The Shape of Smooth Objects and the Way Contours End

Jan J. Koenderink & Andrea J. van Doorn

Rijksuniversiteit Utrecht, The Netherlands

1.0 Shape of Smooth Objects in Mathematics, Perception, and the Theories of Academic Art

Geometrically a smooth object can be defined as a connected, bounded portion of space, such that the boundary possesses a unique tangent plane everywhere. More generally, an object may be smooth except for a finite number of points (e.g. the apex of the right circular cone, the corners of a cube) or curves (e.g. the edges of a cube). If we take a sufficiently small surface patch (so that it is 'of one shape'), then what are its possible shapes? Quite different answers are provided by mathematics (differential geometry) and theories of academic art—and the later no doubt voices convictions gained by visual perception.

In mathematics the basic dichotomy is between isoclastic and anticlastic curved patches (also called elliptic and hyperbolic patches), as shown in Figure 1 (Hilbert & Cohn-Vossen, 1932). The importance of this basic division was clearly outlined by Gauss in his classic paper *Disquisitiones generales circa superficies curvas* (1827) although the elementary facts were already known at that time (e.g. Meusnier's law). In the case of an elliptic patch the object lies completely on one side of the tangent plane, whereas in the case of a hyperbolic patch the surface cuts the tangent plane. Thus elliptic patches 'enclose' space: the outside of an egg is everywhere an elliptic surface ('ovoid'), but so is the *inside* of an eggshell. Both the inside and the outside are positively curved (elliptic), although one speaks of a convex patch in the former case and of a concave patch in the latter case. Thus, elliptic patches bound either material (convex) or pockets of air (concave). On the other hand the hyperbolic patch cannot enclose anything: as shown in Figure 1, it is a saddle-shaped surface. (There exist no smooth objects that are completely enclosed in a hyperbolic 'skin': there *must* be elliptic patches, or else the surface contains creases.)

In general, the surface of a smooth object can be divided into elliptic and hyperbolic areas. The dividing curves are called parabolic lines. (On such a line the surface is of a cylindrical nature.) The parabolic lines are generically a family of nested, closed curves, which outline the elliptic patches or 'object-like' (in the sense of space-enclosing) parts: bulges of material and pockets of air.

In both theory and practice of plastic art the basic dichotomy is not that of elliptic versus hyperbolic, but of convex versus concave (both elliptic!) whereas the hyperbolic patch is scarcely noticed. We present some examples from both theory and practice.

Perhaps the earliest well articulated text on European academic art theory that treats the subject in depth was writ-

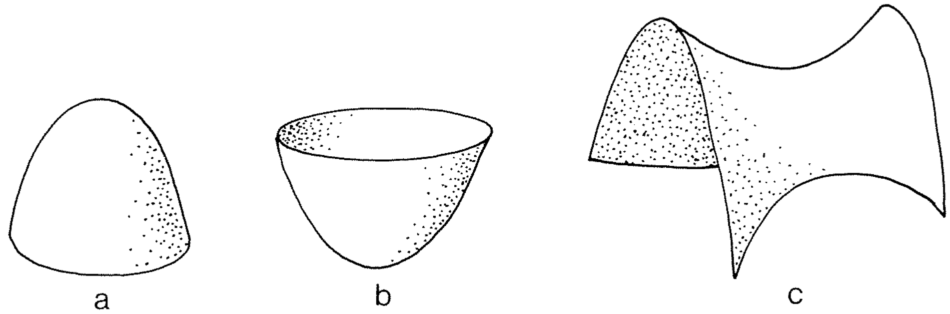

Figure 1 (a) A convex, elliptic patch; (b) a concave, elliptic patch—a bowl-like shape; (c) a hyperbolic patch—or saddle shape.

ten by Leon Battista Alberti (1435). His summing up of the possible forms that a surface patch can assume reads as follows:

"We have now to treat of other qualities which rest like a skin over all the surface of the plane. These are divided into three sorts. Some planes are flat, others are hollowed out, and others are swollen outward and are spherical. To these a fourth may be added which is composed of any two of the above. The flat plane is that which a straight ruler will touch in every part if drawn over it. The surface of the water is very similar to this. The spherical plane is similar to the exterior of a sphere. We say the sphere is a round body, continuous in every pat; any part on the extremity of this body is equidistant from its center. The hollowed plane is within and under the uttermost extremities of the spherical plane as in the interior of an eggshell. The compound plane is in one part flat and in another hollowed or spherical like those of the interior of reeds or on the exterior of columns."

Remarkably, the possibility of the hyperbolic patch is not included. This list has been repeated through the centuries up to our time. For instance, in the important book of Kurt Badt (1963) on plastic art we read (our translation):

"There are three basic possibilities: plane, convex, and concave curvatures. Of these three only the convex is—as a sign and expression of the visibly surging vital force—fundamentally plastic."

Thus, little has changed a century and a half after Gauss's paper!

Experimental psychologists seem to be better acquainted with the theory of art than with mathematics. For instance, Arnheim (1956) in a paragraph on plastic form discusses convex patches but completely neglects the hyperbolic case, whilst Gibson (1979) discusses under 'surface geometry' the space-enclosing properties of convex and concave patches, and also does not mention other possibilities.

In the practice of plastic art we encounter the same phenomenon. Delacroix (quoted in Rawson, 1969) once remarked that you can spot a piece of genuine classic sculpture among pieces of renaissance art by the fact that the ancients grasped form 'par des mileux'. He must have meant the quatrocento, because from the early sixteenth century onward European artists have done the same thing (see Figure 2): they treated form as a conglomerate of

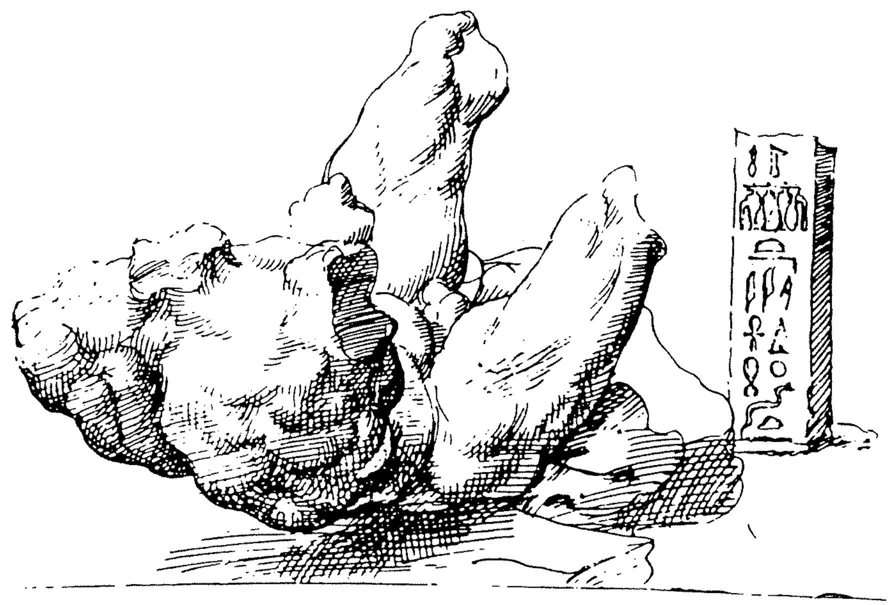

Figure 2 An example of a conglomerate of ovoids. Maarten van Heemskerk: Torso.

ovoid shapes. In its crudest form this leads to the treatment of a human body as a 'bag of melons' [cf. Cellini's (1728) famous derogatory description in his *Vita* of a piece by Bandinelli]. In drawing, the torso acquires a knotty appearance. The draftsman indicates the ovoids by strongly curved lines (Rawson, 1969), mostly in pairs with the concave sides facing each other, and the object is depicted by a concatenation of a great many of such pairs indicating bulges next to bulges and inside bulges, etc. In fact only the elliptic patches are indicated. The sculptor does the same: bulges border on bulges in such a way that the hyperbolic part in between dwindles to a V-shaped furrow, as in Figure 3. Thus art theory and practice go hand in hand (Badt, 1963).

In naturalistic art it does not matter how you deform the hyperbolic parts as long as the elliptic patches are sufficiently articulated. Thus the elliptic parts have a 'figure'-like (thing-like), the hyperbolic parts a 'ground'-like (no-thing-like), character. The hyperbolic patches are perceptually only regions of transition—the 'glue' that keeps things together but is of itself of little interest.

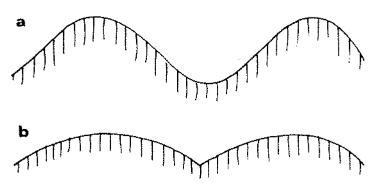

Figure 3 Schematic sections through an undulating surface (a) and its common expression in plastic art (b).

Perhaps the mathematician Felix Klein had the same idea when he sought the key for aesthetics in plastic art in the pattern of parabolic lines. We cannot be sure because Klein published no material on this subject. But from a secondary source (Hilbert & Cohn-Vossen, 1932) we know that he had the parabolic lines drawn on a bust of the Apollo of Belvedere for this reason (Figure 4). (This bust survived the wars: we succeeded in tracing it to the Institute of Mathematics of the Göttingen University.) The result is not particularly enlightening and it seems that Klein gave up this idea. It is our aim to show here that the role of the parabolic lines in the perception and depiction of solid shape is nevertheless of key importance.

2.0 The Contours of the Projection of Objects

It is not possible to see the entire surface of an opaque smooth object simultaneously: parts of the surface are necessarily occluded by the object itself. The curve on the object that divides potentially visible from nonvisible parts is called the *rim* in this paper. By 'potential visibility' we mean visibility except for 'contingencies' like occlusion by a second body or distant parts of the same body. It helps to think of the body as made of 'tinted air'. Then you can see through its volume and 'contingencies' do not count any more. The rim is then the locus of points where the visual direction grazes the surface of the body. The rim consists of a family of smooth loops on the surface. The border of the object's projection in the visual field is called its *contour*. Like the rim, the contour is a closed curve (possibly a nested collection of loops). However, the contour is not necessarily smooth everywhere. Remember that the contour is in the visual field (the space of all visual directions) whereas the rim is on the object. In practice only part of the contour is visible (except when the body is made of tinted air), because of interposition. In a drawing, the body is demarcated by its *outline*. Thus rim, contour and outline are different concepts in the context of this paper and should not be confused. The contour is the visible projection of the rim. The outline is a subjective expression of the visible part of the contour and what the draftsman knows and feels about the object. If the draftsman aims at 'photographic veridicality' we would expect a similarity between outline and the visible part of the contour.

Figure 4 Bust of Apollo.

In the case of an ovoid—an object bounded by an elliptic surface only—the contour is a single, smooth, closed curve. In such cases the visible contour has no end. In the general case the visible part of the contour *can* end. It can be shown (see Appendix) by mathematical means that visible contours always end

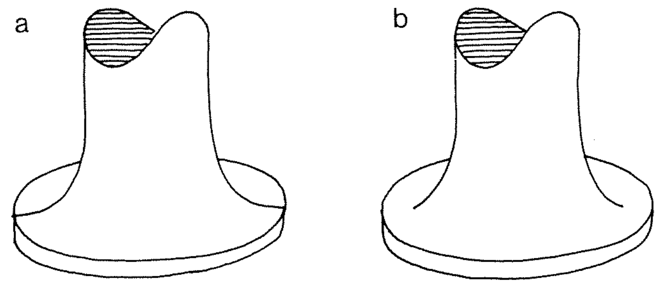

Figure 5 Wrong (a) and right (b) depiction of the contours of a pillar base.

in a simple, lawful manner. The visible contour must be concave at its end points, and the corresponding part of the rim must lie inside a hyperbolic patch. This simple natural law seems to be essentially unknown. Even in texts that aim at a didactic exposition of types of contours the fact is not only not noticed, but there are even illustrations with contours that end in an impossible way (Kennedy, 1974).

Most draftsmen seem also reluctant to concede this natural law: in fact most contours in drawing end in a *convex* way. In cases where a convex ending would show up strangely the artist often devises devious ways to avoid the ending contour. The base of the cylindrical (not fluted) pillar (Figure 5) is a nice case in point: the artist can avoid the ending contour by taking the contour to the edge of the base. This oddity is quite common.[1]

3.0 What Do People Draw?

We have already offered the hypothesis that vision grasps a shape as a conglomerate (or hierarchical structure) of pot-like space enclosures, much like a bunch of grapes. This was exemplified by the case of the 'ovoid-method' in drawing and sculpture. How could one define such ovoids geometrically? Here a simple method presents itself.

Clearly the ovoid is an isolated elliptical patch. On the body it is demarcated by a closed parabolic line. In the visual field the ovoid is demarcated by the projections of the visible portion of this parabolic line and the visible portion of the rim where the rim traverses the elliptic patch. Thus the contour is *not* necessarily present, nor is the projection of the parabolic line (see Figure 7). If both are simultaneously present, then the contour and the projection of the parabolic line join smoothly, without a 'corner'. Thus the boundary of the elliptic patch in the visual field is a closed, smooth loop that consists of the projection of the parabolic line, the contour, or both. Presumably the draftsman indicates as much of this loop as is needed to suggest the ovoid.

Thus in naive drawings the female breast viewed *en face* is often completely encircled (thus stressing its 'thing-like' nature), whereas a photograph would often show no contour at all; in the case where there is no contour the draftsman indicates the projection of the parabolic

[1] A good example is provided by Dürer's woodcut from the series *Seventeen Cuts from the Life of the Virgin*, 1505 (cat. no. Bartsch 88). 10

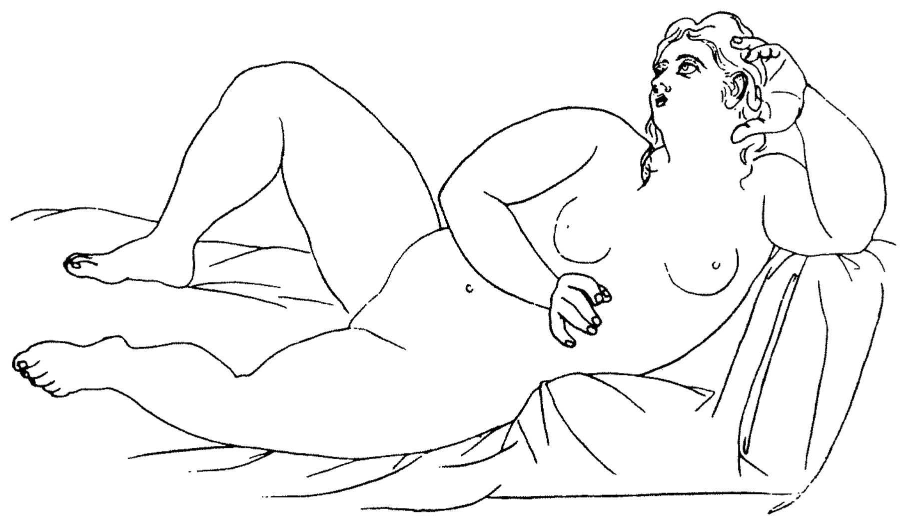

Figure 6 Picasso: *Nu couché* 1920.

line. In three-quarter view part of the contour and part of the projection of the parabolic line are drawn.

The importance of the parabolic line for visual perception and depiction is also clear from yet another consideration. At the parabolic line one of the principal curvatures of the surface changes from convex to concave. As a result, the angle of a fixed direction (e.g. the principal direction of light) with the tangent plane reaches an extreme value at the parabolic line (see Figure 8). Thus it is understandable, and it can be proved mathematically (Koenderink & van Doorn, 1980), that the extrema of the illumination, the highlights and shadows, cling to the parabolic lines.

Consequently the art of shadowing (chiaroscuro) also necessarily stresses the parabolic lines. The drawing of parabolic lines may even be regarded as a rudimentary way of shadowing: some study reveals that the draftsman tends to draw the parabolic line on the shadow side, and leave it out on the light side.

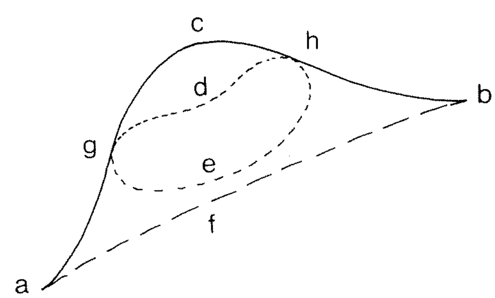

Figure 7 Typical example of an elliptic bulge in three-quarter view: acb, visible part of contour; afb, invisible part of contour; geh, visible part of projection of parabolic line; gdh, invisible part of projection of parabolic line. The 'ovoid' would be bounded in projection by the loop $gche$.

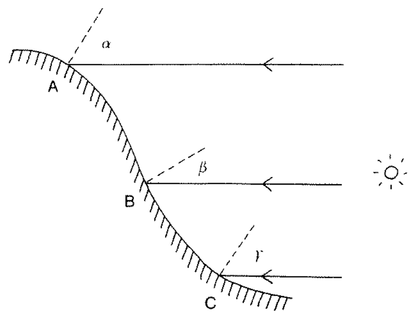

Figure 8 In a section of the object, B is at a parabolic line. The angle of a fixed direction (e.g. of sunlight) with the normal direction to the surface reaches an extreme value at B: $\alpha > \beta$ and $\gamma > \beta$.

4.0 Discussion

Felix Klein was certainly right when he suspected the importance of the parabolic lines for our visual appreciation of shape. He merely sought in the wrong direction, probably expecting to find some infallible rule for the placement of these curves as a recipe for beauty. Such a rule—like the Golden Section—would fall outside the realm of science. But a strong case can be made for the hypothesis that vision grasps shape as a hierarchical structure of elliptic patches, and this hierarchy is identical with that of the family of nested parabolic loops.[2]

It is a sobering thought that in the scientific study of shape perception the same ignorance of the basic geometrical facts is common: here the stimulus is very much defined in terms of the response.

References

Alberti, L.B. (1435) *Della Pittura.* English translation by J.R. Spencer *On Paint-*

[2]A referee pointed out to us that we should draw no conclusions about how people *see* things from how they *draw* things. We believe this to be a subjective standpoint. Since the point is debatable, we add here a few lines to make our standpoint.

Both draftsmen and sculptors in many European and Asiatic stylistic periods build their shapes from 'ovoids' (in academic practice), that is elliptically bounded pot-like volumes. This is seen most directly from the exclusion of certain facts of nature from artistic artifacts, such as the concavely ending contour in drawing and the direct juxtaposition by way of V-shaped furrows of bulges in sculpture. The fact that academic theory neglects hyperbolic shape and that all talk about 'positive' and 'negative' shapes is only about convex and concave *elliptic* patches, corroborates academic practice.

The referee's standpoint is common, perhaps originating with Michelangelo who is reported to have said that he drew with his head (*col cervello*) instead of with his hands (*con la mano*). By way of writers on art (like Zuccaro, Bellori) the problem attracted philosophers and it became known as the problem of 'Raphael without hands' (Lessing). We believe that the philosopher of art Conrad Fiedler (1887) definitively demolished this fiction. The interested reader should consult his lucid writings, as we can only state the argument in too few lines here. Suppose someone manifests himself to you through generally incoherent utterances. You would certainly consider his discursive thinking to be defective. Now suppose someone is only able to offer you incoherent drawings: we would say that his visual thinking deficient! (Remember that you cannot put your visual experience into words!)

ing. London: Routledge & Kegan Paul (1956).

Arnheim, R. (1956) *Art and Visual Perception.* London: Faber & Faber.

Badt, K. (1963) Wesen der Plastik. In *Raumphantasien und Raumillusionen.* Köln: M. DuMont Schauberg.

Cellini, B. (1728) *La Vita.* First edition. Naples: Antonio Cocchi. There have been many subsequent editions in most European languages.

Fiedler, C. (1887) *Der Ursrung der künstlerischen Tätigkeit.* reprinted in 1977 *Schriften über Kunst* (with an introduction by H. Eckstein). Köln: DuMont Buchverlag.

Gauss, C.F. (1827) *Disquisitiones generales circa superficies curvas.* Translated into English *General Investigation of Curved Surfaces.* New York: Raven Press (1965).

Gibson, J.J. (1979) *The Ecological Approach to Visual Perception.* Boston: Houghton Mifflin.

Hilbert, D. & Cohn-Vossen, S. (1932) *Anschauliche Geometrie.* Berlin: Springer.

Kennedy, J.M. (1974) Icons and information. In *Media and Symbols: The Forms of Expression, Communication, and Education*, Seventy-third Yearbook of the National Society for the Study of Education. Chicago, Ill.: The University of Chicago Press, chap. 9, pp. 211 - 240.

Koenderink, J.J. & van Doorn, A.J. (1980) Photometric invariants related to solid shape. *Optica Acta*, 27:981 - 996.

Rawson, P. (1969) *Drawing.* London: Oxford University Press.

Appendix: The Ending Contour

For initiates of the catastrophe theory or differential geometry a very general and elegant proof (without formulae!) consisting of only a few sentences would suffice.[3] Such knowledge is not assumed here and a 'pedestrian' proof is offered. This cannot be done without some formulae. Figure A1 illustrates the steps.

Consider Cartesian coordinates (x, y, z). Let the points $(x, y, F(x, y))$ describe the points of the surface, such that the 'material' is at points $z \leq F(x, y)$ and 'air' at points $z > F(x, y)$. Let the origin be a point of the surface, and let the $x - y$ plane be a tangent plane at this point. The eye is thought to view the object from a large distance (effectively from infinity), say from $x = +\infty$, $y = z = 0$. It is always possible to choose the axes such that this situation applies.

In the neighborhood of the origin $F(x, y)$ is approximated by its Taylor series. Terms up to the third order have to be retained for the description of an ending contour. No generality is lost by neglecting higher-order terms. Thus, set

$$F(x,y) = \frac{1}{2}(ax^2 + 2bxy + cy^2) + \frac{1}{6}(ex^3 + 3fx^2y + 3gxy^2 + hy^3) .$$

The condition for the rim is

[3] The projection of the surface of the body in the visual field has as generic singularities only folds and cusps. The folds are the smooth part of the contour, the cusps belong to points where one looks in an asymptotic direction. Thus only one of the branches is visible and it is concave.

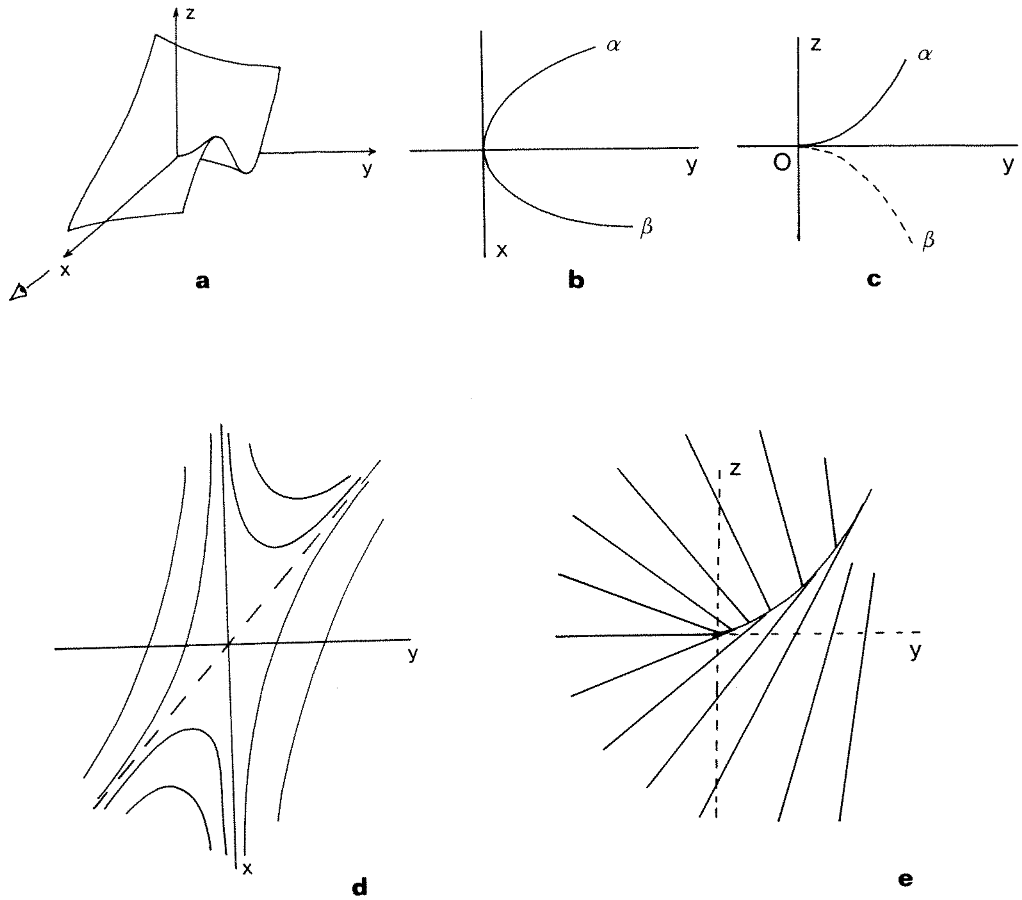

Figure A1 Geometrical properties of a three dimensional object.
(a) General impression of the object in case of an ending contour. In this example the surface is: $F(x, y) = xy + y^2 - \frac{1}{3}x^3$.
(b) The rim projected on the ground plane. The rim is, of course, a space curve. In this example the rim is $y = x^2$.
(c) Projection of the rim on the $y - z$ plane is similar to the contour in the (two-dimensional) visual field. The branch $O\alpha$ is visible, the branch $O\beta$ not. In this example the contour is: $z = \pm\frac{2}{3}y^{3/2}$.
(d) Equal height, $F(x, y) = k$, curves in the $x - y$ plane. The x-axis is an asymptote. In this example we have $y(x + y) = \epsilon$.
(e) Curves of equal distance to the eye projected on the $y - z$ plane (visual field). these curves touch the contour. In general the contour is the envelope of the equidistant curves. In this example the equation of the equidistance lines is $z = x_0 y - \frac{1}{3}x_0^3$.

$$\frac{\partial F}{\partial x} =$$
$$ax + by + \frac{1}{2}(ex^2 + 2fxy + gy^2)$$
$$= 0 \, .$$

This is a parabola. If we stipulate that the contour must end at the origin, and must be visible for $y > 0$ only, then this parabola must be of the form

$$y - kx^2 = 0 \;\; (k = \text{a positive constant}).$$

This yields the following conditions for the coefficients:

$$a = f = g = 0; \;\; e/b < 0 \, .$$

Elimination of x yields the contour, that is the projection of the rim in the $y - z$ plane:

$$z = \pm A y^{3/2}$$
$$\text{where } A = \frac{2}{3} b \left[-\frac{2b}{e} \right]^{\frac{1}{2}}$$

(higher-order terms are irrelevant).

This is a semicubic parabola with a cusp at the origin. Only one of the branches (plus sign) is visible. Thus the contour ends concavely. The level lines of $F(x, y)$ are $F(x, y) = \epsilon$, or

$$bxy + \frac{1}{2}cy^2 +$$
$$\text{higher order terms} = \epsilon.$$

Near the origin the higher-order terms may be neglected, and we obtain

$$y(bx + \frac{1}{2}cy) = \epsilon,$$

a pair of hyperbolae with asymptotes $y = 0$ and $bx + \frac{1}{2}cy = 0$. Thus we reach the following conclusions:

(i) the patch is hyperbolic for the surface lies on both sides of the tangent plane $z = 0$;
(ii) at the ending contour one looks directly in the direction of the asymptote $y = 0$.

We may think of the asymptotes as 'drawn on the surface'; they are fixed whenever the shape is given. The contour must end whenever the rim (a curve that depends on the viewing position) touches such an 'asymptotic line'.

The sketches in Figure A1 clarify the geometry of the ending contour.

Reprinted from *Perception*, 11:129–137 (1982).