



Structure and Perception in *Herma* by Iannis Xenakis

Robert A. Wannamaker



KEYWORDS: Xenakis, *Herma*, set theory

ABSTRACT: This paper presents a detailed analytical discussion of *Herma* (1961) for solo piano by Iannis Xenakis. The model of the work as a set-theoretic demonstration, advanced in the writings of the composer, is discussed. A statistical analysis reveals numerous inconsistencies between the published score and this model. The aesthetic consequences of these are explored, and an attempt is made to develop an alternative understanding of the music as it is perceived by listeners in terms of temporal gestalt formation.

Received 7 November 2000

[1] Introduction

[1.1] *Herma* (1961) for piano was Xenakis's first composition for a solo instrument. It was commissioned in 1961 by pianist/composer Yuji Takahashi, whom Xenakis met on a trip to Japan in April of that year. The Greek title may be translated as "bond," but also as "foundation" or "embryo," and perhaps reflects an intuition on the composer's part that this was to be a seminal work insofar as it was his first departure from purely stochastic means of composition.

[1.2] This technically very difficult piece makes unprecedented demands upon the performer, who must play complex rhythmic figures involving huge leaps with perfect evenness of articulation. In a good performance, the effort is repaid by the creation of a sense of seething, amorphous energy and a powerful forward momentum. The experience is unlike that associated with any prior piece in the piano literature. The ear strains to absorb it, but does not succumb to frustration or numbness. The frenzied sonic activity maintains a surprising sense of perceptual lucidity, never becoming muddy.

[1.3] The composer himself provides a detailed theoretical discussion of *Herma's* construction in his book, *Formalized Music*.⁽¹⁾ Referring to *Herma* as an example of "Symbolic Music," Xenakis advances a model of the piece which involves the exemplification of specific mathematical relationships between certain pitch sets. The macro-structure of the composition is clearly outlined as well. Details regarding the specific pitch and rhythmic choices made are not included, however, and, while the composer makes indications of what *ought* to be listened for in the piece, a typical listening unavoidably raises certain questions which the composer does not address regarding what *actually is* heard. In particular, the ability of listeners to recognize the proposed model in the music, and the ability of the model to account for important aspects of the music as heard, both beg examination.

[1.4] This paper explores the relationship of the score to its set-theoretic model, and the relationship of both to the music as it is perceived. [Section 2](#) introduces the composer's published model of the composition, expanding on his treatment. [Section](#)

3 subjects the musical materials of *Herma* to statistical analysis, which is appropriate given the composer’s stochastic means of pitch selection. The statistical deployment of pitches is compared both to the published model and to the perceptual organization of musical details as revealed through the close reading of a representative score excerpt from the viewpoint of gestalt perception. Detailed examination of the score for consistency with the model is undertaken in Section 4 and certain significant discrepancies are identified. The implications of these are considered in Section 5 within a broader discussion of aesthetic issues raised by the relationship of the model to listeners’ perceptions and by the writings of the composer. An Appendix provides a brief introduction to the necessary mathematical background.

[2] *Herma* As a Set-theoretic Demonstration

[2.1] The composer adopts as a universal pitch set⁽²⁾ R , the set of all 88 keys on the standard piano keyboard. From R he selects three subsets, A , B , and C , with non-empty intersections. The central conceit of the piece is to present these pitch sets and others derived from them by means of elementary set operations in a fashion which will demonstrate that a given target set, F , can be expressed in terms of A , B , and C using two different set-algebraic forms (one of which is the disjunctive normal form discussed in the Appendix).

[2.2] **Figure 1** is redrawn following *Formalized Music* (p. 176) with minor corrections.⁽³⁾ It illustrates diagrammatically how the set F may be expressed in two different fashions:

$$F = ABC + A\bar{B}\bar{C} + \bar{A}B\bar{C} + \bar{A}\bar{B}C \quad (1)$$

$$= GC + \bar{G}\bar{C} \quad (2)$$

where I have introduced the convenient shorthand

$$G = AB + \bar{A}\bar{B}.$$

The figure is divided in half by a bold horizontal line, the upper portion (labeled “PLANE 1”) corresponding to Equation 1 and the lower (labeled “PLANE 2”) to Equation 2. Each “PLANE” is further subdivided into two rows of Venn diagrams, each row having a different dynamic marking to its right-hand side.

[2.3] The line of diagrams marked *fff* illustrates the four atomic sets appearing in the disjunctive normal form of F , while the line labeled f shows selected intermediate stages in the computation of these as intersections of A , B , C , and their complements. Together these two rows of diagrams are apparently intended to demonstrate how the four relevant atomic sets are computed and united to yield F . The solid arrow-tipped lines connecting the diagrams are evidently meant to suggest the temporal flow of an imagined mathematical construction, consisting in essence of the line marked *fff* with lemmata interjected from the line marked f .

[2.4] The two lines which “PLANE 2” comprises similarly illustrate the computation of Equation 2, demonstrating diagrammatically that the result is the same as that specified by the disjunctive normal form. The development begins with the “construction” of $AB + \bar{A}\bar{B}$ on the line marked *ppp*, before branching to illustrate the parallel constructions of GC (marked *ff*) and $\bar{G}\bar{C}$ (marked *ppp*), whose union is F . This illustrates the manner in which the figure preserves information about the structures of Equations 1 and 2, insofar as the rules of operator precedence (“order of operations”) required to compute the right-hand-sides of these equations partially dictate the order in which the Venn diagrams are arranged as read from left to right. Of course, the arrangement depicted is not the only one which might be employed to this end, since there is mathematical flexibility regarding the order in which the intermediate quantities are computed, so long as the order of operations is observed.⁽⁴⁾

[2.5] The broken arrow-tipped lines in the diagram seem mostly to represent supporting observations traversing the two PLANES. Thus two broken lines lead from ABC and $\bar{A}\bar{B}\bar{C}$ in PLANE 1 to their union, $(AB + \bar{A}\bar{B})C$, in PLANE 2. Some of the sets connected by broken lines, however, lack a simple algebraic relationship. For instance, $AB + \bar{A}\bar{B}$ is not immediately relatable to $\bar{A}\bar{B}C$.

[2.6] The composer describes the above set-theoretic considerations as being *outside-time* (p. 170), indicating that no reference to temporality is made or required. He observes that, to find employment in music, entities *outside-time* must somehow be combined with a *temporal* event schema to yield events *in-time*.

[2.7] In *Herma*, a pitch set is articulated *in-time* as a sequence of pitches drawn from the set. (Indeed, the introduction of each

new pitch set is indicated directly in the score with the name of the set.) The top line of **Figure 2** indicates the succession of different sets presented. Thus *Herma* commences with a substantial passage composed using notes selected from the set R (i.e., the entire piano keyboard). A section using only the notes of the pitch set A follows, with successive sections drawn from the contents of the sets \bar{A} , B , \bar{B} , C , and \bar{C} in that order. Then the derived pitch sets corresponding to each Venn diagram in Figure 1 are introduced with the dynamic markings given in the figure and with new sets being introduced at the indicated dynamic and roughly in the order shown (reading from left to right). Some sets are given multiple expositions with subsequent presentations being marked *rappel* (“recall”) in the score. Different pitch sets are sometimes sounded simultaneously, with contrasting dynamics and temporal densities being employed to aid in their discrimination. Sometimes, as in the A and B sections, the *same* pitch set is articulated in two simultaneous layers with contrasting dynamics and temporal densities, consisting of, in the composer’s terms, a “linear” (i.e., temporally sparse and dynamically loud) layer and a “cloud” (i.e., temporally dense and, usually, dynamically soft) layer, with the sustain pedal depressed whenever a “cloud” is being sounded.

[2.8] Individual pitches are drawn at random from the given sets without registral preference (except at one point in the \bar{C} section: cf. [paragraph 3.6](#)). This employment of stochastic procedures, retained throughout the composition, is intended to avoid melodic or harmonic patterns that would serve to distract the listener from the perception of the pitch sets and relationships between them *as such*. In the words of the composer:

If we want to be free the sounds should follow without any melodic law, independently of one another. So we have to play them at random. . . . How can I demonstrate the elements of the sets? By playing them. But in order to remain neutral I have to play them at random. ⁽⁵⁾

The elements of each class are presented stochastically, that is unrestrictedly, in order not to disturb the basic plan of operations and of logical relationship between classes. ⁽⁶⁾

[2.9] Each randomly selected pitch is associated with an attack randomly placed in time. The mean attack density per unit time does not change continuously with time, but assumes a sequence of constant values specified by the composer, with the abrupt change to each new density being indicated at the appropriate point in the score (e.g., as “0.8 s/s” or 0.8 sounds per second) and in *Formalized Music* (p. 177). Note durations appear to be randomly selected as well, although this is made definite neither by the score nor by the composer’s writings about the piece. The opening R section exhibits a wide variety of rhythmic figures, but thereafter values divisible by a sixteenth-note or quintuplet sixteenth note (5:6) are employed almost exclusively. The work is notated mostly in 12/8 time with the exception of three brief occurrences of 6/8 in the second half of the work and the opening R section, which begins in 4/4 and visits a wide variety of time signatures. These markings notwithstanding, the composer’s preface to the score indicates that, “The whole piece is to be played without accents, the bar-lines serving merely as divisions in time.”

[2.10] All pitch sets are musically represented strictly by means of displaying their members, randomly selected. Set algebraic operations (i.e., intersection, union, complementation), relationships between sets (e.g., equality), and rules of logic (e.g., implication) are not represented as such. The composer contends that

If the observer, having heard A and B , hears a mixture of all the elements of A and B , he will deduce that a new class is being considered, and that a logical summation has been performed on the first two classes. This operation is the *union*. . . .

We have not allowed special symbols for the statement of the classes; only the sonic enumeration of the generic elements was allowed. . . .

. . . we have not allowed special sonic symbols for the three operations . . . ; only the classes resulting from these operations are expressed, and the operations are consequently deduced mentally by the observer. In the same way the observer must deduce the relation of equality of the two classes, and the relation of implication based on the concept of inclusion (pp. 171–172).

[2.11] Here, then, we have a model of *Herma* as an exemplification using pitch sets of the equivalence of two set-theoretic expressions. Pitch sets are represented by sequences of their members, randomly selected, with the succession of pitch sets corresponding to an exhibition of selected intermediate quantities that would be produced in a step-by-step computation of the expressions in question. The listener is left to deduce the relationships between different pitch sets and the meaning of the succession (i.e., to recognize that something is being exemplified, and what that particular something is). The composer

aply compares the function of musical time here to the expanse of a chalkboard:

The role of time is again defined in a new way. It serves primarily as a crucible, mold, or space in which are inscribed the classes whose relations one must *decipher*. Time is in some ways equivalent to the area of a sheet of paper or a blackboard (p. 173).

[2.12] From the mathematical model of *Herma* discussed above, we now turn to questions regarding the perception of the piece by listeners. Requirements for a listener to be capable of following the set-theoretic exposition in question would have to include the following:

1. complete and accurate characterization of each pitch set involved;
2. an ability on the part of the listener to discriminate and compare the contents of different pitch sets, requiring accurate recollection of the contents of pitch sets previously presented.

The next two sections address in detail the constitutive materials of *Herma* and the auditory perception thereof in light of the above criteria, and will attempt to identify any potential obstacles to the perception of the underlying set-theoretic model.

[3] Musical Segments in *Herma*

[3.1] In describing the musical segmentation of *Herma*, I will use the language of temporal gestalt formation as developed by James Tenney.⁽⁷⁾ A detailed introduction to the subject of temporal gestalt formation in music is beyond the scope of this article, and we must restrict ourselves to a few relevant definitions. For a thorough discussion of the topic, the reader should consult Tenney's book. This terminology and conceptual framework are introduced in order to identify correspondences between the set-theoretic model and the musical segmentation as I hear it (and as I believe most listeners will hear it).

[3.2] Following Tenney, we will adopt the term *holarchical* instead of *hierarchical* in order to avoid any connotation of relative importance. For our purposes, we will regard the lowest holarchical gestalt level as that of the note and the highest as that of the piece. A gestalt at the holarchical level immediately above that of the note is referred to as a *clang* and constitutes "a sound or sound-configuration which is perceived as a primary musical unit or aural gestalt."⁽⁸⁾ A gestalt at the next highest level is referred to as a *sequence* and consists of "a succession of clangs which is set apart from other successions in some way, so that it has some degree of unity and singularity, constituting a musical gestalt on a larger perceptual level or temporal scale. . . ."⁽⁹⁾ Finally, a temporal gestalt all of whose component, next-lower-level temporal gestalts exhibit the same overall statistical properties (mean, variance, etc.) in some particular sonic parameter are said to be *ergodic* in that parameter.⁽¹⁰⁾

[3.3] The second line in Figure 2 shows the temporal gestalt units associated with *Herma* at the first and second holarchical levels below that of the piece as I hear them. I will refer to these as *sections* and *subsections*, respectively.

[3.4] Inspection of the figure shows that sectional and subsectional boundaries are correlated with changes in temporal density, dynamic and timbral envelope (i.e., use of the sustain pedal). The composer systematically deploys such changes to aid in discrimination of the pitch sets. During the first 250 seconds of the piece, these three parameters exhibit simultaneous changes which produce clearly demarcated sections and which are correlated with changes in pitch set. Throughout the work as a whole, gestalts at the sectional level are defined primarily by drastic changes in temporal density, with periods of silence or undisturbed sustained tones separating sections. These delineations are frequently supported by large changes in dynamic, changes in the use of the sustaining pedal, or both. Less drastic and coordinated changes in temporal density, dynamic, or sustain are responsible for temporal gestalt formation at the level of subsections.

[3.5] During the later portion of the work, when pitch sets change rapidly and temporally overlap one another, successive pitch sets are not separated by periods of inactivity and are distinguished only by changes in dynamic or pedaling. As a result, the holarchical level at which changes of temporal gestalt unit and changes of pitch set are correlated falls from that of the section to that of the subsection. This occurs around 250 seconds into the piece, at which point the first sets corresponding to intersections are introduced. After this, sections, which are still demarcated by hiatuses, no longer serve to define single pitch sets. This change itself, however, is not a perceptually salient event. In summary, we observe that all subsection boundaries coincide with pitch-set boundaries and *vice versa*, with the sectional/subsectional schema serving to embed the temporalized pitch-sets within a perceptual holarchy.

[3.6] With one significant exception, the active pitch range remains the entire range of the piano keyboard throughout the composition, and the random selection of pitches is apparently uniformly distributed over the entirety of the active pitch set.

The single brief exception occurs during the \bar{C} section between 230 and 250 seconds into the piece. At this point, the range engaged narrows towards the middle of the keyboard, culminating in a dramatic, fortissimo, chordal passage preceding the introduction of the first pitch sets derived from intersections. Otherwise, the pitch materials of the piece are ergodic at the holarchical level of section. That is, sequences within a subsection are distributed uniformly among all registers of the piano and this situation persists throughout the piece with the exception noted. (Our discussion will later return to the apparent aesthetic contradiction between a compositional approach based primarily on the articulation of specific pitch distributions and a perceived ergodicity in pitch materials.)

[3.7] At lower holarchical levels, the situation is more complex. **Figure 3** shows the frequency of occurrence of each pitch on the piano keyboard during the articulation of each pitch set. As in hexadecimal notation, the decimal numbers 10, 11, 12 . . . are represented by the capital letters A, B, C . . . so that all entries in the table are single characters. We observe that all pitch classes are represented somewhere within each of A , B and C , thus ensuring a fairly chromatic musical texture. We also note that each of these sets consists largely of clustered “clumps” of pitches separated by considerable intervals, a tendency which is inherited by their complements and intersections. Otherwise the sets exhibit no simple structure. The composer describes them as “amorphous.”⁽¹¹⁾

[3.8] It should be clear that randomly sounding notes that are uniformly distributed over such “clumpy” sets may sometimes produce an impression of multiple polyphonic lines (i.e., *auditory streaming* according to similarity in pitch⁽¹²⁾). This effect is perceptible throughout much of the piece, although it is usually complicated by other factors. Generally speaking, events stream according to complex interactions between their proximity in register (determined partially by the macro-structure of the pitch sets and partially by chance), their proximity in time, and their similarity in dynamic.

[3.9] **Figure 4** shows a typical excerpt from the score: the final four measures of the exposition of set A , where “linear” and “cloud” versions of the set are simultaneously articulated. Circles indicate *clangs* as I typically hear them, with sequences of *clangs* connected by ties. The initial G4 is segregated from subsequent events by its dynamic, but such segregation is not consistent throughout the passage, as is immediately illustrated by the perceived grouping of the next three notes despite differences in dynamic. Here this is largely a product of the temporal proximity of the attacks, but also arises in part due to the continuous use of sustain, which blurs together temporally clustered events in the bass more effectively than ones in the treble. The following sequence in the treble is segregated by difference in pitch from the middle register tones, reflecting the macro-structure of A (i.e., the distinct clumps of pitches in A in the ranges F5–E6 and F4–G4, respectively). Indeed, each sequence identified in Figure 4 falls into a single one of the following clumps suggested by Figure 3 (except in the highest register, where the two-highest pitched clumps are regularly traversed): F1–F2, B2–D3, G3–G \sharp 3, D4–G4, F5–E6, and A6–D7.

[3.10] This should not, however, be regarded as evidence that the macro-structure of the pitch sets is the primary determining factor in low-level temporal gestalt formation. Many of the intervallic skips observed in this example between notes within a single clump are larger than the pitch intervals separating distinct clumps, so other factors clearly also contribute. As already indicated, these include similarity in dynamic and temporal proximity, but also subtler matters of accentuation and articulation which may vary from one performance to another. (For instance, the D2 in the third of the four measures shown in Figure 4 is inaudible in Takahashi’s performance.) Furthermore, the attitude of the listener makes an important contribution. In casual listening, for instance, one may attend primarily to the fortissimo notes, ignoring the pianissimo ones and thus perceive a much different set of gestalts.

[3.11] Perhaps the most important perceptual consideration regards the sheer complexity of the sonic information presented. In producing Figure 4 I proceeded in a rather traditional manner by attempting to identify distinct melodic voices one at a time, an approach which already presumes a very particular strategy for listening (i.e., one of attending to only to activity in a certain register in an attempt to identify a coherent melodic voice). It frankly seems impossible to grasp the seething entirety confronted as a whole, and when I try to do so my ear flits from one sequence to another, ignoring certain events in favor of others. The events perceived as salient seem to be as much the product of how one listens (possibly favoring certain registers, dynamics, etc.) as of the auditory stimulus, and the sense of the music as being “too much to take in” is indispensable to its overall musical effect.

[3.12] To a large extent, seemingly, it is this perceptual complexity and instability which lends the work its considerable musical richness, producing a kaleidoscopic profusion of fractured pitch sequences in different registers between which the listener’s attention tends to skip. The uniform distribution of activity over the entire keyboard ensures that the density of sound in any given register remains relatively sparse. This produces a certain sense of lightness and lucidity, as opposed to the

sense of density which would ensue if such a rapid succession of attacks were confined to a narrower pitch range.

[4] Discrepancies Between Theory and Realization

[4.1] A number of significant issues arise upon detailed comparison between the set-theoretical conception of *Herma* and its realization as a musical score. The first of these becomes apparent upon inspection of the R pitch set tabulated in Figure 3. It is clear that, although R nominally contains all 88 pitches playable at the standard piano keyboard, several are not actually represented in the introduction to the piece where R is articulated. In particular, $D\sharp 1$, $E1$, $F\sharp 1$, $B2$, $G\sharp 3$, $D\sharp 4$, $G4$, and $C5$ are never sounded. This is presumably the consequence of using a random drawing procedure to determine each sounding pitch in succession, with pitches drawn remaining eligible to be drawn again. The exposition of R in *Herma* comprises only 205 notes, but, using such a random selection scheme, a much larger number of draws will usually need to be made before all 88 different possible pitches have been obtained. Consider, for instance, a situation in which 87 different pitches *have already been drawn*. The final pitch remains one among 88, so on average a further 88 draws will be required before it is drawn. Indeed, a computer simulation employing ten million trials of the pitch lottery has shown that an average of 446 draws are required to obtain all 88 possible pitches (with a standard deviation of 110 draws), and that the probability of obtaining all 88 pitches in 205 or fewer draws is extremely small (about 3 in 100000). It seems reasonable to suppose that all 88 pitches are not represented in the introduction of the piece because the composer wished to curtail its duration without modifying his method of pitch selection.

[4.2] Nonetheless, the fact that certain pitches in any given set may not be acoustically articulated seems problematic. If the object of the composition is to demonstrate certain set-theoretic facts, then a complete representation of the sets of interest would seem to be a necessary prerequisite. An obvious solution to the problem of sectional duration would be to not return pitches drawn to the lottery. Then if one begins with a lottery containing n representatives of each pitch, one is assured that after $88n$ draws each pitch will have sounded precisely n times. In any event, no such procedure was adopted in the work at hand.

[4.3] By definition of set complementation (see [Appendix](#)) we know that notes appearing in the A section of *Herma* ought theoretically to be absent from the \bar{A} section. Inspection of Figure 3 indicates that, in fact, there are twelve instances of identical pitches occurring in both sections. It is unlikely that these anomalies are the result of artistic license on the part of the composer, for two reasons. The first of these is the unambiguous commitment to the idea of a set theoretic demonstration cast in pitch which Xenakis makes in *Formalized Music*, a commitment which would be thoroughly betrayed by compromising the integrity of the pitch sets once they are defined. The second reason is that no aesthetic motivation for the introduction of the specific discrepancies is evident either in the score or upon listening. The perplexing notes do not seem to be specifically demanded in any obvious way by their musical context.⁽¹³⁾

[4.4] We must consider the possibility that notes found in common between a given set and its complement are errors arising either at the time of composition or in the typesetting of the score. The latter seems unlikely on the following grounds. *Herma* was written in 1961 and premiered by Takahashi a few months thereafter. The Boosey and Hawkes score (see [note 6](#)), the only score of the piece which has been published, is copyrighted 1967 and the data of Figure 3 is derived therefrom. However, the Denon recording of *Herma* by Takahashi was made in 1972.⁽¹⁴⁾ The latter recording features at least some of the pitch discrepancies mentioned above. (Not all have been verified by listening, but at least $B2$, $B3$, $D4$, $D\sharp 4$, $F5$, $F\sharp 5$, and $C6$ are clearly audible in both A and \bar{A} .) Takahashi was in frequent contact with Xenakis throughout the 1960's,⁽¹⁵⁾ and so the composer would have had ample opportunity to indicate corrections to the score if there were any to be made. Furthermore, since Takahashi had performed the work prior to its publication, he would likely have been working not from the Boosey and Hawkes score when he recorded it, but rather from a copy of Xenakis's manuscript. This suggests that the anomalies in the published score may have survived from the original manuscript.

[4.5] Comparing A , B , and C with their complements, 33 separate instances are observed of pitches appearing both in a set and in its complement. Proceeding for the sake of argument on the assumption that such a characterization is appropriate, we can attempt to identify which pitch occurrences are "erroneous." Most of the discrepancies are of the sort in which a pitch occurs several times in a given set and only once in its complement. In such cases, the single occurrence may be assumed to represent the "error." In other cases, a pitch may be relegated to one set or another after the examination of sets occurring later in the score. For example, the appearance of $C8$ in $\bar{A} \bar{B} C$ means that it must belong to \bar{B} rather than to B . Not all contradictions are so easily resolved, however. For instance, the pitch $B1$ occurs twice in A and twice in \bar{A} . The only subsequent sets in which $B1$ is repeatedly observed are G and $\bar{G} \bar{C}$. Neither observation definitively places $B1$ in either A or \bar{A} and, worse, the sets in question are themselves disjoint! In the end, we surmise that $B1$ belongs in \bar{A} on the weak evidence

that other pitches in A sound several times during the exposition of that set (i.e., significantly more than twice). The question of whether $B1$ belongs in B or \bar{B} presents similar difficulties. We place it in \bar{B} based on the fact that it appears in neither the exposition of $\bar{A} B \bar{C}$ nor in that of F .

[4.6] Eliminating problematic pitches from A , B , and C allows the generation of what I will call “best-guess” sets. These can be used to algebraically compute the contents of pitch sets appearing later in the work. The “best-guess” sets and sets derived from them using algebraic computation software are shown in Figure 3 as shaded areas. These are to be compared with the numeric entries representing the observed frequency of occurrence of each pitch in the score. (Occurrences in the sections marked *rappel* have been included.)

[4.7] The agreement between observed and computed sets is excellent, with a few exceptions. Some expositions of sets (such as those of $\bar{B} \bar{C}$ and \bar{G}) do not exhibit all of the expected pitches, perhaps only because they are too brief. Indeed, \bar{G} is expected to contain 27 different pitches, but its exposition comprises only 12 notes.

[4.8] In two cases a set is presented at length in the score but shows little agreement with the predicted set. In the case of $\bar{G} \bar{C}$ fewer than half of the scored notes fall within the predicted set. Curiously, the notes that fall outside the computed set occur almost exclusively within the first exposition of $\bar{G} \bar{C}$ while the notes heard in the four occurrences of this set marked *rappel* agree almost perfectly with the prediction. It is almost as if the first exposition is not of $\bar{G} \bar{C}$ at all, but is of some other set which is mislabeled in the score and in *Formalized Music*. Apparently this is not the case, however, for the following reason. $A\sharp 2$ and $A3$ are definitely members of the atomic set $\bar{A} \bar{B} \bar{C}$. Thus, if the “unknown set” can be expressed in terms of intersections, unions and/or complementations of A , B , and C , then it must contain $\bar{A} \bar{B} \bar{C}$ as a subset (see the [Appendix](#) regarding disjunctive normal form). There are *very many* other members of $\bar{A} \bar{B} \bar{C}$, however, which are not presented in the exposition in question, so it would seem that the “unknown set” heard in the first exposition marked $\bar{G} \bar{C}$ cannot be expressed in terms of the sets A , B , and C .

[4.9] The other set that agrees poorly with the predictions is $\bar{A} \bar{C}$. In this case, the contents of the initial presentation of the set and the single presentation marked *rappel* agree with each other, although it would also appear that the set observed cannot be expressed in terms of A , B , and C .

[4.10] Where agreement between prediction and observation is generally good, a few discrepancies usually remain. Particularly troublesome is the pitch $D\sharp 3$, which occurs three times in \bar{A} , five times in \bar{B} , and ten times in \bar{C} . This would imply that it is a member of $\bar{A} \bar{B} \bar{C}$ and therefore ought not to occur in F (see Figures 8 and 9 in the [Appendix](#)). Nonetheless, it is observed to occur nine times there. One possible explanation is that the composer deliberately introduced this pitch so that F would contain at least one representative of each pitch class. Indeed, $D\sharp$ would not be represented in F if the offending $D\sharp 3$ were excised. This explanation is partial at best since $D\sharp 3$ is also present in pitch sets occurring earlier in the work where it ought not to be found, such as in $\bar{A} \bar{B} C$ and $G C$. Alternatively, the fact that $D\sharp 3$ is observed twice in C might prompt us to include it in C in spite of its repeated observation in \bar{C} which would then account for its presence in $\bar{A} \bar{B} C$, $G C$ and F .

[4.11] Whatever the case may be, we must conclude that the mapping of the proposed set-theoretic model of *Herma* onto its score does not satisfy the first criterion for the audible intelligibility of the model given at the end of Section 2. First, the sets of interest are not completely represented in the score. It has already been remarked that this deficiency is manifest in R . The problem recurs, however, with respect to various pitch sets throughout the score. In each case there exist pitches that are represented in the score neither within the set in question nor within its complement.

[4.12] Furthermore, it cannot be said that the sets of interest are accurately represented from a mathematical perspective. The appearance of a pitch both within a given set and within its complement baldly contradicts the definition of complementarity. Inconsistencies appear to persist throughout the score based on the set-algebraic computations discussed above, in which significant disagreement between computed sets and their representations in the score is repeatedly observed.

[4.13] The second intelligibility criterion from Section 2, requiring that listeners be able to discriminate and compare the pitch sets of interest, also raises concerns. Psychological studies investigating the ability of subjects to name musical tones with different pitches have found that they can reliably do so only when the number of tones is less than 5–6 (although subjects with absolute pitch perform better).⁽¹⁶⁾ A , B , and C each have more than twenty members and their complements, of course, have considerably more. Even allowing for absolute pitch to assist in the identification of different pitches, the capacity of short-term memory for items differentiated in a single dimension is usually taken to be 7 +/- 2 items so that a

listener should not be able to retain such large harmonically amorphous pitch sets in short-term memory.

[4.14] It might be hoped that the clumpy macro-structure of the pitch sets could be of use in distinguishing them insofar as the emergence of voices occupying slightly different registers might be used to identify particular sets and transitions between them. In addition to the fact that set macro-structure is only one of several factors influencing stream formation (cf. [Section 3](#)), other impediments to such identifications would exist. Due to the substantial number of different sets presented, the time required for a comprehensive sampling of their members to sound, and the considerable overlap between their contents apparent in [Figure 3](#), it would be necessary to remember the detailed contents of several streams over a significant period of time in order to perceive the “clumps” with sufficient completeness and certainty to distinguish between sets. This would again overtax the usual limits of short-term memory.

[4.15] Theoretically, it might be possible with sufficient listening to commit each note of the piece to long-term memory, but this would represent an extremely unusual listening situation. It must be noted that Takahashi, after practicing the piece for months, wrote to the composer a letter containing the following passage: “Theoretical conceptions of such great interest. Would like to have technical explanations. Let me know about Symbolic Music.”⁽¹⁷⁾ If a practiced performer with a score in hand cannot grasp the work’s construction, is it reasonable to assume that a listener should be able to do so?

[5] A Critical Summary Regarding *Herma*

[5.1] Ultimately, the numerous inconsistencies between the score and its set-theoretic model which are discussed above remain, perhaps, a minor issue. In my opinion, they detract little from the effectiveness of the piece. This observation in itself, however, highlights certain deeper paradoxes associated with the compositional process.

[5.2] The dramatic surface of *Herma* conceals a deep-seated contradiction between the technical approach of the composer and its musical products. Xenakis is clearly serious about his conception of the work as a temporal blackboard upon which is inscribed a set-theoretic argument demonstrating the equivalence of two different expressions for the target set *F*. Nonetheless, this “argument” is entirely impenetrable to the listener. It is unrealistic to suppose that even musically astute listeners can apprehend and retain in memory large, harmonically amorphous pitch sets which are subjected to purely stochastic expositions. From a perceptual standpoint, the music is ergodic in pitch at the highest holarchical gestalt levels. Certainly no sense is ever produced of an argument cast in pitch proceeding towards a conclusion. If it were otherwise, one supposes that the set-theoretic inconsistencies discussed above would not be performed and recorded by musicians of the highest competence.

[5.3] It seems clear that no sense of forward motion or cathartic resolution is communicated by the pitch content of the piece. Rather, the momentum of the music is entirely a product of its rhythmic and dynamic intensity and the tremendous overt technical virtuosity which must be displayed by its performer. Similarly, the richness of the musical fabric is a product not of its specific pitch content, but rather of the complex and unstable perceptual streaming produced by the “clumpy” macro-structure of the pitch sets and the use of simultaneous contrasting dynamics and temporal densities (cf. [Section 3](#)). This result is ironic insofar as dynamics, rhythm, and timbre are all compositionally subordinated to specific pitch content, purportedly serving as devices to temporally demarcate the expositions of different pitch sets. Thus a fundamental contradiction exists between the compositional approach and the perceived result.

[5.4] This would be less remarkable if Xenakis had not previously been harshly critical of certain serialist composers on just such grounds. In his well-known article “The Crisis of Serial Music,”⁽¹⁸⁾ the composer upbraided serialists for

1. the fundamental contradiction between their compositional methods (involving the manipulation of musical lines) and the results thereof (a perception of mass or surface, with no audible sense of “line”);
2. subsuming all other musical parameters to that of pitch.

It would seem that both of these criticisms could be as easily leveled at Xenakis’s own compositional approach in *Herma*, insofar as the set theoretic construction is not audibly manifested despite the marshalling of dynamic, temporal density and timbral envelope (i.e., sustain) to aid in the discrimination of the different pitch sets.

[5.5] Within Xenakis’s compositional output as a whole, it may be that *Herma* is best viewed as a seminal/transitional work, as suggested by its title. Its straightforward set-theoretic construction is unique in Xenakis’s oeuvre. Immediately subsequent compositions returned to purely stochastic considerations (*ST/10*, *Atrees*, and *ST/48* of 1962) or group theory (*Akrata*, *Nomos Alpha*, and *Nomos Gamma*, composed between 1964 and 1968). The latter pieces also mark the appearance of Xenakis’s

theory of scales (or *sieves*, in his terminology; see *Formalized Music* 194–200), which probably represents the most direct descendent of the set theoretic techniques adopted in *Herma*. This assertion is corroborated by Xenakis’s own comments:

In *Herma* I chose sub-sets from the chromatic scale—that is, I chose some of the points on the straight line. After that I put to myself the following question: How can one carry out this process on a more general level which would comprehend all the scales used in the past and all those that may come into use in the future? The sieve theory gives the answer. This answer may not be complete, but it’s certainly effective and many-sided.⁽¹⁹⁾

Sieves are constructed with the aid of standard set-algebraic operations applied to pitch sets, but these primitive sets display definite regularities (unlike the “amorphous” *A*, *B*, and *C* sets of *Herma*). Thus the products of these operations can be made more easily apprehensible and discriminable to listeners.

[5.6] In my opinion, none of these considerations detract from the fine musical qualities to be found in *Herma*, among which the following must be counted. First, its texture seems almost amorphous and yet it maintains tremendous forward momentum (although it supplies no reassuringly obvious goal). The impression created is perhaps reminiscent of certain natural (volcanic or meteorological) processes. In this light, the inscrutability of its internal logic seems less troubling insofar as nature also operates by laws which are often concealed, although the composer’s writings seem to suggest that this is not the light in which he would prefer to have the piece viewed (*Formalized Music*, 170–177).

[5.7] Furthermore, *Herma* refreshingly manages to eschew almost every established cliché of piano writing, although it cannot be said that it is entirely without precedent. Xenakis studied under Messiaen during the 1950s and it is reasonable to assume that *Herma* is informed by Messiaen’s “Mode de Valeurs et d’Intensités” from *Quatre Etudes de Rythme* (1949–50). The latter piece represents the first example of “total organization” composed in Europe, and as such is the product of a quite different compositional approach from that employed in *Herma*. Nonetheless, as Xenakis himself has observed regarding total serial organization (see [note 18](#)), the perceived results can be similar to those obtained using random procedures and the qualitative similarities between the two pieces in question are easy enough to hear. There is no question, however, that *Herma* goes far beyond previous models in unreservedly embracing the notion of a seething mass of sound as a viable alternative to linear writing. Hearing a good performance such as Takahashi’s seems like nothing so much as getting caught in a storm.

[5.8] In the end it may seem that *Herma* is successful for listeners on grounds other than those which facilitated its construction by the composer. What is the proper object of analysis in such a case? As in much music, there may exist subtly or markedly different analyses germane to the composer, the performer and listener. (Perhaps even to each listener.)⁽²⁰⁾

[A] Appendix: Elementary Set Theory

[A.1] A *set* or *class* is a collection of objects, called *members* or *elements* of the set. Let *R* denote a *universal set* consisting of all elements of interest. A set *A* is called a *subset* of *R* if every element in *A* is also a member of *R*. Given such a subset *A* we may define its *complement*, \bar{A} , as the (possibly empty) set of elements which are in *R* but not in *A*. Such relationships among sets are often illustrated by means of Venn diagrams of the sort shown in **Figure 5** in which the rectangular area represents the universal set, *R*, with subsets thereof represented by areas demarcated within the rectangle. In **Figure 5** a set *A* is represented by the area within a circle, while the area outside of the circle but within *R* represents \bar{A} .

[A.2] The *intersection*, *AB*, of two sets, *A* and *B*, is that set which consists of all elements which belong to both *A* and *B*. Similarly, one may define the *union*, *A+B*, of two sets, *A* and *B*, as that set which consists of all elements which belong to either *A* or *B*. The intersection and union of two sets are illustrated by the filled regions in **Figures 6(a)** and **6(b)**, respectively.

[A.3] Two sets which have no elements in common are said to be *disjoint*, and their intersection is represented by a forward slash superimposed on a circle, the *empty set*. Two such disjoint sets are illustrated in **Figure 7**.

[A.4] The set operations defined above can be applied to three or more sets. **Figure 8** illustrates how the universal set, *R*, may be *partitioned* into eight disjoint sets equal to intersections between three sets *A*, *B*, *C* and their complements.

[A.5] Observe that any set formed from unions, intersections, and/or complementations of *A*, *B*, and *C* can be expressed as a union of the disjoint sets which are illustrated. (These disjoint sets are sometimes referred to as *atoms*.) For instance, the

complex looking set, F , of **Figure 9** may be expressed as

$$F = ABC + A\overline{B}\overline{C} + \overline{A}B\overline{C} + \overline{A}\overline{B}C.$$

Any set thus written is said to be expressed in *disjunctive normal form*. In general, there are other ways in which a given set may be expressed, and it is this fact that provides Xenakis with a conceptual starting point in the composition of *Herma*.

Robert A. Wannamaker
Music Department
York University
4700 Keele Street
Toronto, ON M3J 1P3
CANADA
rob@audiolab.uwaterloo.ca

Footnotes

1. Iannis Xenakis, *Formalized Music*, rev. ed. (Stuyvesant, N.Y.: Pendragon Press, 1992). Page number references to this book are given in the text henceforth.

[Return to text](#)

2. Throughout this article, we concern ourselves with pitches, as opposed to pitch classes, unless otherwise indicated.

[Return to text](#)

3. Two minor errors in PLANE 2 of the diagram as given by Xenakis have been corrected in **Figure 1**. First, the region corresponding to ABC in the leftmost figure on the *ff* line has been shaded, where it was blank in the original. Second, a missing overline has been added to the character C in the rightmost figures on each line (both *ppp* and *ff*). Also, I have deduced that the broken line between $\overline{A}\overline{B}C$ and $(AB + \overline{A}\overline{B})C$ should have its arrowhead pointing at the latter rather than the former.

[Return to text](#)

4. Xenakis observes that the form of Equation 2 is more computationally efficient than the disjunctive normal form of Equation 1 in the sense that it involves fewer total union, intersection, and complementation operations. The difference is ten operations versus fourteen, assuming that intermediate results are available for multiple re-uses once computed. In particular, observe that once $AB + \overline{A}\overline{B}$ has been computed, it can be used to compute \overline{C} by means of a single complementation, as opposed to computing \overline{C} from scratch using A , B , and C . This fact has little practical implication for the composition, however, since not all of the different operations involved are sequentially illustrated. PLANES 1 and 2 each contain fewer than the respective fourteen and ten diagrams, with the \overline{C} diagram, furthermore, being repeated. Xenakis indicates that seventeen operations, rather than fourteen, are required to compute the right-hand side of Equation 1 (*Formalized Music*, 173). This would be the case if the complements \overline{A} , \overline{B} , and \overline{C} could not be multiply re-used in computations. Such re-use of intermediate results is, of course, necessary to achieve the computational efficiency associated with the alternative form specified by Equation 2.

[Return to text](#)

5. Bálint A. Varga, *Conversations with Iannis Xenakis* (London: Faber & Faber, 1996).

[Return to text](#)

6. Iannis Xenakis, *Herma* (London: Boosey & Hawkes, 1967).

[Return to text](#)

7. James Tenney, *META ≠ HODOS: A Phenomenology of 20th-century Musical Materials and An Approach to the Study of Form; and META Meta ≠ Hodos*, 2nd edition (Oakland: Frog Peak Music, 1992).

[Return to text](#)

8. *Ibid.*, 87.

[Return to text](#)

9. Ibid., 94.

[Return to text](#)

10. Ibid., 113.

[Return to text](#)

11. Varga, *Conversations*, 85.

[Return to text](#)

12. Albert S. Bregman, *Auditory Scene Analysis: The Perceptual Organization of Sound* (Cambridge, Mass.: MIT Press, 1990).

[Return to text](#)

13. One might be tempted to explain the apparent inconsistencies by invoking the notion of “fuzzy sets,” first introduced into the scholarly literature in 1965 by engineer and mathematician Lotfi A. Zadeh as a generalization of the classical set concept (Lotfi A. Zadeh, “Fuzzy Sets,” *Information and Control* 8 (1965): 338–353). Is it possible that Xenakis anticipated this concept in 1961?

A “fuzzy set” is one whose elements can possess degrees of confidence of membership between zero (in which case the element certainly *is not* in the set) and one (in which case the element certainly *is* in the set). In a classical (or “crisp”) set, on the other hand, the degree of membership of any element is either *precisely* one or *precisely* zero.

If the degree of membership of an element x in a given fuzzy set A is, for example, 0.8, then the degree of membership of x in \overline{A} is $1 - 0.8 = 0.2$. Thus, one would expect to observe x in a randomly drawn sample set whether it was drawn from A or \overline{A} (albeit with greater frequency in a sample set drawn from A).

Nonetheless, I maintain that a fuzzy set model is not appropriate for *Herma*, for the following reasons. First, nowhere in his detailed theoretical discussion of *Herma* in *Formalized Music* does the composer introduce any concept relatable to fuzzy sets. Indeed, his description of complementation clearly assumes classical sets: “If class A has been symbolized or played to [the listener] and he is made to hear all the sounds of R except those of A , he will deduce that the complement of A with respect to R has been chosen” (page 171).

Furthermore, it should be noted that it is impossible in principle for a listener or analyst to determine with certainty the precise degree of membership of any given pitch in a fuzzy pitch set from a finite random sampling of the set in question. That is, if it were given that fuzzy sets were employed in *Herma*, it would not be possible to definitively characterize them either upon hearing the work or upon inspection of the score. The composer unequivocally advances a model of the work as a set-theoretic argument cast in pitch. The introduction of fuzzy sets would not only represent an unmotivated complication of the exposition, but would prevent the terms in that exposition from being determinable with perfect confidence.

[Return to text](#)

14. Iannis Xenakis, *Herma*, Yuji Takahashi, piano. Denon 33CO–1052 [1972]. Compact disc.

[Return to text](#)

15. Nouritza Matossian, *Xenakis* (New York: Taplinger, 1986), 147.

[Return to text](#)

16. Brian J. C. Moore, *An Introduction to the Psychology of Hearing*, 4th Ed. (San Diego: Academic Press, 1997), 246.

[Return to text](#)

17. Matossian, *Xenakis*, 151.

[Return to text](#)

18. Iannis Xenakis, “La crise de la musique serielle,” *Die Gravesaner Blätter* 1 (1955): 2–4. Cf. Matossian, *Xenakis*, 85–86.

[Return to text](#)

19. Varga, *Conversations*, 96.

[Return to text](#)

20. I would like to thank Profs. David Lidov and James Tenney for their encouragement and guidance during the writing of

this article. As well, I would like to thank two anonymous reviewers for their comments and MTO Editor Eric Isaacson both for his helpful suggestions and for performing some computer simulations which instigated those reported in Paragraph 4.1.

[Return to text](#)

Copyright Statement

Copyright © 2001 by the Society for Music Theory. All rights reserved.

[1] Copyrights for individual items published in *Music Theory Online* (*MTO*) are held by their authors. Items appearing in *MTO* may be saved and stored in electronic or paper form, and may be shared among individuals for purposes of scholarly research or discussion, but may *not* be republished in any form, electronic or print, without prior, written permission from the author(s), and advance notification of the editors of *MTO*.

[2] Any redistributed form of items published in *MTO* must include the following information in a form appropriate to the medium in which the items are to appear:

This item appeared in *Music Theory Online* in [VOLUME #, ISSUE #] on [DAY/MONTH/YEAR]. It was authored by [FULL NAME, EMAIL ADDRESS], with whose written permission it is reprinted here.

[3] Libraries may archive issues of *MTO* in electronic or paper form for public access so long as each issue is stored in its entirety, and no access fee is charged. Exceptions to these requirements must be approved in writing by the editors of *MTO*, who will act in accordance with the decisions of the Society for Music Theory.

This document and all portions thereof are protected by U.S. and international copyright laws. Material contained herein may be copied and/or distributed for research purposes only.

Prepared by Brent Yorgason and Tahirih Motazedian, Editorial Assistants