

Visual Information About Rigid and Nonrigid Motion: A Geometric Analysis

James T. Todd
University of Connecticut

A mathematical analysis is presented that attempts to describe the available visual information about rigid and nonrigid motion and the three-dimensional structure of rigidly moving objects. Unlike other approaches, the analysis is based on the geometric relations among a set of trajectories defined over an extended region of space-time. Two experiments are reported in which observers viewed computer simulations of moving objects and were required to judge whether the observed motion appeared to be rigid or nonrigid. The results suggest that the mathematical limitations of a trajectory-based analysis of visual information are consistent with the perceptual limitations of actual human observers.

In the mathematical analysis of visual information, perceptual theorists have tended to assume that the only permissible alterations an object can undergo are rigid displacements—changes in position that preserve an object's size and shape. The reasons for this assumption are twofold: First, rigid transformations are probably the most common type of change that one is likely to encounter in a terrestrial environment; second, the assumption of rigidity imposes several constraints that simplify a mathematical analysis. It is important to keep in mind, however, that there are a large number of easily recognizable styles of change such as bending, stretching, twisting, and flowing that do not preserve an object's rigidity. Thus, if the human visual system is able to exploit the intrinsic constraints on rigid motion, as is typically assumed, then there must be some well-defined property of optical stimulation that is uniquely associated with that particular class of transformations.

The need for an analysis of visual information that is capable of distinguishing between rigid and nonrigid motion is clearly revealed by a growing body of evidence that human observers are able to perceive this distinction even under minimal or unusual viewing conditions. The most frequently studied category of rigid motion is rotation in depth about a fixed axis. Human observers

can easily recognize this abstract style of change whether it is applied to a continuous closed contour (Gibson & Gibson, 1957; Wallach & O'Connell, 1953), an irregular pattern of separate elements (Gibson & Gibson, 1957; Metzger, 1934; von Fieandt & Gibson, 1959; Wallach & O'Connell, 1953; White & Mueser, 1960), a single straight line (Jansson & Borjesson, 1969; Johansson & Jansson, 1968; Wallach & O'Connell, 1953), or an isolated configuration of only two or three points (Borjesson & von Hofsten, 1972, 1973; Johansson, 1974b). Other experiments have demonstrated that observers can recognize their own egomotion (Lee & Aronson, 1974; Lee & Lishman, 1975; Lishman & Lee, 1973; Warren, 1976) or the translation in depth of external objects (Borjesson & von Hofsten, 1972, 1973; Johansson, 1950, 1964; Schiff, 1965). There is also evidence that observers can recognize more complicated rigid motions such as rotation about an axis that is translating (Duncker, 1929/1937; Johansson, 1964, 1973, 1974a; Proffitt, Cutting, & Stier, 1979) or rotating (Johansson, 1958, 1974b; Rubin, 1927).

The simplest example of nonrigid change is the relative motion of two rigid objects. There have been several demonstrations that human observers can correctly identify the relative translation of two textured, transparent planes (Balch & Shaw, 1978; Gibson, Gibson, Smith, & Flock, 1959; Mace & Shaw, 1974), the relative rotation of two textured, transparent cylinders (Ullman,

Requests for reprints should be sent to James T. Todd, who is now at the Department of Psychology, Brandeis University, Waltham, Massachusetts 02254.

1979), or the relative motions of individual points within a moving frame of reference (Johansson, 1950, 1958; Restle, 1979). Other styles of nonrigid change include visco-elastic deformations of an object. The evidence suggests that observers can recognize the bending or stretching of a line of points (Jansson, 1977), the elastic compression of a fishnet pattern (von Fieandt & Gibson, 1959), or the bending of a rectangle (Jansson & Johansson, 1973; Jansson & Runeson, 1977). Some varieties of nonrigid change can be correctly identified as the movement of biological organisms. From the relative motions of an unrecognizable configuration of elements, observers can recognize a human form that is walking, jogging, or dancing¹ (Cutting, 1978; Cutting, Proffitt, & Kozlowski, 1978; Johansson, 1973, 1975, 1976; Kozlowski & Cutting, 1977), the expressions of a human face (Bassili, 1978), or intentional, social interactions such as affection or aggression (Bassili, 1976; Heider & Simmel, 1944).

Although there is convincing evidence that human observers are able to recognize a wide variety of rigid and nonrigid motions, the available data do not reveal the underlying information on which these distinctions must ultimately be based. To adequately describe that information, it is necessary to analyze the abstract mathematical relations between moving objects in a natural environment and the structure of light at a point of observation. Let us now examine a specific technique for performing the required analysis.

A Trajectory-Based Analysis of Visual Information

The mathematical analysis of visual information presented below is based on an assumption that the primitive units of human motion perception are global trajectories in an optic array defined over an extended region of space-time. Within this framework the different classes of motion such as rigid and nonrigid are defined by specific geometric relations among a set of trajectories. For example, consider a rigid object that is rotating about a fixed axis. By definition, (a) all points on the object must

move through three-dimensional space along circular trajectories; (b) the centers of these trajectories must all lie along a single straight line that is perpendicular to each plane of rotation; and (c) all of the points must traverse their trajectories at the same frequency. Now consider a projection of this event onto a planar surface such as a windowpane or a television screen. It is easy to prove using elementary projective geometry that (a) the projected images of all points on the object must move about the projection surface along elliptical trajectories (see Figure 1; Eriksson, 1974; Johansson, 1974b); (b) the minor axes of these trajectories must all lie along a single straight line; and (c) all of the points must traverse their trajectories at the same frequency. These three properties of optic structure are necessary consequences of rotation about a fixed axis independent of viewing distance. However, there is one additional constraint on the eccentricities of the elliptical trajectories that is affected by the amount of perspective in any given projection. At an infinite viewing distance (parallel projection) all of the trajectories must have the same eccentricity (see Figure 2A), whereas at shorter viewing distances (polar projection) the eccentricities must increase monotonically along a line connecting the centers of each ellipse (see Figure 3). As is demonstrated in Figure 1, an elliptical trajectory has seven degrees of freedom: phase (α), frequency (ω), eccentricity (ϵ), orientation (θ), size (A), the x-intercept of a line formed by extending the minor axis (B), and the distance between the x-intercept and the center of the ellipse (C) (cf. Restle, 1979). The constraints on rotation about a fixed axis described above affect frequency, orientation, eccentricity, and the x-intercept. The other three degrees of freedom uniquely determine the three-dimensional structure of an object but are irrelevant to the distinction between rigid and nonrigid motion. (A specific procedure for computing the three-dimensional structure of any object rigidly rotating about a fixed

¹ Although each individual limb moves rigidly in a human gait, the overall configuration is nonrigid because the relative orientation of different limb segments is constantly changing.

axis is provided in the Appendix for both parallel and polar projection.)

Figure 2 shows five different classes of motion that are defined by this analysis. Each ellipse represents the parallel projection of the endpoint of a rod moving in three-dimensional space over an extended period of time; the number of points in each ellipse represents the number of time intervals needed to traverse the depicted trajectory. (The moving rods are represented by solid lines.) Case A is a typical example of a pattern of trajectories satisfying the constraints of rotation about a fixed axis. Cases B-E provide specific examples of how these constraints can be violated.

A primary difficulty with the analysis described in the Appendix is that it cannot at present deal with more complicated rigid motions such as rotation about an axis that is translating or rotating. When an object rotates about a moving axis the trajectory of its projected image falls within a general

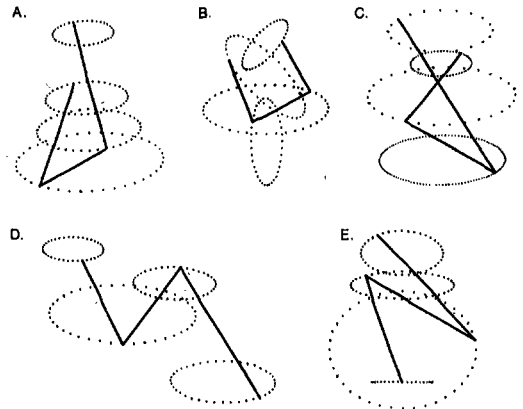


Figure 2. Five different classes of motion that are defined by a geometric analysis of visual information. (Each ellipse represents the image trajectory under parallel projection of the endpoint of a straight line segment moving in three-dimensional space; the number of points in each ellipse represents the number of time intervals needed to traverse the depicted trajectory. [A] Case A is a typical example of rigid motion because all of the trajectories have the same orientation, frequency, x-intercept, and eccentricity. [B] Case B is nonrigid because each trajectory has a different orientation. [C] Case C is nonrigid because each trajectory has a different frequency. [D] Case D is nonrigid because each trajectory has a different x-intercept. [E] Case E is nonrigid because each trajectory has a different eccentricity.)

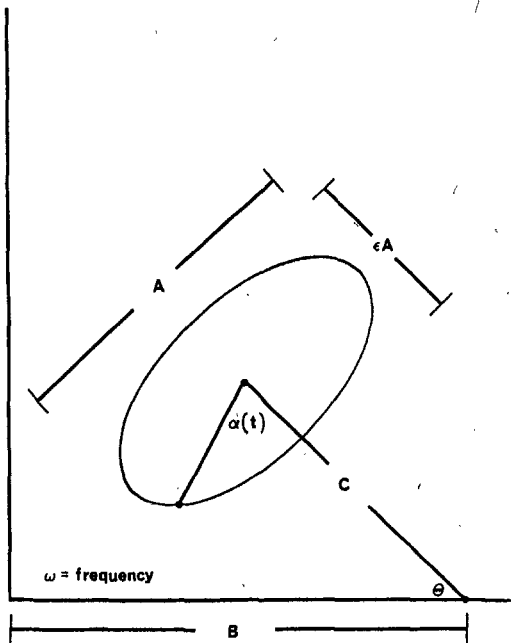


Figure 1. The parallel projection of a single point rotating about a fixed axis produces an elliptical trajectory that has seven degrees of freedom: phase (α), frequency (ω), eccentricity (e), orientation (θ), size (A), the x-intercept of a line formed by extending the minor axis (B), and the distance between the x-intercept and the center of the ellipse (C).

class of forms that are referred to in the present article as hypertrochoids.² For example, consider the projected motion of a single point rotating about an axis that is itself rotating about a fixed point. This type of motion is typically observed in a spinning top or a gyroscope. The first component of rotation is *spin*; the second component is *precession*. Several examples of the projected trajectories that can be generated by such a motion are given in Figure 4.

The problem of a moving axis of rotation could potentially be avoided in a variety of ways. If the axis is moving parallel to the

² Geometers have long been interested in the trajectories generated by moving points within different types of mechanical systems (cf. Lockwood, 1961) and have developed a precise terminology for describing these trajectories. For example, the trajectory of a closed curve that rolls without slipping on a fixed curve is a *roulette*. If the rolling curve is a circle then the resulting trajectory is a *trochoid*. When a circle rolls along a straight line, the trajectory of a point on the circle is a *cycloid*. The term *hypertrochoid* is used here to refer to the projected image of a trochoid on a planar surface.

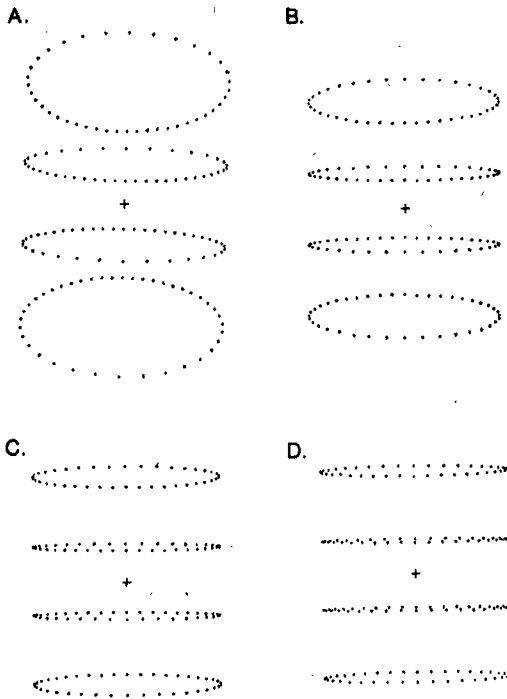


Figure 3. The image trajectories under polar projection of four points on the surface of a transparent cylinder that is rotating about an axis that is parallel to the projection surface. (The viewing distance in [A] is related to that of [B], [C], and [D] by scale factors of 2, 4, and 8, respectively. Unlike trajectories under parallel projection, the eccentricity of each ellipse increases monotonically with its vertical distance from the point of observation, which is represented by a small cross. The ellipses are also assymetrical, since an image point moves slower on the half of its trajectory that is proximal to the point of observation. This is represented in the figure by a greater density of points. The proximal half of each ellipse is the projected image of the portion of an element's three-dimensional trajectory that is farthest away in depth.)

projection plane then the problem could be avoided by adopting a moving frame of reference in which the axis is fixed. There is considerable evidence that human observers do indeed adopt a moving frame of reference when observing rotation about a moving axis, provided that there are a sufficient number of elements in the display (e.g., Proffitt et al., 1979). The classic example of this phenomenon is the perception of rolling motion along a planar ground surface. If observers view a single spot of light on the perimeter of a rolling wheel, the spot will appear to be moving along a cycloidal tra-

jectory. As additional spots of light are added, however, an observer will eventually perceive a set of elements all moving along circular trajectories within a moving frame of reference. More complicated examples involving hierarchically nested frames of reference have also been described in the literature (e.g., Johansson, 1973; Restle, 1979).

If the axis of rotation has a component of translation perpendicular to the projection plane, then the image trajectories must satisfy two additional constraints that are necessary though not sufficient conditions for rigid motion: (a) all of the trajectories must intersect at a common focus (the vanishing point); and (b) the velocity of each image point must gradually approach zero as it moves closer and closer to the point of intersection (cf. Lee, 1974). Two different patterns of trajectories that satisfy these constraints and one pattern of trajectories that does not satisfy these constraints are shown in Figure 5. There is no known procedure

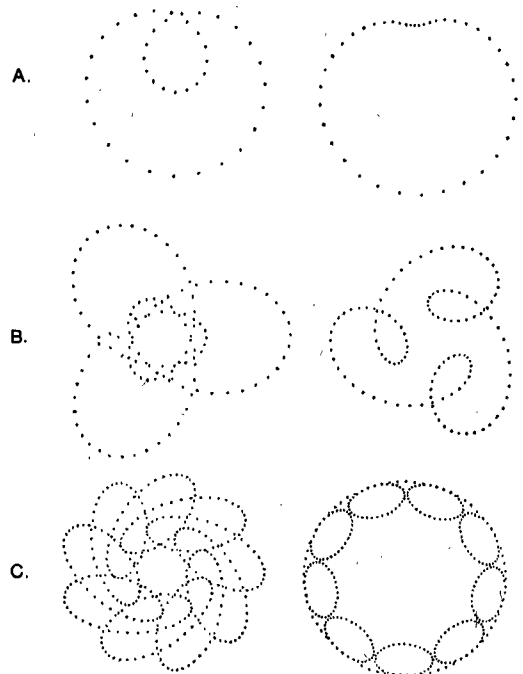


Figure 4. The parallel projection of a single point rotating about an axis that is undergoing precession. (The symmetry of the resulting patterns is uniquely determined by a ratio between the frequency of spin [ω] and the frequency of precession [ϕ]: [A] $\omega/\phi = 1$; [B] $\omega/\phi = 3$; [C] $\omega/\phi = 9$.)

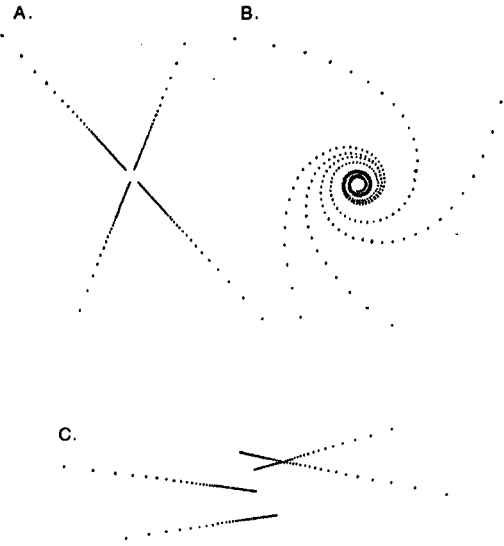


Figure 5. The image trajectories under polar projection of (A) a rigid configuration of four points translating in depth, (B) a rigid configuration of four points translating in depth and rotating about an axis that is perpendicular to the projection surface, and (C) a nonrigid configuration of four points translating in different directions.

based on relative trajectories that is capable of determining the three-dimensional structure of an arbitrary rigid object that is translating in depth.

Empirical Implications

The trajectory-based analysis presented above is an attempt to describe the visual information by which human observers are able to distinguish between rigid and nonrigid motion and to perceive the three-dimensional form of rigidly moving objects. The analysis can be applied under either parallel or polar projection, but it cannot at present deal with all possible varieties of motion. When an object spins about a moving axis of rotation, for example, it is necessary to assume that the observer adopts a moving frame of reference in which the axis is fixed.

It is important to keep in mind when evaluating this analysis as a psychological model that there are other possible techniques for distinguishing between the optic projections of rigid and nonrigid motion and for determining the three-dimensional structure of

rigidly moving objects. Previous investigators have conceived of an object's motion as a sequence of static snapshots (Ullman, 1977, 1979) or as an array of vectors defined at an instantaneous moment in time (Lee, 1974; Longuet-Higgins & Prazdny, 1980). Each of these conceptual frameworks has its own advantages and disadvantages from a purely mathematical point of view; but there are almost no data available to suggest which analysis is most consistent with the actual capabilities and limitations of human perception.

The research reported in the present article examines the abilities of human observers to distinguish between the rigid and nonrigid motion of rotating wire figures under parallel projection. The goal of this research was to discover the parameters of an object's motion that can affect performance and thus to provide an empirical basis for selecting a model of visual information that is psychologically valid.

Experiment 1

Experiment 1 examines the abilities of naive observers to distinguish between rigid and nonrigid motion without the benefit of practice or feedback.

Method

Subjects. Ten naive observers, all graduate students at the University of Connecticut, participated in the experiment. None of the observers was familiar with the mathematical issues being investigated; they were not informed of the specific visual information for distinguishing between rigid and nonrigid motion until after the experiment was completed.

Apparatus. Stimuli were presented on a Tektronix 611 cathode ray tube, refreshed every 5 msec by a Nova minicomputer. The displays were viewed binocularly at a distance of approximately 2.5' (76.2 cm) from a 6.5" × 8.5" (16.5 × 21.6 cm) display screen. Head movements were not restricted.

Stimuli. The displays consisted of three connected line segments (each composed of 11 collinear points) whose endpoints moved along randomly generated elliptical or hypertrochoidal trajectories. Phenomenally, this appeared as an object moving in three dimensional space. Observers were instructed to press one response key if the observed motion appeared to be rigid or a second response key if it appeared to be nonrigid.

There were three separate conditions in which the parameters of motion were varied. Condition A (slow rotation) included all of the categories of motion represented in Figure 2. The possible values of A were 1.52,

2.06, 2.59, 3.12, and 3.06 cm; the possible values of C were -3.05 , -1.52 , 0 , 1.52 , and 3.05 cm; and the possible values of α were 36° , 72° , 108° , . . . 360° . The values of these parameters were selected at random for each trajectory without replacement. All of the parameters that are naturally constrained by rigid motion were assigned fixed values ($\epsilon = .375$; $B = 0$ cm; $\omega = .275$ Hz; $\theta = 0^\circ$). Nonrigid motions were generated in the same way except that one of the constrained parameters was assigned a different value for each trajectory: In Case B, the possible values of θ were 0° , 45° , 90° , and 135° . In Case C, the possible values of ω were $.092$, $.183$, $.367$, and $.55$ Hz. In Case D, the possible values of B were -3.81 , -1.27 , 1.27 , and 3.81 cm. And in Case E, the possible values of ϵ were 0 , $.25$, $.5$, and $.75$. Condition B (fast rotation) was identical to Condition A except that all of the frequencies were doubled. Condition C (precession) was also identical to Condition A except that the entire configuration of points was rotated about the center of the screen with a frequency (ϕ) of $.275$ Hz.³ Given all possible permutations of these different parameters, any one of slightly more than 21 trillion distinct displays could appear on any given trial.

The displays were generated according to the following series of equations:

$$X1 = A \cos(\omega t + \alpha)$$

$$Y1 = \epsilon A \sin(\omega t + \alpha)$$

$$X2 = X1 \cos(\theta) + Y1 \sin(\theta) + B$$

$$Y2 = Y1 \cos(\theta) - X1 \sin(\theta) + C$$

$$X = X2 \cos(\phi t) + Y2 \sin(\phi t)$$

$$Y = Y2 \cos(\phi t) - X2 \sin(\phi t),$$

where X and Y are actual screen coordinates and $X1$, $X2$, $Y1$, $Y2$ are dummy variables used during the computation.

Each display was presented at a rate of 44 frames/sec. The simulated object would rotate for 2.73 sec in one direction and then reverse to rotate for an equal amount of time in the opposite direction. In Conditions A and C, the reversals occurred every three quarter-cycles; in Condition B, they occurred every three half-cycles. The displays continued to oscillate in this manner until a response was recorded, at which time the trial was terminated.

Procedure. The concept of rigid motion was carefully explained at the beginning of an experimental session using a wire coat hanger to demonstrate both rotation and precession. Nonrigid motion was also demonstrated by bending and twisting the coat hanger. (Stretching motions were described but not demonstrated.) The observers were then given 18 trials of practice. The actual experiment consisted of 180 randomly arranged trials: 20 trials of rigid motion (Case A) and 10 trials of each category of nonrigid motion (Cases B-E) for each of the three conditions. An experimental session was divided into three blocks of 60 trials each. No feedback of any sort was given until after the experiment was completed.

It is important to recognize that the methodology of this experiment differs somewhat from the methods that have traditionally been employed in the study of human motion perception. There are several reasons why this particular methodology was adopted. A fundamental

problem in studying complex visual displays is that the number of variables needed to generate a display can be very large, making it impossible to investigate each variable independently within a factorial design. The usual method for coping with this complexity is to restrict the investigation to a relatively small number of individual stimulus configurations (e.g., Borjesson & von Hofsten, 1972, 1973). This approach is adequate for demonstrating sufficient conditions for producing a perceptual effect, but it does not allow one to decide which properties of the display are necessary for obtaining the effect and which ones are not. A better procedure for testing a specific model is to generate as many displays as possible, even if it means that different observers see different stimulus configurations. For example, the trajectory-based analysis of visual information described earlier defines a broad class of displays that ought to be perceived as rigid motion, and a complementary class of displays that ought to be perceived as nonrigid motion. If this analysis is psychologically valid, then it ought to be possible to sample these classes at random to obtain the predicted perceptual effects. This type of procedure provides a more powerful test of the model than a limited investigation of a few specific exemplars.

It should also be pointed out that there is at least one potentially confounding variable in this experiment that could conceivably be used to identify some of the nonrigid displays. Note in Figure 2 that the object represented in Case D is noticeably larger than the objects represented in the other cases. This is an unavoidable result of separating the minor axes of the different ellipses and cannot be controlled without making those ellipses noticeably smaller than the ones used for Cases A, B, C, and E. In order to use this information in the absence of feedback, however, it would be necessary to know exactly how the displays were generated. All of the observers who participated in the experiment were carefully screened to ensure that they had no knowledge of the methodological details of the experiment or the theoretical issues being investigated. As an added precaution, the observers were questioned after the experiment was completed about their strategies for performing the task. Had any of them reported that their responses were based on size or some other irrelevant variable, their data would have been discarded. This was not necessary.

Results and Discussion

The results of the experiment (see Table 1) provide convincing evidence that human observers are highly sensitive to the available

³ In generating the precession trials, each point was moved along an elliptical trajectory, the center of which was rotated in a circular path about the center of the display screen. The radius of this circle for any given point was equal to the distance between the center of its nested elliptical trajectory and the center of the display screen (i.e., $\sqrt{B^2 + C^2}$). Two examples of the possible trajectories that could be generated by this procedure are shown in Figure 4A.

Table 1
*Percentage of Correct Responses for 10
 Observers to Each Category of Motion in
 Experiment 1*

Category of motion	Rotation		Precession
	Slow	Fast	
Case A (rigid)			
%	96.5	96.5	60.0
<i>n</i>	200	200	200
Case B (nonrigid)			
%	95.0	99.0	81.0
<i>n</i>	100	100	100
Case C (nonrigid)			
%	98.0	95.0	84.0
<i>n</i>	100	100	100
Case D (nonrigid)			
%	95.0	95.0	81.0
<i>n</i>	100	100	100
Case E (nonrigid)			
%	60.0	59.0	48.0
<i>n</i>	100	100	100
<i>M</i>	90.2		69.0
<i>N</i>	1,200		600

Note. Overall *M* for Cases A-D = 88.6 (*N* = 1,500). For Case E, *M* = 55.7 (*N* = 300).

information about rigid and nonrigid motion under parallel projection. The observers responded "rigid" to 84.3% of the displays that were mathematically correct projections of rigid motion, and they responded "nonrigid" to 82.5% of the displays that were mathematically correct projections of nonrigid motion. There were large differences, however, among the various cases and conditions. For Cases A, B, C, and D, the level of performance was 96.3% in the two rotation conditions, but dropped to only 73.2% for the precession condition. The level of performance was lowest of all for Case E, where the displays violated the rigidity constraint on eccentricity. The percentage of correct responses was 59.5 and 48.0 for rotation and precession, respectively.

After an experimental session was completed, the observers were asked to describe their strategies for performing the task. All of the observers insisted that their responses were based on the phenomenal impression of rigid or nonrigid motion in three-dimen-

sional space, and it was clear from their descriptions that their ability to perform the task did not involve an explicit understanding of how these classes are distinguished within a two-dimensional visual projection. The observers were also asked to describe verbally some randomly selected exemplars from each of the 15 categories of motion. As it turned out, this proved to be a surprisingly difficult task. They had little trouble describing the rigid rotations, but their descriptions of nonrigid motion were for the most part unintelligible. When prompted by the experimenter to look for changes in the three-dimensional length of the different line segments, or changes in angle between line segments, they generally reported a change in angle for every example of nonrigid motion. Changes in length were sometimes reported for Case D, but these changes were often interpreted as translation in depth. Another observation reported by several observers was that the objects appeared to "bounce" off an invisible obstruction when they abruptly reversed their direction of rotation.

Most of the existing models of the available information about rigid and nonrigid motion would have a difficult time accounting for these results. The traditional hypothesis, first proposed by Wallach and O'Connell (1953) and later by Johansson (1964), is that three-dimensional rigid motion will be perceived whenever the projected contours of an object exhibit simultaneous change in length and angle, and that changes in length alone or angle alone will be perceived as nonrigid motion in the projection plane. This hypothesis cannot be extended to the present experiment, since all of the displays, both rigid and nonrigid, involved simultaneous changes in the lengths and angles of the projected contours.

A model of information based on Ullman's (1977) rigidity test could account for the observers' overall ability to distinguish between rigid and nonrigid motion, but it could not explain the detrimental effect of a moving axis of rotation, or why the displays of Case E were more difficult to recognize than the other examples of nonrigid motion. Ullman's rigidity test involves solving a set of simultaneous equations. If there is a solution, then the observed motion is rigid. If

there is no solution, then the observed motion is nonrigid. This test does not distinguish among different categories of nonrigid motion, and it can easily be applied to any configuration of five or more moving elements (under parallel projection only), regardless of whether the axis of rotation is moving or stationary. These would be desirable properties for optimal performance of this particular task, but they do not appear to be characteristic of actual human observers.

In contrast to these other approaches, a model of visual information based on relative trajectories is able to accommodate all of the differences that were observed during the present experiment. The detrimental effect of precession is consistent with the requirement that all motions be defined within a frame of reference in which the axis of rotation is fixed. Although the model does not suggest how a frame of reference is determined, it does not seem surprising that a frame of reference that is stationary with respect to the display screen is easier to adopt than one that is rotating with respect to the display screen. The large number of errors for the displays of Case E is also compatible with a trajectory-based analysis. The constraint on the relative eccentricities of the elliptical trajectories, which was violated in Case E, is logically independent of the other constraints on rigid motion, which were violated in Cases B, C, and D. Moreover, this particular constraint is less general than the others because it is significantly affected by viewing distance (see Appendix).

Experiment 2

Experiment 2 examined the abilities of practiced observers to distinguish between rigid and nonrigid motion under more impoverished conditions, with immediate feedback after every trial.

Method

Subjects. Three highly practiced observers participated in the experiment and were paid \$3/hour for their services.

Apparatus and general procedure. The apparatus and general procedure were roughly equivalent to those used in Experiment 1. As before, the displays rotated for a certain amount of time in one direction and then reversed to rotate for an equal amount of time in the

opposite direction. In this case, however, the amplitude of oscillation was systematically varied using possible values of 18°, 90°, 180°, and 270°. The purpose of this manipulation was to systematically reduce the amount of information available in order to assess the observer's sensitivities.

There were three separate conditions: Condition A simulated rotation about a fixed axis with a frame rate (*FR*) of 11 frames/second; Condition B simulated rotation about a fixed axis with *FR* = 44 frames/sec; and Condition C simulated rotation about a moving axis with $\omega = \phi = .275$ Hz and *FR* = 44 frames/sec. The initial orientation of the axis of rotation was selected at random on each trial.

Exclusion of cases E and D. Two of the categories of nonrigid motion that were used in Experiment 1 were excluded from the set of possible displays. Case E was not included because it is significantly more difficult to detect than the other possible violations of rigidity constraints. Case D was also eliminated because the images produced in any given frame were noticeably larger than for the other cases. When the observers were given feedback in a pilot version of this experiment, they quickly noticed that the size of the displayed object could reliably indicate if its motion was nonrigid, especially at low amplitudes of oscillation. Thus, Case D was excluded from the final version of the experiment to eliminate size as a potentially confounding cue.

Experimental session. Each experimental session consisted of 300 trials using only one condition and a single amplitude of oscillation. Immediate feedback was always presented after each response. Following a correct response, a small + was presented on the screen for 1 sec, and following an incorrect response, a small "-" was presented. Within a given condition, the sessions were arranged in order of increasing difficulty. The observers first completed Condition A, then Condition B, and, finally, Condition C.

Results and Discussion

The results are presented in Figure 6. These data clearly demonstrate that observers are highly sensitive to violations of rigidity constraints even when they are allowed to view only a small portion of the symmetry period of each trajectory (i.e., 180° for an ellipse). For rotation about a fixed axis, the observers' performance was nearly perfect when the amplitude of oscillation was 180° or 270°. The level of performance was greatly reduced when the oscillation amplitude was only 18°, but it was still significantly greater than chance. (In Condition A, the 18° oscillations consisted of just three frames.) As in Experiment 1 there was a dramatic drop in performance in Condition C when the simulated objects rotated about a moving axis of rotation. This indicates that the detrimental effect of

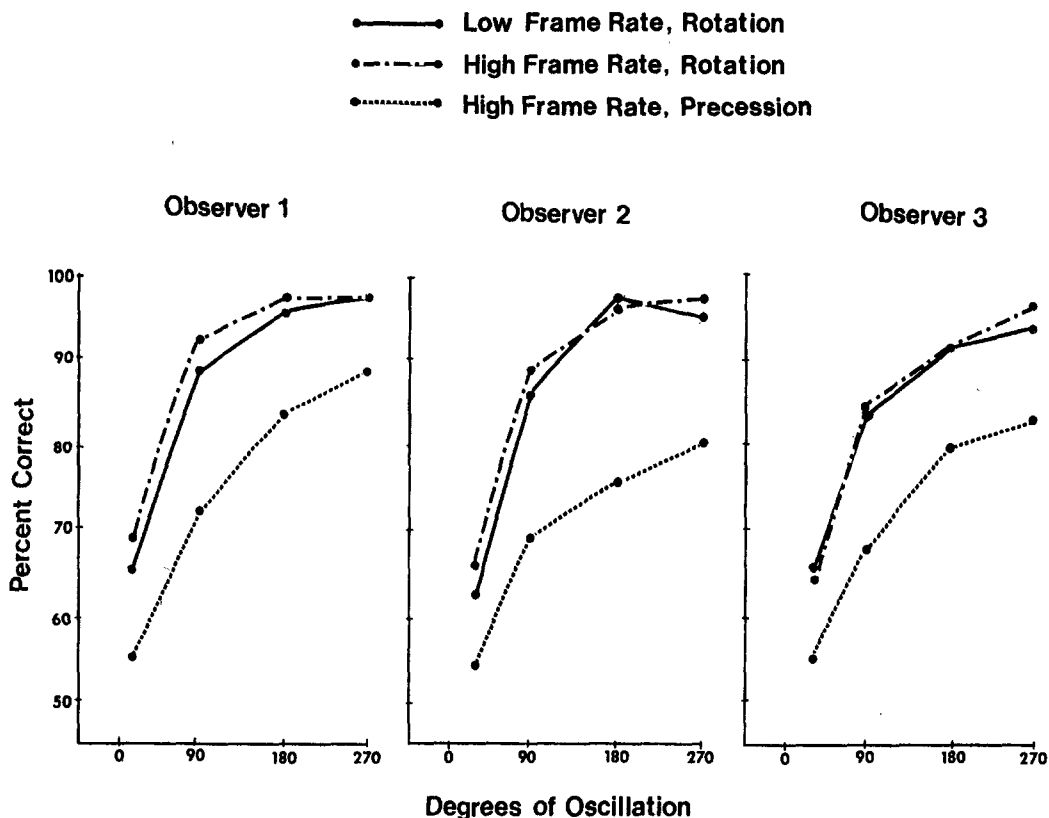


Figure 6. The percentage of correct responses in distinguishing rigid from nonrigid motion for three practiced observers as a function of oscillation amplitude.

precession does not go away with practice or feedback.⁴

Ullman (1977) has proven mathematically that under ideal conditions, two distinct views of five identifiable points are sufficient for distinguishing rigid from nonrigid motion, but a greater number of views or points may be necessary in practice to average out any error in the registration of each point's position. The results of the present experiment suggest that human observers may require considerably more information than one might suspect on the basis of Ullman's analysis. When observers were presented with 9 frames of 31 distinct points over an easily discernible period of 200 msec, and covering an easily discernible angular displacement of 18°, the average level of performance did not rise above 67%, even when the axis of rotation was fixed. This lack of performance is even more significant considering that the observers were highly prac-

ticed, were given feedback on every trial, and were allowed to view the displays for as long as they desired. (This unlimited viewing time ought to have facilitated any averaging process that might have been necessary.) The level of performance did not approach perfection until the observers were presented with 80 frames of 31 points over a period of 1.8 sec covering 180°. The additional observation that an easily noticeable fourfold reduction in frame rate had virtually no effect on performance would lead one to conclude that contrary to Ullman's model, the number

⁴ The present finding that rotation about an arbitrary fixed axis at an angle to the display screen is easier to recognize as rigid motion than rotation about a moving axis contrasts with the results of Green (1961). Using a similar set of conditions, Green found that observer's subjective rigidity judgments were unaffected by a moving axis of rotation. Unlike the present experiment, however, Green's displays included no examples of nonrigid motion.

of distinct frames is not a critical variable for the perception of rigidity.

From the standpoint of a trajectory-based analysis, the critical variable for distinguishing between rigid and nonrigid motion is the amount of each element's trajectory that an observer is allowed to view. In a frame-by-frame presentation such as the one used in the present experiment, the frame rate must be rapid enough to provide a reasonable approximation of the projected paths of motion, but the number of frames per se is not relevant to the analysis. What percentage of an object's trajectory should be necessary for optimal performance? The answer to this question depends on the symmetry period of the trajectory, which in the case of an ellipse is 180° . Since an entire elliptical trajectory can be generated from an arbitrary 180° segment by a simple rotation, it follows that any segment greater than 180° provides redundant information. From a practical point of view, this suggests that the accuracy of a geometric analysis should begin to deteriorate whenever the trajectory segments available to be analyzed are smaller than 180° . This prediction is quite consistent with the three sets of performance curves shown in Figure 6.

Other research in the field of human motion perception has revealed that the amount of information required for specifying different properties of an object's structure and motion can vary considerably from task to task. For example, Lappin, Doner, and Kottas (1980) have shown that two successive frames of a computer-generated visual display provide sufficient information for observers to detect the coherence and three-dimensional structure of a transparent sphere composed of 512 points, which is rotated around a vertical axis through its center. Because this type of display reveals only a small portion of each element's trajectory (5.6°), the accuracy of the observer's judgments seems to contradict the results of the present experiment, in which the level of performance for an 18° rotation was only barely above chance. It is important to keep in mind, however, that the displays used by Lappin et al. contained 512 points on a continuous smooth surface. Although each point was displaced only a short distance, the en-

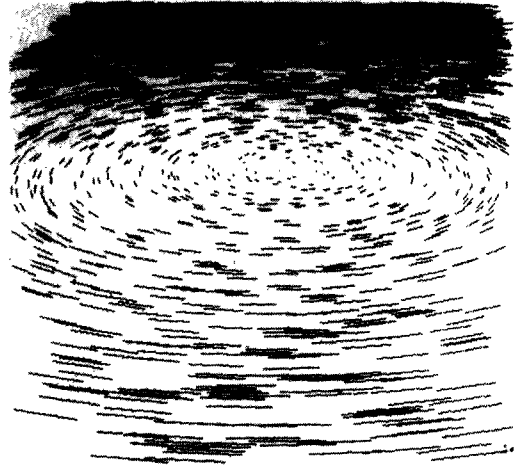


Figure 7. The polar projection of a planar ground surface containing 3,000 random points that are rotated 5° about a vertical axis. (Although each point is displaced only a short distance, the entire set of elliptical trajectories is uniquely specified by the overall field structure.)

tire set of 512 points formed a vector field that uniquely specified the entire elliptical trajectory of each element in the set (see Figure 7). This suggests that a given pattern of trajectories can be specified in two ways: by a few elements moving over a long interval of time or by an entire field of elements moving over a short interval of time. In either case, the trajectories will specify a particular object undergoing a particular style of change.

General Discussion

This article has attempted to develop a mathematical analysis of visual information that can account for an observer's ability to perceive the three-dimensional form of a moving object and to distinguish between rigid and nonrigid motion. The analysis was based on the geometric relations among space-time trajectories on a visual display screen, under either parallel or polar projection. Three different categories of motion were considered with varying degrees of success: rotation about a fixed axis (see Figures 1, 2, and 3), rotation about an axis that is moving parallel to the display screen (e.g., precession; see Figure 4), and translation in depth (see Figure 5). For rotation about a fixed axis, it was formally demonstrated that

the rigidity and three-dimensional form of a moving object are completely specified. This analysis could not be extended to rotation about a moving axis, however, without making additional assumptions about establishing a moving frame of reference. For translation in depth, the analysis described some necessary conditions for rigid motion, but it did not address the problem of how the three-dimensional form of a moving object might be specified.

In an attempt to provide some empirical data relating to the proposed analysis, two experiments were reported in which observers judged whether the motion of a three-dimensional object appeared to be rigid or nonrigid. Experiment 1 examined the performance of 10 naive observers without the benefit of practice or feedback. Experiment 2 examined the performance of three highly practiced observers who were provided with immediate feedback after every trial. The results indicated the following: (a) The ability to distinguish between rigid and nonrigid motion is significantly affected by whether the axes of rotation are moving or stationary. (b) Some violations of rigidity constraints are easier to detect than others. Those constraints that vary with viewing distance (e.g., relative eccentricity) are perceptually less salient than other constraints that are independent of viewing distance (e.g., relative frequency and orientation). (c) Within the limits of the procedures used in these experiments, the frame rate and speed of rotation of a simulated object have no effect on performance. (d) Performance is significantly affected by the amount of each element's trajectory that an observer is allowed to view, however. All of these effects are generally compatible with a trajectory-based analysis of visual information, but they are not easily accommodated by other models based on apparent motion (e.g., Ullman, 1977, 1979) or instantaneous visual flow (e.g. Lee, 1974; Longuet-Higgins & Prazdny, 1980) that have been proposed in the literature.

The contrasting empirical implications of these alternative analyses of visual information may be partially due to the fact that they are designed to operate at fundamentally different scales of observation. For a

trajectory-based analysis, the primitive units are global trajectories defined over an extended region of space-time—a coarse scale of observation. For analyses based on apparent motion (e.g., Ullman, 1977, 1979) or instantaneous visual flow (e.g., Lee, 1974; Longuet-Higgins & Prazdny, 1980), the primitive units are fixed points or motion vectors defined over an infinitesimal region of space-time—a fine scale of observation. The difference between these two approaches to the analysis of visual information is analogous to the difference between looking at an object using normal vision and looking at the same object with the aid of an electron microscope. The two views of the same object may be so different that they bear no recognizable relationship to one another (cf. Mandelbrot, 1977).

It is interesting to note that trajectory-based analyses of visual information have been devised independently by several researchers in the field of human motion perception. Indeed, the critical insight that elliptical and trochoidal trajectories are appropriate units of analysis for the study of visual information was first suggested by Johansson (1973, 1974a, 1974b) and Eriksen (1974). This type of analysis has been applied successfully to a variety of different phenomena. For example, Johansson (1973, 1975, 1976) has analyzed the complex pattern of motion in a point-light walker display as a nested set of pendular trajectories. Cutting (1978) has extended this analysis by demonstrating that the sex of a point-light walker can be specified by a geometric relation between the projected elliptical trajectories of the shoulder and hip within a moving frame of reference (see Barclay, Cutting, & Kozlowski, 1978; Cutting et al., 1978). A similar approach has also been adopted by Restle (1979). His coding theory, which is based on the parameters for generating an elliptical trajectory, can accurately predict how a set of moving elements will be perceived within hierarchically nested frames of reference. The research described in the present article is a natural extension of these earlier investigations. Its primary contribution is to describe how the geometric relations among a set of projected trajectories can be used to specify the three-dimen-

sional structure of an object and to define perceptually meaningful categories of motion such as rigid and nonrigid.

References

- Balch, W., & Shaw, R. E. The role of perceptual organization in the depth perception of kinetic lattice displays. *Perception & Psychophysics*, 1978, 23, 493-498.
- Barclay, C. D., Cutting, J. E., & Kozlowski, L. T. Temporal and spatial factors in gait perception that influence gender recognition. *Perception & Psychophysics*, 1978, 23, 145-152.
- Bassili, J. N. Temporal and spatial contingencies in the perception of social events. *Journal of Personality and Social Psychology*, 1976, 33, 680-685.
- Bassili, J. N. Facial motion in the perception of faces and emotional expression. *Journal of Experimental Psychology: Human Perception and Performance*, 1978, 4, 373-379.
- Borjesson, E., & von Hofsten, C. Spatial determinants of depth perception in two-dot motion patterns. *Perception & Psychophysics*, 1972, 11, 263-268.
- Borjesson, E., & von Hofsten, C. Visual perception of motion in depth: Application of a vector model to three-dot motion patterns. *Perception & Psychophysics*, 1973, 13, 203-208.
- Cutting, J. E. Generation of synthetic male and female walkers through manipulation of a biomechanical invariant. *Perception*, 1978, 7, 393-405.
- Cutting, J. E., Proffitt, D. R., & Kozlowski, L. T. A biomechanical invariant for gait perception. *Journal of Experimental Psychology: Human Perception and Performance*, 1978, 4, 357-372.
- Duncker, K. Induced motion. In W. D. Ellis (Ed.), *A sourcebook of gestalt psychology*. London: Routledge & Kegan Paul, 1937. (Originally published, 1929.)
- Eriksson, S. E. A theory of veridical space perception. *Scandinavian Journal of Psychology*, 1974, 15, 225-235.
- Gibson, J. J., & Gibson, E. J. Continuous perspective transformations and the perception of rigid motion. *Journal of Experimental Psychology*, 1957, 54, 129-138.
- Gibson, J. J., & Gibson, E. J., Smith, O. W., & Flock, H. Motion parallax as a determinant of perceived depth. *Journal of Experimental Psychology*, 1959, 58, 40-51.
- Green, B. F. Figure coherence in the kinetic depth effect. *Journal of Experimental Psychology*, 1961, 62, 272-282.
- Heider, F., & Simmel, M. An experimental study of apparent behavior. *American Journal of Psychology*, 1944, 57, 243-259.
- Jansson, G. Perceived bending and stretching motions from a line of points. *Scandinavian Journal of Psychology*, 1977, 18, 209-215.
- Jansson, G., & Borjesson, E. Perceived direction of rotary motion. *Perception & Psychophysics*, 1969, 6, 19-26.
- Jansson, G., & Johansson, G. Visual perception of bending motion. *Perception*, 1973, 2, 321-326.
- Jansson, G., & Runeson, S. Perceived bending motion from a quadrangle changing form. *Perception*, 1977, 6, 595-600.
- Johansson, G. *Configurations in event perception*. Uppsala, Sweden: Almqvist and Wiksell, 1950.
- Johansson, G. Rigidity, stability, and motion in space. *Acta Psychologica*, 1958, 13, 359-370.
- Johansson, G. Perception of motion and changing form. *Scandinavian Journal of Psychology*, 1964, 5, 181-208.
- Johansson, G. Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 1973, 14, 201-211.
- Johansson, G. Vector analysis in visual perception of rolling motion: A quantitative approach. *Psychologische Forschung*, 1974, 36, 311-319. (a)
- Johansson, G. Visual perception of rotary motion as transformations of conic sections—A contribution to the theory of visual space perception. *Psychologia*, 1974, 17, 226-237. (b)
- Johansson, G. Visual motion perception. *Scientific American*, 1975, 232(6), 76-88.
- Johansson, G. Spatio-temporal differentiation and integration in visual motion perception. *Psychological Research*, 1976, 38, 379-393.
- Johansson, G., & Jansson, G. Perceived rotary motion from changes in a straight line. *Perception & Psychophysics*, 1968, 4, 165-170.
- Kozlowski, L. T., & Cutting, J. E. Recognizing the sex of the walker from a dynamic point-light display. *Perception & Psychophysics*, 1977, 21, 575-580.
- Lappin, J. S., Doner, J. F., & Kottas, B. Minimal conditions for the visual detection of structure and motion in three dimensions. *Science*, 1980, 209, 717-719.
- Lee, D. N. Visual information during locomotion. In R. B. MacLeod & H. Pick (Eds.), *Perception: Essays in honor of James Gibson*. Ithaca, N.Y.: Cornell University Press, 1974.
- Lee, D. N., & Aronson, E. Visual proprioceptive control of standing in human infants. *Perception & Psychophysics*, 1974, 15, 529-532.
- Lee, D. N., & Lishman, R. Visual control of locomotion. *Scandinavian Journal of Psychology*, 1975, 1, 87-95.
- Lishman, R., & Lee, D. N. The autonomy of visual kinesthesia. *Perception*, 1973, 2, 287-294.
- Lockwood, E. H. *A book of curves*. Cambridge, England: Cambridge University Press, 1961.
- Longuet-Higgins, H. C., & Prazdny, K. The interpretation of a moving retinal image. *Proceedings of the Royal Society of London*, 1980, 208, 385-397.
- Mace, W. M., & Shaw, R. E. Simple kinetic information for transparent depth. *Perception & Psychophysics*, 1974, 15, 201-209.
- Mandelbrot, B. B. *Fractals: Form, chance, and dimension*. San Francisco: Freeman, 1977.
- Metzger, W. Tiefenercheinungen in optischen bewegungsfeldern. *Psychologische Forschung*, 1934, 20, 195-260.
- Proffitt, D. R., Cutting, J. E., & Stier, D. M. Perception of wheel-generated motions. *Journal of Experimental Psychology: Human Perception and Performance*, 1979, 5, 289-302.
- Restle, F. Coding theory and the perception of motion configurations. *Psychological Review*, 1979, 86, 1-24.
- Rubin, E. Visuell wahrgenommene wirkliche Bewegungen. *Zeitschrift für Psychologie*, 1927, 103, 384-392.

Schiff, W. Perception of impending collision. *Psychological Monographs*, 1965, 79(11, whole No. 604).
 Ullman, S. The interpretation of visual motion. Unpublished doctoral dissertation, Massachusetts Institute of Technology, 1977.
 Ullman, S. *The interpretation of visual motion*. Cambridge, Mass.: MIT Press, 1979.
 von Fieandt, K., & Gibson, J. J. The sensitivity of the eye to two kinds of continuous transformation of a shadow-pattern. *Journal of Experimental Psychology*, 1959, 57, 344-347.

Wallach, H., & O'Connell, D. N. The kinetic depth effect. *Journal of Experimental Psychology*, 1953, 45, 205-217.
 Warren, R. The perception of ego motion. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, 2, 448-456.
 White, B., & Mueser, G. Accuracy of reconstructing the arrangement of elements generating kinetic depth displays. *Journal of Experimental Psychology*, 1960, 60, 1-11.

Appendix

A Trajectory-Based Analysis of Structure and Rigidity

The analysis of this section will consider a point, P1, rotating at a constant angular velocity about an arbitrary axis in three-dimensional space; an arbitrary point of observation, P0; and an arbitrary projection plane a unit distance from P0. Let P2 be the center of the circular trajectory formed by the motion of P1. We will construct a coordinate system such that the XY plane passes through P2 and is parallel to the projection surface. As shown in Figures A1 and A2, the angle of intersection between the circular trajectory of P1 and the XY plane is designated by λ , and the two points of intersection are designated P3 and P4, respectively. (This assumes that $\lambda \neq 0$.) The Z axis is determined by a line passing through P0 that is perpendicular to the XY plane, and the Y axis is oriented so that it is perpendicular to the line connecting P3 and P4. Additional parameters for defining the position of P1 at any given mo-

ment include the distance R between P1 and P2, the distance X between P2 and the Y axis, the distance Y between P2 and the X axis, the distance Z between P0 and the origin, the frequency ω and the phase angle α between $\overline{P1P2}$ and the Y axis at some arbitrary moment in time, t.

Now consider the image of this event on the projection plane (see Figure A3). As is demonstrated by Eriksson (1974), the image of P1 (designated by P1') will gradually trace out an elliptical trajectory whose minor axis lies along the projection of the axis of rotation. Let (t0) be the moment when P1' crosses the top of the minor axis. Since time (t) is not affected by the projection, frequency and phase are uniquely specified by the following equations:

$$\omega = \omega' \tag{A1}$$

$$\alpha = \omega't, \tag{A2}$$

where ω' is the frequency of motion on the pro-

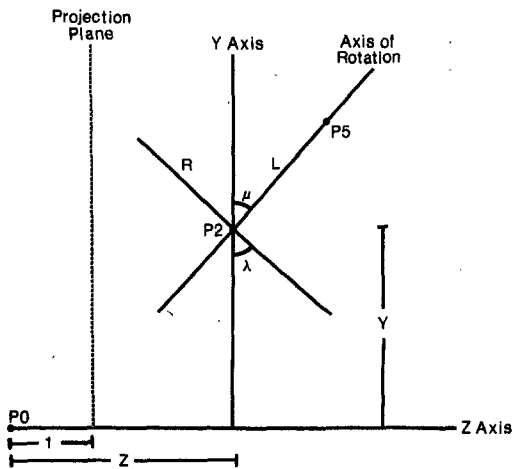


Figure A1. The side view of a circular trajectory in three-dimensional space at an angle to the projection plane.

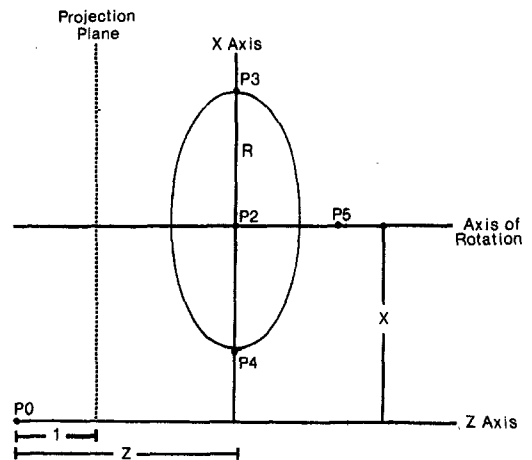


Figure A2. The top view of a circular trajectory in three-dimensional space at an angle to the projection plane.

jection plane. It follows from Equations A1 and A2 that P3' is the position of P1' when $\omega't = \pi/4$; P4' is the position of P1' when $\omega't = 3\pi/4$; and P2' lies midway between P3' and P4'. Additional variables measured on the projection plane include the distance, a , between P2' and the top of the minor axis; the distance, b , between P2' and the bottom of the minor axis; and the distance, c , between P2' and P4'. As shown in Figure A3, the method of projection is uniquely specified by the relation between a and b : If $a = b$, then it is a parallel projection; if $a \neq b$, then it is a polar projection.

Under parallel projection, the trajectory parameters R and λ are optically specified by the following equations:

$$R = c \tag{A3}$$

$$\lambda = \arccos a/c \tag{A4}$$

(Note that λ is only specified up to a reflection.) These same parameters are also specified under polar projection, although the geometric relationships are slightly more complicated.

$$a = (Y + R \cos \lambda) / (Z - R \sin \lambda) - Y/Z$$

$$a = (R \cos \lambda + (YR \sin \lambda) / Z) / (Z - R \sin \lambda) \tag{A5}$$

$$b = Y/Z - (Y - R \cos \lambda) / (Z + R \sin \lambda)$$

$$b = (R \cos \lambda + (YR \sin \lambda) / Z) / (Z + R \sin \lambda) \tag{A6}$$

$$c = (R + X) / Z - X/Z$$

$$c = R/Z \tag{A7}$$

Thus, Equation A7 demonstrates that R is specified within a scale factor. The value of λ is determined by forming a ratio from Equations A5 and A6 and substituting Equation A7 where appropriate to obtain $b/a = (1 - c \sin \lambda) / (1 + c \sin \lambda)$. Finally, by rearranging terms we get

$$\lambda = \arcsin ([1 - b/a] / [c + bc/a]). \tag{A8}$$

(In this case, λ is specified uniquely, since the point on the projected trajectory that is closest to P2' will always correspond to the point on the three-dimensional trajectory that is farthest away in depth.)

To summarize briefly, the foregoing analysis has shown that under parallel projection R , ω and α are specified uniquely, and λ is specified up to a reflection. Under polar projection, ω , α and λ are specified uniquely, and R is specified within a scale factor. In either case the orientation of the

axis of rotation in the projection plane is uniquely determined by the minor axis of the elliptical trajectory of P1'.

We now have enough information to decide whether any two points are rigidly rotating about a fixed axis, based solely on their optic projections. To satisfy the definition of rigidity, they must share the same axis of rotation and be rotating at the same frequency. The relative values of R and α need not be the same. Moreover, all that is necessary to define the three-dimensional structure of the two points is to specify the distance L that separates them along the axis of rotation (see Figure A1).

With this in mind, we will define a point P5 that is L units away from P2 along the axis of rotation. (P5 is assumed to be the center of a circular trajectory.) Let P5' be the image of P5 on the projection plane and let d be the distance between P2' and P5' (see Figure A3). As is demonstrated in Figure A1, $\mu = \pi/2 - \lambda$. For parallel projection, L is specified simply by

$$L = d / \cos \mu. \tag{A9}$$

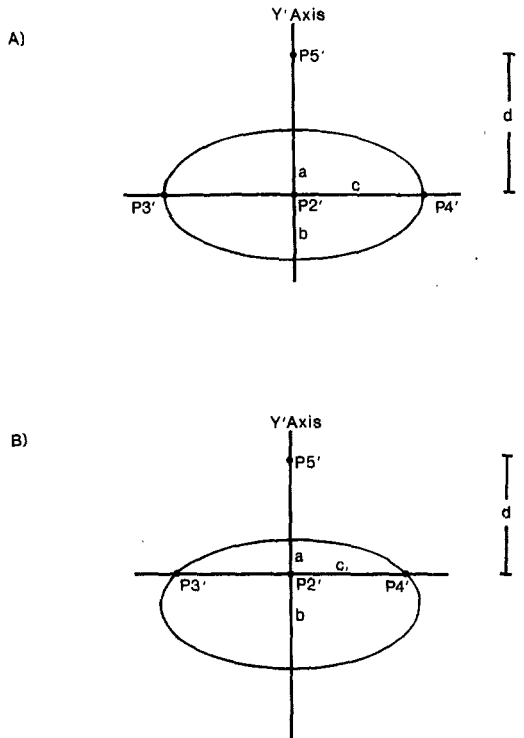


Figure A3. The image trajectory of an object rotating about a fixed axis under parallel projection (A) and polar projection (B).

For polar projection the specification of L is considerably more complicated. We begin from Figure A1 with $D = (Y + L \cos \mu)/(Z + L \sin \mu) - Y/Z$. Rearranging terms we get

$$L/Z = d/(\sin \mu [\cot \mu - d - Y/Z]). \quad (\text{A10})$$

To find the ratio Y/Z , we rearrange the terms of Equation A5 and substitute Equation 7 where appropriate to obtain

$$Y/Z = (a/c \sin \lambda) - a - \cot \lambda. \quad (\text{A11})$$

By substituting Equation A11 into Equation A10, we find that L is specified within a scale factor.

Thus it is demonstrated that when two points rotate about a fixed axis their rigidity and three-dimensional structure are specified by their projected trajectories in the optic array. Under parallel projection the structure is specified up to a reflection. Under polar projection, the structure is specified within a scale factor. It is important to point out, moreover, that the necessary conditions for carrying out this analysis are all optically specified. Rotation about a fixed axis is specified by the presence of elliptical trajectories in the optic array and the method of projection is specified by the relationship between a and b .

Received March 24, 1981 ■

Correction to Petzold

In the article "Distance Effects on Sequential Dependencies in Categorical Judgments" by P. Petzold (*Journal of Experimental Psychology: Human Perception and Performance*, 1981, Vol. 7, No. 6, pp. 1371-1385), the phrase "negative distance effect" should be replaced by the phrase "inverse distance effect" in two places on page 1377. The last sentence of the first complete paragraph should read, "Using the mean judgments R , the inverse distance effect can be described by the differences. . ." The last sentence on the page should read, "This result verifies the existence of the inverse distance effect in the case of sequential contrast produced by previous stimuli (*Property 6*)."

In addition, a change should be made to an equation on page 1378. The first equation in the second column of the page should begin with a lowercase p .

A final error: The texts of the captions for Figures 10 and 12 on pages 1381 and 1382 were inadvertently reversed. The caption of Figure 10 on page 1381 should read, "Differences $d = R(\mu, f_1, \hat{\mu}) - R(\mu, f_2, \hat{\mu})$ as a function of the absolute difference $|\mu - \hat{\mu}/S|$ averaged over several pairs (f_1, f_2) ." The caption of Figure 12 on page 1382 should read, "Conditional probabilities for the response $f = 3$ in dependence on the previous response f for several 'stimuli' $\mu(s)/S$ and for the preceding 'stimulus' $\hat{\mu}(s)/S = 5$."