The Perspective Representation of Functions of Two Variables

B. KUBERT, J. SZABO, AND S. GIULIERI

Aerospace Corporation, San Bernardino, California*

ABSTRACT. A perspective transformation is developed whereby an object in space as viewed from an arbitrary point can be projected into a plane and plotted. Families of curves which can be used to define such an object are discussed and examples are given. An algorithm which eliminates the plotting of hidden portions of the object is discussed.

KEY WORDS AND PHRASES: perspective, perspective representation, perspective plotting, illustrations, graphics, computer graphics

CR CATEGORIES: 3.20, 3.41, 4.41

1. *Introduction*

There is a well-known Chinese saying that states, "A picture is worth a thousand words." That this proverb holds today in the fields of physical science and engineering is demonstrated in the abundance of illustrations and figures contained in technical papers and reports. Many of these figures depict mathematical functions of one or two variables. Functions of one variable are easily plotted in that the two dimensions of a sheet of paper naturally accommodate the two variables (independent and dependent) associated with such functions. If functions of two variables are to be represented some artifice must be used.

The two most common methods of representing a function of two variables are contour plots and planar projections. Whereas contour plots emphasize quantitative aspects and are certainly most convenient if numerical information must be extracted from the figure, a planar projection depicts the function in such a manner that its properties are most easily visualized. For example, a topographic map, which is a contour plot of the altitude of a portion of the earth's surface, is useful for obtaining altitude differences and slopes but a photograph gives a much better qualitative description of the area.

The utility of such qualitative descriptions of mathematical functions is not restricted to textbooks and technical reports. Their application in connection with the solution of problems in physics and engineering helps the scientist and engineer to obtain insight into the nature of the functions with which he is working.

Published literature concerning computer-generated planar projections has previously dealt with the representation of objects constructed of relatively simple surfaces (i.e. planar or quadric surfaces [1, 2]). This paper is given to a rather simple technique for the representation of arbitrary continuous single-valued functions of two variables of bounded variation. A number of illustrations of the results are

* San Bernardino Operations

Journal of the Association for Computing Machinery, Vol. 15, No. 2, April 1968, pp. 193-204.

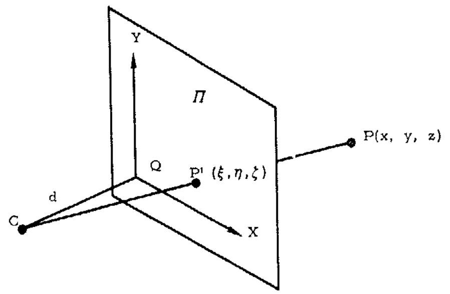


FIG. 1

given. Currently an alternate method of representing such functions is being investigated by Di Leonardo [3]. An example of his work can be found in [4].

2. *Perspective Transformation*

The planar projection that best represents real objects as viewed by the naked eye through a homogeneous medium is perspective. This type of projection maps an arbitrary point of space, P, to a point, P′, on a plane, Π, such that all lines $\overline{PP'}$ intersect in a common point C. The point C is called the center of projection and corresponds to the eye of the observer. The plane, Π, is called the plane of projection and is oriented so that its normal is parallel to the line of sight.

Let the function to be plotted, say $f(x, y)$, be given in terms of a set of rectangular coordinates (x, y, z) where

$$z = f(x, y). \tag{1}$$

(See Figure 1.) Suppose that the observer's eye is situated at the point C with coordinates (c_x, c_y, c_z) referred to the same rectangular coordinate system, and further, that the line of sight makes angles α, β, and γ with the x-, y-, and z-axes, respectively. Let a distance, d, be given and define $Q(q_x, q_y, q_z)$ as a point such that $\overline{CQ}$ is the direction of sight and $|\overline{CQ}| = d$. Let the plane of projection, Π, be constructed through Q and normal to $\overline{CQ}$. The straight line from C to an arbitrary point in space $P(x, y, z)$ will intersect Π in some point $P'(\xi, \eta, \zeta)$. P′ is the perspective image of P in Π with respect to C.

From the geometry it follows that

$$q_x = c_x + d(\cos\alpha), \quad q_y = c_y + d(\cos\beta), \quad q_z = c_z + d(\cos\gamma) \tag{2}$$

and

$$\frac{\xi - c_x}{x - c_x} = \frac{\eta - c_y}{y - c_y} = \frac{\zeta - c_z}{z - c_z} = K, \tag{3}$$

where K is the common ratio and varies with P. By definition

$$\overline{CQ} = d(\cos\alpha\mathbf{i} + \cos\beta\mathbf{j} + \cos\gamma\mathbf{k}). \tag{4}$$

Since P′ lies in the plane Π it must satisfy

$$\overline{CQ}\cdot\overline{QP'} = 0. \tag{5}$$

But $\overline{CQ}\cdot\overline{QP'} = \overline{CQ}\cdot(\overline{QC} + \overline{CP'}) = 0$ so that $\overline{CQ}\cdot\overline{CP'} = \overline{CQ}\cdot\overline{CQ}$ or

$$(\xi - c_x)\cos\alpha + (\eta - c_y)\cos\beta + (\zeta - c_z)\cos\gamma = d. \tag{6}$$

Substitution of eqs. (3) into eq. (6) gives

$$K = d/[(x - c_x)\cos\alpha + (y - c_y)\cos\beta + (z - c_z)\cos\gamma], \tag{7}$$

which, together with eqs. (3), defines P′ as

$$\xi = c_x + K(x - c_x), \qquad \eta = c_y + K(y - c_y), \qquad \zeta = c_z + K(z - c_z). \tag{8}$$

It remains to express P′ in terms of a two-dimensional system of coordinates in Π. The equation of Π (i.e. eq. (5)) can be written

$$(\xi - q_x)\cos\alpha + (\eta - q_y)\cos\beta + (\zeta - q_z)\cos\gamma = 0. \tag{9}$$

The line formed by the intersection of the horizontal plane $\zeta - q_z = 0$ and Π is

$$\frac{\xi - q_x}{\cos\beta} = \frac{\eta - q_y}{-\cos\alpha} = \frac{\zeta - q_z}{0}. \tag{10}$$

One of the new coordinate axes in Π (the X-axis) will be defined along this line.

A unit vector, in the positive X-direction, in terms of the original coordinates, is given by

$$\mathbf{U}_X = S_1[(\cos\beta)\mathbf{i} - (\cos\alpha)\mathbf{j}]/\sin\gamma, \tag{11}$$

where $S_1 = \pm 1$. In order to define the other coordinate axis in Π (the Y-axis) a unit vector, $\mathbf{U}_Y$, in Π must be found such that

$$\mathbf{U}_X\cdot\mathbf{U}_Y = 0. \tag{12}$$

Let

$$\mathbf{U}_Y = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}. \tag{13}$$

Equation (12) requires that

$$a(\cos\beta) - b(\cos\alpha) = 0. \tag{14}$$

Since $\mathbf{U}_Y$ is in Π it must be normal to $\overline{CQ}$. Thus

$$a(\cos\alpha) + b(\cos\beta) + c(\cos\gamma) = 0. \tag{15}$$

Furthermore, since $\mathbf{U}_Y$ is a unit vector,

$$a^2 + b^2 + c^2 = 1. \tag{16}$$

The simultaneous solution of eqs. (14), (15), and (16) yields

$$\begin{aligned} a &= S_2\cos\alpha\cos\gamma/\sin\gamma,\\ b &= S_2\cos\beta\cos\gamma/\sin\gamma,\\ c &= -S_2\sin\gamma, \end{aligned} \tag{17}$$

where $S_2 = \pm 1$.

The signs of S_1 and S_2 will depend on the directions of the positive X- and Y-axes. If the observer is oriented such that up corresponds to an increasing component of z and the usual convention is followed (i.e. positive X to the right, positive Y up), $\mathbf{U}_Y$ must have a positive component in the positive z-direction. Thus,

$$\mathbf{U}_Y \cdot \mathbf{k} > 0. \tag{18}$$

Furthermore, $\mathbf{U}_X \cdot \mathbf{U}_Y$ must be parallel to the line of sight and must point to the observer's side of Π.

$$\mathbf{U}_X \cdot \mathbf{U}_Y = -\frac{1}{d} \cdot \overline{\mathrm{CQ}}. \tag{19}$$

Substitution of eqs. (4), (11), (13), and (17) into the relations (18) and (19) yields

$$S_2 \sin \gamma < 0, \qquad S_1 S_2 = -1.$$

Since $0 < \gamma < \pi$, $\sin \gamma > 0$ and $S_1 = 1$, $S_2 = -1$.

If the new coordinate system has its origin at Q, the components of P′ in terms of the new axes will be $(\overline{\mathrm{QP'}} \cdot \mathbf{U}_X, \overline{\mathrm{QP'}} \cdot \mathbf{U}_Y)$. It follows that

$$\begin{aligned} X &= [(\xi - q_x) \cos \beta - (\eta - q_y) \cos \alpha]/\sin \gamma, \\ Y &= (\zeta - q_z)/\sin \gamma. \end{aligned} \tag{20}$$

To recapitulate: Give the center of projection, (c_x, c_y, c_z), the direction of sight, $(\cos \alpha, \cos \beta, \cos \gamma)$, and the distance of the plane of projection, d, to find the projection of a point in space, $P(x, y, z)$, compute

(1) $q_x = c_x + d(\cos \alpha)$, $q_y = c_y + d(\cos \beta)$, $q_z = c_z + d(\cos \gamma)$.

(2) $K = d/[(x - c_x) \cos \alpha + (y - c_y) \cos \beta + (z - c_z) \cos \gamma]$.

(3) $\xi = c_x + K(x - c_x)$, $\eta = c_y + K(y - c_y)$, $\zeta = c_z + K(z - c_z)$.

(4) $X = [(\xi - q_x) \cos \beta - (\eta - q_y) \cos \alpha]/\sin \gamma$, $Y = (\zeta - q_z)/\sin \gamma$.

Note that step (4) requires $\sin \gamma \neq 0$, the transformation being singular otherwise. This singularity occurs because the two planes whose intersection defines the X-axis are identical. If it is required that the line of sight be vertical (i.e. $\sin \gamma = 0$) the X-axis can be defined as the intersection of Π and the vertical plane $\eta - q_y = 0$. If, furthermore, up corresponds to an increasing y-component, positive X is to the right, and positive Y is up, then the perspective transformation is

$$\begin{aligned} X &= [-(\xi - q_x) \cos \gamma + (\zeta - q_z) \cos \alpha]/\sin \beta, \\ Y &= (\eta - q_y)/\sin \beta. \end{aligned} \tag{21}$$

Equations (21) can be used in place of eqs. (20) for this case.

3. *Generating Curves—Array of Data Points*

From the infinity of points on the surface representing the function in eq. (1), a selection must be made of a subset that will conveniently and effectively represent the surface. Most convenient for this purpose is a two-dimensional rectangular array of (x, y, z) triples, each row and column consisting of points in a space curve on the surface. For example, the (x, y)-coordinates of points in a given column might lie

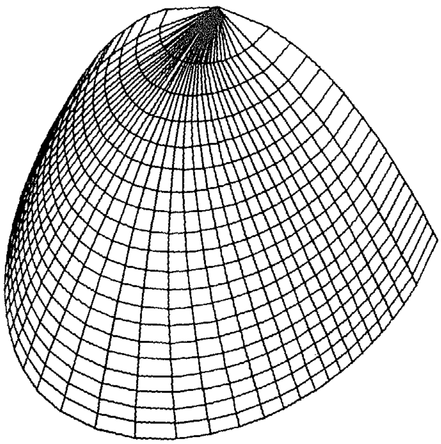


FIG. 2

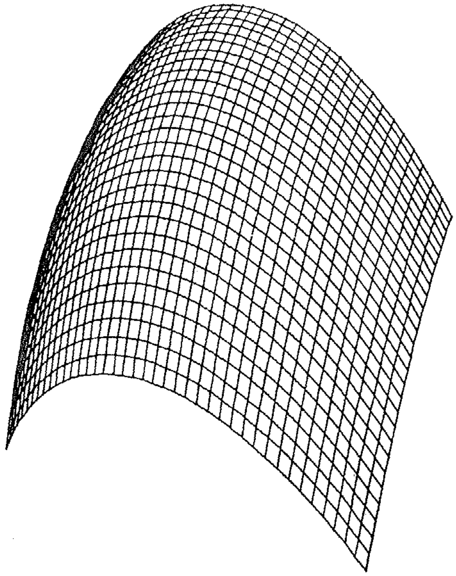


FIG. 3

on the planar curve

$$u = \phi(x, y), \tag{22}$$

where u is fixed for the column but varies from column to column, with the z-coordinate, of course, satisfying eq. (1). Similarly, the points in a given row might satisfy

$$v = \psi(x, y), \tag{23}$$

where v is fixed for the row. Figure 2 is an example of a perspective drawing that utilizes such generating curves. It depicts the function

$$z = -0.5x^2 - y^2 \tag{24}$$

by means of the generating curves

$$u = -0.5x^2 - y^2, \qquad -1 \le u \le 0, \tag{25}$$

$$v = y/x^2, \qquad -90° \le \tan^{-1} v \le 90°. \tag{26}$$

The two families given by eqs. (25) and (26) are orthogonal to one another. This orthogonal property is generally desirable, as perspective views generated from such families tend to be well defined and aesthetically pleasing.

The most simple pair of mutually orthogonal families is

$$u = x, \qquad x_{\min} \le x \le x_{\max}, \tag{27}$$

$$v = y, \qquad y_{\min} \le y \le y_{\max}, \tag{28}$$

with generation of the rectangular array of (x, y, z) triples being trivial for this case. Figure 3 shows the function given by eq. (24) represented in terms of these generating curves for $0 \le x \le 1.4$ and $-1 \le y \le 1$.

Functions that give rise to more complicated pictures are shown in Figures 4, 5,

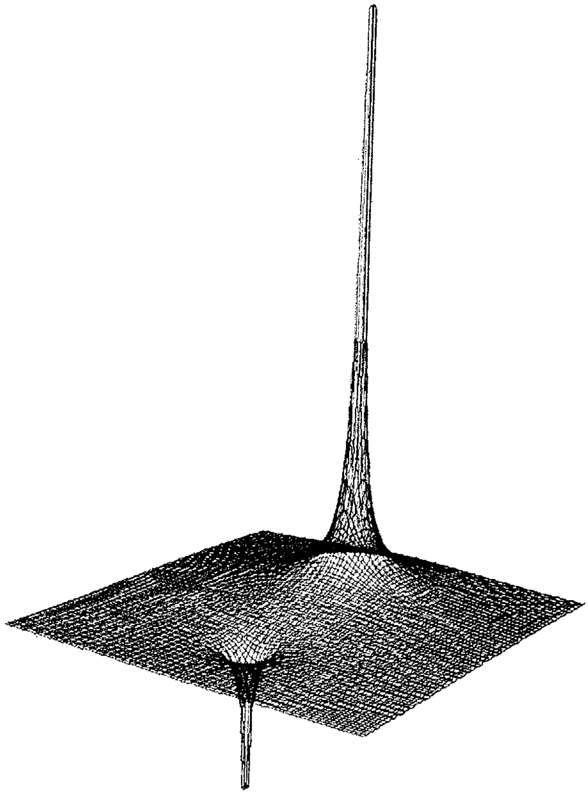


FIG. 4

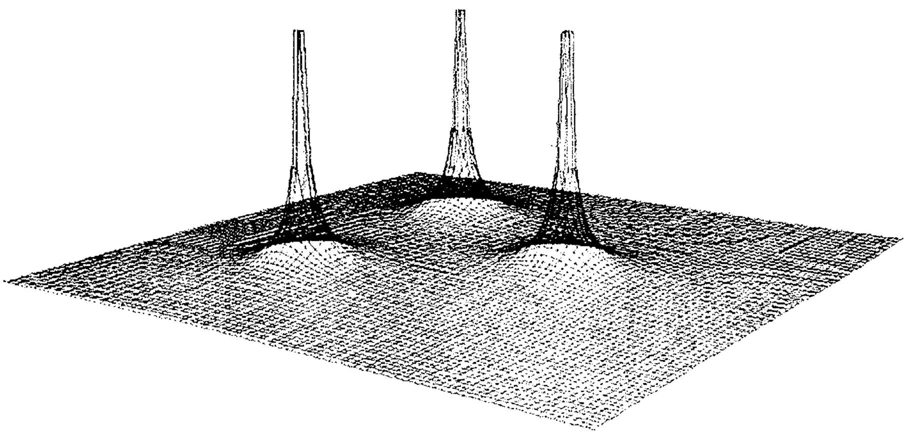


FIG. 5

and 6. Figure 4 represents the function

$$z = 4/\overline{PA} - 1/\overline{PB}, \tag{29}$$

which is the electric potential at a point $P(x, y)$ caused by a positive charge of four units placed at $A(a_x, a_y)$ and a negative charge of one unit at $B(b_x, b_y)$. Figure 5 shows the electric potential at P caused by unit positive charges placed at A, B,

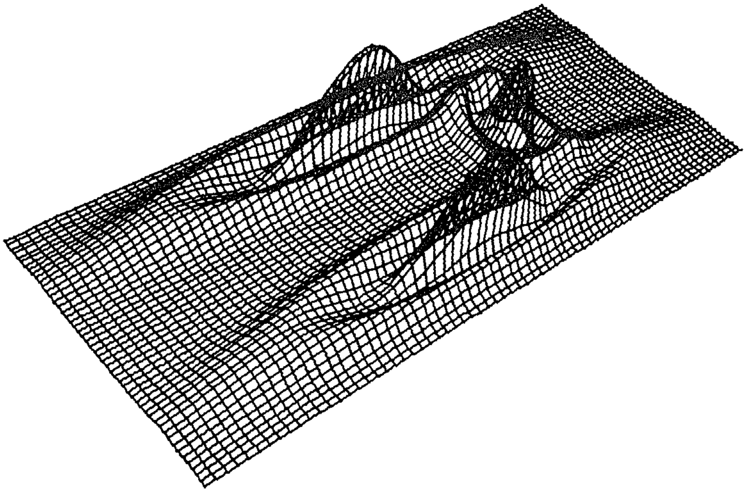


Fig. 6

and C. For this case

$$z = 1/\overline{\mathrm{PA}} + 1/\overline{\mathrm{PB}} + 1/\overline{\mathrm{PC}}. \tag{30}$$

The function shown in Figure 6 is given by

$$z = \sum_{i=1}^{3} w_i^2 e^{-w_i}, \tag{31}$$

where $w_i = (x - h_i)^2/a_i^2 + (y - k_i)^2/b_i^2 - 1$ and the constants are given in the following schedule:

i	a_i	b_i	h_i	k_i
1	$\frac{1}{2}$	1	2	4
2	3	$\frac{1}{2}$	5	2
3	3	$\frac{1}{2}$	5	6

4. *Elimination of Hidden Surfaces—Criterion for Visibility*

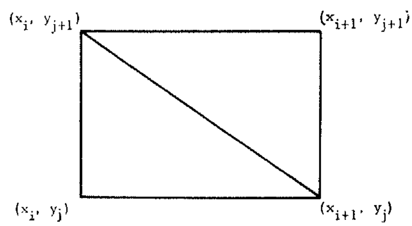
Figures 4, 5, and 6 have the undesirable property that various segments of the surface, S, being represented project into the same segment of Π, the plane of projection. Although this property may be useful in certain applications, in general it causes the picture to be somewhat difficult to interpret. This is especially apparent in Figure 6. The surface representing the function defined in eq. (1) can be thought of as the surface of an opaque solid model of the function. Then, of the various surface segments that project into a single segment of Π, only one, i.e. the closest one, would be visible to the observer. It would be highly desirable to plot only such segments that would be visible.

The outstanding property of a visible point, P, on S is that the line segment $\overline{\mathrm{CP}}$ contains only one point of S, namely P. Conversely, if P were hidden, $\overline{\mathrm{CP}}$ would contain at least one point of S other than P. We utilize this property in determining whether or not to plot P.

Recall that we represent eq. (1) by a rectangular array of points whose x- and y-coordinates satisfy eqs. (22) and (23) along its columns and rows. Let us require that these equations be precisely eqs. (27) and (28).[1] From this finite array of points it will be necessary to construct a function, say $z^* = f^*(x, y)$, that approximates eq. (1) and is defined and continuous throughout the region of interest. In order to make the procedure economically attractive z^* must be obtainable at an arbitrary point with a minimum of computation. This suggests a linear approximation.

Let us define the grid element E_{ij} to be the rectangle bounded by the points (x_i, y_j), (x_i, y_{j+1}), (x_{i+1}, y_{j+1}), and (x_{i+1}, y_j). Since three noncolinear points uniquely determine a plane in space it follows that, in general, a linear approximation satisfying eq. (1) at the four corners of E_{ij} does not exist. We can, however, by choosing three of the four corners, obtain a linear approximation that will be applicable to half of the area of E_{ij}. In a like manner we can approximate the remaining half. This, of course, can be done in two different ways, each of which will yield a slightly different approximation to eq. (1) in E_{ij}. The two possible approximating functions will, however, agree along the straight line segment connecting any two adjacent corners of E_{ij}.

Let us arbitrarily divide E_{ij} into two segments by a line connecting (x_i, y_{j+1}) and (x_{i+1}, y_j). Then a linear approximation to eq. (1) can be obtained for each of these segments. Let R be the rectangular region bounded by the lines $x = x_{\min}$, $x = x_{\max}$, $y = y_{\min}$, $y = y_{\max}$. For an arbitrary point $(\tilde{x}, \tilde{y})$ in R the following procedure then holds:



(i) Find i, j such that $x_i \leq \tilde{x} < x_{i+1}$, $y_j \leq \tilde{y} < y_{j+1}$.

(ii) Compute $m_{ij} = \dfrac{y_{j+1} - y_j}{x_i - x_{i+1}}$, $m = \dfrac{\tilde{y} - y_j}{\tilde{x} - x_{i+1}}$.

(iii) Set $k = \begin{Bmatrix} 0 \\ 1 \end{Bmatrix}$ if $\begin{Bmatrix} m \leq m_{ij} \\ m > m_{ij} \end{Bmatrix}$.

(iv) $z^* = A\tilde{x} + B\tilde{y} + C$, where A, B, and C satisfy

$$z(x_i, y_{j+1}) = Ax_i + By_{j+1} + C,$$

$$z(x_{i+1}, y_j) = Ax_{i+1} + By_j + C,$$

$$z(x_{i+k}, y_{j+k}) = Ax_{i+k} + By_{j+k} + C.$$

The resulting approximation, z^*, is defined and continuous for all points $(\tilde{x}, \tilde{y})$ in R.

In order to answer the question whether P is or is not visible from C, it is necessary to refer to points on the line segment $\overline{CP}$. Let $\tilde{P}$ be an arbitrary point on $\overline{CP}$.

[1] This restriction is not required nor is it desirable in the general case. It does, however, lead to a fairly efficient algorithm for testing the visibility of an arbitrary point on S.

Then if (x, y, z), $(\tilde{x}, \tilde{y}, \tilde{z})$, and (c_x, c_y, c_z) are the coordinates of P, $\tilde{\mathrm{P}}$, and C, the following equations hold:

$$\frac{\tilde{x} - c_x}{x - c_x} = \frac{\tilde{y} - c_y}{y - c_y} = \frac{\tilde{z} - c_z}{z - c_z}. \tag{32}$$

Visibility of P will be determined by examining the behavior of

$$\delta(\mathrm{P}, \tilde{\mathrm{P}}) = \tilde{z} - z^* \tag{33}$$

as $\tilde{\mathrm{P}}$ traverses that part of the line segment $\overline{\mathrm{CP}}$ that is contained in R. Figure 7 depicts $\tilde{z}$ and z^* for some typical situations. In Figure 7 (a) and (b), P is visible since $\tilde{z}$ and z^* never intersect (except at P). In Figure 7 (c), P is hidden since an intersection different from P does occur. Such considerations lead to the following:

Criterion for Visibility. If $\delta(\mathrm{P}, \tilde{\mathrm{P}})$ remains always positive or always negative at all points of $\overline{\mathrm{CP}}$ in R, then P is visible. Otherwise it is hidden.

Note that z^* is piecewise linear. Thus it will suffice to examine $\delta(\mathrm{P}, \tilde{\mathrm{P}})$ at its vertices only. These vertices occur along three families of lines, i.e. (1) $x_i =$ constant, (2) $y_j =$ constant, and (3) the lines connecting (x_i, y_{j+1}) and (x_{i+1}, y_j). In that we are using only lines of constant x or y as generating curves, it is unimportant to consider vertices of type three. In the few cases where consideration of such vertices would affect the plot, the effect of ignoring them would be merely a slight flattening of the approximating surface. (See Figure 8.)

In order to establish the visibility of a point $\mathrm{P}(x, y, z)$ the following steps are executed for every constant x (column) and every constant y (row) in the data array:

(i) For fixed $\tilde{x}$ (or $\tilde{y}$), compute $\tilde{y}$ (or $\tilde{x}$) and $\tilde{z}$ from eq. (32).

(ii-a) If $\tilde{x}$ was fixed, compute z^* from

$$\frac{z^* - z(\tilde{x}, y_j)}{z(\tilde{x}, y_{j+1}) - z(\tilde{x}, y_j)} = \frac{\tilde{y} - y_j}{y_{j+1} - y_j}, \qquad y_j \le \tilde{y} \le y_{j+1}.$$

(ii-b) If $\tilde{y}$ was fixed, compute z^* from

$$\frac{z^* - z(x_i, \tilde{y})}{z(x_{i+1}, \tilde{y}) - z(x_i, \tilde{y})} = \frac{\tilde{x} - x_i}{x_{i+1} - x_i}, \qquad x_i \le \tilde{x} \le x_{i+1}.$$

(iii) Compute $\delta(\mathrm{P}, \tilde{\mathrm{P}}) = \tilde{z} - z^*$.

If for every $\tilde{\mathrm{P}}$ thus tested $\delta(\mathrm{P}, \tilde{\mathrm{P}})$ has the same sign, then P is visible. Otherwise it is hidden.

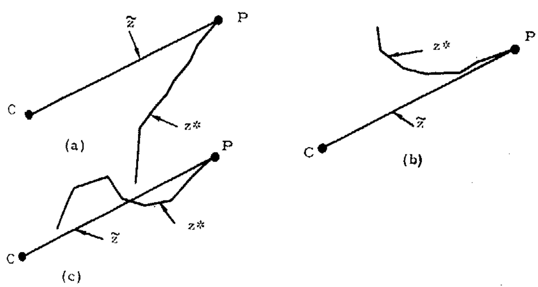


FIG. 7

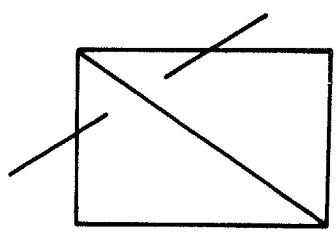


FIG. 8

5. *Connection of Visible Points*

Our considerations on the connection of visible points are facilitated by defining a new function of P as follows:

$$\phi(\mathrm{P}) = +1 \quad \text{if P is visible and} \quad \delta(\mathrm{P}, \tilde{\mathrm{P}}) > 0,$$

$$\phi(\mathrm{P}) = -1 \quad \text{if P is visible and} \quad \delta(\mathrm{P}, \tilde{\mathrm{P}}) < 0,$$

$$\phi(\mathrm{P}) = \quad 0 \quad \text{if P is hidden.}$$

The procedure outlined in Section 4 suffices to determine the value of $\phi(\mathrm{P})$ for each point in the data array. The question remains, just how should these points be connected? When it is determined that two points should be connected, this is done by means of a straight line segment in the plane of projection. Two points in the data array are connected only if they are adjacent. By adjacent is meant that the two points are in the same row (or column) and adjacent columns (or rows).

In general, the projection of a visible point is represented by the intersection of two curves, one of constant x and one of constant y. The projection of a hidden point is not plotted. Let P and Q denote adjacent points in the data array. Then four cases must be examined.

5.1 $\phi(\mathrm{P}) = \phi(\mathrm{Q}) = \pm 1$. In this case both points are visible. It is true that there may be hidden points on the line segment $\overline{\mathrm{PQ}}$. However, a general method to

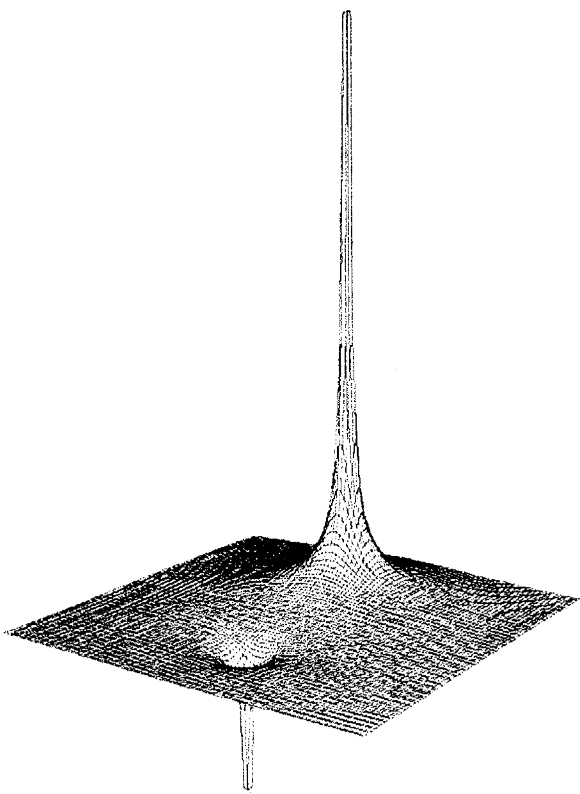


FIG. 9

discover all such hidden points cannot be accomplished with a finite number of calculations. Most typically in this situation all points on the line segment $\overline{PQ}$ are visible. Therefore in this case we always connect the adjacent points. If the resulting plot has a flaw due to the ignoring of hidden points on $\overline{PQ}$, the situation can be corrected by increasing the number of rows and columns of data points in R.

5.2 $\phi(P) = \phi(Q) = 0$. Here both points are hidden. Analogous to the case discussed in Section 5.1, there may be visible points on $\overline{PQ}$. However, we ignore this possibility and draw no line between P and Q.

5.3. $\phi(P) = \pm 1$, $\phi(Q) = 0$. In this case P is visible and Q hidden. We assume that a point, say P′, exists on the line segment $\overline{PQ}$ such that all points interior to $\overline{PP'}$ are visible whereas all points interior to $\overline{P'Q}$ are hidden. Any flaw in the resulting plot due to the invalidity of this assumption can be remedied by sufficiently increasing the number of rows and columns of data points in R.

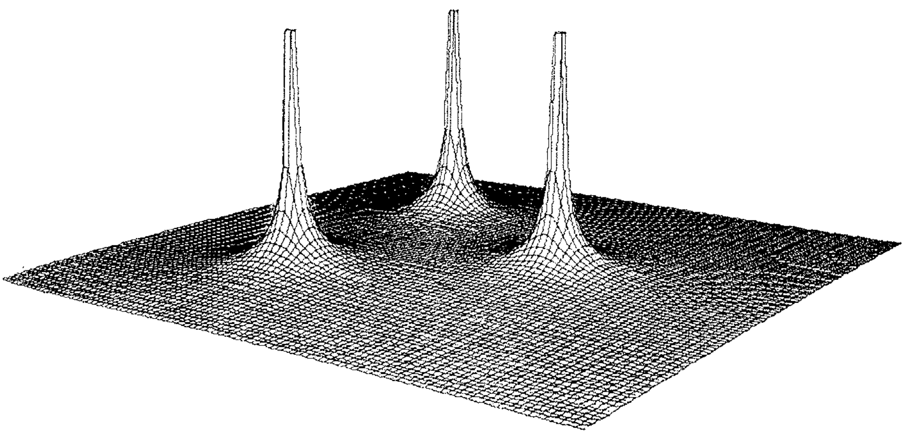


FIG. 10

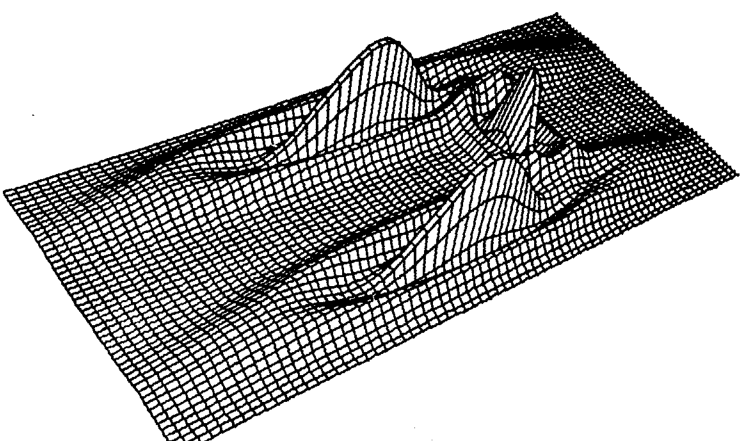


FIG. 11

The point P′ is located to any desired accuracy by means of a binary search (i.e. successive halvings of the interval PQ). Then $\overline{PP'}$ can be drawn.

5.4. $\phi(P) = -\phi(Q) = 1$. In this case both P and Q are visible but a segment of the interior of $\overline{PQ}$ is hidden. We assume here that two points, P′ and Q′, exist on the line segment $\overline{PQ}$ with the following properties: At an arbitrary point on the interior of $\overline{PP'}$, say P^*, the function $\phi(P^*) = +1$. At an arbitrary point on the interior of $\overline{Q'Q}$, say Q^*, the function $\phi(Q^*) = -1$. At an arbitrary point on the interior of $\overline{P'Q'}$, say $\bar{P}$, the function $\phi(\bar{P}) = 0$. Thus $\overline{PP'}$ and $\overline{Q'Q}$ are visible while $\overline{P'Q'}$ is hidden. Both P′ and Q′ can be located by means of a binary search. Then $\overline{PP'}$ and $\overline{Q'Q}$ can be plotted.

Figures 9, 10, and 11 are pictures of the functions shown in Figures 4, 5, and 6 with the additional feature that the hidden surfaces have been eliminated. The gain in clarity is obvious. The plots in Figures 4–6 and 9–11 were all made automatically by the IBM 1627 plotter from a tape produced by the IBM 7094 computer.

REFERENCES

1. ROBERTS, L. G. Machine perception of three-dimensional solids. Tech. Rep. No. 315, Lincoln Laboratory, MIT, Lexington, Mass., May 1963.
2. WEISS, R. A. BE VISION, A package of IBM 7090 FORTRAN programs to draw orthographic views of combinations of plane and quadric surfaces. *J. ACM 13*, 2 (April 1966), 194–204.
3. Private correspondence between L. Sashkin of Aerospace Corp. and D. J. Di Leonardo of Westinghouse Electric Corp.
4. The annual computer art contest of *Computers and Automation*. *Computers and Automation 14* (Aug. 1965), 10–11.

RECEIVED JULY, 1966; REVISED AUGUST, 1966