by resolving each auxiliary tone to the principal tone in whose R_k it falls, we see that, (as is traditional) when a tonic triad is used the leading tone (7th degree) leads to the tonic (1st degree), the supertonic (2nd degree), leads to the mediant (3rd degree), the subdominant (4th degree) leads to the mediant, and the submediant (6th degree) leads to the dominant (5th degree).

Clearly, when harmony changes in the use of the major scale, a different proper (perhaps strictly proper) subset temporarily supplies principal tones and the tendencies of the remaining tones can be determined as above. As we shall see later, the proper and strictly proper subsets of the major and minor scales together with the respective ranges (and R_k) of each of their elements supply the materials and rules of the traditional "figured bass" system characteristic of Western music of the Baroque and Classical periods.

Now if we turn our attention to the minor mode of the major scale (Example 1, still), the sixth degree becomes the first degree, and all other elements assume degree names in cyclic order. Again if we select the minor triad composed of the resulting first, third, and fifth degrees to supply principal tones (Example 4) we will arrive at the same pattern of degree resolution as we did for

---

[25]Note that adding F$^\#$ or B$^\flat$ might confuse the key. Not counting these strengthens the argument.

the major scale (which is in agreement with traditional theory and practice).

If we examine the $\bar{R}_k$ in Example 1, we will observe that $G^{\#}$ is in the range of G and that $F^{\#}$ is in the range of F. The substitution of $G^{\#}$ for G results in the "harmonic minor" scale (Example 5), and the *subsequent* substitution of $F^{\#}$ for F (which remains valid) results in the "melodic minor" scale (Example 3). (Substituting $F^{\#}$ for F *initially* results only in another major scale). Note that (in Example 1) $G^{\#}$ is common to the $\bar{R}_k$ corresponding to both G and A (since we here consider 12 tone equal temperament, $G^{\#} = A^{\flat}$). Hence it is not surprising that we discover that a majority of the "intervals" in the harmonic minor scale (which derives from a substitution of this tone only) are ambiguous, rendering it unsuitable for use with temporarily changing tonics (temporary tonics). This is not the case with the melodic minor scale, although it has many more ambiguous intervals than the "natural" minor scale (the mode on the sixth degree of the major scale). Thus we see that by substituting $F^{\#}$ for F and/or $G^{\#}$ for G we are still within the ranges of elements of a single scale (reference frame), and the common practice of using the melodic minor scale in ascending musical passages and the natural minor scale in descending musical passages serves to avoid the exposure of ambiguous intervals. In general, the interpretation of a musical scale in terms of a set of ranges ($\bar{R}_k$) instead of a set of precise ptiches, obviates all explanations for (and resulting confusions of) the common 18th century practice of using any or all of the twelve tones of the chromatic scale with a major or minor scale. Also, the practice in the Baroque period of ending a composition in the minor scale with a major triad consists only of a valid substitution for the third degree. This is true regardless of which version of the minor scale is used.

Note that in the "Chinese pentatonic" scale (Example 6), both from the viewpoint of reinforcement by different tones and ranges of tones, much freedom of choice of tonic exists. This is consistent with musical experience.

In order to understand the application to improper scales, we will reverse the argument used in interpreting the major scale. Suppose tones which are contained in one of its proper modifications are added to the major scale and an improper scale results. These added tones could then act as auxiliary tones to those of the major scale in much the same way in which elements of the major scale did for those of the major triad. Since the proper major scale would supply principal tones from which all intervals could be measured, the impropriety of the scale containing it would not, in this case, present a problem. Thus one method for using an improper scale (other than those already discussed) is to select a (strictly) proper subset (to act as principal tones) such that remaining elements are contained within a proper modification of such a subset (and act as auxiliary tones). The tendencies of such auxiliary tones would be determined by the R_k in which they are contained. The "Hungarian minor" scale (C, D, $E^{\flat}$, $F^{\#}$, G, $A^{\flat}$, B) (Example 7) is easily interpreted in such fashion. Since the union of the blur ($\underline{R}$) and any lower range ($\bar{R}_k$—choose the underside of $E^{\flat}$) is a proper modification, we see that all tones of the scale are contained in a proper modification of the minor triad (C, $E^{\flat}$, G), and the customary "resolutions" of B to C, D to $E^{\flat}$, $F^{\#}$ to G and $A^{\flat}$ to G result.[26]

[26]It should also be mentioned that ornamental tones may be added to an improper scale so that a proper scale results. At present no examples where such is clearly the case have been studied.

It is of interest to note that all tones of the 12-tone equal temperament system are covered by the ranges of the elements of each of the scales and chords in Examples 1-4. This is consistent with customary "chromatic" usage of these materials in Western music. It has been mentioned that in Western music (prior to the twentieth century), strictly proper and proper subsets of the major and minor scales form the "figured bass" system of harmony. This system determines, at each musical instant, the division of the elements of the scale into principal and auxiliary tones (harmonic and non-harmonic tones). Thereby the same tones may alternatively function (for given time periods) as principal and auxiliary tones. (Such changes of function are usually accompanied by changes of temporary tonic.) This is not the case when improper scales are used, where the partition of scale elements into principal and dependent tones is far less flexible.[27] Both the inclusiveness of the ranges of the elements of each of the proper subsets used and the compatibility of all the resulting temporary tonics (to be discussed in the next paper in this series) must be considered. Alternatively, the tonic may be fixed, thereby imposing greater restriction. Thus it is not surprising that "heterophony" (and "homophony") rather than "harmony" in the Western sense characterizes the music of cultures using improper scales (e.g., Java and Bali).

Earlier, much was said of Helmholtz's condition that "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". Clearly, we propose that proper scales have this property because of the consistency of measurement they provide. Hence we expect that "modal" melodic sequences in melodies would be immediately recognizable as such when proper scales are used (provided that the measurement of ambiguous intervals by intervals containing an adjacent endpoint is not frustrated by the motivic structure of the sequence). We expect this to be less the case when improper scales are used (except when the motif being sequenced forms compatibly measured sets of intervals at each of its repetitions). Indeed, in Western music, whereas motivic sequences characterize the use of the major, minor, "whole tone" and "twelve tone" scales (all proper), such is rarely the case when the "Hungarian minor" or other improper "synthetic" scales are used. In Java, where there are two systems of scales, "Slendro" and "Pelog", all Slendro scales thus far examined are strictly proper while all Pelog scales are improper. We have observed that the *ornamentation* of melodies in the Slendro scales often utilize motivic sequences, while this has not been observed in Pelog melodies. Mr. Surya Brata (B. Yzerdraat) of the Institute for Cultural Research, Musicological Project, and the Music Department of the Ministry of Education and Culture, Jakarta, was consulted, and confirmed these observations and commented as well on the "hestiating" and "unstable" quality of Pelog compositions

[27]Note that while it is sometimes possible to select many different proper subsets of improper scales, the continuous alternation between such subsets (as sets of principal tones) results in much greater "reference frame" changes than when this is done with proper scales. This is likely because of the greater weakness of the improper scale as a self-contained (independent of the tonic used) "reference frame". The reader can verify that changing harmonies which are proper subsets of an improper scale (e.g., the "Hungarian minor" scale), when playing melodies in that scale, tends to produce the effect of a modulation.

as opposed to Slendro. More will be said of this and of improper scales in the following papers in this series.

Consider how much a given proper scale, P, may be "mistuned" and still remain proper. If all elements of P are fixed and only p_k is altered, clearly $\underline{\bar{R}}_k$ limits this alteration. However, in any musical performance the tunings of *all* the tones in a scale vary. S-proper modifications are appropriately interpreted as providing limits on the simultaneous (or sequential) variation in tunings of all the elements of a proper scale. That is, since no pairs (x,y) selected from S-proper modifications are contradictory, any set of tunings of P which fall within such a modification may be simultaneously entertained without affecting the propriety of P. This is applicable only to proper sets when used with changing temporary tonics and to the proper subsets of both proper and improper scales when such function as reference frames. When S-proper modifications are applied to a proper subset of a scale they should not be simultaneously applied to the scale as a whole. For example, when C is the tonic of the C-major scale (the C-major triad forms the principal tones), and when B is moving to C, B may be raised. When B does not move to C, however, raising B sounds distinctly "out of tune". This is to be expected since no S-proper modification of the C-major scale exists which contains a raised B, but many S-proper modifications of the C-major triad contain a raised B. (All lower half-modifications (see Section 5) of the C-major triad contain a raised B; no half-modification of the C-major scale does, and since it is not strictly proper, neither do any of its shifted half-modifications—see Section 6, Examples 1 and 2). Another interesting application is to the Javanese "Slendro" scales. These contain five elements, are strictly proper, and are generally played on fixed pitch instruments (such as xylophones). The class of all Slendro scales examined thus far conforms to the limits imposed by $\underline{\bar{R}}$ (on any member of the class) and each vocal performance (studied thus far) of a particular Slendro tuning is limited by an S-proper modification.

Other limitations of tuning will be subsequently discussed which apply both to proper and improper scales (such as the "Hungarian minor" scale or the Javanese "Pelog" scale). More stringent limitations frequently result when improper scales are used.

We have not yet discussed why musical scales should be unequally divided. Our discussion of improper scales has been cursory and we have not indicated reasons why some scales exist in one or another musical culture and others neither are to be found anywhere nor appear amenable to use as "synthetic scales" (i.e., when experimentally constructed by composers). The next paper in this series will deal with these questions after developing quantitative measures by which scales may be evaluated with respect to the criteria discussed here as well as new criteria pertinent to their "information carrying" capacity. "Cadences" and related properties of Western music will also be discussed.

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* **5** (1961), 2–31.

3. J. C. R. LICKLIDER, Three auditory theories. In *Psychology: a study of a science* (ed. by S. Koch), 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 Javenese 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, D. Rothenberg, Technical Reports to Air Force Office of Scientific Research 1963-1969 (grants and contracts AF-AFOSR 881-65, AF49(638)-1738 and AF-AFOSR 68-1596).
8. G. M. STRATTON, Vision without inversion of the retinal image, *Psych. Rev.* **4** (1897). 341–60; 463–81.
9. R. H. THOULESS, Phenomenal 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*, Leipzig, Barth, 1932.
11. CARROL C. PRATT, Comparison of tonal distance and Bisection of tonal intervals larger than an octave, *Jour. of Experimental Psychol.* **11** 1928, 77–87, 17–36.
12. H. MUNSTERBURG, Vergleichen der Tondistanzen, *Beitrage zur Experimentelle Psychology* **4** 1892, 147–177.
13. H. POINCARE, *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, 175 Rue Saint-Honore, 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, Pembroke College, England, unpublished report, 1958.

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

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]

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

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

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

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 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

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

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

(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 α_{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 $\frac{n(n-1)}{2}$ elements of the first $\left[\frac{n}{2}\right]$ 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 $\frac{n(n-1)}{2}$ elements (Part 1, Sec. 6).

$$\psi(P_1)=\begin{matrix}2 & 2 & 3 & 2 & 3\end{matrix}\atop\begin{pmatrix}4 & 5 & 5 & 5 & 5\\ 7 & 7 & 8 & 7 & 7\\ 9 & 10 & 10 & 9 & 10\end{pmatrix}\qquad \psi(P_2)=\begin{matrix}3 & 3 & 4 & 3 & 4\end{matrix}\atop\begin{pmatrix}6 & 7 & 7 & 7 & 7\\ 10 & 10 & 11 & 10 & 10\\ 13 & 14 & 14 & 13 & 14\end{pmatrix}$$

$$\underset{\|\alpha_{ij}\|}{\text{equivalence class}}=\begin{bmatrix}1 & 1 & 2 & 1 & 2\\ 3 & 4 & 4 & 4 & 4\end{bmatrix}\qquad \bar{K}=9$$

($\bar{K}=9$ 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 $\frac{n(n-1)}{2}$ elements of the first $\left[\frac{n}{2}\right]$ rows) is ambiguous or not. If it is, since

$\alpha_{n-i,i+j}=\overline{K}-\alpha_{ij}$, $\overline{K}=2\max(\alpha_{ij})$; if it is not, $\overline{K}=2\max(\alpha_{ij})+1$. 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 $\overline{K}=2\alpha_{22}+1=9$).

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

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

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 δ_{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 $\overline{K}=2\max(\alpha_{ij})$ (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 $\delta_{1,1}=4$, $\delta_{1,2}=6$, $\delta_{1,3}=7$, $\delta_{1,4}=9$, $\delta_{1,5}=8$,

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

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:

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

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

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]

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

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

[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 $\underline{R}_k$ 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 $\underline{R}_k$, 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,

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

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 $p_{i+1} \equiv x + \delta_{1,j}$, $p_{i+2} \equiv x + \delta_{1,j} + \delta_{1,j+1}$, 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 mod 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 mod 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.

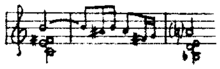
Figure 1 lists all $\|\delta_{ij}\|$ corresponding to proper sets when $m=12$ and S is the integers mod 12. Notice that all $\psi(P)$ with less than five terms and stability $>2/3$ (with the exception of $\Psi(P)=(5,1,5,1)$[14], appear as "chords" in music texts specifying standard "figured bass" notation. 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 $T_i \equiv \delta_{i+1} - \delta_i$, $T \equiv \min T_i$. 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 $P_c(x)$. 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)→(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.

Math. Systems Theory 12, 73–101 (1978)

Mathematical
Systems Theory

A Model for Pattern Perception with Musical Applications*
Part III: The Graph Embedding of Pitch Structures

David Rothenberg

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

Abstract. This is the third paper of a series which begins by treating the perception of pitch relations in musical contexts and the perception of timbre and speech. The preceding papers dealt with those properties of musical scales which allow them to function as reference frames which provide both for the measurement of intervals and for the identification of their elements as scale degrees. The effect of these properties upon the perceptibility of various musical relations and properties has been discussed. Here we extend the treatment to systems of different scales (as exist in many musical cultures) where a listener's recognition of any one scale in the system interacts with his ability to recognize the others. Reading of the two previous papers is required.

16. The Directed Graph, G

Thus far we have assumed that a listener has learned only one scale (measuring set) and will classify (measure intervals in) any stimulus (set of pitches) by embedding it in a key of this scale. Now we will consider the classification of stimuli by a listener who has learned *many* scales (e.g. a sophisticated Western listener or an Indian familiar with the enormous number of distinct "ragas" in use). That is, we have dealt with the problem of determining x in $P_v(x)$ (Part 2, Sec. 14), given a set of points which is a subset of some $P_v(x)$. Now we consider the problem of determining both v and x in $P_v(x)$ when the given points may be a subset of *any* of several given $P_v(x)$, $v = v_1, v_2 \ldots$.

Note that $P_u(x)$ may be a subset of $P_v(y)$, $u \neq v$, where x may or may not equal y. Indeed, if we denote by $P_1(0)$ that $P_v(x)$ corresponding to all the points

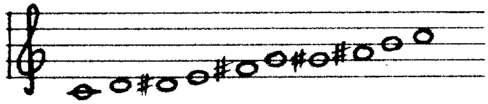
*This research was supported in part by grants and contracts AF-AFOSR 881-65, AF 49(638)-1738 and AF-AFOSR 68-1596.

0025/5661/78/0012-0073$5.80

of S, then for all $v, x, P_v(x) \subset P_1(0)$. That is, the partial orderings of all $P_v(x)$ by means of set inclusion, define a directed graph, G, in which connection indicates set inclusion. This graph consists of a Hasse diagram with $S = P_1(0)$ at the top.

Example 11. Let $m = 12$, $\psi(Q) = (4,2,4,2)$, $\psi(R) = (2,1,1,2,1,1,2,1,1)$, $Q = P_u(0)$, $R = P_v(0)$ where 0 corresponds to C.

$Q =$ 

$R =$ 

For what x,y is $P_u(x)$ (the chord Q transposed to begin on the xth note, counting C as the 0th note) included in $P_v(y)$ (the scale R transposed to begin on the yth note)?

Q has 2-fold symmetry so we need only consider $x = 0,1,2,3,4,5$; $P_u(6) = P_u(0)$ etc.

R has 3-fold symmetry so we need only consider $y = 0,1,2,3$; $P_v(4) = P_v(0)$ etc.

Then

$P_u(0) \subset P_v(0), P_v(2)$

$P_u(1) \subset P_v(1), P_v(3)$

$P_u(2) \subset P_v(0), P_v(2)$

$P_u(3) \subset P_v(1), P_v(3)$

$P_u(4) \subset P_v(0), P_v(2)$

$P_u(5) \subset P_v(1), P_v(3)$

If a line connecting 4 on the lower line to 0 on the upper line indicates that $P_u(4) \subset P_v(0)$, the above may be displayed as follows:

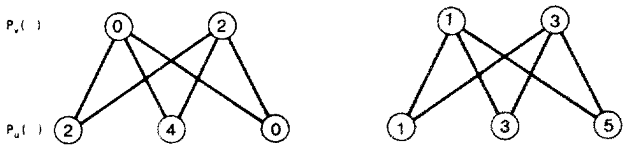


Fig. 1.

Figure 2 is a chart specifying the graph of all proper sets with stability $\geqslant 2/3$ (where S is the integers and $m = 12$). In the digital computer program which computed these results, the first point of S was called s_0 and hence $s_m = s_{11}$. (For musical interpretation, $s_0 = \text{C}$, $s_1 = \text{C}^{\#}$ or $\text{D}^{\flat}$, $s_2 = \text{D}, \ldots$ etc.) The chart is read as follows: All sets are listed according to the descending form $\psi(P)$; inverse sets are adjacent and indicated by brackets; each entry, y,

Fig. 2. Graph of all proper sets with $S \geqslant 2/3$ (where S is the integers and $m = 12$)

$P_u(x)$	$P_v(x)$; v =	24 (5151)	23 (4341)	22 (4332)	21 (4233)	20 (4323)	19 (4242)	18 (3333)	17 (5241)	16 (5142)	15 ([illegible])
33	552								7	0	[illegible],5,10
32	543		7	7		7					3
31	534		11		4	4					[illegible]
30	444										
29	[illegible]	[illegible],8								6	
28	[illegible]	[illegible],11							[illegible]		
27	[illegible]				6		2,6			[illegible]	
26	[illegible]			4			4,10		11		
25	633			10	0			[illegible]			
24	5151										
23	4341										
22	4332										
21	4233										
20	4323										8
19	4242										
18	3333										
17	5241										
16	5142										
15	32322										
14	33231										
13	33132										
12	33222										
11	331131										
10	222222										
9	322221										
8	312222										
7	321321										
6	312312										
5	[illegible]										
4	[illegible]										
3	[illegible]										
2	[illegible]										
1	111111111111										

$P_u(x)$	$P_v(x)$; v =	14 (32231)	13 ([illegible])	12 ([illegible])	11 ([illegible])	10 ([illegible])	9 ([illegible])	8 ([illegible])	7 ([illegible])	6 ([illegible])
33	552			[illegible]			7	10		
32	543	3,6	[illegible]	[illegible]	[illegible],7,11		[illegible]	3	[illegible]	
31	534	[illegible]	7,10		[illegible],7,11		[illegible]	10		4,10
30	444				[illegible]		[illegible]	[illegible]		
29	[illegible]	0						4	0,[illegible]	4,10
28	[illegible]		[illegible]				[illegible]		5,11	[illegible],9
27	[illegible]		[illegible]	[illegible]		[illegible]	[illegible],11	[illegible],6,10	[illegible],10	0,6
26	[illegible]	0				[illegible]	[illegible],5,11	[illegible],4,10	3,9	4,10
25	633	6	4	7			[illegible]	6	0,3,6,9	0,3,6,9
24	5151								[illegible]	4
23	4341				0,4,8					
22	4332	8		8			5	8	5,11	
21	4233		7	[illegible]			[illegible]	6		0,6
20	4323	11	3							
19	4242					0,2,4	6	0	[illegible]	0
18	3333								[illegible]	[illegible]
17	5241									
16	5142							10		
15	32322									
14	33231									
13	33132									
12	33222									
11	331131									
10	222222									
9	322221									
8	312222									
7	321321									
6	312312									
5	[illegible]									
4	[illegible]									
3	[illegible]									
2	[illegible]									
1	111111111111									

$P_u(x)$	$P_v(x)$; v =	5 ([illegible])	4 ([illegible])	3 (1212121)	2 ([illegible])	1 (111111111111)	$P_u(x)$
33	552	2,4,6,9,11	4,6,11		2,6,10	0 - 11	33
32	543	[illegible],7,9	9,11	[illegible],3,6,9	2,3,6,7,10,11	0 - 11	32
31	534	4,6,11	4,6	0,3,6,9	2,3,6,7,10,11	0 - 11	31
30	444		0		[illegible],3	0,1,2,3	30
29	[illegible]	0	0	[illegible],3,6,9	0,4,9	0 - 11	29
28	[illegible]	5	[illegible]	2,5,8,11	0,4,8	0 - 11	28
27	[illegible]	6	[illegible],6,8	[illegible],5,8,11	0,2,4,6,8,10	0 - 11	27
26	[illegible]	6	[illegible],4,8	[illegible],3,6,9	0,2,4,6,8,10	0 - 11	26
25	633	1	0,2	0,2,3,5,6,8,9,11	2,6,10	0 - 11	25
24	5151			0,3		0,1,2,3,4,[illegible]	24
23	4341	[illegible]			0,3,4,7,8,11		23
22	4332	2	[illegible],4	2,5,8,11	0,4,8	0 - 11	22
21	4233	0	0,2	[illegible],8,11	2,6,10	0 - 11	21
20	4323	0,1,7	2	2,5,8,11	3,7,11	0 - 11	20
19	4242		0	2,6	0,2,4	0,1,2,3,4,5	19
18	3333			[illegible]		0,1,2	18
17	5241	7	9		3,7,11	0 - 11	17
16	5142	4	6		2,6,10	0 - 11	16
15	32322	5,[illegible],11	6			0 - 11	15
14	33231			0,3,6,9	0,4,8	0 - 11	14
13	33132			2,5,8,11	0,4,8	0 - 11	13
12	33222	5	5,8		0,4,8	0 - 11	12
11	331131				0,3	0,1,2,3	11
10	222222				0	0,1	10
9	322221		9		3,7,11	0 - 11	9
8	312222		8		0,4,8	0 - 11	8
7	321321			[illegible]		0,1,2,3,4,5	7
6	312312			[illegible]		0,1,2,3,4,5	6
5	[illegible]					0 - 11	5
4	[illegible]					0 - 11	4
3	[illegible]					0,1,2	3
2	[illegible]					0,1,2,3	2
1	111111111111					0	1

Each entry, y, indicates that $P_u(y) \subset P_v(0)$ where u is indicated by the row label and v by the column label. For example, the entry 10 in row 16, column 8, indicates that the only transposition of chord number 16 (i.e., C F G^b B^b) which is included in scale number 8 on C (i.e., C D$^\#$ E F$^\#$ G$^\#$ A$^\#$) is that transposition which has note number 10 (A$^\#$) for its bass note (i.e., A$^\#$, D$^\#$, E$^\natural$, G$^\#$).

indicates that $P_u(y) \subset P_v(0)$ where u is indicated by the row label and v by the column label. Note that because the periodic case is used, $P_u(y) \subset P_v(0)$ implies that for all k, $P_u(y+k) \subset P_v(k)$ where the arguments of P_u and P_v are reduced modulo the periods of the respective scales.

An example of a diagram displaying a small portion of the above graph is:

$$\psi(P_1(x)) = (1,1,1,1,1,1,1,1,1,1,1,1);\ \psi(P_3(x)) = (2,1,2,1,2,1,2,1);$$

$$\psi(P_{13}(x)) = 3,3,1,3,2;\ \psi(P_{21}(x)) = (4,2,3,3);\ \psi(P_{31}(x)) = (5,3,4).$$

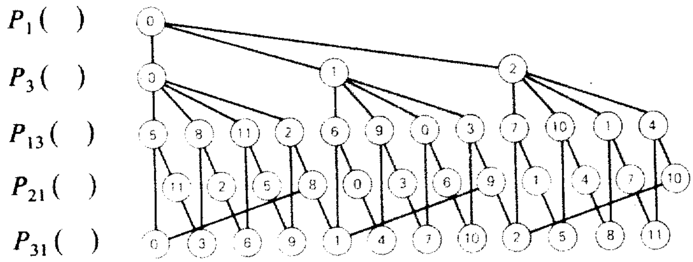


Fig. 3.

To draw a diagram displaying the entire chart would be cumbersome; hence, for convenience, a node on the diagram on the following page (Figure 4) will represent a collection of sets, $\{P_u(x),\ x = 0,1,\ldots\}$, for all permissible values of x (all keys of scale $P_u(x)$). The numbers along the vertical connecting lines between nodes indicate those values of x for which $P_u(x) \subset P_v(0)$ ($P_u(x)$ appears below $P_v(x)$). Since $P_u(y+k) \subseteq P_v(k)$ there is no loss of information in this new representation. Inverses are indicated by dotted horizontal lines with arrowheads at each end. Only sets of cardinality five and greater are shown in this diagram.

17. Graph Equivalence[1]

Note that all subsets of S (including those not on graph G) are subsets of at least one set on graph G.[2] Henceforth let a point of graph G be denoted by "$P_v(x)$", and let subsets of S which may or may not be on the graph be denoted by H. Then, given H one can determine all v and y for which $H \subset P_v(y)$.

To each H let there correspond a set:

$$V(H) = \{P_v(y) \mid H \subset P_v(y)\}.$$

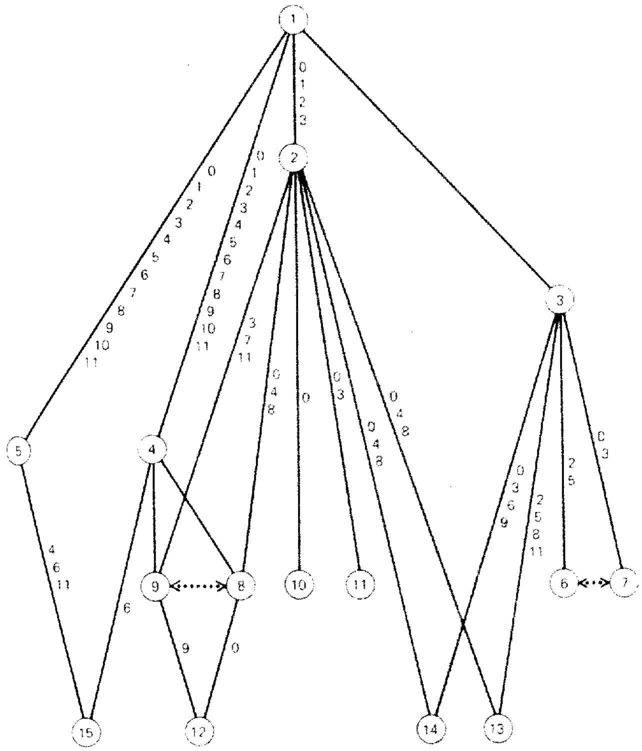
Two sets, H_1 and H_2 will be called *graph equivalent* iff $V(H_1) = V(H_2)$. To each H let there correspond a number

$$I(H) = M - \text{card}(V(H))$$

where M is the total number of points on graph G.

[1]The following discussion applies to the directed graph, G, not to any of its abbreviated versions (as in Figure 3 and 4).

[2]$S \equiv P_1(0)$ is assumed to be on graph G (see above).



$\psi(P_1(x)) = (1,1,1,1,1,1,1,1,1,1,1,1)$

$\psi(P_2(x)) = (2,1,1,2,1,1,2,1,1)$

$\psi(P_3(x)) = (2,1,2,1,2,1,2,1)$

$\psi(P_4(x)) = (2,2,2,2,1,2,1)$

$\psi(P_5(x)) = (2,2,2,1,2,2,1)$

$\psi(P_6(x)) = (3,1,2,3,1,2)$

$\psi(P_7(x)) = (3,2,1,3,2,1)$

$\psi(P_8(x)) = (3,1,2,2,2,2)$

$\psi(P_9(x)) = (3,2,2,2,2,1)$

$\psi(P_{10}(x)) = (2,2,2,2,2,2)$

$\psi(P_{11}(x)) = (3,1,3,1,3,1)$

$\psi(P_{12}(x)) = (3,3,2,2,2)$

$\psi(P_{13}(x)) = (3,3,1,3,2)$

$\psi(P_{14}(x)) = (3,3,2,3,1)$

$\psi(P_{15}(x)) = (3,2,3,2,2)$

Fig. 4.

$I(H)$ will be called the information value of H with respect to graph G. It is equal to the number of points on graph G to which H does not belong.

Evidently $I(H)$ is invariant under musical transposition of H in the above examples. This is because G is complete in the sense that $P_v(y) \in G \rightarrow P_v(y') \in G$ for all v, y, y'.

Intuitively, information values count those points of graph G which need *not* be considered when classifying a given subset H of points in S. If all points of graph G (with the possible exception of $P_1(0) \equiv S$)[3] are interpreted as learned "mental reference frames", any of which may be used by a listener to measure (classify) the intervals of a "signal" or "stimulus" consisting of a string of musical tones,[4] then the information value $I(H)$, corresponds to the number of such "reference frames" which are eliminated as possibilities for classifying a "signal" when that portion of the "signal" specified by H has been heard.

Similarly, graph equivalent sets may be interpreted as different stimuli that can be classified by identical subsets of those "reference frames" which are known to a listener (clearly, graph G contains only those "reference frames" which are known to the listener, $P_1(0) \equiv S$ excepted).[5]

18. Graph-Sufficient Sets, Graph and Node-Efficiency

A *graph-sufficient set* for graph point $P_v(y_1)$ is defined as a subset H of $P_v(y_1)$ such that

$$H \subset P_w(z) \Rightarrow P_v(y_1) \subset P_w(z) \tag{1}$$

i.e., hearing H tells us we are in $P_v(y_1)$ or above it; the only scales (chords) on the graph which contain H are $P_v(y_1)$ and its superscales (superchords).

A *node-sufficient set* for a graph point $P_v(y_1)$ is a subset H of $P_v(y_1)$ such that

$$H \subset P_w(z) \Rightarrow P_v(y_1) \subset P_w(z) \quad \textit{or} \quad P_w(z) \subset P_v(y_1) \tag{2}$$

Note that (2) differs from (1) only in that it permits H to be a subset of some $P_w(z)$ below $P_v(y_1)$. A node-sufficient set distinguishes a particular graph point from all graph points incomparable with it.[6] A set can be graph-sufficient for only one graph point but it can be node-sufficient for several. (E.g., all sets are node-sufficient for $P_1(0) \equiv S$).

Example 12. Let H_1 be a subset of exactly those $P_v(x)$ on the graph shown below which are blackened and let H_2 be a subset of exactly those which are

---

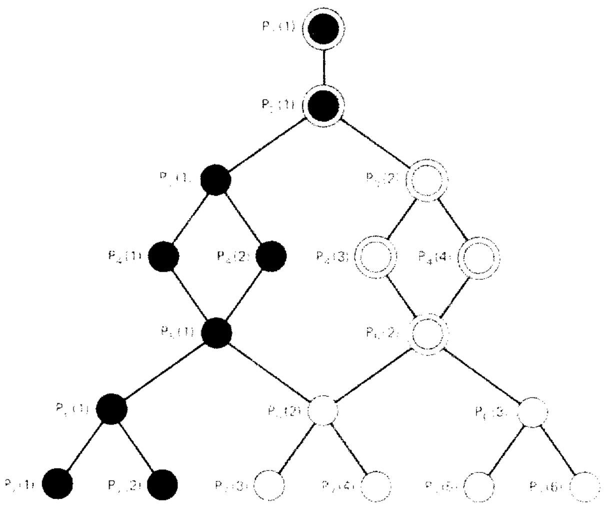
[3]When $P_1(0)$ does not represent a learned classifier and $I(H) = M - 1$ (i.e., H is a subset of no $P_v(y)$ other than S), the intervals in H will likely be heard as "mistunings" of elements of some $P_v(y)$ whose selection from the other graph points will probably depend upon preceding stimuli (see previous discussion of tuning).

[4]This "stimulus" may consist of musical tones simultaneously heard as well as in sequence, provided that these tones are clearly distinguished.

[5]$P_1(1) \equiv S$ is always included in graph G for convenience, since all graph points are connected to it and since its uniqueness causes it to have no effect on the relative values of different $I(H)$.

[6]Here and henceforth "incomparable" means "with respect to inclusion".

surrounded by a circle:



For the above graph (which is not related to the graph of Figure 2) H is node-sufficient for exactly $P_6(1)$, $P_5(1)$, $P_3(1)$, $P_2(1)$ and $P_1(1)$, and it is not graph-sufficient for anything. H is graph-sufficient for exactly $P_5(2)$ and is node-sufficient for exactly $P_5(2)$, $P_3(2)$, $P_2(1)$ and $P_1(1)$.

Note that if all $P_v(y)$ are eliminated from graph G except those which are transpositions of one particular P and their subsets and supersets, those H's which are node-sufficient for each such transposition are sufficient sets as defined in part II, Sec. 14. These will henceforth be called *key-sufficient sets*.

A listener's habits may be such that he will classify a stimulus by the lowest graph point possible (i.e., when $\bar{n}$ (the number of scale tones per octave) is minimal); e.g., a Western musician listening to classical music would most likely hear a set of tones belonging to a key of a major scale as a subset of such a key although these same tones are also a subset of the twelve-tone chromatic scale. In such a case his classifications will depend upon the graph-sufficient sets in the stimulus. However, the same listener may sometimes classify a stimulus by graph point which is not the lowest possible point on the graph. E.g., the same Western musician would hear a set of tones belonging to keys of both a major and a pentatonic scale as a subset of the key of the major (unless nothing but keys of the pentatonic had been heard in the composition for a long stretch of time). In this case his classification will be according to a node-sufficient (rather than graph-sufficient)[7] set. More will be said of this later.

Also of interest are sets which are subsets of at least one transposition of a particular P and of no point on graph G incomparable with all transpositions of P. These determine v in $P_v(y)$, but need not uniquely fix y and are called

[7]Note that a node-sufficient set has lower information value than a graph-sufficient set for the same point.

scale-sufficient for $\psi(P)$. Scale-sufficient sets are applicable to Indian music where any of a large number of scales (the tones in the "Ragas") may be used, but only one mode of each key (that whose tonic coincides with the drone) is used. In the "Alap" portion of an Indian classical performance, scale-sufficient sets are often avoided for a considerable period of time so that the resolution of doubt as to the identity of the "Raga" used assumes dramatic significance.

Graph, node and scale-minimal sets are defined (as previously) as sufficient sets of the same type, no proper subset of which satisfies the identical sufficiency condition.

The definitions of *graph efficiency*, E^G, and *node efficiency*, E^N, for a graph point $P_v(x)$, are similar to the definition of efficiency (cf. part II, Sec. 15) except that the number of graph or node-sufficient sets for $P_v(x)$ appears in the numerator in place of the number of key-sufficient sets. When G is *complete* (i.e., closed under transposition) E^G is the same for all transpositions of a given P (as is E^N).[8]

19. A Sample Graph

In the graph specified by Figure 2, Section 16, S is the integers, $m = 12$, all $P_u(x)$ with stability $\geqslant 2/3$ are on the graph. Figure 5 (below) shows the containment pattern of all 3, 4 and 5 element subsets H of S in those of these graph points with more than six elements or with six elements and a stability of 1. These are familiar to contemporary Western musicians; other points on the graph of figure 2 are subsets of at least one of these and are customarily used as "chords" rather than "scales" in Western music. Hence, were $P_2(x)$ eliminated from consideration, all subsets of only one of the remaining selected graph points[9] would be node-sufficient[10] for that point. P_2 is relatively low in stability and has a high gradient with respect to the whole-tone scale (P_{10}), the twelve-tone scale (P_1) and also P_{11}. Hence P_2 is easily perceived in terms of other (proper) scales. In deference to the contemporary composer, Olivier Messiaen who consciously uses P_2[11], it is reluctantly included in Figure 5. All node-sufficient sets are underlined including those which are node-sufficient only if no $P_2(x)$ is considered (e.g. 6141). Other underlinings are only of *graph*-sufficient sets for $P_1(0)$ or some $P_2(x)$. (It is assumed that listeners will classify node-sufficient sets for $P_1(0)$ or some $P_2(x)$ by some lower graph point e.g. some $P_{11}(x)$ of $P_{10}(x)$).

Each row corresponds to an $H_u(0)$ whose $\psi(H_u(0))$ is shown in the leftmost column. The headings of columns 2–7 specify $\psi(P)$ corresponding to $P_2(x)$,

[8]When G is complete E^G and E^N may be computed without finding corresponding sufficient sets (which in this case, can be easily extracted). See subsequent computation paper.

[9]None of which is now contained in another.

[10]But not necessarily graph-sufficient (e.g., $P_{24}(0)$ (see Figure 2), for which $\psi(P) = (5, 1, 5, 1)$, is graph-sufficient for itself, but not for $P_3(0)$ ($\psi(P_3) = (2, 1, 2, 1, 2, 1, 2, 1)$), for which it is node-sufficient.

[11]In his book, "The Technique of My Musical Language" [15], he lists $P_2(x)$ as one of his "modes of limited transposition". It is not clear, however, that it is heard as a "scale" as defined here (i.e., a "mental reference frame").

$P_3(x)$, $P_4(x)$, $P_5(x)$, $P_{11}(x)$ and $P_{10}(x)$ respectively. $P_1(0)$ is not shown since all $H_u(x)$ are subsets of it. Each table entry, y, indicates that $H_u(0) \subset P_v(y)$ where u is determined by the row label and v by the column label. (Note that $H_u(0) \subset P_v(y) \Leftrightarrow H_u(z) \subset P_v(y+z)$ with appropriate reductions of z and $y+z$). The number in the rightmost column is the total number of entries in all columns except that corresponding to $P_2(x)$. These numbers roughly correspond inversely to information values.

	211211211	21212121	2222121	2221221	313131	222222	
1011	0						0
921	1	0	0 3	0 5			5
912	2	1	110	310			5
831	0 1	0	0		0		3
822	0 2		4 6 8	1 6 8		0	7
813	1 2	0	0		1		3
741	0 1		3	0 5	0		4
732	0	1	1 810	1 3 810			8
723	1	1	1 310	0 3 510			8
714	0 1		8	1 8	0		4
651	0	0	0	0			3
642	0 2	1	4 610	6		0	6
633	2	0 1	010	0			5
624	0 2	0	0 4 6	6		0	6
615	0	1	10	0			3
552	2		1 6 8	1 3 6 810			8
543	1 2	0	1 3	3 510	1		7
534	1 2	0	6 8	1 6 8	1		7
444	0 1 2		0 4 8		0 1	0	6
9111							0
8211	0						0
8121	1	0	0				2
8112	2						0
7311	0						0
7221	1		3	0 5			3
7212		1	110	310			5
7131	0 1				0		1
7122	0		8	1 8			3
7113	1						0
6411	0						0
6321		0	0	0			3
6312	2	1	10				2
6231	0	0	0				2
6222	0 2		4 6	6		0	4
6213	2	0	0				2
6141	0			0			1
6132	0	1	10				2
6123		1	10	0			3
6114	0						0
5511							0
5421	1	0	3	5			3
5412	2		1	310			3
5331	1	0					1
5322	2		6 8	1 6 8			5
5313	1 2	0			1		2
5241	1		3	5			2
5232			1 8	1 3 810			6
5223	1		1 3	3 510			5
5214	1		8	1 8			3
5151		0					1
5142	2		6	6			2
5133	2	0					1
5124	2	0	6	6			3
4431	0 1		0		0		2
4422	0 2		4 8			0	3
4413	1 2		0		1		2
4341	0 1			0 5	0		3

Fig. 5.

	211211211	21212121	2222121	2221221		313131	222222
4332	0	1	810	10			4
4323	1	1	10	0 510			5
4242	0 2	1	410				4
4233	2	1	010	0		0	4
3333		0 1					2
81111							0
72111							0
71211	0						0
71121	1						0
71112							0
63111							0
62211	0						0
62121		0	0				2
62112	2						0
61311	0						0
61221				0			1
61212		1	10				2
61131	0						0
61122	0						0
61113							0
54111							0
53211							0
53121	1	0					1
53112	2						0
52311							0
52221	1		3	5			2
52212			1	310			3
52131	1						0
52122			8	1 8			3
52113	1						0
51411							0
51321		0					1
51312	2						0
51231		0					1
51222	2		6	6			2
51213	2	0					1
51141							0
51132							0
51123							0
51114							0
44211	0						0
44121	1		0				1
44112	2						0
43311	0						0
43221	1			0 5			2
43212		1	10	10			3
43131	0 1				0		1
43122	0		8				1
43113	1						0
42411	0						0

Fig. 5. (cont'd.)

	211211211	21212121	2222121	2221221	313131	222222	
42321			0	0			2
42312	2	1	10				2
42231	0		0				1
42222	0 2		4			0	2
42213	2		0				1
42141	0			0			1
42132	0	1	10				2
42123		1	10	0			3
41412	2			10			1
41331	1						0
41322	2		8				1
41313	1 2				3		1
41232			8	10			2
41223	1			510			2
41133	2						0
33321		0					1
33312		1					1
33231	0	0					1
33222	0		4 6	6			3
33132	0	1					1
32322			6	1 6 8			4

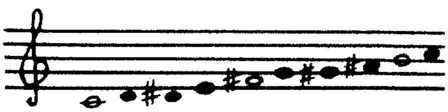
Fig. 5. (cont'd.)

Illustration of the Use of the Table. Let it be required to find those "scales" (=transpositions of $P_1 - P_5$ and $P_{10} - P_{11}$) of which (1) $\{C, F^{\#}, B\}$ and (2) $\{B, D, F^{\#}, G\}$ are subsets.

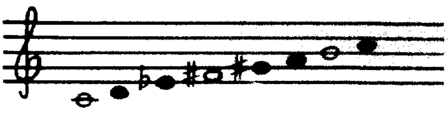
Ad (1). The $\psi(P)$ for $CF^{\#}B$ is 651. Since this is lexicographically later than its cyclic permutations 516 and 165, it's already in descending form, and we look for the row (row 11) of the table which has 651 standing in its leftmost column. (Entries in this column are arranged in order of length, and in reverse lexicographic order between entries of the same length.) Then (reading row 11 from left to right)

$$\{C, F\#, B\} \subset P_2(0), P_3(0), P_4(0) \text{ and } P_5(0)$$

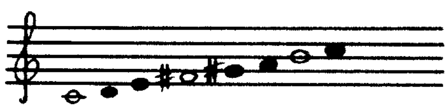
i.e., it is included in Messiaen's mode P_2 starting on C:



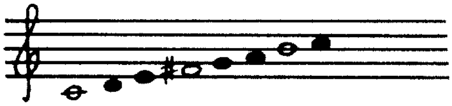
P_3 ("string of pearls") starting on C:



P_4 starting on C (a cyclic permutation of A melodic minor):



P_5 starting on C (C Lydian, a cyclic permutation of G major):



and trivially P_1 (one chromatic scale, omitted from the table because everything is included in it).

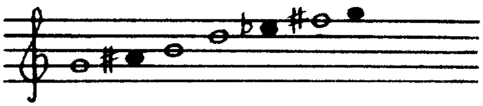
The number 3 written in the rightmost column of line 11 of the table indicates that there are 3 possible $P_u(x)$ in which {C F# B} is included (not counting the trivial P_1's and the dubious P_2's). If there is only one such $P_u(x)$ (as in rows 5331 and 61221 of the table) then H is node-minimal for that one (if P_2 is excluded from the graph); if there are more (as in rows 53211 and 61311) it is node-minimal for P_1 ("atonal" in a precise sense): in either case the row is underlined. The numbers in the right-hand column roughly correspond inversely to information values.

Ad (2). The $\psi(P)$ for B D F# G is 3 4 1 4. Its descending form is 4 3 4 1 which we find in its place in column 1 of the table. This corresponds to the position G B D F#, so 0≡G. Reading the entries opposite 4 3 4 1 from left to right, we see:

Under P_2, 0 and 1, i.e., Messiaen's mode on G or A♭.

Under P_5, 0 and 5, i.e., the Lydian mode on G=0 or C=5 (=D and G major respectively).

Under P_{11} (a symmetric mode unknown to Messiaen) 0, i.e., P_{11} on G



On the table below the first two entries are graph efficiencies[12] for $P_1(0)$ and any $P_2(x)$; the rest are node efficiencies for any $P_3(x)$, $P_4(x)$, $P_5(x)$, $P_{11}(x)$ and $P_{10}(x)$. (ψP)'s are shown on the left):

$\psi(P)$	if all $P_2(x)$ are on graph	if all $P_2(x)$ are excluded from graph
(1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1)	.5517	.4646
(2, 1, 1, 2, 1, 1, 2, 1, 1)	.6190	—
(2, 1, 2, 1, 2, 1, 2, 1)	.7464	.6786
(2, 2, 2, 2, 1, 2, 1)	.9592	.8844
(2, 2, 2, 1, 2, 2, 1)	.9524	.9143
(3, 1, 3, 1, 3, 1)	.8000	.8000
(2, 2, 2, 2, 2, 2)	1.0000	1.0000

If we consider a graph that consists only of $P_1(0)$ and all points whose $\psi(P)$ corresponds to the "major scale" (2, 2, 2, 1, 2, 2, 1) and to the "melodic minor

[12]Note that a graph-sufficient set for $P_v(y)$ cannot be classified by subsets of $P_v(y)$. Hence graph efficiencies are used instead of node efficiencies for $P_1(0)$ and all $P_2(x)$.

scale" (2,2,2,2,1,2,1), node efficiencies are as follows:

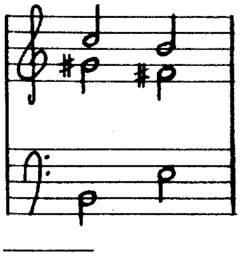
(2,2,2,2,1,2,1)	.7932
(2,2,2,1,2,2,1)	.9143

Note that $\psi(P)$ which are low in efficiency (computed using key-sufficient sets—see Figure 2) may be high in graph or node efficiency. Hence scales which may be too redundant (Part II, Section 15) to be greatly used in diatonic music (e.g., (2,2,2,2,2,2), (2,1,2,1,2,1,2,1)) may be extensively used in music which freely makes use of all twelve tones and their subsets. Examples can be found in Ravel, Stravinsky, Frank Martin, Bartok, etc. (Ravel mixes the "whole tone scale" with the "major" and "minor" scales; Stravinsky's first movement of the "Symphony of Psalms" uses (2,1,2,1,2,1,2,1) extensively, etc.)

Observe that among the $H_u(x)$ with three elements (triads), those whose $\psi(P)$ are (10,1,1), (8,3,1), (6,5,1) have the highest information values. The frequency of use of such triads in twentieth century music is well documented; hence "minjor" chords and chords in fourths. A particularly startling example is Anton Webern's "Piano Variations" (Opus 27). The entire composition consists of a succession of graph-sufficient sets for the "twelve tone scale" (1,1,1,1,1,1,1,1,1,1,1,1), although the "serial" technique of composition does not guarantee this.[13]

If we retain the notion that a "cadence" must uniquely determine key and tonality (Section 15 at the end), and add that it must also determine scale, "cadences" for "non-diatonic" music (i.e., using keys of all scales on the graph) can be constructed. Here a sample cadence for $P_4(0)=\text{C,D,E,F}^{\#},\text{G}^{\#},\text{A,B}$ ($\psi(P)=(2,2,2,2,1,2,1)$), when E is the tonic, will be constructed. Notice (from Figure 5) that $\{\text{C,E,F}^{\#},\text{G}^{\#},\text{B}\}$ (4,2,2,3,1) is as small a minimal node-sufficient set as can be found for $P_4(0)$. Let us also take advantage of the fact that the interval ($\text{G}^{\#}$,C) is ambiguous in this scale, and therefore "resolve" it to an unambiguous interval (as F–B to E–C in C major). The final chord will contain E and B so that the tonic, E, is reinforced by the resulting difference tone (assuming conventional timbres). Notice that $P_4(0)$ is conventionally heard as "A minor", not as a mode with E as a tonic, as in this example:

Example 13



[13]A careful analysis of the rules of the "serial" (i.e., "twelve tone") technique will show that they prejudice the composer in favor of writing successions of graph-sufficient sets for the "twelve-tone scale" and avoiding tonal centers (modes). Examination of works of well-known "serial" composers (e.g. Webern) will show that such is done in excess of the demands of the rules of the technique.

Similar examples can be constructed using the chart in Figure 5. Of course, if a node-sufficient set contains more than one node-minimal subset the identification of the scale is strengthened. A common method is to use all the tones of a scale in a cadence. Oliver Messiaen's music abounds with such examples. Here is an excerpt from the third song, "Dance du Bebe-Pilule" of his "Chants de Terre et de Ciel".[14] The scale is $P_3(1) = \mathrm{D}^b, \mathrm{E}^b, \mathrm{E}, \mathrm{F}^\#, \mathrm{G}, \mathrm{A}, \mathrm{B}^b, \mathrm{C}$ $(\psi(P) = (2,1,2,1,2,1,2,1))$ where E^b is the tonic:

Example 14.

Notice that, with appropriate labeling of node-sufficient (or graph-sufficient) sets, it is possible to extend the traditional "figured bass" system to apply to non-diatonic music.

On such a basis new methods of teaching musical "ear training" can also be suggested.

Condition (c) at the end of Section 15 pertains to "color changes" between "chords" in diatonic cadences. The following discussion is relevant to such differences and similarities between sets of tones in non-diatonic music.

20. Image Distance

Consider two sets, H_1 and H_2, which are not graph equivalent. If the graph G is altered by the removal of certain points, these sets will become graph equivalent. The minimum number, X, of such points which must be removed from graph G for $H_u(x)$ and $H_v(y)$ to become graph equivalent is equal to the cardinality of the symmetric difference between their respective $V(H_1)$ and $V(H_2)$ (Section 17, at the beginning):

$$X = \mathrm{card}(V(H_1) \cup V(H_2)) - (V(H_1) \cap V(H_2)).$$

In order to arrive at a measure for the disturbance which must be induced in that portion of the graph G which is pertinent to the classification of H_1 and H_2, we divide X (above) by the cardinality of that portion (i.e., $\mathrm{card}(V(H_1) \cup V(H_2))$). A number between zero and one results which we call the *image*

[14]Elkan-Vogel Co., Philadelphia, Pa. (U.S.A.). Copyright by Durand & Cie, 1939.

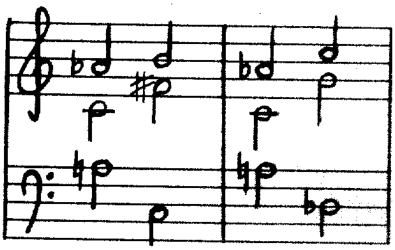
distance, $\bar{I}(H_1, H_2)$, of H_1 and H_2:

$$\bar{I}(H_1, H_2) = 1 - \frac{\text{card}(V(H_1) \cap V(H_2))}{\text{card}(V(H_1) \cup V(H_2))}. \tag{1}$$

In computing the cardinalities in the above expression all elements of the sets being evaluated are, in effect, assigned equal weight (i.e., 1). This corresponds to assigning equal likelihood to all sets on the graph. This is rarely the case in actual music, where some scales and/or keys are more likely to occur than others. However, (1) above is easily modified by assigning different weights (altering the cardinality function so that a set may be counted more than once) to each point on the graph. Such weighting, however, will not alter the ordering of the image distances between any two pairs, $H_u(x)$ and $H_v(y)$, in the examples which will follow. Note that image distance provides a more comprehensive relation between "chords" than the common criteria of "the number of common tones" and "the relation between roots". By use of the chart in Figure 5 the reader can verify that the well-known order of similarity relations between triads in the same and different keys can be derived from image distance (e.g., {CEG} is less similar to {F# A# C#} than to {DFA}). If such similarity relations are intended to apply only to diatonic music, all scales except the major ($\psi(P) = (2,2,2,1,2,2,1)$) and melodic minor ($\psi(P) = (2,2,2,2,1,2,1)$) should be eliminated from the graph. For most nondiatonic music it is probably appropriate to eliminate $P_2(x)$ ($\psi(P) = (2,1,1,2,1,1,2,1,1)$) for all x.

More sensitive discriminations than the familiar criteria for similarity between triads are thus provided. The following less obvious examples are chosen so that the number of common tones in each corresponding pair is the same, so that "positions" of corresponding "chords" are identical, and so that the effect of tonality is minimized. $P_2(x)$ for all x is eliminated from the graph (although its inclusion does not alter the ordering of the different $\bar{I}$), $P_1(0)$ is included (hence $\bar{I} \neq 1$ and distinctions result between pairs of sets whose intersections belong only to $P_1(0)$), and no weighting of different graph points is used. In both examples the second pair has a larger image distance than the first. Example 15 is about as subtle a case as can be constructed in twelve-tone equal temperament:

Example 15.



$\alpha\ \beta$ $\alpha\ \gamma$

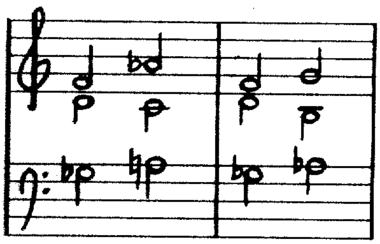
$\bar{I} = 2/5$ $\bar{I} = 6/7$

$(\alpha) = (8,1,3)$ on C

$(\beta) = (6,5,1)$ on C

$(\gamma) = (6,5,1)$ on D^b

Example 16.



$\alpha\ \beta$ $\alpha\ \gamma$

$\bar{I} = 2/3$ $\bar{I} = 6/7$

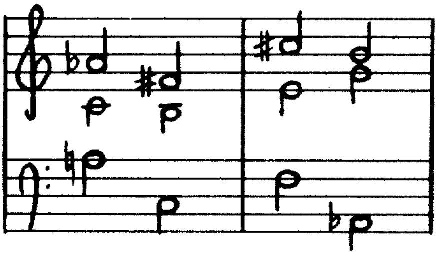
$(\alpha) = (8,3,1)$ on G^b

$(\beta) = (8,1,3)$ on C

$(\gamma) = (8,1,3)$ on B

The first pair of Example 15 cannot be compared with the first pair of Example 16 without alteration, because both pairs are subsets of $P_3(0)$ ($\psi(P_3) = (2,1,2,1,2,1,2,1)$) so that hearing one pair facilitates classification of the next. This can be overcome by transposing (changing the key memberships) of one of the pairs, but then difficulties in avoiding differences between the pairs with respect to dissonance and tonal implication arise. Also, sequences must be avoided, since these generate expectations which facilitate acceptance of the last chord in the sequence. One must attempt to equalize the pairs with respect to all properties except image distance. About the best that can be done in this respect when comparing the first pair of each of the above examples is shown below. However, the voice leading in the first pair is not so smooth as in the second, thereby tending to increase the sense of difference between the first pair relative to the second. (This prejudicial condition can be reversed by exchanging the position of the B and C in the second chord of the first pair).

Example 17.



$\alpha\,\beta$ $\quad\bar{I} = 2/5$

$\gamma\,\alpha$ $\quad\bar{I} = 2/3$

$(\alpha) = (8,1,3)$ on C, $(\beta) = (6,5,1)$ on C: example 15 first pair

$(\gamma) = (8,3,1)$ on F, $(\alpha) = (8,1,3)$ on B: example 16 first pair transposed

Of course, $\bar{I}$ remains relevant if each of a pair, H_1 and H_2, appears in a melodic sequence instead of in a chord progression.

21. Graph Distance

Just as the image distance between a pair, H_1 and H_2, derives from those points by which each of such pairs may be classified, the graph distance between a pair of points on the graph, $P_u(x)$ and $P_v(y)$, derives from those graph points which may be classified by each of such pairs. Graph distance is a relation between keys of scales and is thus relevant to musical "modulation".

A graph point (a particular key of a scale), $P_u(x)$ is often selected as a classifier for some $H = H_1 \cup H_2 \cup \ldots$; each of which H_i may be classified by some graph point which is a subset of $P_u(x)$. (This will occur when the "stimuli", $H_1, H_2, \ldots$, are contiguous, occupy little time, and when the union of graph points by which each of the H_i are classified forms a set which is graph-sufficient for $P_u(x)$.) Hence a graph point is characterized by the other graph points which are subsets of it.[15] Accordingly, we define $W(u,x)=\{P_v(y)|P_v(y) \subset P_u(x)\}$. Note that $W(u,x)$ derives from set inclusion in the direction opposite to that of $\bar{V}(H)$ (Section 17).

Then the *graph distance*, $\bar{D}(u,x,w,z)$, between two graph points, $P_u(x)$ and $P_w(z)$, is defined (analagously to $\bar{I}(H_1,H_2)$) as

$$\bar{D}\,(u,x,w,z)=1-\frac{\text{card}(W(u,x)\cap W(w,z))}{\text{card}(W(u,x)\cup W(w,z))}.$$

Graph distances are intended to correspond to degrees of perceived "similarity"[16] between keys of different musical "scales". For example, the "major scale" (C, D, E, F, G, A, B) appears to be more similar (in this sense) to the "melodic minor scale" (A, B, C, D, E, F$^{\#}$, G$^{\#}$), than does the "whole tone scale" (C, D, E, F$^{\#}$, G$^{\#}$, A) (although each pair has the same number of common tones). Since our notion of "scale" corresponds to "equivalence class" $\|\alpha_{ij}\|$, and since a given $\|\alpha_{ij}\|$ may have more than one member $\psi(P)$, each with corresponding points on graph G, for certain applications it may be necessary to restrict graph G to sets with a single $\psi(P)$ in each $\|\alpha_{ij}\|$. That is, when two distinct sets are the same *key*, *mode*, and *scale* (but different *tunings*) (Section 12), they are perceived as "mistunings" of each other. It is thus inappropriate to consider both such sets as lying on the same graph.[17] Note, however, that when two such sets are low[18] on graph G and when the image distance between them is large, it is unlikely that they will be heard as "mistunings" of each other.[19] Thus it is inappropriate to place two distinct P in the same $\|\alpha_{ij}\|$ on graph G when they are high on the graph and are of the same key and mode. This occurs in its worst form when graph G is complete (i.e. closed under transposition) and two distinct $\psi(P)$ in the same $\|\alpha_{ij}\|$ have points on the graph.

Sometimes it is useful to speak of the "distance" between $\psi(P)$ and $\psi(Q)$ rather than between P_i and Q_i, some pair of their respective graph points. (i.e. "Is *the* major scale more similar to the melodic minor scale than *the* "whole tone scale?") Thus the *scale distance*, $\bar{S}(u,v)$ is defined as

$$\bar{S}\,(u,v)=\min_{x,y}\bar{D}\,(u,x,v,y)$$

The number (or numbers), $(y-x)$, which corresponds to the above minimum, shows the relative keys which correspond to maximal similarity. (When

[15]For example, the "major scale" is characterized by its triads, seventh chords, etc.

[16]Note that our conception of "scale" makes no distinction between "modes" (tonal centers) of a key of such a scale. Hence the "similarity" which is referred to above is distinct from that similarity which results from relations between tonics of keys of scales (except when the discussion to follow applies).

[17]Such problems often arise when $m>12$ (e.g., $m=31$).

[18]We use the convention that sets appear above their subsets on graph G (see Figure 3).

[19]There is one such case when $m=12$; $P=(\text{C},\text{E},\text{G}^b,\text{B}^b)$ and $Q=(\text{C},\text{F},\text{G}^b,\text{B})$ (for which $\Psi(P)=(4,2,4,2)$ and $\Psi(Q)=(5,1,5,1)$) are in the same equivalence class, but there is little perceived similarity because classification ordinarily occurs higher on the graph by different graph points in each case.

graph G is complete, x (or y) may of course be arbitrarily set without affecting the value of $\bar{S}(u,v)$.) In the musical application we may also choose which tonic ("mode")[20] should be assigned to each of such a pair of keys of scales to achieve greatest similarity: When there is more than one tone common to both (keys of scales), assign the tonic of *both* to that tone which is best supported by the difference tones and harmonics generated by the tones of each of the pair of keys (cf Part I, Section 2). (These harmonics and difference tones are dependent upon the timbres being used.)

It is now known that a wide range of inharmonic residues have definite pitch[21] (i.e., a sensation of pitch occurs even when the partials are not integer multiples of a fundamental), and in such cases extremely unfamiliar intervals may sound "pure" or "constant"[22] and familiar "consonant" intervals may sound "dissonant" or "impure".[23] Sometimes a tone which is well supported as a tonic in both of the keys (above) cannot be found, such that footnote 24, Section 7 of Part I is pertinent, and the above procedure does not apply.

Note that implicit in the above interpretation is the assumption that the perception of differential similarities between different pairs of scales (learned "mental reference frames") is always dependent upon those other scales which have been learned by the listener and such of these whose use he may anticipate in a particular situation; e.g., the same listener's expectations will differ (as will graph G) when listening to classical Western music and to twentieth century Western music.

22. The Effect of the Graph upon the Tuning of Scales

Previous to this, we have discussed restrictions on "mistunings" of the elements of a proper subset such that it retains its propriety (Part 1, Section 8) and also restrictions on such mistunings so that any proper or improper scale retains its identity (Part 2, Section 10). The relevance of these different restrictions to a particular $P_u(x)$ depends upon the structure of the graph in which it is embedded.

Consider a graph which contains one and only one point, P, which is a proper set. Note that when axioms 2.2 and 2.3 apply (as they do in Western music) and the elements of the scale are distributed so that adjacent pairs form equal "intervals" (in $P \times P$), $\bar{R}$ (the union of the ranges of all the elements) covers S ((9) of Part 1, Section 4).[24] Note also that the cardinality of a proper P and its corresponding code are identical. Thus it is possible to classify any

[20]"Mode" is here used to indicate the particular tonic of a scale in the sense that the different "church modes" indicate different tonics in the major scale.

[21]See J. F. Schouten, R. J. Ritsma and B. Lopex Cardozo, "Pitch of the Residue", [16] and also J. E. Evetts, "The Subjective Pitch of a Complex Inharmonic Residue", [17].

[22]We are here referring to "acoustical consonance" as described by Helmholtz (in terms of coincidence of harmonics and beats), not to qualities of intervals deriving from context, (e.g., ambiguity, etc.).

[23]In fact, consistent alterations in timbre (especially of this type), when coupled to change in pitch, can alter the initial ordering (see Part 1, Section 2).

[24]In fact, it is sufficient that these "intervals" form a repeating sequence of two magnitudes only ((8), ibid)).

element of P as a "mistuning" of an element of another proper set, $\bar{P}$, with the same cardinality whose points are equally spaced. When several elements of P are simultaneously presented, S-proper modifications apply, and the interpretation of elements of P as mistuned elements of $\bar{P}$ is restricted. Note, however, that R (the union of all S-proper modifications)$\equiv \bar{R}$ (Section 5), and since $\bar{R} \equiv S$, great latitude remains. Thus, when tonality is a factor, it is usually possible to adjust the tuning of the elements of $\bar{P}$ so that each temporary tonic is reinforced (by different tones and harmonics) as it occurs (this often happens in performances of "free twelve-tone" music). In this fashion nearly all unfamiliar proper scales with the same number of elements may appear on first hearing to be different tunings of a single proper scale (provided, of course, there is only one point on the graph). This phenomenon is familiar to Western musicians who have experimented with exotic tunings of seven tone scales—when such scales are proper they often tend at first hearing to sound like mistuned major scales (or modes of such scale).[25] This also accounts for why it is sometimes erroneously stated by unsophisticated listeners that the pattern of the type of pentatonic scale in China and Thailand can be represented by the black keys of the piano. (The Thai pentatonic scale is extracted from a seven tone equal temperament system, not a twelve.)

Suppose P is no longer the only point on the graph. Since it now is necessary to be able to distinguish between P and the other graph points, the above discussion no longer applies. Even if all the graph points are different keys of the same scale, sufficient sets for each such key must be identifiable. That is, each mode of the scale must be distinct and hence each column of $\|\alpha_{ij}\|$ must be distinct. Hence, in music where one of many scales and keys (or modes) of such scales may be used, mistunings of elements of a particular P are restricted according to E-ranges or SE modifications both when P is proper and improper (Section 13). When graph points are keys of different scales, it is necessary to express each $\psi(P)$ in the same number of units per cycle, m (the cardinality of $P_1(0)$). The larger m is, the more stringent may be the restrictions on the mistunings of the elements of the individual graph points. That is, all graph-sufficient sets must now maintain their distinctness. (It is worth noting that, when experimenting with unfamiliar synthetically constructed musical scales, apparent resemblances between different proper scales of a given cardinality tend to disappear as soon as different keys of such scales are used in a musical composition).[26] Thus the entire graph of mental reference frames available to a listener and relevant to his expectations at a given time influences the limits of permissable deviations in tuning.

23. Propriety, Redundancy, the Graph and Musical Form

In general, musical form derives from a number of symmetries between different portions of a composition ranging from very small to very large units (e.g.,

[25]Of course, this effect can be easily overcome if sufficiently exotic tone colors are chosen—especially those utilizing inharmonic partials.

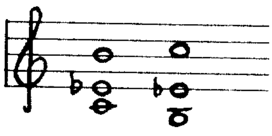
[26]This has been reported to me by musicians who have experimented with equipment for producing such synthetic scales.

motifs, phrases, sections, etc.). To a large extent the types of such symmetries are bounded by the materials used which determine the properties with respect to which symmetries (equivalence) can occur. Helmholtz writes concerning the consequences of "definiteness and certainty in the measurement of intervals for our sensation":[27]

> Upon this reposes also the characteristic resemblance between the relations of the musical scale and of space, a resemblance which appears to me of vital importance for the particular effects of music. It is an essential character of space that at every position within it like bodies can be placed, and like motions can occur. Everything that is possible to happen in one part of space is equally possible in every other part of space and is perceived by us in precisely the same way. This is the case also with the musical scale. 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.

Such a property (which amounts to a set of tones retaining their identity as a "Gestalt" when "inversion" (permutation of the order of the elements) occurs, as we have seen, is possessed only by proper scales, in fact only by strictly proper scales. Hence Western music using the major scale makes extensive use of modal transpositions of motifs. Since the major scale is also very high in efficiency, melodies sometimes depend upon phrases where doubt is not immediately resolved. Also, since the major scale has many proper subsets (see Part I, Section 8 and Figure 1, Part II, Section 13) and many intervals with roots (see the discussion preceding Figure 1) melodies make use of principal-auxiliary tone relationships determined harmonically (harmonic and non-harmonic tones), rhythmically, and/or by the ambiguity of the tritone. Notice again that the "figured bass" system is composed of the proper subsets of the major and minor scales. Hence these retain their identity when "inverted". This is not the case with improper sets. E.g.

Example 18.



Note again that a strictly proper set which is low in efficiency (or graph-efficiency, if appropriate) retains motivic symmetry as one of its principal compositional resources. Hence it is not surprising that the twelve-tone system uses motivic material as its chief resource and that the historical progression from modal to chromatic music in the West has been characterized by an increased dependence upon harmonic material and motivic symmetry.[28] It is

---

[27]Herman Helmholtz, *On the Sensation of Tone*..., [1], page 370.

[28]Clearly, music written purely in the whole-tone scale would be even more motivically dependent, since harmonic resources are sparser.

also of interest to note that the existence of an ambiguous interval in a proper scale with high stability compensates for a loss in similarity between certain few modal transpositions of motifs by facilitating partition into principal and dependent tones and by increasing the effect of cadences (Section 16 and the remarks following Figure 5).

In the case of improper scales not all the above resources are available. Motivic similarities are not possible between all portions of the scale without severe distortion (which, of course, can be deliberately utilized). Also, the requirement of instantaneous measurement of intervals necessitates a partition into principal and dependent tones and/or the firm fixing of a tonic (Section 8). Since sufficient sets with as few elements as possible facilitate identification of the tonic and/or principal tones, low efficiency (high redundancy) is extremely important for the use of improper scales.[29] In fact, low efficiency is of greater significance to the use of improper scales than high efficiency is for proper scales, since only a small part of the compositional resources of proper scales are sacrificed by low efficiency. However, failure to identify the tonic and/or principal tones of an improper scale until many tones had been heard would decimate compositional materials. That is, the principal resources of melodic form when improper scales are employed are those symmetries which depend not upon similar interval relations, but upon similarities between sequences of principal and dependent tones. (Also, the non-invertability of improper scales and the "tense" quality of contradictory intervals have expressive potential which can be exploited by the use of symmetries and distinctions derived from such properties.) The above speculations are confirmed by the ethno-musicological investigations undertaken thus far.

The symmetries referred to above can be quite subtle, such as harmonic sequence in Western music and nuclear theme in Javanese music.[30] Many more abstract relations (and relations between relations) can be found in sophisticated music. Just as the choice of materials determines those symmetries which are recognizable at lower hierarchical levels (e.g., the motivic level), the choice of such symmetries in turn circumscribes the relations at the next higher level (e.g. phrase). Upon such decisions the musical characteristics of a culture and of individual style depend. By way of analogy it is of interest to note that many of the properties of spoken languages are consequences of the choice of materials. For example, the fact that Chinese words generally consist of only one syllable coupled with the variety of words in the commonly used vocabulary may necessitate the superimposition of sliding tones (which Chinese uses) in order to avoid ambiguity.

Here we have studied those relations which apply at the "phonemic level" of musical languages and some relations which apply at the next hierarchical levels (i.e., "motif" and "phrase" level). Many of the relations which may apply between units at the level above the phonemic level are circumscribed by the points on the graph *G*, of expected "reference frames". As we have seen, the graph *G*, is utilized to define equivalences between sets of tones with respect to properties defined on the graph. The removal or addition of graph points alters

[29]Unless, of course, the tonic is fixed in advance by a drone.
[30]See Mantle Hood, "The Nuclear Theme..." [5].

such equivalences (upon which available techniques of musical composition depend).

Note that at any hierarchical level in a musical composition the identification of the units themselves (e.g., "motif", "phrase", "sentence", "section") depends upon the properties with respect to which symmetries or equivalences occur. In unfamiliar music these properties (as well as the identity of the units dependent upon such properties) are often difficult to determine. (For example, it is very difficult for a Western listener to extract the "nuclear theme" from Javanese music.) Work is in progress which treats the problem of symmetries at all musical hierarchical levels. This consists of an adaptive model and algorithm which attempts to reduce actual music to a symbolic form (which represents the relevant symmetries, relations and properties) which, in turn, is used for synthesis. The reconstructed music is subject to feedback provided by a listener who indicates where the synthesis has been successful. The algorithm then alters the symbolic representation and convergence is attempted by repeating the procedure iteratively. This work will be described in another paper.

Computer programs which perform the computations described in this paper are available on request, as are tables of such computations for unfamiliar "reference frames". These are discussed in the next paper in this series.

24. Description of Equipment and Proposed Experiments

Only a very brief description will be given here of experiments and equipment for testing the theory. The equipment is nearly complete and experiments are expected to begin shortly. Ethnomusicological testing of the theory, however, has already yielded positive results and such information is available on request.

Several reference structures ("keys" of "scales" where the tonic is not fixed —henceforth called "structures") can be constructed in which a particular interval is acoustically identical in all, but is ambiguous in some cases and unambiguous in others. It is predicted that, for a single subject who learns two structures, the perception of such an interval common to both will be more difficult and less accurate after manipulating and listening to that structure in which the particular interval is ambiguous. A keyboard controlled device is presented to the subject which produces all tones in a particular structure and *no others*. The keyboard is arranged so that no information other than a correspondence between direction and the raising and lowering of pitch is provided. The structures used are, of course, unfamiliar to the listener (the keyboard device is capable of producing *any* set of tones with less than thirty-two elements within an octave). The subject is asked to manipulate the device until he feels he is able to anticipate any pitch from the surrounding pitches he produces. This is checked by asking him to tune (adjust the knob on) an oscillator with continuously variable frequency to a given pitch when deprived of the freedom to produce it on the keyboard. This is accomplished by disconnecting the lever on the keyboard from the output of its corresponding oscillator. Alternatively, the subject may be tested by being required to select the missing pitch from several alternatives presented to him.

Several tests are made with different frequencies eliminated in this manner. When these tasks are successfully performed the subject is removed from the keyboard. The experimenter then produces a sequence forming an interval which is either "ambiguous" or "unambiguous" for the particular structure. One of the two component tones is sustained and the subject is asked to match the other tone by adjusting a variable-frequency oscillator initially set at the frequency of the sustained tone. Alternatively, he is tested by being required to select the tone from among several alternatives. His performance is measured for accuracy and speed. The experiment is later repeated with the structure exchanged for another in which the particular interval under consideration is now ambiguous if it previously was unambiguous or vice versa. The structures being used are selected so that their stabilities ($\bar{S}$) are as nearly equivalent as possible.

In another series of tests, tones *not* in the structure are produced, and the subject is required to match them on a variable frequency oscillator. We expect errors to occur in the direction of those tones of the structure in whose range the tone being matched lies, and such errors are expected to decrease as the boundaries of the range is approached.

The results of the above test may be checked against those of another series of tests wherein the subject is required to identify the tones not in the structure with tones in the structure which are "most similar". Such identification is expected to differ when the same tone is presented in the context of two distinct structures (even when tones in both structures which are adjacent to the tone to be identified are identical in both cases), provided the tone to be identified lies in the range of different tones in each structure.

Of course, the timbre of the tones used in the above experiments are of central importance. The equipment will be capable of producing a large variety of timbres. Nearly all the potentialities of a large electronic organ and of an "electronic music synthesizer" are available simultaneously by the adjustment of keyboard controls. Inharmonic partials (non-integer multiples of the frequency of a given tone) can also be produced by the depression of a single key on the keyboard. Such resulting timbres can differ for each tone in a key of a scale. In this way the construction of cases where axiom 2.3 is violated will be attempted (see the fifth paper in this series). In the experiment timbre will be adjusted for each structure (which will often contain "irrational" frequency ratios) so as to minimize its resemblance to other familiar structures (e.g., any key of a "major scale") and so that the construction of the initial ordering (as described in Part 1, Section 2) is as easy as possible.

Of course, all predictions tested by the experiments must take into account a listener-dependent minimum pitch discrimination, ε. In a series of tones each of which differs from the previous tone by less than ε, intransitivities in the relation of apparent equivalence (between tones) may occur, (e.g., $x \equiv y, y \equiv z$ and $x \not\equiv z$). Techniques for dealing with this situation are developed in the following section and are applied to the computation of experimental predictions.

The mathematical model also predicts a discrete alteration in the stability ($\bar{S}$) of a structure and in the identity of its ambiguous intervals when certain amounts of mistuning of the tones forming the structure has taken place (see

Section 25, which follows). Thus the experiment previously described is repeated with the different structures replaced by such different tunings of the same structure ($\bar{S}$ now differs for both tunings).

Another set of experiments will be performed to check predictions of the model concerning the function of sufficient sets. A particular structure (both proper and improper sets are used) is connected to the keyboard and the subject is instructed to manipulate and listen to the keyboard as before. He is then removed from the keyboard and informed that he will be presented with a sequence of tones belonging to this structure. However, he is also told that the absolute pitch location of the structure will be altered between each sequence (i.e. it may be in any key of the tuning of the scale). He is asked whether another tone added at the end of the sequence belongs to the same structure (key of tuning of the scale) as the preceding tones. He is presented with some sequences which, exclusive of the last tone, contain sufficient sets for a key and to which this tone does indeed belong, and some to which it does not. Some sequences are also presented which do not contain sufficient sets. The subject's responses are checked for identity with the predictions of the model.

These experiments are repeated with a particular element of the reference frame mistuned *after* the subject is removed from the keyboard. The limits of such mistuning which permit identifications of final tones of the sequence are checked against the computed E-range. These E-ranges are computed first for the key of the scale used and then are enlarged as much as possible so that the sufficient sets for each key remain distinct from those of other keys (which, by definition, are on the graph). Both mistunings are tested, as are mistunings which exceed such limits.

The experiments are again repeated with *several* elements of each structure mistuned after the subject is removed from the keyboard. This mistuning is varied as each sequence is presented. The variations are at first within the limits defined by SE-modifications and then exceed such limits. The cases where the subject produces both correct and incorrect identifications are correlated with the various mistunings.

A final repetition of the experiment is performed when more than one scale and its keys are on the graph. That is, before being tested, the subject is permitted to learn more than one unfamiliar structure (scale). He is then interrogated by being presented with sequences selected from keys of any such scale.

One of the by-products of the acceptance test for a subject (wherein he is required to produce missing tones in a reference structure to demonstrate that he has learned it) is the length of time required for such learning to take place. This will be correlated with the stability and propriety of the reference structures learned, and with the component intervals and the tone colors used.

Note that in these experiments, equipment is used which permits the subject to learn by *producing* stimuli. This is deliberate and derives from well-known theories and experimental results indicating that both visual and language learning is facilitated by such methods. Musical examples are well known. The teaching of musical dictation is much facilitated if the student is previously taught to sight-sing. Note also that in the acceptance test, when a subject is asked to produce a tone that has been eliminated from a learned reference frame, he is provided with feedback to his response other than just "correct" or "incorrect". He may subsequently check himself by producing the correct tone (which is immediately reconnected to the keyboard).

The equipment used in the experiments also is useful as a musical instrument capable of exploiting new musical materials. The variety of pitches and tone colors available as well as the construction of its keyboards are suited to this purpose. These aspects of the equipment will be described in a subsequent paper covering details of the musical application.

25. The Use of a Listener—Dependent ε.

For application to the perception of tones we must consider that it is impossible for a listener to order two intervals (or determine that they are distinct) when these differ by less than some small ε (dependent upon the listener and the timbre being used). This ε is an interval selected from $S \times S$ and may therefore be compared with any interval in $P \times S$. Accordingly when addition is defined as in Part I, (2.4) we define P to be *strictly ε-proper* when

$$|\alpha_{i+1,j} - \alpha_{i,k}| \geqslant \varepsilon \text{ for all } i,j,k.$$

Analogously the *ε-stability* of P is defined as

$$\bar{S}_\varepsilon = 1 - \operatorname{card}\left\{(i,j)\Big|\,|\alpha_{ij} - \inf_k(\alpha_{i+1,k})| < \varepsilon \vee |\alpha_{ij} - \sup_k(\alpha_{i-1,k})| < \varepsilon\right\}/n(n-1)$$

$$(i=1,\ldots,n-1; j=1,\ldots,n)^{31}$$

(n is the period and m is the number of units per octave.)

If $\|\alpha_{ij}\|$ is strictly proper and $\varepsilon > T$ (T in the sense of Part I, Section 3) clearly $\|\alpha_{ij}\|$ is not strictly ε-proper. However, ε will usually be used when a specific $\|\delta_{ij}\|$ has been selected, and for such cases only we define the *average tolerance*

$$\bar{T} = \left(\sum_{i=0}^{n-1} T_i\right)/mn$$

Note that if $\bar{S}_\varepsilon$ is computed with a fixed ε, $\bar{S}_\varepsilon$ will be decreased from this computed value when ε is increased to the point where $\varepsilon \geqslant T_\varepsilon$, where T_ε is defined as follows:

If $T_i < \varepsilon$, T_i^ε is computed thus: Define $V_i = \max_j(\alpha_{ij})$, $D_{i+1} = \min(\alpha_{i+1,j})$. Delete from the row $i+1$ all entries $\alpha_{i+1,j}$ for which $\alpha_{i+1,j} - V_i < \varepsilon$; call the

[31] Note that a single element of row i may be within ε of elements in both rows $i-1$ and $i+1$. Hence the symbol "$\vee$" (or) appears in the formula for $\bar{S}_\varepsilon$ so that such elements are not counted twice.

smallest one remaining $\alpha_{i+1,\bar{j}}$ and set $b=\alpha_{i+1,\bar{j}}-V_i$. Now delete from the row i all entries α_{ij} for which $D_{i+1}=\alpha_{ij}<\varepsilon$; call the largest remaining one $\alpha_{i\bar{j}}$ and set $c=D_{i+1}-\alpha_{i\bar{j}}$. $T_i^\varepsilon=\min(b,c)$. If $T_i \geqslant \varepsilon$, $T_i^\varepsilon=T_i$, $T_\varepsilon=\min(T_i^\varepsilon)$.

T_i^ε will be called the *ε-row tolerance* and T_ε the *minimum ε-tolerance*. It is to be understood that T_i^ε and T_ε will be intervals in $P\times S$ unless a particular $\|\alpha_{ij}\|$ is specified, in which case they will assume numerical values.

T_ε corresponds to the pitch discrimination required of a particular listener for the computed value of $\bar{S}$ to apply to his perceptions.

For the application to the perception of tones, which usually involves an octave cycle, we will assume that all $\|\delta_{ij}\|$ which are members of the same $\|\alpha_{ij}\|$ are drawn from sets,[32] S, such that the same interval corresponds to an octave cycle for all $\|\delta_{ij}\|\in\|\alpha_{ij}\|$ and ε is the same interval for all $\|\delta_{ij}\|\in\|\alpha_{ij}\|$. Then ε is a constant fraction of an octave; that is for all $\|\delta_{ij}\|\in\|\alpha_{ij}\|$, ε/m is constant.[33]

Hence the structure of the equivalence classes is affected by considering two intervals which differ by less than ε as "equal" (note that this is *not* an equivalence relation since it is possible that $i_1=i_2$, $i_2=i_3$ and $i_1\neq i_3$):

(a) All $\|\alpha_{ij}\|$ will have a finite number of members since when m is so large that

$$\varepsilon>\min_{ijkl}(\delta_{ij}-\delta_{kl}|\delta_{ij}\neq\delta_{kl}) \tag{1}$$

the ordering of the δ_{ij} is altered so that $\|\delta_{ij}\|$ is no longer a member of $\|\alpha_{ij}\|$.

(b) Those $\|\alpha_{ij}\|$ where $\psi(P)^*$ (K and m are minimum) is such that (1) applies will be empty. (See Section 11, Part II.)

(c) Some $\|\alpha_{ij}\|$ will acquire additional members in cases where m is such that marginally unequal intervals become equal so as to cause the ordering to coincide with that of $\|\alpha_{ij}\|$.

Since, when we consider two intervals which differ by less than ε as equal, we are no longer dealing with an equivalence relation, the equivalence class notation used thus far fails (e.g. $\psi(P)=(3,2,2,2,2)$ and $\varepsilon=1$).

To remedy this we define two scales as *ε-equivalent* if and only if whenever two intervals in the first scale differ by ε, so also do the corresponding ones in the second and vice versa. The following notation realizes this in a way convenient for automatic computation and the classes remain equivalence classes; i.e.,

$$(\|\alpha_{ij}^1\|\equiv\|\alpha_{ij}^2\|)\wedge(\|\alpha_{ij}^2\|\equiv\|\alpha_{ij}^3\|)\Rightarrow\|\alpha_{ij}^1\|\equiv\|\alpha_{ij}^3\|.$$

Let $\beta(P)$ be a vector containing the subscripts of all the terms of the first $n-1$ rows of $\|\delta_{ij}\|$. The order in which these subscripts appear will be the same as the ascending order of the values (δ_{ij}) to which each such subscript corresponds. When $\delta_{ij}=\delta_{kl}$, ij precedes kl if $i<k$ or if $i=k$ and $j<l$. All such subscripts corresponding to equal values in $\|\delta_{ij}\|$ will be enclosed by brackets.

[32]In most cases it may be assumed that all $\|\delta_{ij}\|\in\|\alpha_{ij}\|$ are drawn from the same set S.

[33]It may well be that ε is a function of $\bar{n}$ (the number of scale-notes per octave). In such case, $\bar{n}$ is constant for all members of an equivalence class, and ε may be experimentally determined.

When not separated by brackets, terms of $\beta(P)$ will be separated by dots (".").

Example 19. $\psi(P)=(3,2,3,2,2)$, $\varepsilon=0$

$$\|\delta_{ij}\|=\begin{bmatrix} 3 & 2 & 3 & 2 & 2 \\ 5 & 5 & 5 & 4 & 5 \\ 8 & 7 & 7 & 7 & 7 \\ 10 & 9 & 10 & 9 & 10 \end{bmatrix} \qquad \|\alpha_{ij}\|=\begin{bmatrix} 2 & 1 & 2 & 1 & 1 \\ 4 & 4 & 4 & 3 & 4 \\ 6 & 5 & 5 & 5 & 5 \\ 8 & 7 & 8 & 7 & 8 \end{bmatrix}$$

$$\beta(P)=(1,2.1,4.1,5)(1,1.1,3)2,4(2,1.2,2.2,3.2,5)(3,2.3,3.3,4.3,5)$$
$$3,1(4,2.4,4)(4,1.4,3.4,5)$$

If we use the symbol " $\underset{=}{\varepsilon}$ " to denote that two intervals differ by less than ε, $[(p_ip_j)\ \underset{=}{\varepsilon}\ (p_kp_l)]\wedge[(p_kp_l)\ \underset{=}{\varepsilon}\ (p_qp_r)]\wedge[(p_ip_j)\ \underset{\neq}{\varepsilon}\ (p_qp_r)]\Rightarrow[(p_ip_j)\leqslant(p_kp_l)\leqslant(p_qp_r)]$ $\vee[(p_qp_r)\leqslant(p_kp_l)\leqslant(p_ip_j)]$. Hence, $\beta(P)$ can be modified so that brackets enclose all terms which are $\underset{=}{\varepsilon}$ and the properties of an equivalence class are retained. In Example 19, if $\varepsilon=1$,

$$\beta(P)=(1,2.1,4.1,5(1,1.1,3)(2,4)2,1.2,2.2,3.2,5)$$

$$(3,2.3,3.3,4.3,5(3,1)(4,2.4,4)4,1.4,3.4,5)$$

where the first left-hand bracket pairs with the first right-hand bracket, the second left-hand with the second right-hand, etc.

Using this notation different $\psi(P)$'s can be mapped into corresponding $\beta(P)$'s using the relation of " $\underset{=}{\varepsilon}$ " and these $\beta(P)$'s mechanically compared for identity.

When $T_\varepsilon\geqslant 0$ only the first $\frac{n(n-1)}{2}$ terms of $\beta(P)$ are needed,[34] and the absence of one bracket in the final pair will indicate equality of the last group of terms to a term not shown. Also, the row subscript of δ_{ij} can be omitted in these cases if a partition sign "|" is used to separate rows. Hence the above example becomes:

$$\beta(P)=(2.4.5(1.3)|(4)1.2.3.5).$$

If any particular element of P, say P_k, is altered by an amount, φ_k, there will exist some maximum positive and negative values of φ_k for which P will remain in the same equivalence class, i.e., E-range, see Section 10. If, as above, the relation of "$\underset{=}{\varepsilon}$" is used instead of "$=$" in the mapping, these maxima will be altered.

Note that if φ_k is added to p_k only certain elements of δ_{i_j} are altered. Since $\delta_{i_j}=p_{i+j}-p_j$, these are δ_{ik} and $\delta_{i,k-i}$ for all $i<n$. If φ_k is positive, all δ_{ik} are

[34] $T_\varepsilon\geqslant\varepsilon$ guarantees that no terms from the remainder of $\|\delta_{ij}\|$ enter into the total ordering of the first $\frac{n(n-1)}{2}$ terms—see Part I, Section 6.

reduced and all $\delta_{i,k-i}$ are increased. If φ_k is negative, the reverse is true. Let φ_k^+ and φ_k^- represent the maximum positive and negative values respectively which φ_k can assume without altering the equivalence class membership of P. An examination of all possibilities yields the following formula:

$$\varphi_k^{\pm} = \begin{matrix}\min\\ \max\end{matrix}\left\{\begin{array}{l} \begin{matrix}\min\\ \max\\ i,l\end{matrix}\left\{\begin{array}{l} 1/2(\delta_{lk}-\delta_{i,k-i}\pm\epsilon)\,\big|\,|\delta_{i,k-i}-\delta|\leqslant\epsilon \\ 1/2(\delta_{lk}-\delta_{i,k-i}\mp\epsilon)\,\big|\,\delta_{lk}-\delta_{i,k-i}>\epsilon \end{array}\right. \\ \\ \begin{matrix}\min\\ \max\\ i,j,l\end{matrix}\left\{\begin{array}{l} \delta_{lj}-\delta_{i,k-i}\pm\epsilon\,\big|\,|\delta_{i,k-i}-\delta_{lj}|\leqslant\epsilon \\ \delta_{lk}-\delta_{ij}\pm\epsilon\,\big|\,|\delta_{ij}-\delta_{lk}|\leqslant\epsilon \\ \delta_{lj}-\delta_{i,k-i}\mp\epsilon\,\big|\,\delta_{lj}-\delta_{i,k-i}>\epsilon \\ \delta_{lk}-\delta_{ij}\mp\epsilon\,\big|\,\delta_{lk}-\delta_{ij}>\epsilon \end{array}\right. \end{array}\right.$$

where the upper sign ("+" or "−" or "min" or "max") corresponds to φ_k^+ and the lower to φ_k^-. The formula[35] is deliberately written redundantly to indicate the procedure for efficient computation.

The above indicates that each equivalence class when the relation "$\underset{=}{\varepsilon}$" is used has both a maximum and minimum pitch discrimination necessary for its apprehension. ε corresponds to this maximum; i.e., intervals which differ by less than ε must *not* be distinguished. Let $\Phi = \min_k(|\varphi_k^+|, |\varphi_k^-|)$. Then Φ corresponds to the minimum; i.e., intervals which differ by more than this amount *must* be distinguished.

We now define an ε-sufficient set Q_ε, as previously, except that the meaning of "subset" is altered: A set of points Q^* is called an ε-subset of P if all points of Q^* correspond to points of P biuniquely in such a way that the difference between any pair of points in Q^* and the difference between the corresponding pair of points in P differ by no more than ε.

An ε-*minimal* set is an ε-sufficient set which contains no proper subset[36] which is ε-sufficient.

ε-efficiency, E_ε is defined similarly to efficiency except that ε-sufficient sets are used.

Also, E_ε^G and E_ε^N are the same as E^G and E^N respectively, except that ε is involved in the determination of the graph and node-sufficient sets respectively.

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* **5** (1961), 2–31.

---

[35]The above formula, strictly speaking, in some cases should be "$\varphi_k^+ <$ some quantity", or "$\varphi_k^- >$ some quantity," indicating that the conditions are not satisfied at equality. However, for simplicity this is omitted.

[36]"Proper" in its usual meaning in terms of set inclusion!

3. J. C. R. LICKLIDER, Three auditory theories, in *Psychology: a study of a science*, (ed. by S. Koch) McGraw-Hill, New York, 1959, 41–144.
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, Djalarta, 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-AFOSR881-65, AF49(638)-1738 and AF-AFOSR68-1596)
8. G. M. STRATTON, Vision without inversion of the retinal image, *Psych. Rev.* **4** (1897), 341–360 and 463—481.
9. R. H. THOULESS, Phenomenal regression to the real object, *Brit. Jour. Psychol.*, **21** (1931), 339–359.
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 *and* 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. POINCARE, *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), Part 2, 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 received August 29, 1977.