EWK

Reprinted from JOURNAL OF MATHEMATICAL PSYCHOLOGY
All Rigts Reserved by Academic Press, New York and London

Vol. 30, No. 3, September 1986
Printed in Belgium

A Theory of Information Structure
II. A Theory of Perceptual Organization

MICHAEL LEYTON

Harvard University

A theory of perceptual organization is offered that accords with the Description Postulate presented in Paper I; i.e., the theory claims that perceptual organization is defined by the input group of a machine which has a decomposition sequence prescribed by an algebraic stability ordering. An extensive analysis of the decomposition sequence appears to explain and predict a number of organizational phenomena: For example, the analysis appears to show (1) that the group sequence is divided at some point along it into two components, one determining the internal relational structure of a form and the other determining the form's extrinsic structure; (2) that the interaction between the internal and external aspects of a form is defined by a specific product operation at this point in the group sequence; (3) that a rigorous theory of grouping can be obtained by the analysis of certain subsequences within the internal sequence; (4) that the perceived hierarchical structure of a form is determined by the ordering on those subgroup subsequences; (5) that perceptual cartesian reference frames are also group sequences which conform to the Description Postulate; (6) that the impression of the imposition of a cartesian reference frame, on a form, is determined by subsequences of the group decomposition sequence of the form.

14. THE MEANING OF PERCEPTUAL ORGANIZATION

14.1. *Introduction*

The cognitive principles elaborated in Part I will now be applied to yield a unified theory of perceptual organization.

The 20th century has seen a number of major approaches to the study of perceptual organization. These approaches are quite different and can, in fact, exist independently of each other because they attempt to handle apparently separate content domains. For example, Binford (1971) and Marr and Nishihara (1978) have concerned themselves with complex natural shapes such as that shown in Fig. 20a, or artifacts such as the goblet shown in Fig. 20c. These researchers have claimed that the figures are given by the assignment of local symmetry axes (e.g., to the limbs of the man) which in turn allow the shapes to be described as a concatenation of roughly cylindrical modules, where the symmetry axes become the rotational axes of the cylinders. Marr and Nishihara further propose that any such

Requests for reprints should be sent to Michael Leyton, Departments of Psychology and Computer Science, State University of New York, Buffalo, NY 14260.

0022-2496/86 $3.00

shape is given by a hierarchy of successively greater detailing. Thus, for example, at the first most rough stage of detail or resolution, the entire man in Fig. 20a, can be approximated by a single upright cylinder shown in Fig. 20b. At a second finer stage, the man can be given by a cylindrical torso with attached cylindrical limbs —which is the diagram shown in Fig. 20a. Thus there is a progression from Fig. 20b to Fig. 20a. At yet a third stage of resolution, a set of finer cylinders, the toes and fingers, are discerned. Note that the main cylinder in Fig. 20b can progress to the goblet in Fig. 20c, where the latter resolution reveals a different structure of detail than the former. Also working in this school of analysis, Blum (1973; Blum & Nagel, 1978) and Brady (1983) have examined more general shapes such as Fig. 20d and developed algorithms to extract smooth, curved, trajectories of symmetry axes.

With a very different set of concerns, the German Gestalt movement (e.g., Wertheimer, 1923) provided several criteria (regularity, symmetry, etc.) by which stimulus sets are formed into groupings. For example, Fig. 20e is grouped into rows

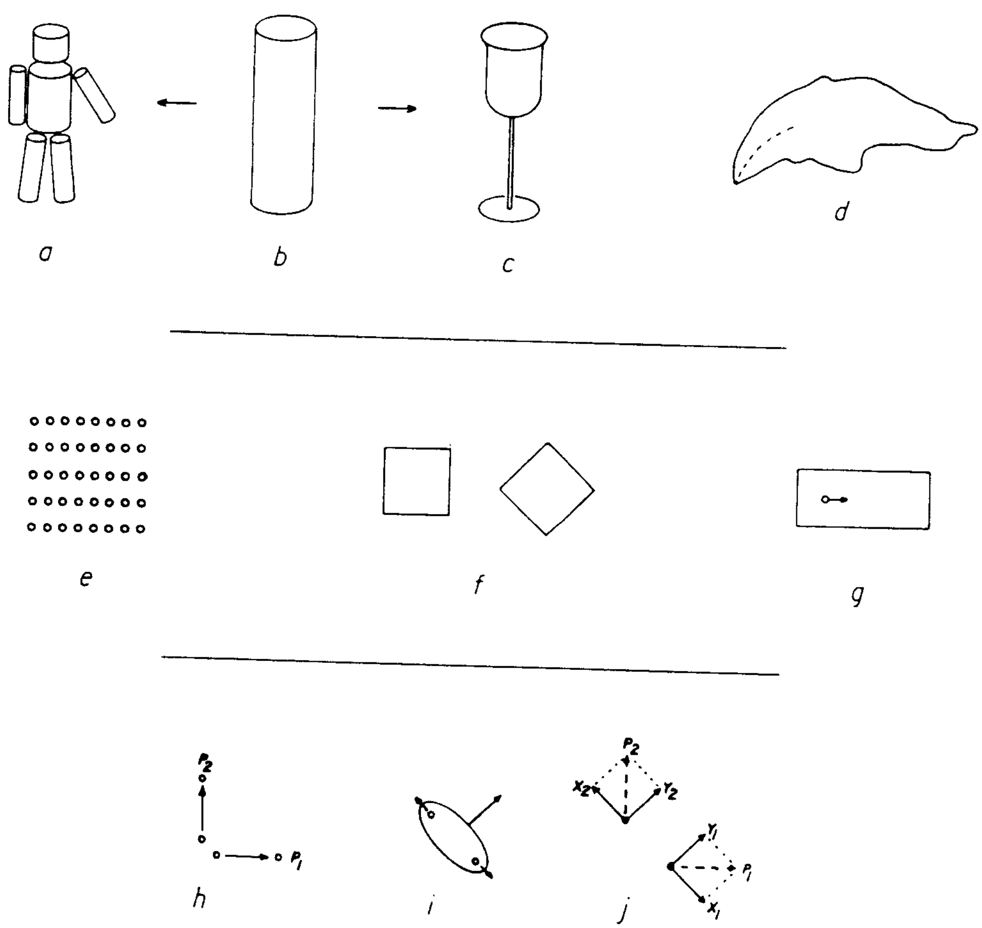

FIG. 20. Some of the examples to be given a unified explanation in this paper: (a–c) complex shape hierarchies; (d) complex shapes with curved symmetry structure; (e) grouping; (f) orientation and form phenomena; (g) induced motion; (h) relative motion.

rather than columns because the dots in the rows are closer together than the dots in the columns. Another concern of the school was the orientation-and-form problem, of which Fig. 20f (discovered by Mach, 1897) is the most famous: One orientation of a regular four-sided figure can look like a square while another looks like a diamond. In another area of investigation, Duncker (1929) showed that if a dot is fixed while a rectangular surround is moved (Fig. 20g), the resulting percept is that of a fixed surround and a moving dot. The larger object was argued to be the reference frame.

In a quite different perceptual school, Johansson (1950) provided an elegant theory of the following type of phenomenon: Johansson had discovered that if two dots, as depicted in Fig. 20h, are moved perpendicularly and in phase, they are not perceived as such. Instead they are perceptually organized as the ends of a diagonal invisible rod which is being stretched and contracted along its length while moving in the opposite diagonal direction. The explanation offered by Johansson was that the motion vectors P_i (Fig. 20h) are split into two distinct pairs: (X_1, X_2) which define the *relative* motion of the dots away from each other, and (Y_1, Y_2) which define the *common* motion of the pair.

Each of the above theories has been strongly corroborated. And yet it is evident, when one examines them, that they have a very different explanatory character from each other, being particularly attuned to the class of phenomena for which each was derived. To make any of the explanations applicable to any of the other domains, one would have to introduce some generalizations or ad hoc arguments.

The present paper will attempt to provide a unified theory of these domains. We shall not, however, proceed by generalizing any of the above explanations or introducing ad hoc arguments. We will start by extending the theory of information structure developed in Part I. A mathematical exposition of the argument has previously appeared in Leyton (1984b), and a neuronal model has appeared in Leyton (1985).

14.2. *The Full Organizational Group*

In this section, I will introduce the main proposal of this theory of perceptual organization.

To motivate the discussion we look first at a simple example. Consider a square: With respect to this percept, one can define two input-space descriptions. (1) In Sections 8 and 9, I tried to show that a square is cognitively embedded in a transformation space defined as SL_2R; i.e., the square is perceived as capable of being stretched, sheared and rotated in particular ways, with the associated stability ordering on the subgroups. Therefore, SL_2R describes a structure of inputs that can be *externally* applied to the figure. The term *external* will refer to the fact that this group *alters* the percept. External groups will be denoted by the symbols G_E. (2) On the other hand, we can give the square exactly the same type of description that we gave to a hexagon in Fig. 9. In the latter, the 12 elements of D_6 distributed themselves around the sides of the hexagon, two elements per side. They described how the hexagon can be generated *internally* from only one side. The term *internal* is used to denote the fact that the generation is from *inside* the hexagon, that is, from a proper subset of the figure—a side. Correspondingly, the square can be internally generated by the group D_4. This input group generates the shape from a single side, e.g., the top side. D_4 has eight elements, and thus they will be distributed around the square, two per side, in exactly the same way as D_6 was distributed around the hexagon. Internal groups will be denoted by the symbols G_I.

To summarize, it has been claimed here that the square has two input-space descriptions, one from the group SL_2R which is external because it alters the square, and the other from D_4 which is internal because it generates the percept from a subset of the square. I claim that the *percept*, the square, is given by the *interaction* of the descriptions from SL_2R and D_4. Two laws which appear to determine this interaction will be offered later.

The illustration that has just been given is very simple. However, I will attempt to show that a wide variety of perceptual organizations are the interaction between an external and an internal group; and that the recognition of this leads to the solution of a number of previously unsolved problems. Thus the basic postulate of this paper is.

PROPOSAL 14. *A* **perceptual organization** *of a stimulus set, S, is the interaction of two descriptions in terms of machines*:

(1) *A description from an* **external group, G_E**, *which describes how the set S can change, and*

(2) *A description from an* **internal group, G_I**, *which generates S from a subset of itself.*

A central claim that I am making is that the two groups interact. In fact, I will attempt to show that they form a *sequence*, $G_E \cdot G_I$, which accords with the principles that were developed in the general cognitive theory. In the present case, the product "·" will describe also the *interaction* of G_E and G_I. That is, one can define:

DEFINITION 15. The **full group of a perceptual organization** is the group sequence

$$G = G_E \hat{\imath} G_I$$

where G_E and G_I are the external and internal groups, respectively, and $\hat{\imath}$ is the product describing their interaction.

In what follows, the proposal will be initially applied to simple shapes. However, as more theory is developed of the constructs just defined, the proposal will be applied to obtain a theory of *perceptually ambiguous shapes* (Sect. 17), of *various motion phenomena* (Sect. 18), and of *complex, natural and abstract, shapes* (Sect. 19). Although most of the discussion will be of two-dimensional perceptual organizations, we shall describe, in the lengthy Footnote 11, how the same analytic tools can be generalized to handle three-dimensional representations. The analysis of complex *scenes* will be discussed in the paper on planning, because the analysis to be offered depends on a theory of trajectory planning.

15. The External Group

Because there was much discussion, in Part I, of SL_2R as the external group of a square, little needs to be added here except the following two comments.

15.1. *Why the Stability Ordering $A \cdot N \cdot SO_2$?*

Let us first try to understand more fully why subjects gave the stability ordering of SL_2R to be rotations followed by shears followed by pure deformations, rather than an alternative ordering on those three subgroups.

A concept from linear algebra will be important: An *eigenspace* is a straight subspace, e.g., a straight line, which passes through the origin and remains over itself under a transformation of the overall space (e.g., Hoffman & Kunze, 1961). The concept can be illustrated by returning to the transformations in SL_2R.

(1) *Rotations.* Under any rotation, the entire space shifts around the origin (see Fig. 21a). Every line is moved off itself. Thus there is no eigenspace.

(2) *Shears.* When a shear acts, there is one and only one line through the origin which stays on itself. For example, in Fig. 21b, the invariant line is the bold line, the x axis. Thus in that figure, the x axis is the eigenspace of the shear transformation shown.

(3) *Pure deformations.* When a pure deformation acts, there are exactly two straight lines through the origin which stay on themselves. In Fig. 21c, these are the two bold lines shown; that is, the x and y axes. Note that any other line, through the origin, moves off itself. Thus, in Fig. 21c, the x and y axes are each eigenspaces of the transformation shown.[1]

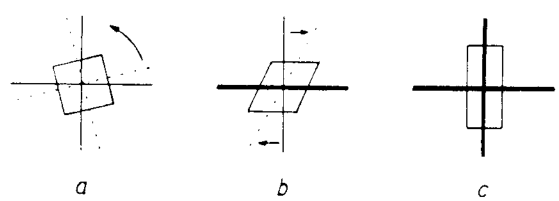

Fig. 21. The eigenspace lines (bold lines) of (a) the rotation group (none), (b) the shear group (one), and (c) the pure deformation group (two).

[1] $A \cdot N \cdot SO_2$ is the standard Iwasawa decomposition of SL_2R. In this classical decomposition, A is the set of unimodular diagonal matrices. Because the matrices are diagonal, the entries are eigenvalues. Because the matrices are unimodular and hence non-singular, there must be two linearly independent eigenvectors and hence two invariant lines.

The importance of the eigenspace concept is that it is a *stability* phenomenon. An eigenspace is an invariant. An invariant is a property which remains unchanged, i.e., stable. Thus the size of eigenspace is one measure of the amount of stability associated with a transformation. That is, a transformation with more eigenspace is more stable.

The crucial thing now to observe is that the algebraic stability ordering chosen by the subjects is actually the order of the sizes of the associated eigenspaces. That is, as can be seen from the above discussion, *the successive groups—rotations, shears, and pure deformations—have, respectively, zero, one, and two eigenspace lines; i.e., increasing stability using the eigenspace criterion.* Therefore, it seems reasonable to suppose that the particular algebraic stability ordering, chosen by subjects for the subgroup decomposition of SL_2R, is based on the phenomenon of eigenspaces. Corroboration for this view will be offered during the course of Part II.

15.2. *The External Cartesian Reference Frame*

It is virtually impossible to do any theoretical research on perceptual organization without soon requiring the notion of a cartesian reference frame. The square/diamond problem (Fig. 20f) is only one example of the importance of this notion. One claim to be made and supported in Part II is that the cartesian frame has three groups of transformations that are cognitively associated with it.

If one gives someone a blank sheet of paper, and asks them to draw a cartesian frame on the paper, they will almost certainly draw the frame (1) in the horizontal–vertical orientation (which will be called the *gravitational* orientation to simplify the discussion), (2) such that the horizontal and vertical axes are perpendicular to each other, and (3) such that there is a sense that a unit distance along the horizontal equals a unit distance along the vertical.[2] To summarize: a frame is drawn that is (1) in the gravitational orientation, (2) perpendicular, and (3) square.

In terms of representing data (e.g., graphically) in a cartesian frame, none of these conditions need hold. One has an infinite number of frames of varying orientation, shearness, and non-squareness which can adequately structure data. However, only one is usually chosen.

The explanation offered here, for this, is that the following three groups are cognitively associated with the frame, in accord with theory presented in Part I.

(1) The rotation group, SO_2, highlights the property of orientation. This group allows the cognitive system to reference frames of any orientation to the gravitationally aligned one. This is done by imaging the rotation group onto the collection of such frames. Thus, each frame is identified with a particular rotation. Therefore, one can consider there to be an input-space description (map) from the rotation group onto the collection of frames of any orientation. Recall that it was

[2] Note that these choices could be determined by educational factors. However, what will be important to recognize is that the choices are cognized as the most stable. Stability values can be set either culturally or by evolution.

argued in Section 9.2 (Proposal 9) that reference is the factorization of instability. This implies that the rotated frames are all perceived as dynamically tending to the gravitationally aligned one, which is perceived as the most stable case. Therefore, the set of all such frames forms a dynamical system. Again, in accord with the Principle of Stimulus Identity (Sect. 12), the gravitational frame, being the stable or non-input state, is chosen to represent the entire reference system of frames of any rotation; i.e., it is the *machine object*, in the sense of Definition 4.

(2) The shear group, N, highlights the important frame-property that one can call *rectangularity*. The group allows one to perceptually reference skewed frames to the rectangular case. One does this by imaging the shear group onto the collection of such frames; that is, any frame is identified with a particular shear transformation. Thus there is an input-space description which maps the shear group onto the collection of such frames. Furthermore, because, reference, as I have argued, is a stability phenomenon (Proposal 9 and Definition 12), the sheared frames are all perceived as dynamically tending to the rectangular one, which is perceived as the most stable case. Thus the set of all such frames forms a dynamical system, i.e., machine. Observe that, in accord with the Principle of Stimulus Identity (Sect. 12), the rectangular frame, being the stable or non-input state, is chosen to represent the entire reference system of frames of any shearness; i.e., it is the *machine object*.

(3) The pure deformation group, A, identifies and highlights the property of the cartesian frame which one can call *squareness*. The group can be understood as the means whereby one references frames, in which the width/height ratio is not unity, to the case where the ratio is unity. An input-space description from the group onto the set of such frames causes them to be structured as a dynamical system (machine) tending to the square case, which receives the identity element and is therefore the machine object. Again, by the Description Postulate, the latter frame is the most stable case, and by the Principle of Stimulus Identity, any other frame is identified as essentially that object with deformation.

Thus it seems that the cognitively salient aspects of the cartesian reference frame imply the action of SL_2R. Recall now that I have argued that a reference frame is a group with an algebraic stability ordering (Definition 10). The above discussion seems to have established the cartesian reference frame as an example of this. The group SL_2R can be considered to be imaged on the entire system of frames, labeling each frame with a member of the group and ordering the group members by algebraic stability. This view of the perception of the Cartesian reference frame is captured by:

DEFINITION 16. The **External Cartesian Reference Frame** is SL_2R with the algebraic stability ordering defined by increasing eigenspace dimension in the group's usual decomposition. Given a set, X, an assignment of an External Cartesian Reference Frame is an input-space description

$$C_E\colon SL_2R \to X \cup \{0\}.$$

In what follows, the term Cartesian Reference Frame, with upper case initials, will refer to the concept which I have defined here; and the term cartesian reference frame, with lower case initials, will refer to the conventional construct. My claim is that Definition 3 embodies what is *cognitively* understood by a cartesian reference frame, and that the conventional definition represents this incorrectly (despite being useful, for example, in mathematical descriptions, etc.).

16. The Internal Group

Proposal 14 states that a perceptual organization of a stimulus set S is the interaction of two groups: G_E, the external group, describing the *allowable changes* of S; and G_I, the internal group, describing a way S can be *generated* from a subset of itself. A theory of the internal group is now offered.

16.1. *The Decomposition of the Internal Group*

I will first argue that the *internal* structure of a perceptual organization conforms to the Description Postulate (Sect. 11). That is, I propose:

PROPOSAL 17. *The internal organization group is perceptually factorized into a product of groups*

$$G = G_1 \cdot G_2 \cdots G_n$$

which has an algebraic stability ordering.

In what follows, the meaning of Proposal 17 will first be examined and then experimental corroboration will then be offered. In the subsequent sections, the structure defined in the proposal will be explored still further.

Let us consider the internal group of a square. So far we have regarded this group as D_4. However, the experiments described later will apparently show that the group is perceptually larger, in this way: D_4 merely generates the square from one of its sides. However, let us now look at the internal structure of a side itself. It is perceptually generated from one of its points by applying, to that point, the group, R, of translations along a line (the group was discussed in Sect. 4.2.3, Example 2i). Thus, for example, if the central point[3] of a particular side receives the identity element of R, then any point in the side becomes labeled by the translation, g,

[3] In Leyton (1984c), I argue that a stimulus set is given a *class* of input-space descriptions. I define a *G-description subclass* to be a subclass where the domain of each of the members is one particular group G. In the case of the side, the end points can also be referents for the side (i.e., in addition to the central point—the referent demonstrated in Sect. 16.2). Thus, there is an R-description subclass where three of the members of the class have, respectively, the three referents just listed. To simplify the discussion, I will confine considerations to only the member that maps the identity element to the central point.

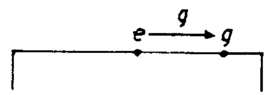

FIG. 22. The action of the group R along a polygon side.

which obtains the point from e. Figure 22 illustrates this. Thus one can consider there to be an input-space description D: $R \to S \cup \{0\}$ where S is a side. This string of symbols simply reads: the description D takes the group R of translations and images it onto a side S. Thus, for example, the particular translation g, in R, is imaged onto a particular point on that side. (Note that the ends of the line can be specified here by letting D map all of the translations past the ends of the side to the set $\{0\}$.)

With respect to a square, we now have two groups: (1) D_4, which describes a square as generated from a side; and (2) the group R which describes a side as generated from a point. My claim is that the two groups, R and D_4, are fitted together in the following way. It will be helpful to consider Fig. 23a. Choose any point in any of the sides of the square, for example the point x_3 in Fig. 23a. The point is referenced, via the group R, to x_2, the central point on that side. Then, assuming that the most stable side in the D_4 description is the top side, the point x_2 is referenced, via the group D_4, to the point x_1 on the top side.[4] Thus the internal group is factorized into the decomposition sequence $D_4 \cdot R$.

Consider what this view implies about the perceptual structure of a square. The result of successively applying the two groups is that each point in the square can be regarded as perceptually given by a pair of coordinates, (g_1, g_2) in $D_4 \cdot R$; the first coordinate g_1 being the member of D_4 which obtains the side from the top side, and the second coordinate g_2 being the translation which obtains the point from the central point of that side. Figure 23b illustrates this view by exhibiting the coordinates associated with the factorization sequence given above. The figure takes the most stable point in the entire square to be the central point on the top side. Algebraically, this point therefore has only identity elements for coordinates. That

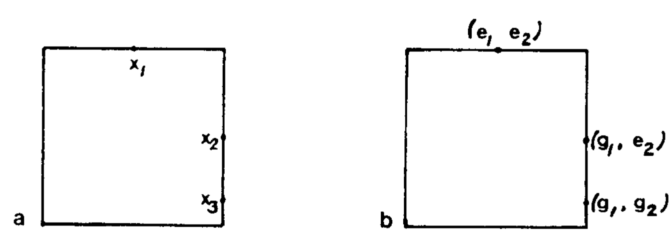

FIG. 23. (a) A successive reference sequence on a square. (b) Its coordinization in $D_4 \cdot R$.

[4] In fact, I found (Leyton, 1984a) that either the top or the bottom are chosen as referents. Thus, my claim (in Leyton, 1984c) is that one assigns a D_4-description subclass to the square, where one member of that class is a D_4 map starting at the top side of the square, and another member is a D_4 map starting at the bottom.

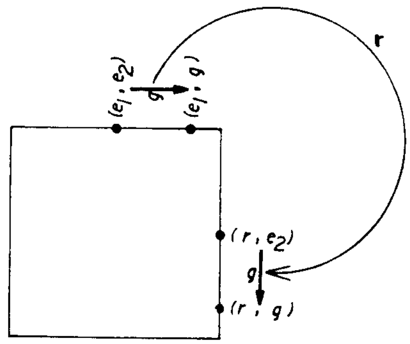

FIG. 24. D_4 acting as a control group with respect to R; that is, mapping, via r, the translation g on one side to the translation g on another.

is, it is given by the coordinate pair (e_1, e_2), where e_1 is the identity element in D_4 and e_2 is the identity element in R. By applying an operation g_1 from D_4, one obtains the central member of another side, e.g., the right side in Fig. 23b. The latter central point therefore has coordinates (g_1, e_2). Finally, one obtains the point x_3 by applying the translation g_2 in R. This gives x_3 the coordinates (g_1, g_2).

Observe that the process can continue like this: If the stimulus set were a set of squares arranged in a two-dimensional grid, the input group would be $Z^2 \cdot D_4 \cdot R$, where the final reference factor, Z^2, is the input group which generates the grid of squares from a single square. (The group Z^2 was discussed in Sect. 4.2.3, Example 2ii.) This last factor would reference any square in the grid to the most stable square.[5]

Let us conclude this section by considering here a very simple feature illustrated by the above description of the square. As we shall see later, this feature is the consequence of a profound phenomenon. The feature is this: The action of D_4 takes the structure of generation of one side (i.e., created by R) and images it onto the structure of generation of every other side. Thus, for example, consider the following two points on the top side: The central point (e_1, e_2), and an arbitrary point (e_1, g) on that side. The relationship between the two points is a translation g given by the top horizontal arrow shown in Fig. 24. The pair of points are mapped via the rotation r (in D_4) onto the pair of points (r, e_2) and (r, g) on the right side. But these latter points are related by the translation g in R. Thus, the translation g in R, on the top side, is transferred via an operation r in D_4 to the translation g on the image side.

16.2. *Preliminary Corroboration*

Because this section describes a set of supporting experiments, the reader who wishes to follow the argument without a break should go straight to Section 16.3.

[5] Note that experiments performed by Goldmeier (1937/1972, Chap. 3) indicate that various factors, such as proximity, can alter the ordering of the groups just mentioned. In Leyton (1986a), I have presented a view of the Gestalt criteria in terms of the machine-theoretic approach developed in these papers, arguing that the criteria maximize the amount of machine structure (algebraic connectivity and stability) in any description. This view will be further developed in the last paper in the present series.

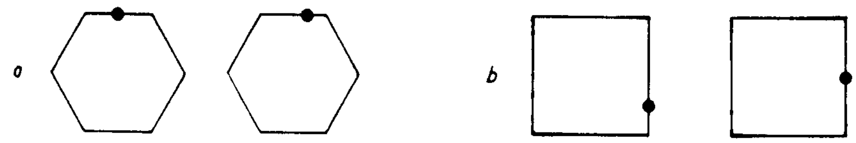

FIG. 25. Examples of comparison pairs given to subjects who were asked which point was the referent or the more stable.

Empirical corroboration for the above view will be offered from several directions as the theory unfolds, and its consequences are analyzed, in the following sections. However, as a preliminary corroboration, a succession of very simple experiments are described. They are ordered so that they correspond approximately to the successive levels of reference which have been described in the discussion above. The first experiment attempts to check that the central point of a side is a referent for that side. The next tries to show that subjects use the operations from D_n to relate the sides. The experiment also attempts to show that the transformations preserve the internal structure of each side. The final experiment tries to demonstrate that the D_n descriptions use the top or the bottom sides as the identity.

The experiments might be regarded as virtually trivial were it not for the fact that *the simple results seem to contradict the only other extensive formal theory of perceptual organization which exists in perceptual psychology today—the coding theory of Leewenberg (1971).* In the latter theory, any point is chosen as the initial point of the generating code; any side has exactly the same code letter and thus the same status as any other side; and there is no fully explicit hierarchical structure in the code. These important aspects of coding theory have not so far been challenged.

(1) *The central point of a side is a referent.* In Leyton (1984a, Experiment 6), subjects were presented with pairs of the same polygon (see Fig. 25 for examples). Two positions were identified (using dots) on a common side of the polygon pair. Half of the subjects were asked which of the two positions was "that position with respect to which the other position is judged; i.e., the referent." The other half of the subjects were asked which of the two positions was the "more stable or at rest."[6]

The results were that subjects chose the central point significantly in both conditions [$n = 48$ per comparison pair; expected mean = 24 choosing center; worst score = 45, yielding $t(11) = 7.0$, $p < .0005$, one-tailed]. Observe that the fact that the same replies were given in the two conditions (the reference and stability questions) seems to corroborate the claim, in the Description Postulate, that the algebraic ordering is a stability ordering. This thereby appears to substantiate Proposal 9, that reference is the factorization of instability. Note that the result can be regarded as violating coding theory, which does not consider that there is a privileged origin on a line. Thus the present experiment must rule out most of the codes predicted by coding theory, because those codes can have any point as an origin.

[6] There were 48 subjects. Each subject was presented with 24 pairs—half the total number of possible pairs. Thus, there were four groups of subjects with 12 in each group.

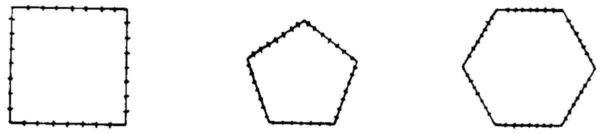

FIG. 26. The polygons used in the task of creating a one–one mapping of one side to another.

(2) *The action of* D_n. This experiment tries to corroborate the proposal that D_n is indeed the relevant group in the perceptual analysis of an n-sided polygon. Essentially, the experimental task here is to check that the generators of D_n—which are r, the smallest rotation, and t, the reflection—are indeed the perceived relations between the sides of a regular polygon.

However, it is necessary to also test another aspect of the theory of internal groups. In Section 16.1, it was claimed that, because the internal group of the polygon decomposes into factor subgroups, where the decomposition sequence describes the successive action of the factors, the internal relationships of a side are mapped onto the internal relationships of the adjacent sides.

In an attempt to show simultaneously that the D_n generators act between adjacent sides and that there is the within-side invariance of the type just described, the following procedure was used (Experiment 7 in Leyton, 1984a). The subjects were given a square, a pentagon, and a hexagon, each side of which was marked with seven equally spaced points Fig. 26. Each subject was assigned a pair of adjacent sides. The task of the subject was to select an arbitrary point from one of the designated sides, and then choose the point in the other side which perceptually corresponded with the impression of first point chosen. The subject was asked then to return to the first side and repeat the procedure, until all points were used up—that is move backwards and forwards between the two sides choosing arbitrary points from the first, and exactly corresponding ones from the second. Essentially therefore the subject was constructing a mapping from the first side to the second. Note that there are an enormous number of ways (5040) in which the seven points on the first side can be mapped (one-to-one) onto the seven points on the second.[7]

The results showed that, with considerable statistical significance, subjects chose one of the generators of D_n, to relate the sides, and that, in so doing, they preserved the within-side relationships. [176 results out of a total of 180 scores accorded with this.] Figures 27a and b show an example where reflection and rotation, respectively, were used. Subjects often said that they had difficulty over which of the two operations to choose. A distributional analysis of the choices reveals interesting biases at different vertices of the figures. Note that these results seem to violate coding theory. The latter, being essentially concerned with *coding* (i.e., ordering stimulus sets into strings), does not allow the relationship of the description to the stimulus set to be a many-one structure.

[7] There were 60 subjects, each given a side pair from a square, a pentagon, and a hexagon. A total of 20 subjects were tested on each pair (10 in each pair direction).

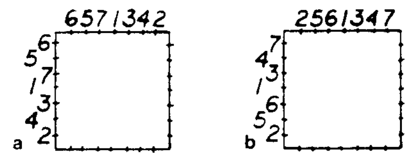

FIG. 27. A reflection and rotation mapping constructed by one subject.

(3) *The horizontal sides are referents.* In Leyton (1984a, Experiment 8), each subject was presented with a square, a pentagon, and a hexagon, and was told to regard each polygon as a collection of sides at different angles. The subject was then asked which side, in the presented figure, would s/he "normally judge the orientations of the other sides with respect to." Another set of subjects were presented with the same figures and asked which side was "most in a state of equilibrium or balance."[8]

In both conditions, the subjects gave one of the horizontal sides as the reply [$n = 12$ per subject group; expected mean = 6 choosing horizontal in a group; worst score = 11, yielding $t(11) = 5.0205$, $p < .0005$, one-tailed]. For example, the replies for the hexagon were split almost evenly between the top and the bottom side. This allows one to regard these sides as identity elements in two different input-space descriptions, and reinforces the claim of Leyton (1984a) that what is assigned to a stimulus set is a class of descriptions. Observe that the fact that both conditions (the reference and stability judgements) yielded the same replies seems to provide yet another corroboration of the claim made in the Description Postulate that the algebraic ordering corresponds to the stability ordering. This can further be regarded as validating Proposal 9, that reference is the factorization of instability. Note that the result seems to violate coding theory, which does not consider that there are privileged referent sides. Thus the experiment rules out most of the codes predicted by that theory, because those codes can start from sides that are not chosen as referents.

16.3. *Theory of Grouping: Structural Theory*

A fundamental problem that has faced perceptual psychologists is that of finding out what a *grouping*, or *unit*, or *part* is. Why does a figure partition more naturally in some ways and not in others? No substantive definition of a unit has to my knowledge ever been given. In this section an attempt is made to develop a rigorous definition of a grouping (unit or part). In fact, two very different definitions will be offered and an argument that they are equivalent will be presented.

We ended Section 16.2 by saying that, in a square, D_4 images the internal structure of one side onto the internal structure of another. This apparently trivial

[8] The term *equilibrium* was used to pool more of the members of the D_4-description subclass. Pilot studies had indicated that, if the phrase "at rest" were used, the bottom side would to be chosen more frequently.

phenomenon involves the notion of *control*—that is, mapping one machine (or dynamical system) onto another. However, the phenomenon also involves the notion of *grouping*, or *part*, because each side is a part. What I will attempt to show, in this section, is that the notion of *control* and that of *grouping* are crucially related.

Before we can begin, however, two concepts require clear recollection.

(1) It is necessary first for the reader unfamiliar with group theory to recall carefully the notion of *coset*. Re-reading the short Section 4.2.1 is crucial if the rest of the section is to be understood. We saw, for example, that D_6 is divided into two halves: (i) the subgroup of six rotations Z_6, and (ii) the set, tZ_6, of reflections, each of which is a rotation taken from Z_6, multiplied by the vertical reflection t. The set tZ_6 is called a *coset* of Z_6 and t is called a *coset leader*. Generally, given a subgroup H of a group, the cosets are all of the form g_1H, g_2H, g_3H,..., where g_i are the coset leaders. The set of cosets of a subgroup H entirely partition the overall group, G; just as Z_6 and tZ_6 partition D_6.

(2) The second area which needs careful recollection is that concerning the Principle of Nested Control. For the main claim to be made now is that *the Principle of Nested Control leads to a theory of grouping.* Again, if the reader does not have a clear recollection of the concepts introduced in the initial discussion of this principle, re-reading the first half of Section 10, at least until the term *stability* is introduced, is crucial if the reader is to understand the argument to be presented. To remind the reader here *only* of the statement of the Principle, recall that, given a group decomposition sequence, $G = G_1 \cdot G_2 \cdots G_n$, the *i*-subsequence of the component G_j is its right-hand sequence $G_{j+1} \cdot G_{j+2} \cdots G_n$. The Principle of Nested Control states that the cognition of an algebraic stability ordering, on this sequence, forces any component G_j to be cognized as moving its *i*-subsequence around, or equivalently as moving dynamics of the form $G_{j+1} \cdot G_{j+2} \cdots G_n$, onto other dynamics of that form.

It is now possible to derive the main proposal of the theory of grouping to be offered in this paper. The derivation (which first appeared in Leyton, 1984b) proceeds as follows. First, it would be reasonable to define a grouping in this way: *A grouping is a set of elements which occur in a common perceptual structuration.* Recall, however, that the Description Postulate claims that *any structuration equals a machine state space.* The two statements just given lead to the following proposal:

G1: A grouping is a set that is cognitively described as a state space.

However, the Description Postulate also proposes that *a state space is cognized as having an input-space description* and that *an input-space description is cognized as having an algebraic stability ordering.* This point is crucial, because it allows us to use the Principle of Nested Control which states that *every input element g, in an algebraic stability ordering, is cognized as acting to the right of itself,* i.e., moving a machine state space labeled by its *i*-subsequence, $G_{j+1} \cdot G_{j+2} \cdots G_n$, onto any other state space which also receives that labeling (e.g., like the gravitational trajectories in Sect. 10). Therefore, one must conclude that *any cognized state space is of the form* $g \cdot G_{j+1} \cdot G_{j+2} \cdots G_n$; that is, an action of g on $G_{j+1} \cdot G_{j+2} \cdots G_n$. But the string $g \cdot G_{j+1} \cdot G_{j+2} \cdots G_n$ is a coset of $G_{j+1} \cdot G_{j+2} \cdots G_n$ (the coset leader being g). Thus we are lead to this crucial result:

G2: Any cognized state space is a coset of an i-subsequence.

Finally, by putting together the two above statements, G1 and G2, we obtain the central proposal of our theory of grouping:

PROPOSAL 18. *A* **grouping** *is a coset of an i-subsequence.*

Furthermore, as a corollary, we have:

COROLLARY 19. *A perceptual organization is partitioned (perceptually) into the set of cosets of each i-subsequence of its internal group decomposition.*

Let us illustrate these two statements. Consider first a square. As was argued in Section 16.1, its internal group decomposition appears to be $D_4 \cdot R$. Observe now that any reference sequence can be considered to be preceded by the identity element, the trivial dynamical system, i.e., a point $\{e\}$. Therefore, the internal group decomposition sequence of a square can be regarded as $D_4 \cdot R \cdot \{e\}$. By Proposal 18, the perceptual parts of the square are obtained, therefore, by elaborating the cosets of each of the *i*-subsequences of $D_4 \cdot R \cdot \{e\}$. We do this below.

Table 1 shows the *i*-subsequences and their cosets in $D_4 \cdot R \cdot \{e\}$. The first *i*-subsequence is $\{e\}$, because it is the right-most subgroup in the sequence. Its cosets are all of the form $g_i \cdot g_j \cdot \{e\}$, where g_i is an element in D_4, and g_j is an element in R. The coset leader is the element $g_i \cdot g_j$. However, by the discussion in Section 16.1, the multiple $g_i \cdot g_j$ is simply the coordinate pair defining an arbitrary point in the square (recall Figs. 23 and 24). Thus the set of cosets, $g_i \cdot g_j \cdot \{e\}$, is the set of points of a square. Therefore, the factorization of the internal group into those cosets corresponds to the partition of the square into its points. Note that the coset leader $g_i \cdot g_j$ is interpreted as a control variable; i.e., it maps points to points ("moves" them around). Thus the situation accords with the Principle of Nested Control.

TABLE 1

The Partition of a Square as $D_4 \cdot R \cdot \{e\}$

i-subsequences	Cosets
$\{e\}$	$g_i \cdot g_j \cdot \{e\}$
$R \cdot \{e\}$	$g_i \cdot R \cdot \{e\}$
$D_4 \cdot R \cdot \{e\}$	$D_4 \cdot R \cdot \{e\}$

TABLE 2

The Partition of a Square as $R \cdot D_4 \cdot \{e\}$

i-subsequences	Cosets
$\{e\}$	$g_i \cdot g_i \cdot \{e\}$
$D_4 \cdot \{e\}$	$g_i \cdot D_4 \cdot \{e\}$
$R \cdot D_4 \cdot \{e\}$	$R \cdot D_4 \cdot \{e\}$

Now consider the next i-subsequence of $D_4 \cdot R \cdot \{e\}$; that is, the sequence $R \cdot \{e\}$. This is a side. Its cosets are $g_i \cdot R \cdot \{e\}$ where g_i is taken from D_4. The process of going through the available coset leaders, g_i, enumerates the set of sides. Thus the partition of the internal group, into the set of cosets $g_i \cdot R \cdot \{e\}$, is the partition of the square into its sides. Observe that the situation conforms to the Principle of Nested Control: The coset leader is interpreted as a control variable mapping dynamical systems of the form $R \cdot \{e\}$ (*a* side) onto dynamical systems of that form.

This example appears at first to be trivial. However, one can in fact show that it crucially involves the Principle of Nested Control in a non-trivial way. To do this, let us reverse the order of the components R and D_4, and show that one does not obtain parts of a square.

Thus, suppose that the square has $R \cdot D_4 \cdot \{e\}$ as its internal organization group sequence (i.e., reversing the first and second factors). Table 2 shows the i-subsequences and their cosets in $R \cdot D_4 \cdot \{e\}$. Without loss of generality, let us take $\{e\}$ to be the top central point. Thus, because D_4 describes the polygon operations, the i-subsequence $D_4 \cdot \{e\}$ would be the four central points of the sides, as shown in Fig. 28a—a cross-like subset. Any coset would be of the form $g_i \cdot D_4 \cdot \{e\}$, a translation of each of these points some distance g_i to the left or right of the center of each side—a skewed cross-like subset. An example is shown in Fig. 28b.

It is true that one can see that these points are visually equivalent. This is due to the algebraic properties of the description map. However, no research, to my knowledge, has ever claimed that subjects regard such a set as a "part" of a square. Conventionally, the parts are regarded as the points and the sides—that is, the sets elaborated earlier in the non-reversed group decomposition. In conclusion, therefore, one finds that reversing the order leads to an empirically incorrect partition.

Observe that this situation is essentially due to the Principle of Nested Control. In the above reversed sequence, we would have to see R as a control variable map-

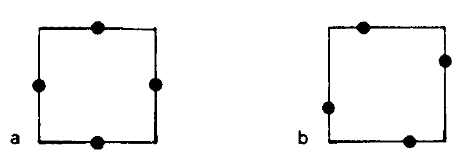

FIG. 28. (a) D_4 applied to the central point of a side. (b) A coset of that set in $R \cdot D_4 \cdot \{e\}$.

ping cross-like subsets, $D_4 \cdot \{e\}$ onto skewed versions of such sets. Instead, we find it much easier to see D_4 as the control variable, mapping the sides $R \cdot \{e\}$ to each other.

The example just described shows the following: The control structure is a structure of observer-assigned *mappings*, which send subsets of stimuli onto other subsets of stimuli. An orbit (i.e., state space) of such mappings creates a grouping. For example, in $D_4 \cdot R \cdot \{e\}$, an orbit of $\{e\}$ creates a point, an orbit of $R \cdot \{e\}$ creates a side, and an orbit of $D_4 \cdot R \cdot \{e\}$ creates a square. As we have seen, the only cognized orbits are *i*-subsequences of the nested structure of mappings.

Another way of showing that the perception of a part depends crucially on the arrangement of the control groups, or equivalently on the *i*-subsequences, is by causing the omission of one of the subgroups in the factorization. To illustrate this, let us consider a simple effect due to Wertheimer (see Rock, 1975, p. 273). In the sequence of dots in Fig. 29, the separation between pairs *ab*, *cd*, *ef*, and *gh* is only slightly less than between *bc*, *de*, *fg*, and *hi*. Depending on whether the observer sees this difference in spacing, a different percept arises. The dots are perceived either as a homogeneous line, or as a line partitioned into pairs. According to the theory being offered, the phenomenological difference is given by differences in the respective internal group descriptions, and subgroup decompositions.

Let us look at the non-homogeneous case first. The appropriate control structure seems to be

$$Z \cdot Z_2 \cdot \{e\}$$

where the levels of control are

$\{e\}$, a single point,

Z_2 (the cyclic group of the light switch) which controls the position of a point and thereby creates a pair, by reflectionally mapping one point onto its associate,

Z, the group of whole numbers, which controls the position of each pair, by translating the pairs along the line onto each other. (Z was discussed in Sect. 4.2.3, Example 1.)

The theory claims that the parts are determined by this system of control, i.e., *mappings*; in fact, that the parts are given by the *i*-subsequences in the mapping system. The first control or mapping *i*-subsequence is the point $\{e\}$. Its cosets are all of the form $g_i \cdot g_j \cdot \{e\}$. The coset leaders $g_i \cdot g_j$ are the maps which act on the points $\{e\}$. The algebraic partition of the internal group, into these cosets (and hence mappings), perceptually partitions the stimulus set into the set of individual points. The

FIG. 29. A grouping example of Wertheimer (cited by Rock, 1975): Pairings *ab*, *cd*, *ef*, and *gh* might be grouped by subjects.

next i-subsequence is $Z_2 \cdot \{e\}$, which is a pair of points. Its cosets are all of the form, $g_i \cdot Z_2 \cdot \{e\}$. The coset leaders g_i are the maps that act on the pairs $Z_2 \cdot \{e\}$, moving them about. The partition of the internal group, into these cosets (and hence mappings), partitions the percept into pairs of points. The final i-subsequence $Z \cdot Z_2 \cdot \{e\}$ is the whole percept, i.e., the unity. All its cosets are itself. The reader can easily check that any other selection of elements from the control or mapping hierarchy, other than the cosets of i-subsequences, does not produce a part. This is because any other selection breaks out of the control structure; i.e., does not accord with the way points are perceptually *mapped* onto each other.

To clarify this further, contrast the above with the homogeneous case. The appropriate control structure seems to be

$$Z \cdot \{e\}$$

where Z controls the position of each point $\{e\}$, mapping it onto each other point. The first i-subsequence is $\{e\}$. Its cosets are all of the form $g_i \cdot \{e\}$. The coset leaders g_i are the maps which act on the point $\{e\}$. Factorization of the full group sequence, into these cosets, yields a partition of the percept into points. However, the next i-subsequence is $Z \cdot \{e\}$, the whole. Thus there is no intermediate control group, *because each point is mapped onto any other point by Z.* Observe that this was not true in the previous case; for example, a right-hand point in a pair was not mapped by Z onto a left-hand point in another pair. Therefore in the present case, there is no intermediate mapping or control level occurring between the control group $\{e\}$ and the control group Z. Thus, according to the theory, there is no set of parts occurring between the lowest level, the points, and the highest level, the whole. This appears to explain the homogeneity of the percept. Table 3 contrasts the cosets elaborated for the non-homogeneous case with those of the homogeneous case. Note that it is exactly such an analysis that explains Gestalt examples such as that shown in Fig. 20e.

Having considered a number of examples which seem to corroborate this theory

TABLE 3

The Partition of Fig. 29 under Two Interpretations

i-subsequences	Cosets
Non-homogeneous case	
$\{e\}$	$g_i \cdot g_j \cdot \{e\}$
$Z_2 \cdot \{e\}$	$g_i \cdot Z_j \cdot \{e\}$
$Z \cdot Z_2 \cdot \{e\}$	$Z \cdot Z_j \cdot \{e\}$
Homogeneous case	
$\{e\}$	$g_i \cdot \{e\}$
$Z \cdot \{e\}$	$Z \cdot \{e\}$

of partitions, I now propose another definition of a part and attempt to show that it is equivalent to that which has been discussed so far.

The whole percept can be viewed as an object. One might therefore be interested in having a useful definition of its *subobjects*, i.e., stimulus subsets that are also perceived as objects. This, I suggest, is provided by the following:

DEFINITION 20. Given an internal group decomposition $G = G_1 \cdot G_2 \cdots G_n$, a **subobject** is a machine object of a group factor G_i (or an image of that machine object).

What is the relationship between a subobject and a part as viewed above? The two definitions are in fact very different. The derivation of Proposal 18 claims that a part, or a grouping, is a *state space*; Definition 20 claims that it is an *object*. However, under the theory being elaborated here, the two turn out to be equivalent as follows: Recall that the Principle of Nested Control says that given an i-subsequence $G_{j+1} \cdot G_{j+2} \cdots G_n$, the group G_j acts as a control group on that sequence, moving it about. Observe that the sequence itself is a state space. In fact, if we let the point $\{e\}$ be considered as the most right-hand factor of the sequence, $G_{j+1} \cdot G_{j+2} \cdots G_n$, then the sequence is the set of states of $\{e\}$. These are obtained by applying all the input groups along the sequence, to $\{e\}$. However, with respect to the *next* factor along, G_j, the sequence $G_{j+1} \cdot G_{j+2} \cdots G_n$ becomes an object, in its own right, which is moved about under the action of G_j. Thus an i-subsequence has two interpretations: (1) a state space on level $j+1$, and (2) a machine object on level j.

The examples we have been considering appear to corroborate this view. However, as a further example, consider a grid of squares, shown on the left in Fig. 30. The group decomposition of this example is

$$Z^2 \cdot D_4 \cdot R \cdot \{e\}.$$

The orbit of the first group factor $\{e\}$ is a single point on the far right of Fig. 30. However, on the next level, this orbit becomes an object acted on by the translation group R. As a object, the state space of the point is a side. However, on the next

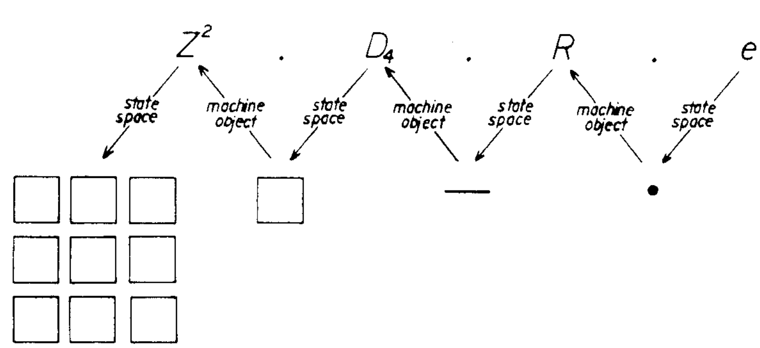

FIG. 30. The machine object on one level becoming a state on the next in $Z^2 \cdot D_4 \cdot R \cdot \{e\}$.

level, the state space itself (the side) becomes an object acted on by the group D_4. As an object, its state space is a square. However, on the next level, the state space itself (the square) becomes an object acted on by the group, Z^2, which generates the grid. These relationships are shown in Fig. 30.

Thus the Principle of Nested Control implies that Definition 20 and Corollary 19 define the same phenomenon. As a group orbit, an i-subsequence is a state space and is thus a *part* in the sense of Corollary 19. However, in becoming a machine object at the next level, it is a part in the sense of being a *subobject* (Definition 20). Thus a set of stimuli is a part in the sense of Corollary 19, at level $j+1$, if it is a part in the sense of Definition 20, at level j.

One can see, therefore, that the parts of a figure are arranged in a *hierarchy*. This set-theoretic hierarchy, on the stimulus set, appears to be strictly determined by the group-theoretic decomposition sequence. All the examples seem to corroborate this view (e.g., in figures, points are the parts of sides and sides are the parts of the figure itself; Table 3 shows the hierarchies responsible for the Wertheimer example discussed earlier, etc.). Thus it seems that parts are structured hierarchically because the successive i-subsequences of the internal group form a *nested structure of control.*

16.4. *The Internal Cartesian Reference Frame*

In Section 15.2, a theory of the extrinsic structure of the cartesian reference frame was offered; that is, the cognized structure of allowable action *on* that frame. In order to understand perceptual organization, I claim that one requires also a theory of the *internal* structure of that reference frame.

Two properties of the internal structure are evident:

(1) The left half of the frame can be generated from the right half by the reflection group. This group consists of the identity element, e, and the reflection operation, t_V; that is, reflection about the vertical. Since the continual application of the latter creates a continual cycle of size 2, backwards and forwards between the two halves of the plane, the group has the same structure as the group Z_2 of the light switch. This group can thus be written as $Z_2 = \{e, t_V\}$.

The same considerations apply to the relationship between the top and the bottom half of the plane. This can be given by the group $Z_2 = \{e, t_H\}$, where t_H is reflection about the horizontal line.

We can take all possible multiples of the operations mentioned above by simply multiplying the two groups. Thus we have the following product group which has four elements:

$$Z_2 \times Z_2 = \{e, t_V, t_H, t_V \cdot t_H\}. \qquad (7.4.1)$$

Note that the expression $Z_2 \times Z_2$ denotes the product of two copies of the reflection group Z_2.

(2) Alternatively, the frame can be generated from a single point in the plane by

applying the two-dimensional group of translations, R^2. This group was discussed in Section 4.2.3, Examples 2ii and 2iii.

In fact, the actions of the two groups, $Z_2 \times Z_2$ and R^2, within the frame, can be combined using a type of product called a *semi-direct product*, as will now be explained. The product represents the fact that the double reflection group $Z_2 \times Z_2$ makes mirror image copies of the translation group R^2. That is, one copy of R^2 has translations going in the positive x and positive y directions. But this copy can be reflected in the vertical axes by t_V. Thus another copy of R^2 has translations going in the negative x direction and positive y direction. In fact, for each of the members, e, t_V, t_H, $t_V \cdot t_H$, of $Z_2 \times Z_2$, there is a differently reflected copy of R^2. This mapping action of $Z_2 \times Z_2$ on R^2 is captured by saying that the product of $Z_2 \times Z_2$ and R^2 is a semi-direct product

$$(Z_2 \times Z_2)_{(\times_{sd})} R^2$$

where $_{(\times_{sd})}$ denotes the product operation. Note that this string of symbols consists quite simply of three successive components: the double reflection group $Z_2 \times Z_2$, followed by the product operation, $_{(\times_{sd})}$, followed by the translation group, R^2. Thus, to summarize, the above has argued that the total group acting within the frame is the sequence, $(Z_2 \times Z_2)_{(\times_{sd})} R^2$.

Because the group $(Z_2 \times Z_2)_{(\times_{sd})} R^2$ can be considered to *internally generate* the cartesian frame, one can take the frame to be the machine which has this group as its input space and the point as its object. That is:

DEFINITION 21. The **Internal Cartesian Reference Frame** is the group sequence

$$(Z_2 \times Z_2)_{(\times_{sd})} R^2.$$

Given a stimulus set X the **assignment** of an internal cartesian reference frame is an input-space description

$$C_I: (Z_2 \times Z_2)_{(\times_{sd})} R^2 \to X.$$

17. Interaction Theory

17.1. *The Interaction Laws*

In Section 14, I proposed that perceptual organization is the interaction between the internal and external input groups of a figure. The internal group generates the figure from a subset; and the external group alters the whole. A theory of the structure of the external group was elaborated in Sections 8, 9, and 15; and a theory of the structure of the internal group was offered in Section 16. Using these two theories, let us now attempt to derive a solution to the problem of the way in which the two respective structures *interact*. As stated in Proposal 14, I will argue that this interaction fundamentally defines the *perception* of the stimulus set.

Recall first that an eigenspace is a straight subspace (e.g., a line) which remains over itself in a transformation. For example, the bold lines in Fig. 21 showed the eigenspaces of rotation (0 eigenspace), shear (1 eigenspace line), and pure deformation (2 eigenspace lines).

Observe now that the pure deformation group is the most *stable* group of the sequence $A \cdot N \cdot SO_2$ applied to a square. So the two bold lines in Fig. 21c are the *eigenspace lines of the most stable external group factor*. However, observe another property of these two lines. They are each *symmetry axes* of the square.

Thus, I propose that the interaction between the internal (generating) and external (altering) group sequences is determined by the following law:

> FIRST INTERACTION LAW: *The symmetry axes of the internal group are eigenspaces of the most stable external groups.*

The law will be extensively corroborated over the next few sections even in highly complicated shapes.

The law equates symmetry axes and eigenspaces. The two phenomena have one important property in common: they are both lines which are *invariant* under transformations. However, besides this, they have entirely different properties and roles. In order to understand how non-trivial the act is of forcing an equation between these two concepts, we consider them first separately.

Symmetry axes. A symmetry axis is a line in which every point is sent to itself under the associated transformation (e.g., a rotation or reflection). This is illustrated in Fig. 31a, with a triangle. While the entire figure has to be reflected to describe its symmetry, point P remains unchanged. Observe that the reflection is a member of the *internal group* of the figure because it is capable of generating the whole figure from only half of it.

Eigenspaces. An eigenspace is a line of points which usually *move* along that line under a linear transformation. Consider Fig. 31b. A pure deformation acts on the triangle, stretching it upwards. The central axis is an eigenspace. Points such as P *move* along that line. The pure deformation is a member of the external group of the triangle; that is, it describes what can happen to the triangle.

To summarize: A symmetry axis comprises an array of *static* points; an

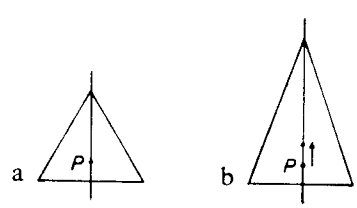

FIG. 31. (a) A reflection transformation maps P to itself. (b) A pure deformation moves P, despite leaving the entire line invariant.

eigenspace typically comprises an array of *moving* points. Conceptually, therefore, one should consider the First Interaction Law as stating:

> **Invariant lines that are rigid, under transformations belonging to the internal group, become lines of flexibility, under transformations in the external group.**

The First Interaction Law is now combined with another phenomenon to yield a theoretical result which will prove to be important to us. Let us return to the square. We saw that the symmetry axes which bisected the sides become the eigenspaces of the pure deformation group, as in Fig. 21c. However, these axes are not the only symmetry axes that a square has. The diagonals are also symmetry axes. Nevertheless, as the experiments in Section 17.7 corroborate, these symmetry axes are less salient than those we looked at before. This means that the most *salient* symmetry axes are the eigenspaces of the most stable external groups.

Quite simply one can argue the following: Some symmetry axes are more perceptually salient than others. It seems reasonable therefore to assume that the more perceptible a symmetry axis is, the more likely it is to become an eigenspace, under the First Interaction Law. But this means that the figure is perceived as more *flexible* in that direction; i.e., that the associated transformation is perceived as more allowable. This result is stated as follows:

> SECOND INTERACTION LAW: *The allowability of transformations in the external group is proportional to the salience of those symmetry axes which are associated with the transformations via the First Interaction Law.*

Observe that the Interaction Laws illustrate a way in which the view offered in this series of papers is different from the current coding-theoretic and feature-analytic approaches. An eigenspace, or direction of possible flexibility, is not a feature that can be extracted from a single figure. Thus, because a coding system is essentially a means of converting the retinal array into a string of letters, the code cannot contain its own flexibility.

Simple as the Interaction Laws may appear, they have, I believe, a profound cognitive basis, which we will now try to uncover. What I will attempt to do, in the next paragraph, is to supply a proof of the Interaction Laws which reveal this cognitive basis. If the reader finds the next paragraph too difficult, it can be omitted.

Derivation of the Interaction Laws. It has been proposed, in this paper, that the internal structure of a figure is a machine. Thus, it would seem reasonable to suppose that if one applied an external transformation to a stimulus set, such that the imposed machine structure remained as unaltered as possible, then the perceptual organization would have been changed as little as possible. Thus, there appears to be an important relationship between the internal group and the *stable* external groups. A stable external group would be one that preserves the perceptual organization, and would thereby preserve the internal machine structure. To understand the action of stable external groups, one should therefore consider how the internal machine structure can be preserved under alteration of the underlying stimulus set. Recall (Sect. 4) that the computational connectivity of a machine is captured in the associated group. Essentially, the group describes the symmetry structure of the machine (Sect. 4.2.1), and this structure is a powerful way of representing the computational connectivity. Supposing, therefore that, in accord with the Description Postulate, a machine structure has been mapped onto a stimulus set. It is clear that when a transformation is applied to that stimulus set such that the symmetry structure has been preserved, then the machine structure will have been preserved. Therefore, let a figure be symmetrical about a particular axis, P; that is, there is an input, g, from the stimuli on one side of P to the stimuli on the other side of P. If an external transformation is now applied such that the axis remains in the final figure, then the same transformation g would still map the stimuli on one side of P to the other. There are three types of axis-preserving transformations: stretches along the axis (i.e., in which the axis is an eigenspace), rotations of the axis, and translations of the axis. The first alternative is the only case where the axis remains coincident with itself. Hence one can consider it to be the most stable. Therefore, an eigenspace of a stable external transformation is a symmetry axis of the internal group. Thus, one obtains the Interaction Laws.

17.2. *Preliminary Corroboration*

Although more complicated corroborations of the Interaction Laws will be presented later in the sections on motion phenomena and on complex shapes, the present section describes two very simple experiments which appear to corroborate the Interaction Laws. In each experiment, a different means was employed of creating axis salience within a figure: (a) by aligning an axis with a gravitational one, and (b) by splitting the figure along an axis. The subject had to decide which of the available axes in the figure most allowed an elongation. Choosing the salient one would confirm the Interaction Laws. All the test figures used were regular polygons because, with these figures, one can exploit the fact that each presents a "pool" of geometrically equivalent axes from which the subject can make a selection, biased only by the experimentally chosen salience variable.

(1) *Salience by gravitational alignment.* In Leyton (1984a, Experiment 9), the gravitational axes of equal-sided polygons were made salient by presenting nine such test polygons, one below the other, in a vertical line. All possible equal-sided polygons were used, from triangles all the way up to octagons. Because increasing the number of sides still further might have destroyed perceptual clarity, only circles were added. Diamonds were included as well as squares, and crystals as well as octagons. I call octagons placed on their tips, crystals. Their relation to octagons is the same as that of diamonds to squares, and they are perceptually distinct from octagons.

For each test polygon, the subject was presented with a set of comparison figures

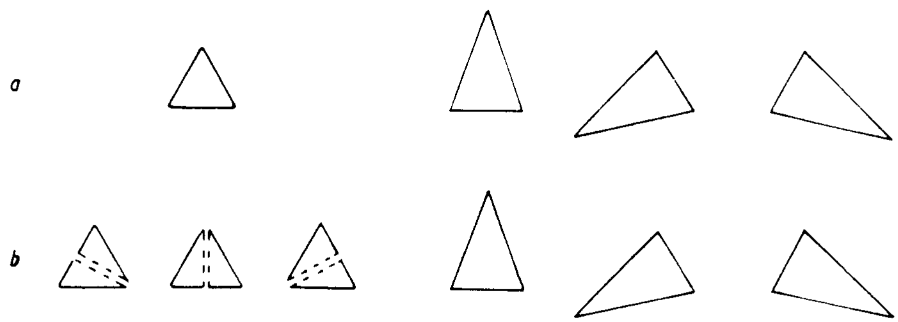

FIG. 32. (a) One of the 15 examples used for an experiment concerning the Interaction Laws, in the case of gravitational salience: Subjects had to choose the elongation which best matched the regular polygon on the left. (b) In the case of salience by division: The divided examples had to be matched with the elongated ones.

where each figure was an elongation of the associated test polygon along one of its axes—see Fig. 32a. Every subject was given 15 sets of comparison figures as follows: (1) triangles; (2 and 3) two sets of gravitationally orientated squares; (4 and 5) two sets of gravitationally orientated diamonds; (6) pentagons; (7 and 8) two sets of hexagons; (9) septagons; (10 and 11) two sets of octagons; (12 and 13) two sets of crystals; (14 and 15) two sets of circles. The subject had to select, from each of these sets, that figure which best matched the impression of the corresponding test polygon. The result was that subjects, with statistical significance, chose the elongation which was gravitationally aligned [$n = 6$, 15 comparisons per subject; expected mean = 7.5 sets in which gravitational alignment is chosen; actual mean = 12.5, $t(5) = 9.444$, $p < .0005$, one tailed].

(2) *Salience by division.* In Leyton (1984a, Experiment 10), subjects were presented with the same sets of elongations, as in the above experiment. However, these had to be compared with the same number of equal-sided versions where each of the latter figures was split in two along an axis—see Fig. 32b. For each elongation, subjects had to select the divided figure which best matched the perceptual impression made by that elongation. With considerable statistical significance, subjects chose the division which had been made along the axis of elongation; thus supporting the Interaction Laws [$n = 6$; 15 comparisons per subject; expected mean = 7.5 sets in which every matching corroborates Interaction Laws; worst score = 12 yielding $t(5) = 14.118$, $p < .0005$, one tailed].

17.3. *The Full Organization Group*

Recall (Sect. 14.2) that the full group of a perceptual organization was defined to be a product of G_E, the organization's external (*altering*) group, and G_I, the internal (*generating*) group. In Part I, I argued that G_E has a decomposition sequence; and in Section 16, I argued that the internal group is structured also by a decomposition sequence. The consequence is this:

PROPOSAL 22. *The full group of any perceptual organization has a decomposition sequence*

$$G = G_{E_1} \cdot G_{E_2} \cdots G_{E_n} \cdot G_{I_1} \cdot G_{I_2} \cdots G_{I_m}$$

where the first n factors are the external transformation groups, and the last m factors give the decomposition sequence of the internal group—i.e., the partition hierarchy of the organization. Both sequences are an algebraic stability ordering.

Thus the claim here is that perceptual organization conforms to the Description Postulate upon which our cognitive theory is based.

17.4. *Perceptual Organization and the Full Cartesian Reference Frame*

Let us now consider how the view being developed here might provide a theory of the relationship between the cartesian reference frame and perceptual organization. In Section 15.2, the External Cartesian Reference Frame was defined to be $SL_2R = A \cdot N \cdot SO_2$, that is SL_2R with its stability ordering (A is the pure deformation group, N is the shear group, and SO_2 is the rotation group). In the previous section, the Internal Cartesian Reference Frame was defined to be $(Z_2 \times Z_2)_{(\times_{sd})} R^2$; that is, the double reflection group $Z_2 \times Z_2$ times the group R^2 of two-dimensional translations. I now define the Full Cartesian Reference frame as the product of the two group sequences, $A \cdot N \cdot SO_2$ and $(Z_2 \times Z_2)_{(\times_{sd})} R^2$. This yields a rather lengthy string of symbols:

$$(A \cdot N \cdot SO_2)\hat{\imath}((Z_2 \times Z_2)_{(\times_{sd})} R^2).$$

However, the reader should note that the sequence is not difficult to understand. First, note that it is split into two halves, $A \cdot N \cdot SO_2$ and $(Z_2 \times Z_2)_{(\times_{sd})} R^2$. Second, let us go through the symbols, one by one, and see that they are easily explicable. The symbols are:

A, the pure deformation group;

N, the shear group;

SO_2, the rotation group;

$\hat{\imath}$, the product operation equating symmetry axes and eigenspaces;

$Z_2 \times Z_2$, the group generated by the horizontal and vertical reflections;

$_{(\times_{sd})}$, the semidirect product operation; and

R^2, the two-dimensional translation group.

Thus we have:

DEFINITION 23. The **Full Cartesian Reference Frame** is the group sequence

$$(A \cdot N \cdot SO_2) \hat{\imath} ((Z_2 \times Z_2)_{(} \times_{sd)} R^2).$$

The **assignment** of the Full Cartesian Reference Frame to a stimulus set X is an input-space description

$$C_E \hat{\imath} C_I \colon (A \cdot N \cdot SO_2) \hat{\imath} ((Z_2 \times Z_2)_{(} \times_{sd)} R^2) \to X.$$

The machine-theoretic view, in fact, provides a theory of what it means to *impose* a cartesian reference frame on a figure. Conventionally, one regards the figure as a structure and the cartesian frame as a cross that is placed over the figure. However, the view that I offer here is crucially different. It is as follows: The Full Cartesian Reference Frame is a product of input groups, $(A \cdot N \cdot SO_2) \hat{\imath} ((Z_2 \times Z_2)_{(} \times_{sd)} R^2)$. However, recall Proposal 22. It says that the full group of a figure's perceptual organization is also a group sequence $G_{E_1} \cdot G_{E_2} \cdots G_{E_n} \cdot G_{I_1} \cdot G_{I_2} \cdots G_{I_m}$. Thus we have two group sequences, one for the frame and one for the perceptual organization. I now claim that people's impression that a cartesian reference frame has been imposed is due to the occurrence of the former sequence as part of the latter sequence. That is:

PROPOSAL 24. *The basis for the impression that a cartesian reference frame has been imposed on a perceptual organization, is this: The sequence of operations defining the frame,*

$$(A \cdot N \cdot SO_2) \hat{\imath} ((Z_2 \times Z_2)_{(} \times_{sd)} R^2)$$

occurs along the group decomposition sequence of the organization,

$$G_{E_1} \cdot G_{E_2} \cdots G_{E_n} \cdot G_{I_1} \cdot G_{I_2} \cdots G_{I_m}.$$

In order to gain an understanding of what this means perceptually, let us consider first the *internal* group sequences of the two sequences. These are, respectively,

$$(Z_2 \times Z_2)_{(} \times_{sd)} R^2$$

of the frame, and

$$G_{I_1} \cdot G_{I_2} \cdots G_{I_m}$$

of the organization. The proposal is that the first sequence of operations must occur along the second sequence. However, by Proposal 16, the latter sequence is the generative sequence of the organization. Therefore, a claim being made in Proposal 24 is this: *The imposition of the cartesian frame is the stipulation that the sequence of operations which generate the figure must include the operations* $(Z_2 \times Z_2)_{(} \times_{sd)} R^2$; *that is, the double reflection group times the two-dimensional translation group.*

Consider now the two external sequences:

$$A \cdot N \cdot SO_2$$

of the cartesian frame, and

$$G = G_{E_1} \cdot G_{E_2} \cdots G_{E_n}$$

of the organization. Proposal 24 claims that the first sequence of operations must occur along the second sequence. However, by Proposal 16, the latter sequence is the flexibility structure of the stimulus set. Therefore, a claim being made in Proposal 24 is this: *the imposition of the cartesian frame is the introduction of the flexibilities $A \cdot N \cdot SO_2$ as a subset of the total flexibility structure of the stimulus set.*

Thus, as an example, consider the D_6 component in the internal sequence of a hexagon. $Z_2 \times Z_2$, part of the internal sequence of the cartesian frame, turns out to be a subgroup of D_6, as predicted by the proposal. This can be seen by making the following correspondences between the elements in

$$Z_2 \times Z_2 = \{e, t_V, t_H, t_V \cdot t_H\}$$

and the elements in

$$D_6 = \{e, r_{60}, r_{120}, r_{180}, r_{240}, r_{300}, t, tr_{60}, tr_{120}, tr_{180}, tr_{240}, tr_{300}\}.$$

Going through the members of $Z_2 \times Z_2$, one by one, the correspondences are:

e in $Z_2 \times Z_2$ is the element e in D_6;

t_V in $Z_2 \times Z_2$ is the element t in D_6;

$t_V t_H$ in $Z_2 \times Z_2$ is the element r_{180} in D_6;

t_H in $Z_2 \times Z_2$ is the element tr_{180} in D_6.

This is shown explicitly in Fig. 33. Therefore $Z_2 \times Z_2$ is embedded in D_6 and is part of the input-space description of the hexagon, thus corroborating Proposal 24.

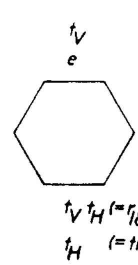

FIG. 33. $Z_2 \times Z_2$ as a subgroup of D_6.

17.5. *The Interaction Laws and the Decomposition of the Organization Group*

Let us now examine the relationship between the Interaction Laws and the organization-group decomposition. Consider, without loss of generality, that subsequence of the organization group which is the Cartesian Frame:

$$(A \cdot N \cdot SO_2)\hat{\imath}((Z^2 \times Z_2)_{(}\times_{sd)} R^2).$$

As we noted, this sequence has two halves. First, there is the $A \cdot N \cdot SO_2$ sequence (pure deformation × shear × rotation); and second, there is the $(Z_2 \times Z_2)_{(}\times_{sd)} R^2$ sequence (double reflection times two-dimensional translation). But now look at the entire sequence. At the point of juncture between the two sequences, there is the symbol $\hat{\imath}$. This means that we consider the Interaction Laws to define the action of the product operation, $\hat{\imath}$, between the internal and external sequence of the frame.

Let us examine this proposal more closely. The Interaction Laws imply that the symmetry axes of the internal sequence $(Z_2 \times Z_2)_{(}\times_{sd)} R^2$ must be eigenspaces of the stable groups of the $A \cdot N \cdot SO_2$ sequence. In order to understand this, let us trace the fate of the symmetry axes and eigenspaces, as they are transformed along the full sequence. First, observe that two symmetry axes are yielded by $Z_2 \times Z_2$. The Interaction Laws state that these must be eigenspaces of the most stable subgroups of $A \cdot N \cdot SO_2$. The most stable subgroup is A; the pure deformation group. It has two eigenspace lines, as we saw in Fig. 21a. Thus according to the First Interaction Law, the two symmetry axes of $Z_2 \times Z_2$ become the two eigenspace lines of the pure deformation group A. Subsequently, one of these eigenspaces is preserved by the shear group N, and the other is sheared (recall Fig. 21b). Finally, the rotation group SO_2 moves both of these off themselves (recall Fig. 21a). That is, the Full Cartesian Frame involves the following sequence of actions:

$$\begin{aligned} Z_2 \times Z_2(2 \text{ sym. axes}) &\to A(2 \text{ eigs}) \\ &\to N(1 \text{ eig}) \\ &\to SO_2(0 \text{ eigs}). \end{aligned}$$

17.6. *A Solution to the Orientation and Form Problem*

One can regard orientation and form research as among the most important in perceptual psychology, for it shows that perception is not a mere *re*-presentation of the external world, but a *description* (Goldmeier, 1937/1972; Rock, 1973). A fundamental principle of Gestalt psychology is that the same stimulus can have several descriptions, each of which constitute completely different percepts. The orientation and form phenomenon is a particular instantiation of that principle: Different orientations of the same stimulus lead to perceptually different objects (see Sect. 8.1 for a review of this phenomenon). However, while the literature has demonstrated that this phenomenon exists, no theory has emerged that *explains* why it occurs. Since its original discovery by Mach in 1897, the phenomenon has remained essentially a mystery. In fact, it will be argued here that, the orientation and form problem has not been fully described. Of course, it is usually the case in science that while a problem remains a mystery it is inadequately described. Only its solution provides a structure which opens up and clarifies the problem. Let us therefore bring together parts of the perceptual theory, which has been developed here, in an attempt to clarify, generalize, and solve the orientation and form problem. In what follows, the solution, which is offered, will be stated first, and then illustrated.

The Solution

The first step towards solving the problem uses the theory, developed in Sections 16.4 and 17.4, of the role of the internal structure of the cartesian frame: This internal structure, I have claimed, is not a simple coordinate system that the perceiver lays over the figure; it is part of the *generative* structure of the figure. It is literally part of the *form* of the figure. The internal frame is the group $(Z_2 \times Z_2)(\times_{sd})R^2$—the double reflection group times the two-dimensional translation group—which is a sequence of operations occurring along the sequence of groups $G_{I_1} \cdot G_{I_2} \cdots G_{I_m}$ that generate the figure. The latter sequence will also be called the figure's *internal geometry.*

The second crucial step is to use the theory, developed in Sections 15.2 and 17.4, of the external structure of the cartesian frame. This structure, I have claimed, is a sequence of groups $A \cdot N \cdot SO_2$ which occurs along the decomposition sequence $G_{E_1} \cdot G_{E_2} \cdots G_{E_n}$ of the external group of the figure, i.e., the structure is a subset of the flexibilities of the figure. The important property to recall now is that, in accord with the Principle of Nested Control, the external sequence *acts on the internal sequence.* However, as was observed in the previous paragraph, the internal sequence is the internal *geometry*, and this geometry contains the operations $(Z_2 \times Z_2)(\times_{sd})R^2$. Thus the Full Cartesian Reference Frame, $(A \cdot N \cdot SO_2)\hat{\imath}((Z_2 \times Z_2)(\times_{sd})R^2)$ should be understood as an action of $A \cdot N \cdot SO_2$ on the internal geometry, $(Z_2 \times Z_2)(\times_{sd})R^2$. Now, for each particular value of stretch × shear × rotation, that is each member $a \cdot n \cdot r$ of $A \cdot N \cdot SO_2$, one has a sequence, $(a \cdot n \cdot r)\hat{\imath}((Z_2 \times Z_2)(\times_{sd})R^2)$. This sequence is the alteration $a \cdot n \cdot r$(stretch × shear × rotation) followed by the internal geometry, $(Z_2 \times Z_2)(\times_{sd})R^2$, to which the alteration is applied. Thus each particular sequence, $(a \cdot n \cdot r)\hat{\imath}((Z_2 \times Z_2)(\times_{sd})R^2)$, defines a particular figure. Therefore, the Full Cartesian Reference Frame, $(A \cdot N \cdot SO_2)\hat{\imath}((Z_2 \times Z_2)(\times_{sd})R^2)$ defines a space of figures, each with its own geometry, $(a \cdot n \cdot r)\hat{\imath}((Z_2 \times Z_2)(\times_{sd})R^2)$.

The next crucial step is to use the Interaction Laws, in the manner described in the previous sections, to trace the fate of the $Z_2 \times Z_2$ axes under the actions of the successive elements $a \cdot n \cdot r$. The two axes of $Z_2 \times Z_2$ are converted into two eigenspaces for the element a; then n shears one of the eigenspaces and preserves the other; and finally r moves the eigenspaces off themselves. In this way, figures are successively *referenced* to each other.

Thus the resulting system of figures, $(A \cdot N \cdot SO_2)\hat{\imath}((Z_2 \times Z_2)(\times_{sd})R^2)$, comprises a reference frame of related figures. I claim that *what has been called the orientation* *and form phenomenon is the case where at least two such systems of figures, are constructed for a particular stimulus set.* Observe that if there are two such systems, then, via the Interaction Laws, each stimulus set receives two possible internal geometries.

An Illustration

One can illustrate this argument by considering Fig. 34a. The percept has phenomenological ambiguity. It can be viewed either as a sheared square (Fig. 34b) or a stretched diamond (Fig. 34c). To make this statement more precise let us assign SL_2R coordinates to the respective figures. Consider first Fig. 34b. Being a sheared square, the figure has no pure deformation. This means that it is at the identity element, e_A, of the pure deformation group. However, it has some non-trivial shear value, n, in N. Furthermore, because it is neither horizontal nor vertical it has a non-trivial rotation value, r_1, in SO_2. Therefore, the perceptual interpretation has coordinates

$$(e_A, n, r_1)$$

in $A \cdot N \cdot SO_2$. That is, these coordinates are the *external* ones of the interpretation.

Alternatively, consider the figure interpreted as a stretched diamond; that is, Fig. 34c. Because it is stretched, it has a non-trivial pure deformation. a. However, because it has no shear, it is at the identity element, e_N, of the shear group. Finally, because it has a different orientation between the vertical and horizontal, it has a different non-trivial rotation value, r_2. Thus, in this interpretation, the percept has coordinates

$$(a, e_N, r_2)$$

in $A \cdot N \cdot SO_2$.

From the above two sets of external coordinates we can determine the associated internal geometries. Recall that the Interaction Laws equate symmetry axes with the eigenspaces of the most stable external groups. Thus, let us consider first the coordinates (e_A, n, r_1) of the figure in Fig. 34b. Recall the sequence at the end of Section 17.5. It says that the eigenspace of the shear coordinate, n, must have been a symmetry axis. Therefore, although the figure presented in Fig. 34b is not symmetrical about the line shown, the figure from which it is derived (by shearing) must have been symmetrical about that line. Furthermore, by Section 17.5, the line denoting the angle of shear (parallel to the slanted sides) must also have been a symmetry axis. Therefore the figure, from which Fig. 34b was derived, must have had a $Z_2 \times Z_2$ structure with symmetry axes as *side bisectors.* By the next section, this means that it must have been a square.

Now consider the coordinates (a, e_N, r_2) of Fig. 34c. Because there is a non-trivial pure-deformation value, a, two eigenspace lines are yielded at this stage of the sequence. By the Interaction Laws, these eigenspaces must have been symmetry axes. However, because the shear value is non-existent (it is e_N), the symmetry axes cannot have been destroyed. Therefore, the original version must have had a

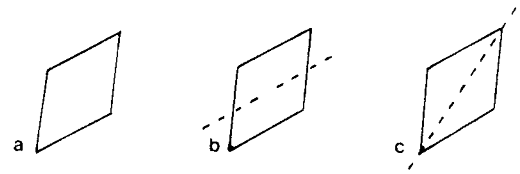

FIG. 34. (a) A perceptually ambiguous shape, seen either as (b) a sheared square, or as (c) a stretched diamond.

$Z_2 \times Z_2$ structure where the symmetry axes were the *angle bisectors* rather than side bisectors as in the previous case. By the next section, the figure must have been a diamond.

The above considerations imply that the perceptual organizations of the two percepts are given, respectively, by the two strings

$$(e_A \cdot n \cdot r_1)\hat{\imath}((Z_2 \times Z_2)_{(}\times_{sd)} R^2)$$

and

$$(a \cdot e_N \cdot r_2)\hat{\imath}((Z_2 \times Z_2)_{(}\times_{sd)} R^2).$$

Recall now that, given a subgroup H, a *coset* is a set of the form gH, where g is the coset leader (Sect. 4.2.1). Observe that each of these strings is a coset. In the first string just defined, the coset leader is $(e_A \cdot n \cdot r_1)$; and in the second string it is $(a \cdot e_N \cdot r_2)$. The subgroup in both cases is $(Z_2 \times Z_2)_{(}\times_{sd)} R^2$. Thus we have concluded that *a figure is a coset.* Furthermore, because the coset leaders are control variables with respect to their i-subsequences, $(Z_2 \times Z_2)_{(}\times_{sd)} R^2$, the cosets conform to the Principle of Nested Control; that is, the coset leaders map i-subsequences to i-subsequences. In the present case, this means that the coset leaders map figures to figures.

17.7. *A Solution to the Square–Diamond Problem*

The Square–Diamond problem was discovered by Mach (1897) and was the first recorded example of the orientation and form phenomenon. Mach had observed that if a square is placed with the pairs of opposite corners aligned along the vertical and horizontal axes, it appears phenomenologically unlike a square, i.e., like a "diamond." In the almost 90 years since Mach's discovery, the phenomenon has remained a mystery. Let us now use the theory developed in the previous section to propose a solution to the problem. (This solution was first published in Leyton, 1982). I argue (a) that the square and diamond are distinguished by their internal generative structure, (b) that they are distinguished by their external group assignments, and (c) that the latter distinction is that predicted from the former by the Interaction Laws. These results are supported in two experiments, as follows.

(1) *The difference between the internal group assignments.* In the previous sections, I argued that the internal cartesian frame is actually part of the generative

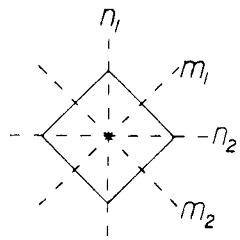

FIG. 35. The available symmetry axes of a regular tetragon.

structure of a shape. It is sufficient to show here that the $Z_2 \times Z_2$ (double reflection) component of the internal frame is assigned differently to the square, from the way it is to a diamond. The difference can be illustrated as follows. Both figures have the same underlying stimulus set which I will call a *tetragon*. Consider Fig. 35, a tetragon on which two pairs of symmetry axes have been drawn: (m_1, m_2) the side bisectors, and (n_1, n_2) the angle bisectors. The following experiment tries to show that $Z_2 \times Z_2$ is assigned to a square such that (m_1, m_2) are its symmetry axes, and $Z_2 \times Z_2$ is assigned to a diamond such that (n_1, n_2) are its symmetry axes.

In Leyton (1984a, Experiment 11), I presented subjects with 32 pairs of tetragons in the form shown in Fig. 36. Each member of a pair was in the same orientation. Each tetragon had a pair of dots as shown in Fig. 36. The dot pairs were created as follows. The two tetragons always had one dot in a common position. The other dot was in a reflectionally symmetric position on the figure. The axis of symmetry used in one figure was a side bisector, and the axis used in the other figure was an angle bisector. For half of the pairs the subjects were asked which member of the pair was more square-like. For the other half of the pairs, the subjects were asked which member of the pair was more diamond-like.

Although the subjects were shown the figures in the conventional square and diamond orientation, half the subjects were asked the "wrong" question: When presented with a pair of conventional squares (e.g., Fig. 36c) they were asked which member of the pair was more diamond-like. Similarly they were asked the "wrong" question for a pair of conventional diamonds.

The results showed that, with considerable statistical significance, the figures which were symmetric about the side bisectors were regarded as square-like, and the figures which were symmetric about the angle-bisectors were regarded as diamond-like [$n = 12$; 32 trials per subject; expected mean = 16 trials accord with experimental hypothesis; actual mean = 31.67; $t(11) = 47.00$, $p < .0005$, one tailed]. This was the case even for subjects given the *wrong question* condition.

FIG. 36. Comparison pairs, where, for each pair, subjects had to choose the member which was more square-like or more diamond-like.

Thus the hypothesis that the square and the diamond are distinguished by the assignment of their $Z_2 \times Z_2$ (double reflection) structures appears to be corroborated.

(2) *The difference between the external group assignments.* In the previous sections, it was argued that the cartesian reference frame is a group decomposition where the external sequence is $A \cdot N \cdot SO_2$. Furthermore, it was argued that this sequence of group factors must occur along the external group sequence of a figure if the latter yields the impression that a frame has "been imposed" on it. I now claim that this sequence is assigned differently for a square than for a diamond. It is sufficient to show a difference in the assignment of the first group factor, A (the pure deformation group).

My claim is that pure deformations are perceived as more *allowable* along the side bisectors in a square and along the angle bisectors in a diamond. Observe that this accords with the Interaction Laws: These directions are the salient axes of the respective internal $Z_2 \times Z_2$ structures. In particular, the Second Interaction Law is relevant here. It states that axis *salience*, in the internal group, is translated into *allowability* of those external transformations determined by the First Interaction Law (i.e., the external transformations which have eigenspaces along the symmetry axes).

In Leyton (1984a, Experiment 12), I presented subjects with pairs of figures, of the type shown in Fig. 37. Each figure was derived from a tetragon in the same orientation, but one figure was a pure deformation along a side bisector of the tetragon, and the other was a pure deformation along an angle bisector. The results showed that, with considerable statistical significance, the former type was regarded as more square-like and the latter as more diamond-like. (Scores had greater statistical significance than in the previous experiment.)

This supports the hypothesis that the square and the diamond are distinguished by the assignments of the external groups, i.e., by their flexibilities. Furthermore the external distinction made is that which is predicted by the Interaction Laws from the internal distinction provided by the preceding experiment.

Thus the theory presented in the previous section appears to provide a solution to the square–diamond problem.

FIG. 37. Comparison pairs, where, for each pair, subjects had to choose the member which was more square-like or more diamond-like.

18. Motion Perception

18.1. *Absolute Motion*

It is clear that the motion of an imaged object on the retina is a combination of the motion of the "actual" object depicted, and the motion of the observer in relation to the object (e.g., the movement of the head and eye). Furthermore, it is clear that the nervous system separates and identifies these two components as such. For example, one is able to turn one's head, while watching a horse race without seeing the horses go faster or slower depending on the direction in which one turns one's head.

What will now be argued is that this phenomenon can be explained by the Cognitive Stability Principle and the concept of an algebraic stability ordering (Sect. 9).

First note that motion can, of course, be represented by vectors (arrows)—a fact exploited in Johansson's careful and extensive analysis of the psychology of motion perception, to be discussed later (Johansson, 1950). Second, note that such vectors form, algebraically, a group under addition of vectors (In fact, generally all vector spaces are defined as groups; e.g., Hoffman & Kunze, 1961). Addition of vectors works by placing the vectors in sequence, end to end, at their respective angles.

There are two motion vectors of which the observer has direct knowledge: (a) The vector, v, which describes the motion of the retinal image; and (b) the motion vector, v_2, created by the observer's eye and head movements. By *subtracting* the latter from the former, the observer obtains a derived vector, $v_1 = v - v_2$. This vector, v_1, is then cognitively defined to be the actual motion vector of the perceived object. Note that, in accord with the phenomenological position underlying this series of papers, the object does not actually possess a velocity. The observer defines the velocity to be the remainder after the factorization process.

The subtraction described in the previous paragraph is a factorization that can be explained by the Cognitive Stability Principle: The observer's eye-movement vector, v_2, is defined as an unstable component of the retinal field and is therefore removed.

The consequence is that the group, G, of retinal vectors, is decomposed into the group G_2, of observer vectors, and the defined group, G_1, of object vectors:

$$\begin{aligned} G &= G_1 \cdot G_2 \\ &= \text{Object group} \cdot \text{Observer group.} \end{aligned}$$

Furthermore, this decomposition is an algebraic stability ordering—that is, G_2 is factorized before G_1.

Thus the above appears to be a corroboration of the Description Postulate, that the stimulus set is described by an input group with an algebraic stability ordering. Furthermore, recall again that the observer's phenomenal impression of the object's velocity corresponds to the zero value in the G_2 group (that is, upon the factorization of G_2). This corroborates the Principle of Stimulus Identity (Sect. 12) which states that a stimulus is identified as that (other) stimulus which is the machine object (identity element) under a common input-space description.

18.2. *Induced Motion*

In this section, we ultimately analyze situations like the following: Duncker pointed out that, in the night sky, as shown in Fig. 38, the moon is seen to be moving relative to the clouds, and the clouds are seen as moving relative to the buildings. In order to develop a precise analysis of this example, we look first at the following example: Duncker (1929) showed that if a point, a, is kept stationary with respect to the observer but a luminous rectangular frame surrounding the point is moved, then the observer has the impression that the point has moved (see Fig. 20g). That is, the rectangular frame acts as a reference for the point.

In Leyton (1982), it was argued that this example can be explained by the Description Postulate as follows: First, note that, although in the literature (Rock, 1975) the rectangular frame is called the reference frame, closer examination appears to reveal that this may not be the case. Indeed, I claim that it is not be configuration of stimuli, forming the rectangle, which acts as the reference anchor, but the *velocity* of that configuration. For, it is the velocities of the point and the rectangle that are being compared.

Thus according to the approach taken in the present series of papers, one constructs first the relevant *state space*, which is the two-dimensional set of velocities. The referencing enters because each velocity is labeled by the generating input that is required to create it from the zero velocity. Thus the set of velocities is parameterized by the input group of a machine, and this defines the relevant cognitive reference frame.

With this concept, one can interpret Duncker's data by saying that the velocity of the rectangular frame provides the reference point, zero, in the velocity space; that is, the rectangle's velocity is the *machine object* of the associated machine. The question now remains as to why the group identity element is assigned to the rectangle rather than the point stimulus, a.

Recall that the stability ordering on the input group requires the group identity element to be assigned to the most stable point encompassed by the reference frame. The fact that stability is the criterion, which decides this assignment, is supported by the literature. For example, it as been found that motion is judged with respect

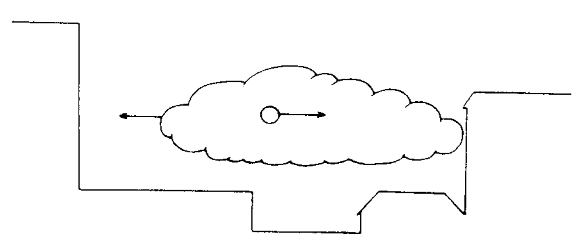

FIG. 38. The cloud–moon illusion.

to objects which are larger, vertically orientated, and more constant (Oppenheimer, 1934, cited in Rock, 1975). Therefore, this class of relative motion examples accords with the claim, in the Description Postulate, that the input group has an algebraic stability ordering.

It is now possible to analyze the situation depicted in Fig. 38. Observe that we have two examples of the rectangle/point structure, one nested inside the other. That is, the buildings/cloud configuration forms one such structure and this contains the cloud/moon configuration which forms another such structure. This yields a three-stage algebraic stability ordering

$$G_1 \cdot G_2 \cdot G_3$$

where each group G_i is the relative velocity group of the system that immediately surrounds the system with the relative velocity group G_{i+1}. Thus the moon's velocity is referred to that of the cloud which is referred to that of the building. Observe that, in accordance with the above discussion, the identity element of G_{i+1} labels the velocity of the immediately surrounding system.

18.3. *Relative Motion*

Having argued that the research on motion perception corroborates the general cognitive principles advanced in the first paper, let us now attempt to show that it corroborates the specific details of the theory of perceptual organization, which has been developed in the present paper.

Johansson (1950) has carried out an extensive analysis of relative motion. Consider Fig. 20h, in the first figure of this paper.[9] It represents two dots being moved perpendicularly to each other in harmonic motion and in phase. Johansson found that, if only one of the dots is presented, it is viewed simply as a dot moving in its line of motion. However, he showed that if the two dots are presented moving at the same time, then they are not perceived as the superposition of the separate movements. Rather, subjects see the dots as if they are moving outward and inward on a common diagonal line. The impression is often verbally described as that of an expanding and contracting diagonal "invisible rod." Furthermore, this line or rod is itself seen moving, as a whole, in the perpendicular diagonal direction, as shown in Fig. 20i. Johansson calls movement along the "rod's" diagonal line, the *relative motion*, and the movement of the line itself, the *common motion*.

Simple vector algebra was used by Johansson to elegantly explain and make rigorous these concepts as follows: Fig. 20j shows the motion vectors P_1 and P_2 decomposed each as the sum of a pair (X_i, Y_i) of perpendicular vectors. The two X_i vectors give the relative motion of the two dots, the $Y_1 = Y_2$ vector gives the com-

[9] Our discussion concerns the class of situations exemplified at the bottom of Fig. 20, that is, the early Johansson paradigm. When we refer to what we will call the "Johansson situation," we will be referring exclusively to this type of situation.

mon motion of the dots. Johansson showed that this analysis fitted a wide variety of examples.

Now let us incorporate these results into the perceptual theory that has been developed in this paper. Observe that, in these examples, one has a decomposition of the absolute motion group G, into two groups: G_2, the relative motion group, and G_1, the common motion group.

$$\begin{aligned} G &= G_1 \cdot G_2 \\ &= \text{Common motion group} \cdot \text{Relative motion group}. \end{aligned}$$

Furthermore, it will now be argued that this group decomposition involves an *external* group and an *internal* group. To see this, first of all observe that the decomposition conforms to the Principle of Nested Control (Sect. 10): G_2 is viewed as an object, an expanding invisible rod. This undergoes the action of a group G_1 which pulls the rod backwards and forwards along the Y direction. Thus, considering the above equation, the group G_1 is perceived as a control parameter acting on the group G_2. Therefore, the relative motion group can be regarded as an internal group.

However, G_2 itself seems to have a *decomposition.* To see this, consider it's internal dynamical structure. Start with one of the end points of the rod, and label it $\{e\}$, the trivial dynamical system. A reflection about the Y axis sends it onto the other end point: that is, yielding the internal group $Z_2 \cdot \{e\}$. In fact, the rod is reflectionally symmetric not only about the Y axis but also about the X axis (see Fig. 20i and j). Thus the internal group is $Z_2 \times Z_2 \cdot \{e\}$.

However, the rod is stretching and contracting. This means that the pure deformation group, A, is applied to the internal group, yielding the generation sequence $A \cdot Z_2 \times Z_2 \cdot \{e\}$. Observe that, in fact, the application of A corroborates the Interaction Laws: The symmetry axes of $Z_2 \times Z_2$ are indeed the eigenspaces of A. Thus the group product appears to be given by $\hat{\imath}$, and one obtains the sequence $A \hat{\imath}(Z_2 \times Z_2) \cdot \{e\}$.

Finally, the common motion group G_1 is applied, translating the rod perpendicularly through space. One can call G_1 the one-dimensional translation group R. Therefore, the full group of the perceptual organization is

$$G = R \cdot A \hat{\imath}(Z_2 \times Z_2) \cdot \{e\}.$$

Thus, what are beginning to emerge, one by one, in this expression, are the group factors defining the Cartesian Reference Frame. There is a simple reason for this: the (X, Y) vectors introduce a cartesian frame in the conventional sense. We appear therefore to have a corroboration of Proposal 24, that the basis of the impression of the imposition of the cartesian frame is the occurrence of the frame's group sequence along the decomposition sequence of the organization group. In fact it would be a simple matter to show that all the factors of the Full Cartesian Frame occur along the decomposition sequence of the organization group. For example, the shear group N should be introduced to describe the extent of shear of the Y axis with respect to the X axis, etc.

Two conclusions appear to be evident from the above argument.

(1) The phenomenon of motion perception seems to be encompassed by the theory of perceptual organization which has been developed in Part II. For example, the phenomenon is explained by the theory of the internal and external group decompositions, by the Interaction Laws, and by the theory of cartesian reference frames.

(2) The perceptual organization of motion appears to conform to the basic cognitive theory in Part I. For example, we have seen that the percept is constructed by a sequence of nested dynamical systems, i.e., each system acting as a control group of its i-subsequence. Thus, in accord with the theory of grouping elaborated in the Section 16.3, the percept partitions into a set of cosets where the coset leaders are values in a control parameter. For example, the cosets of the i-subsequence $Z_2 \times Z_2 \cdot \{e\}$ are the set of separate diagonal occurrences of the rod in the Y direction. Also, in accordance with the theory, the cosets of the opposite sequences of the decomposition (i.e., sequences starting at the left of the decomposition sequence) while forming sets of equivalent points, do not form a perceived partition.

19. The Structure of Complex Shapes

In this section, it will be argued that the theory which has been developed in this paper gives an explanation of complex perceptual shapes such as those in Figs. 20a and d. Ultimately, what will be presented is a theory of the wide variety of shapes in Fig. 41. Such shapes in fact present less theoretic difficulty than the simpler ones we have so far been considering. This is because the simpler ones have pools of geometrically equivalent axes where different contexts determine different salience distributions over the axes.

Let us therefore return to the man shown in Fig. 20a. The Marr–Nishihara theory is that one assigns a hierarchy of successively greater detailing; that is, a multi-resolution hierarchy. The cylinder in Fig. 20b is the first stage; the figure in Fig. 20a is a second stage, etc. As an alternative hierarchy, Fig. 20b can resolve to the goblet shown in Fig. 20c.

The theory developed in Part II allows one, however, to develop a different view. Instead of regarding the hierarchy as one of resolution, I propose that it is one of *deformation.* Thus the progression from the stage where there is only the main cylinder to the stage with arms and legs is given by understanding the former to have a certain structure of flexibility such that it allows stretches at certain points producing the legs and arms. Similarly, another flexibility structure on the initial cylinder allows deformation to a goblet with its own hierarchy. Thus the two hierarchies correspond to the two flexibility structures. What has to be recognized is that the flexibility structure is a view of the main cylinder as a machine object

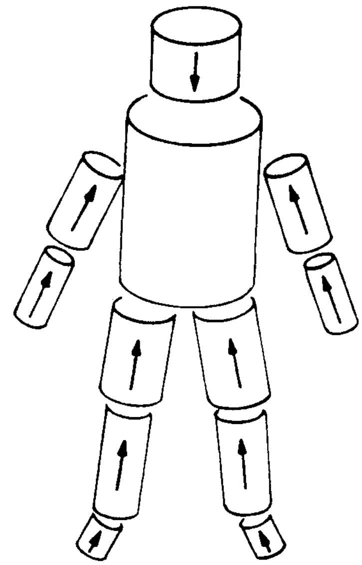

FIG. 39. A man analyzed as a pure-deformation hierarchy.

(identity element, recall Definition 4) with an external group of deformations assigned in a particular way.

This external group *describes* the hierarchy in the sense of Definition 3. That is, there is an input-space description $D: G \to S \cup \{0\}$ which images a deformation group G onto the hierarchy. Any point in the hierarchy, e.g., the shape in Fig. 20a, is the image of some deformation, g, in G. The consequence is that a shape such as Fig. 20a is placed in a pure deformation space of shapes. One can represent the deformation relationships within the hierarchy by a diagram such as Fig. 39. The arrows in that diagram describe the *reversals* of the applied deformations. This means that the arrows also describe the *reference* ordering within the hierarchy.

The view that I have proposed above can be substantiated, in detail, as follows. First, observe that Marr and Nishihara take, as a basic component of their model, the cylindrical module as defined by Binford (1971). This module, called a generalized cone, is diagrammed in Fig. 40a. It is structured by Binford as a circle (the cross section) moved through space along an axis and deformed as it moves.

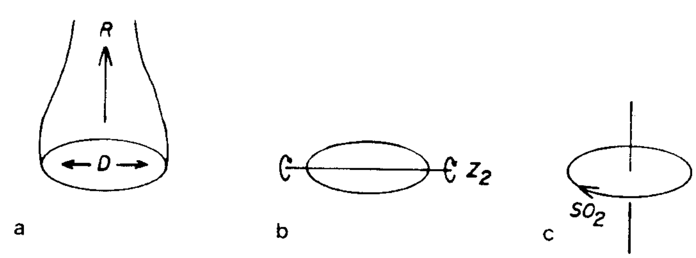

FIG. 40. (a) Binford's concept of a generalized cone. (b) and (c) The cross section analyzed in terms of the group structure: (b) One of the Z_2 axes of the cross section. (c) The SO_2 structure of the cross section.

The axis is called the *spine*, and the deforming action, on the cross section, is called the *sweeping rule*. Let us now see how Binford's proposal conforms to the theory of perceptual organization presented in the present paper. First, the initial cross-sectional circle can be generated by applying the circle group SO_2 to a point $\{e\}$, as shown in Fig. 40c. However, the circle is also reflectionally symmetric, as shown in Fig. 40b. Therefore, the circle is given by a product of the reflection group Z_2, and the rotation group SO_2. However, note that Z_2 makes reflected copies of the rotation group SO_2. Therefore, the product of Z_2 and SO_2 is a semi-direct product (recall the discussion on semi-direct products in Sect. 16.4). That is, one has the product $Z_{2(\times_{sd})}SO_2$ (reflection times rotations). Therefore, the internal group is $(Z_{2(\times_{sd})}SO_2) \cdot \{e\}$.

To generate the entire cylinder from the cross section, one then applies an external group product $R \cdot D$, where D is the dilation group giving changing width (i.e., the sweeping rule), and R is the translation group of the circle through space along the axis (i.e., the spine). This yields Fig. 40a. Thus, one has a full sequence of groups that has $R \cdot D$ as its external sequence, and $(Z_{2(\times_{sd})}SO_2) \cdot \{e\}$ as its internal sequence. That is, the full sequence is:

$$R \cdot D \hat{\imath} (Z_{2(\times_{sd})}SO_2) \cdot \{e\} \qquad \text{Binford–Marr situation.}$$

Now recall the sequence that was given for the Johansson motion experiments:

$$R \cdot A \hat{\imath} (Z_2 \times Z_2) \cdot \{e\} \qquad \text{Johansson situation.}$$

When one compares this sequence, symbol for symbol, with the sequence deduced for the Binford/Marr situation, one finds something remarkable: The two sequences are essentially identical! The one given for the Binford and Marr situation is essentially a three-dimensional version of the sequence that was deduced for the Johansson motion situation. Let us go through these sequences in order from right to left. Trivially, in both cases, one begins with a point $\{e\}$. Then one has the product $Z_2 \times Z_2$ in the motion case and $Z_{2(\times_{sd})}SO_2$ in the Binford/Marr case. These two products are simply different dimensional versions of each other. Because we have three dimensions in the Binford and Marr situation, one of the Z_2 groups in the motion sequence is replaced by SO_2. After this, in both sequences, one comes to a pure deformation factor: A in the motion case and D in the Binford/Marr situation. However, the fact that the dilation group in the Johansson situation was written as A—that is, with unequal stretch in perpendicular directions—was unnecessary. It could have been written as a one-dimensional version of D, the dilation group (with equal stretch in all directions, when multi-dimensional). Finally, one comes to the factor R in each case. Observe that in the motion situation it moves the rod through space, and in the Binford case it moves the cross section through space. Thus the common motion vector, in the Johansson situation, becomes the spine, in the generalized cone; and the relative motion vector, in the Johansson situation, becomes the sweeping rule in the generalized cone. In conclusion, therefore, we observe that the type of analysis that has been elaborated in this paper reveals that the same perceptual structure underlies these two apparently different phenomena.

Let us check further some aspects of the sequence given for the Binford/Marr situation. First, observe that this sequence conforms with the Principle of Nested Control. For example, the external sequence moves the internal sequence about.

Now consider again Fig. 40. Observe that both R and D, in the external group sequence $R \cdot D$, act on the internal sequence via the Interaction Laws: (1) The Z_2 symmetry axes in the plane of the circle (e.g., the horizontal in Fig. 40b) become the eigenspaces of the dilation group, D (e.g., the horizontal line in Fig. 40a); and (2) the rotation axis of the internal group SO_2 (the vertical line in Fig. 40c) becomes the eigenspace of the translation group R (the vertical line in Fig. 40a).

Thus, in particular, when a symmetry axis is perceived along a limb in a figure, the perceptual system interprets it as an eigenspace of R. However, the action of R within the figure is that of extending or shortening the length of a limb. Therefore, it can be understood as a pure deformation, or stretch, in the direction of R.

This brings us from the level of the cylinder to the structure of the hierarchy itself. Corroboration of my proposal that the hierarchy is indeed one of pure deformation was supplied in Leyton (1984a, Experiment 13). Subjects were presented with the 22 drawings of animals, birds, plants, and abstract shapes (shown in Fig. 41). They were asked to assign directions of flexibility to each figure, at four points in each figure, in this way: At each of the four points, there was a star comprised of four lines. One perpendicular pair of lines (in the star) was a pair of local symmetry axes.[10] The other two lines were diagonal to these and were not symmetry axes themselves. The subjects were asked to choose the line describing maximal flexibility of the figure in that region. With considerable statistical significance, subjects chose one of the two local symmetry axes as a line of flexibility [$n = 12$; 88 choices per subject; expected mean $= 44$; actual mean 77.58; $t(11) = 27.228$, $p < .0005$, one tailed]. This implies that the subjects were converting symmetry axes into eigenspaces. Thus the experiment strongly corroborates the proposals (1) that the transitions between the hierarchy levels are created by pure deformations, and (2) that the latter act in accordance with the Interaction Laws.

Eric Grimson (personal communication) has emphasized an important consequence of having the type of hierarchy that has been proposed here, in contrast to the Marr–Nishihara system. Each stage in our hierarchy is defined *explicitly* as the result of a set of actions on previous stages; i.e., in terms used in the first paper, *each stage is defined as its referential relation to the previous stages.* In contrast, any stage in the Marr–Nishihara system is memoryless with respect to its derivation. The explicit derivation memory in our system is of course a consequence of the adherence of our system to the Description Postulate: that is, the set of stages is given an *input-space description*, i.e., a derivation structure; and the description has

[10] The two local symmetry axes were (1) the line coincident with the imaginary line that was equidistant from the two nearest sections of outline; and (2) the line perpendicular to this; i.e., representing a Z_2 symmetry axis of the associated three-dimensional cross section, as in Fig. 40b.

an induced *algebraic stability ordering* which forces a stage to have a direction of derivation reversal. This ensures that a stage is cognized as a point in a reference frame in the strict sense of Definition 12; that is, the stages are internally referenced to each other via an algebraic stability ordering.

Observe that if the pure deformation hierarchy produces the same succession of percepts as the Marr–Nishihara succession, the following perceptual analysis is

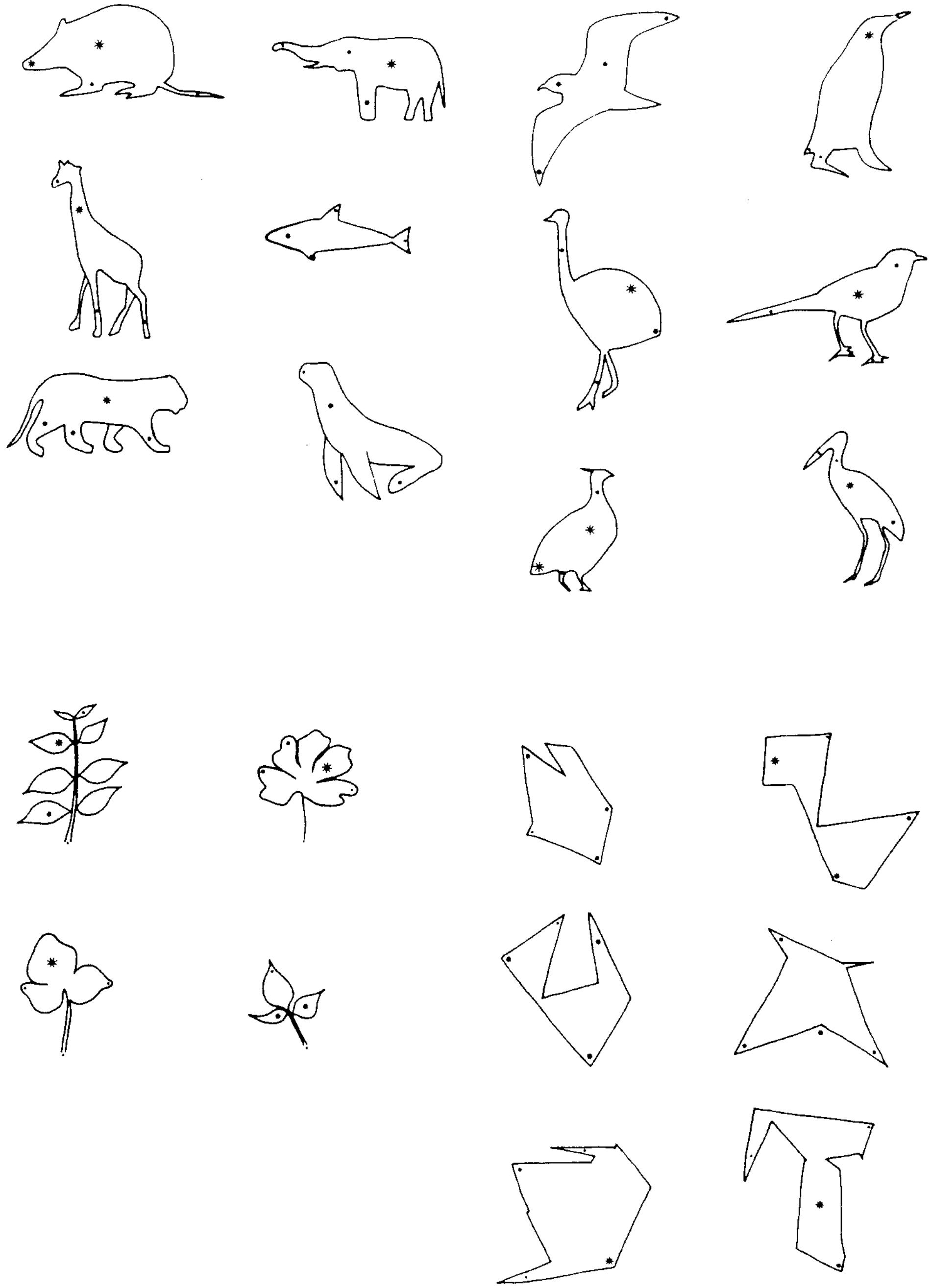

FIG. 41. Twenty-two animals, birds, plants, and abstract shapes, where, at four points in each shape, subjects were asked to choose the direction of maximal flexibility.

implied: (1) Given a figure, the cognitive system assigns local symmetry axes where it can to the figure. (2) In accord with the Interaction Laws, the symmetry axes are converted into eigenspaces; i.e., interpreted as protrusions. (3) The protrusions are removed in order of their size (i.e., length of local symmetry axis), starting with the smallest, because, the smaller the protrusion, the more it is understood as an instability, i.e., a perturbation of the outline of the figure. For example, the ear on the mouse, in the top left of Fig. 41, has (internally) a smaller symmetry axis than the mouse's nose. Therefore, the ear is interpreted more as a perturbation of the figure's outline than the nose is. Thus, the ear is removed first (by contraction) leaving a smooth outline in that region. One then goes onto the next shortest symmetry axis, repeating the same sequence of operations; and so forth. Thus one has an algebraic stability ordering.

Let us now further develop the system described above to provide an analysis of shapes with an underlying curved symmetry structure; e.g., as shown in Fig. 20d. We will incorporate Brady's (1983) proposal that a local straight symmetry axis is propagated along a path (e.g., the dotted line in Fig. 20d); and we shall call the path, the *symmetry line.*

Let us consider the symmetry structure more closely. At each point on the curved symmetry line one can define two axes: (1) the tangent to the line and (2) the cross section. These are illustrated as t_1 and h_1, respectively, in Fig. 42. Together, these axes form a local cartesian frame. This frame undergoes pure deformation, shear, and rotation, as the frame moves along the curved axis. For example, in Fig. 42, besides the frame already discussed, another frame is shown consisting of vectors t_2 and h_2. As the first frame moves along the symmetry line to become the second frame, it clearly undergoes pure deformation, because the cross section narrows. Furthermore, the frame also shears and rotates, because the symmetry line is curved. Thus any protrusion or indentation is given by $(A \cdot N \cdot SO_2)\hat{\imath}(Z_2 \times Z_2)_{(\times_{sd})}R^2)$, the same group sequence which has appeared repeatedly throughout the present paper. The $Z_2 \times Z_2$ component is derived from the pair of axes just described (the local symmetry axis, and the cross section). The $A \cdot N \cdot SO_2$ product represents the fate of this $Z_2 \times Z_2$ structure as it moves along the curved symmetry line. [The full details of this are given in Leyton (1986c), where the isomorphism between the Brady analysis and the SL_2R analysis is established. An earlier version of this isomorphism appears in Leyton, 1985.]

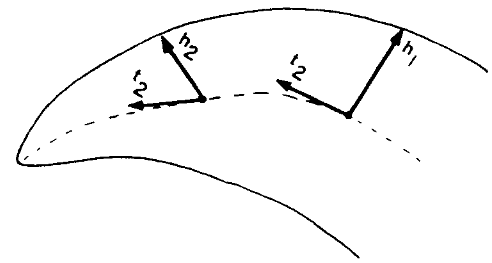

FIG. 42. Two frames on an arbitrary curved axial structure.

It is important to observe that the $A \cdot N \cdot SO_2$ structure acts on the $Z_2 \times Z_2$ structure via the *local* use of the Interaction Laws. That is, the *local* symmetry axes are converted into *local* eigenspaces (local directions of flexibility). For example, the cross section is converted into a direction of local stretch.

Having characterized any curved protrusion by $(A \cdot N \cdot SO_2) \hat{\imath} ((Z_2 \times Z_2)_{(\times_{sd})} R^2)$, it is possible to see how a shape can be successively referenced by the removal, successively of the A, N, and SO_2 factors, as follows: The rotation group is the most unstable in the set of external groups. Therefore it is removed first in the external reference, thus causing the particular section of the figure to straighten. In fact, in Leyton (1986c), I prove a theorem that shows that, for curved structures, rotation and shear have to be removed at the same time due to purely mathematical constraints. Thus one has a reference transition of the type illustrated in going from Figs. 43a to b. Once the axial structure has been straightened in this way, the relevant section of the figure is interpreted as a protrusion, because, according to the Interaction Laws, the axial structure is cognized as an eigenspace structure (a structure of flexibility). Having straightened by removing rotation and shear, one then removes stretch. This factorizarion causes the removal of the protrusions themselves (by shrinking), thus obtaining a more stable, or prototypical shape—illustrated by the oval in Fig. 43c. Observe that this causes a removal of any variation of curvature that is not globally symmetric (e.g., there are no ripples on the oval). The reason for this is given by a theorem which I proposed and proved in Leyton (1986b). This theorem, which will be referred to here as the Symmetry-Curvature Duality Theorem, states that any local curvature variation corresponds to a unique symmetry axis that terminates at the point where the curvature variation is maximal. Thus the removal of the axis by factorizing stretch, causes the removal of the variation.

The disappearance of purely local curvature variation allows the referencing now to become global; for, as is illustrated with the oval in Fig. 43d, the symmetry axes are now global. Thus, by converting the global symmetry structure of the shape (e.g., oval) into a global eigenspace (flexibility) structure, in accord with the Interaction Laws, the global action of the $A \cdot N \cdot SO_2$ sequence can then be locked into place. This sequence of subgroups can be factorized, thus allowing the oval to be referenced to the still more stable figure, the circle; that is, one obtains the transition from Figs. 43c to d.

Let us conclude by stating, explicitly, the successive stages of the reference sequence described in this section. It is important to observe that the stages fall into two halves: the local action of the Interaction Laws, to reference via the local use of the sequence $(A \cdot N \cdot SO_2) \hat{\imath} ((Z_2 \times Z_2)_{(\times_{sd})} R^2)$, recursively, at ever increasing scales;

FIG. 43. A successive reference sequence from a complex shape with a curved symmetry structure.

followed by the global action of the Interaction Laws to reference by the global use of the sequence $(A \cdot N \cdot SO_2)\hat{\imath}((Z_2 \times Z_2)_{(}\times_{sd)}R^2)$. The stages are as follows.

(1) Initially a local symmetry analysis is applied, producing, for example, curved symmetry structures [e.g., of the type described in Blum (1973) or Brady (1983)];

(2) Then the Interaction Laws are used locally to convert the local symmetry axes into local eigenspaces of $SL_2R = A \cdot N \cdot SO_2$.

(3) The shape is then locally prototypified by the removal of the SL_2R factors. However, this successive factorization of $A \cdot N \cdot SO_2$ is itself repeated successively starting from the smallest and progressing to the largest scales. That is, at each referencing level in the recursive referencing process, there is a successive factorization that consists of:

(a) the factorization of rotation and shear, which straightens protrusions;

(b) the factorization of stretch, which removes the protrusions themselves.

(4) At some level, the reference recursion in stage (3) removes all purely local curvature variation, in accord with the Symmetry-Curvature Duality Theorem (Leyton, 1986b). The figure can then be given a global symmetry analysis.

(5) This allows the Interaction Laws to be applied globally, converting global symmetry axes into global eigenspaces of a globally acting $A \cdot N \cdot SO_2$.

(6) Finally, global referencing can occur by the successive removal of these subgroups.

An extension of these principles to three-dimensional representations is discussed in Footnote 11.

[11] Research that I am doing at present extends these techniques to three-dimensional representation. The extension is being made in the following way: Consider an arbitrary smooth three-dimensional shape. Any planar cross section yields a smooth curve with arbitrarily many protrusions and indents. Describe the shape as a family of such (non-intersecting) cross-section curves. This description is a generalization of the generalized cone, because the cross-section curve, in the present case, is allowed to change shape smoothly but arbitrarily. For example, the cross section could start as the complex outline shown in Fig. 43a, and change smoothly through space into differently shaped complex outlines, i.e., with different protrusions and indents.

By the Symmetry-Curvature Duality Theorem, the indents and protrusions (on any cross-section curve) all have unique symmetry lines. Now, as one moves through the family of cross-section curves, the symmetry line of each protrusion and indent will change smoothly, because the protrusions and indents change smoothly. Thus the symmetry lines will form sheets through the shape, and the sheets will smoothly appear and disappear while bending and twisting through the body of the shape. By the Interactions Laws these "symmetry sheets" must be interpreted as sheets of local eigenspaces; i.e., sheets of local flexibilities along which prototypification can occur. To illustrate: let us suppose that a cross section is given by the complex outline in Fig. 43a. This outline has a collection of symmetry lines, and, by our earlier discussion in the paper, each symmetry line implies a local two-dimensional axis structure of eigenvectors (i.e., the cross-section vector and the local symmetry vector). Now, given any such two-dimensional frame on a symmetry line in a cross-section plane, there is a third local axis that points away from the plane and is tanegential to the symmetry sheet that cuts through the plane at that point; i.e., the axis lies along the direction of cutting. Thus, at any point on the symmetry line, there is a local

21. SUMMARY OF PAPER

The basic proposal of the theory developed in this paper stated that a perceptual organization is the interaction of two groups, an external (altering) group, and an internal (generating) group of the set. It was proposed that the cartesian reference frame has an external group which is SL_2R with the algebraic stability ordering defined in the first paper. Furthermore, an argument was given that the algebraic stability ordering, on the SL_2R sequence, is determined by the number of eigenspace lines.

An attempt was then made to show that the internal group also has a decomposition sequence defined by an algebraic stability ordering; that is, it conforms to the Description Postulate. This lead us to a theory of grouping. It was argued that a *grouping* is a coset of an *i*-subsequence of the internal decomposition sequence. An alternative definition of a grouping was proposed, as the machine object of a group factor, and this was argued to be equivalent to that just given.

Following this, the internal structure of a cartesian reference frame was defined.

The major connection between the internal and external structures of a perceptual form was argued as being determined by a product operation along the full decomposition sequence, where this operation was seen as being defined by two "Interaction Laws." These laws essentially equate the symmetry axes of the internal group with the eigenspaces of the external group. One can in fact trace the fate of the symmetry axes as eigenspaces, upwards through the external group factors. The Full Cartesian Reference Frame was also defined as the product of its internal and external group sequences where the product operation is that given by the Interaction Laws. It was then proposed that the basis of the impression of the imposition of a cartesian frame on an organization is the occurrence of the group decomposition factors of the Full Cartesian Reference Frame, along the full group decomposition sequence of the organization.

The consequence of having the operations defining the cartesian frame as members of the decomposition sequence of the organization group of the figure is that the figure is defined as part of a system of figures elaborated by SL_2R, each with an internal generative structure $(Z_2 \times Z_2)_{(\times_{sd})} R^2$. Furthermore, the former acts on the latter via the Interaction Laws.

---

three-dimensional frame; two of the axes are in the plane of the cross section and the third points out of the plane. This three-dimensional frame has a particular stretch, shear, and rotation. Therefore, the state of the frame can be given by a transformation in SL_3R (the three-dimensional version of SL_2R) and we can define the three-dimensional Cartesian Reference Frame as $SL_3R\hat{\imath}((Z_2 \times Z_2 \times Z_2)_{(\times_{sd})}R^3)$ where the three Z_2 axes are the three axes of the frame and the symbol $\hat{\imath}$ converts the axes into eigenvectors of SL_3R, in accord with the Interaction Laws. Prototypfication can now occur by successively removing the SL_3R subgroups, recursively from the most local to the most global level, in the same set of stages as given above for two-dimensional shape. In fact, in the three-dimensional case, the successive effect on any of the cross sections will be similar to the succession of outlines shown in Fig. 43a → b → c → d.

Using a number of the above concepts, it was possible to offer an explanation for the orientation and form phenomenon: The phenomenon occurs when a description map, D, from the full organization group (which includes the frame) can be assigned in two possible ways. The result is that the figure is cognized as being embedded in two alternative systems of figures and, by the Interaction Laws, has different assignments of generative structures. The orientation and form phenomenon was thereby generalized to involve not only orientation but also pure deformation and shear (and, of course, any other subgroup of the external group). Using this theory, it was possible to offer a solution to the square–diamond problem. The square and diamond were shown to be distinguished by their internal group assignments and their external group assignments. Furthermore, it was shown that the latter difference is that predicted from the former by the Interaction Laws.

The discussion then turned to motion perception. First, it was argued that the Description Postulate could explain certain absolute motion phenomena. For example, an attempt was made to show that the collection of retinal motions has a group decomposition with an algebraic stability ordering, i.e., conforming to the Description Postulate. Similarly, it was argued that the Duncker example of induced motion—e.g., in the moon, clouds, building phenomenon—is also given by an algebraic stability ordering. Finally, the Johansson relative motion examples were seen to corroborate the organization principles advanced in this paper; for instance, to be explained by an internal and external group decomposition, by the Interaction Laws, and by the machine-theoretic view of cartesian reference frames.

The discussion concluded by turning to complex natural and abstract shapes. It was argued, using experimental evidence, that the hierarchy identified by Marr and Nishihara is not a resolution hierarchy, as they claim, but a pure deformation hierarchy—one that is an algebraic stability ordering and conforms to the definition of reference frame given in the first paper. Furthermore, the successive stages were seen to accord with the Interaction Laws. In fact, Binford's generalized cone, which underlies such structures, was shown to have exactly the same group decomposition sequence as the Johansson motion phenomenon. Finally, a successive reference structure was proposed for complex shape. The stages progress from the local to the global by recursive use of the Interaction Laws to create successive reference via $(A \cdot N \cdot SO_2)\hat{\imath}((Z_2 \times Z_2)_{(\times_{sd})} R^2)$, at each successive level.

Acknowledgments

This research in visual perception could not have taken place without several penetrating analyses by Jim Higgins concerning modern art. I am immeasurably grateful to Jim for his support and inspiration. I thank Joe Lappin for extensive and detailed feedback on two of the drafts of this paper. Revisions of the paper were supported by NSF grant 1ST-8418164.

References

BINFORD, O. B. (1971). *Visual perception by computer.* Presented at IEEE Systems, Science, and Cybernetics Conf., Miami, FL.

BLUM, H. (1973). Biological shape and visual science (part 1). *Journal of Theoretical Biology*, **38**, 205–287.

BLUM, H., & NAGEL, R. N. (1978). Shape description using weighted symmetric axis features. *Pattern Recognition*, **10**, 167–180.

BRADY, M. (1983). Criteria for Representations of Shape. In A. Rosenfeld & J. Beck (Eds.), *Human and machine vision.* Hillsdale, NJ: Erlbaum.

BRADY, M., & ASADA, H. (1984). Smoothed local symmetries and their implementation. *The International Journal of Robotics Research*, **3**, 36–61.

DUNCKER, K. (1929). Uber induzierte Bewegung. *Psychologishe Forschung*, **12**, 180–259. Translated and condensed in W. D. Ellis (Ed.), *A source book of Gestalt psychology.* New York: Harcourt Brace.

GOLDMEIER, E. (1937). Uber Anlichkeit bei gehenen Figuren. *Psychologische Forschung*, **21**, 146–209. English translation by E. Goldmeier: Similarity in visually perceived forms. *Psychological Issues* (1972, No. 1, Monograph 29). New York: International Universities.

HOFFMAN, K., & KUNZE, R. (1961). *Linear algebra.* New York: Prentice Hall.

JOHANSSON, G. (1950). *Configurations in event perception.* Stockholm, Sweden: Almqvist and Wiksell.

LEEWENBERG, E. L. J. (1971). A perceptual coding language for visual and auditory patterns. *American Journal of Psychology*, **84**, 307–349.

LEYTON, M. (1982). *A unified theory of cognitive reference frames.* Proceedings of the Fourth Annual Conference of the Cognitive Science Society, Universisty of Michigan, Ann Arbor, Michigan, pp. 204–209.

LEYTON, M. (1984a). *A theory of information structure.* PhD thesis, University of California, Berkeley.

LEYTON, M. (1984b). Perceptual organization as nested control. *Biological Cybernetics*, **51**, 141–154.

LEYTON, M. (1984c). *A theory of cognition. II. Perceptual organization.* Submitted for publication.

LEYTON, M. (1985). Generative systems of analyzers. *Computer Vision, Graphics, and Image Processing*, **31**, 201–241.

LEYTON, M. (1986a). Principles of information structure common to six levels of the human cognitive system. *Information Sciences*, **38**, No. 1, entire journal issue.

LEYTON, M. (1986b). Symmetry–Curvature Duality. *Computer Vision, Graphics, and Image Processing*, in press.

LEYTON, M. (1986c). *Local prototypification of complex shape by Lie group action.* Submitted for publication.

MACH, E. (1897). *The analysis of sensations.* English translation (1959), New York: Dover.

MARR, D., & NISHIHARA, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London B*, **200**, 169–294.

OPPENHEIMER, F. (1934). Optische Versuche uber Ruhe und Bewegung. *Psychologische Forschung*, **20**, 1–46.

ROCK, I. (1973). *Orientation and form.* New York: Academic Press.

ROCK, I. (1975). *An introduction to perception.* New York: Macmillan.

WERTHEIMER, M. (1923). Laws of organization in perceptual forms. English Translation (1938) in W. D. Ellis (Ed.), *A source book of Gestalt psychology.* New York: Harcourt Brace.

RECEIVED: February 12, 1986

Printed by the St. Catherine Press Ltd., Tempelhof 41, Bruges, Belgium