

BE VISION, A Package of IBM 7090 FORTRAN Programs to Draw Orthographic Views of Combinations of Plane and Quadric Surfaces

RUTH A. WEISS

Bell Telephone Laboratories, Inc., Murray Hill, New Jersey

Abstract. BE VISION is a package of FORTRAN programs for drawing orthographic views of combinations of plane and quadric surfaces. As input, the package takes rectangular coordinate equations specifying the surfaces plus a three-angle specification of the viewing direction. Output is a drawing on the Stromberg Carlson 4020 Microfilm Recorder. Many views of one scene may be obtained simply by changing the viewpoint.

The various subroutines of the package and their functions are described in this paper. It also gives numerous examples of pictures that were produced by BE VISION. The package has been in use since April 1964.

I. Introduction

The choice of quadric surfaces for the general program was made for practical and esthetic reasons. While there are many geometrical figures consisting of planes alone, the vast majority of objects in the real world contain some curved surfaces. A program for drawing pictures would not be very useful if it did not at least include a provision for the simplest of these, namely quadric surfaces. In drawing the curves of the picture it was found that a great many calculations for each point had to be made. If surfaces of higher degree than two were considered, the storage space and the machine time used would become too large to be practical; however, the program for drawing quadrics has proved versatile enough to produce some quite complicated pictures.

A quadric surface is the locus of the general second-degree equation in three variables :

$$Q(x, y, z) = a_1x^2 + a_2y^2 + a_3z^2 + b_1yz + b_2xz + b_3xy + c_1x + c_2y + c_3z + d = 0.$$

The program can handle any Q including planes. To handle complex objects, not only complete surfaces but also bounded quadric surfaces were included in the vocabulary of the program. Indeed, since many quadrics, such as planes or cylinders, are unbounded, some limits must be provided to obtain a picture at all. Thus, the program accepts two kinds of surfaces as input; namely, (1) the principal equations, which define the surfaces themselves, and, (2) the bounding inequalities which limit the extension of the surfaces of type (1). The type (1) and type (2) equations are alike in that they are of the same form (plane or quadric surfaces), but they also have these important differences:

(a) A *principal surface* may stand alone without boundaries, is defined by an equation, is a visible, opaque surface, and has visible intersections with other principal surfaces and with its own bounds.

(b) A *bounding surface* must be associated with a principal surface, is defined by an inequality, is an invisible, transparent surface, and has visible intersections only with its associated principal surfaces.

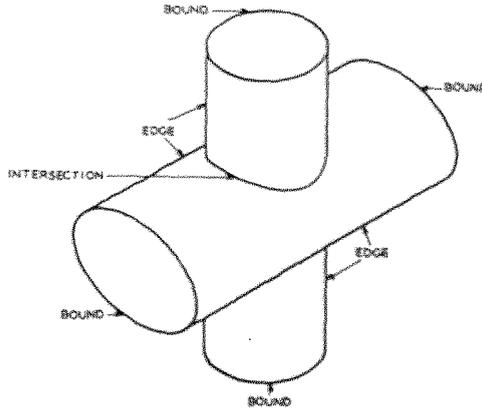


FIG. 1. Intersecting cylinders

The lines appearing in a drawing are of three kinds: the intersection of two surfaces, the intersection of a surface and one of its bounds, and the extremal points of a surface which are not a boundary, referred to hereafter as intersections, bounds and edges, respectively (see Figure 1).

No matter which of the three kinds of curves is being calculated, its coordinates are obtained numerically by incrementing one of the three variables x , y or z from a minimum to a maximum value and solving for the other two; therefore before anything else can be done these minima and maxima have to be found.

The intersections and bounds of a picture remain the same for any view; edges do not. All the intersections and bounds are calculated and stored as ordered lists of the coordinates of the curves. When a particular view is requested, the edges are calculated afresh. Then the lists are tested one at a time. Each point is tested for visibility. If two adjacent points are visible, a segment is drawn connecting their projections in the picture plane.

All the illustrations in this paper were drawn with BE VISION except for labels, arrows and dotted lines.

II. General Explanation of the Program

A. *General Notation*. The following notations are used throughout this paper.

The equation of the general quadric surface is:

$$Q(x, y, z) = a_1x^2 + a_2y^2 + a_3z^2 + b_1yz + b_2xz + b_3xy + c_1x + c_2y + c_3z + d = 0.$$

The quadratic part of Q is:

$$q(x, y, z) = a_1x^2 + a_2y^2 + a_3z^2 + b_1yz + b_2xz + b_3xy.$$

The partial derivatives of Q are:

$$Q_x \equiv Q_1(x, y, z) = 2a_1x + b_3y + b_2z + c_1,$$

$$Q_y \equiv Q_2(x, y, z) = 2a_2y + b_3x + b_1z + c_2,$$

$$Q_z \equiv Q_3(x, y, z) = 2a_3z + b_2x + b_1y + c_3.$$

Two intersecting quadric surfaces are:

$$Q'(x, y, z) = a_1x^2 + a_2y^2 + a_3z^2 + b_1yz + b_2xz + b_3xy + c_1x + c_2y + c_3z + d = 0,$$

$$Q''(x, y, z) = A_1x^2 + A_2y^2 + A_3z^2 + B_1yz + B_2xz + B_3xy + C_1x + C_2y + C_3z + D = 0.$$

B. *Minima and Maxima.* A minimum and a maximum value for each of the three variables, x, y, z must be calculated for every surface in the picture. Any portion of a curve on a surface will be discarded with no further investigations, if it lies outside the limits for that coordinate on that surface.

These limits are obtained either from the bounding equations provided as input or from the natural bounds of the surface itself or from a combination of these two. The simplest bounds which may be placed on a surface are the six inequalities $x_{\min} \leq x \leq x_{\max}$, $y_{\min} \leq y \leq y_{\max}$ and $z_{\min} \leq z \leq z_{\max}$. If any of these is provided as input, it is left untouched. If a minimum or a maximum of a variable has *not* been provided as input, it is obtained in the following way: The equation of the surface itself and the equations of all its bounds are examined for natural limits. From these the lowest value of a variable is taken as its minimum and the highest value of a variable is taken as its maximum.

Let us take one of the variables, say x , and show how its natural limits are determined.

Case 1. If, for example,

$$Q = f(x) = ax + c = 0,$$

then $x_{\min} = -c/a = x_{\max}$.

Case 2. If

$$Q = f(x) = ax^2 + bx + c = 0,$$

then x_{\min} and x_{\max} , respectively, equal the lesser and greater of $-b \pm \sqrt{(b^2 - 4ac)}/2a$.

Case 3. If

$$Q = f(x, y) = a_1x^2 + a_2y^2 + b_3xy + c_1x + c_2y + d = 0,$$

the discriminant is set to 0 and solved for x :

$$(b_3x + c_2)^2 - 4a_2(a_1x^2 + c_1x + d) = 0 \quad (1)$$

where x_{\min} is equal to the lesser of the roots of (1) and x_{\max} , the greater of the roots of (1).

Case 4. If

$$Q = f(x, z) = a_1x^2 + a_3z^2 + b_2xz + c_1x + c_3z + d = 0,$$

the same process is used as was used for case 3.

Case 5. If

$$Q = f(x, y, z) = a_1x^2 + a_2y^2 + a_3z^2 + b_1yz + b_2xz + b_3xy + c_1x + c_2y + c_3z + d = 0,$$

the z discriminant is set to 0 and solved for y , the y discriminant is then set to 0 and solved for x . Then the two roots are x_{\min} and x_{\max} .

The x_{\min} and x_{\max} obtained by processes 3, 4 and 5 are not necessarily the actual

minimum and maximum values of x for the surface Q . They *are* limits for the surface Q but they may well be the values between which the surface Q does not exist (this is the case for a hyperboloid of two sheets). This is useful information for determining visibility but it is not what is desired for minima and maxima. To forestall such cases where no natural limits exist, it is necessary for the user to supply clear limits to the program.

C. Curves of Intersection. An intersection or a bound is a curve of intersection of two quadric or planar surfaces, Q' and Q'' . The points on a curve are obtained by varying one of the coordinates from its minimum to its maximum and solving the equations Q' and Q'' for the other two coordinates at each point. The minima and maxima of x , y and z for the curve are obtained by taking the larger minimum and the smaller maximum for each variable for Q' and Q'' . Every point on the curve is tested for these bounds and eliminated if it does not lie within them. However, these bounds only enclose a rectangular parallelepiped and are not stringent enough. Each point must also be tested to make sure it meets the inequality requirements of all the bounds of the surface or surfaces. If a point does not fall within bounds, it is flagged as invisible.

If Q' and Q'' contain the same variable in the second degree, there will be more than one branch of the curve of intersection. In the case where Q' and Q'' contain at least two of the same variables in the second degree, we wish to eliminate one of the variables, say x , between the equations and get one equation $F(y, z) = 0$.

By using Sylvester's method we obtain an eliminant which furnishes a necessary and sufficient condition that the two equations have a root in common.

If Sylvester's method is applied to the equations for the quadric surfaces, namely, $Q'(x, y, z) = 0$ and $Q''(x, y, z) = 0$ with

$$Q'(x, y, z) = a_1x^2 + a_2y^2 + a_3z^2 + b_3xy + b_2xz + b_1yz + c_1x + c_2y + c_3z + d = 0$$

and

$$Q''(x, y, z) = A_1x^2 + A_2y^2 + A_3z^2 + B_3xy + B_2xz + B_1yz + C_1x + C_2y + C_3z + D = 0,$$

the resultant becomes

$$\begin{vmatrix} a_1 & (b_1y + b_2z + c_1) & (a_3y^2 + a_3z^2 + b_1yz + c_3z + d) & 0 \\ 0 & a_1 & (b_2y + b_2z + c_1) & (a_3y^2 + a_3z^2 + b_1yz + c_2y + c_3z + d) \\ A_1 & (B_1y + B_2z + C_1) & (A_3y^2 + A_3z^2 + B_1yz + C_3z + D) & 0 \\ 0 & A_1 & (B_2y + B_2z + C_1) & (A_3y^2 + A_3z^2 + B_1yz + C_3z + D) \end{vmatrix} = 0.$$

The value of this determinant is the required equation $F(y, z) = 0$. Either y or z , say z , is varied in increments of delta from its minimum to its maximum and substituted in $F = 0$. Since F is in general quartic [5], this gives the possibility of four real solutions for y at each value of z . Four separate lists are kept for these curves if they should exist. At each point, y and z are substituted in $Q' = 0$ and $Q'' = 0$ to get a solution for x ; if $Q' = 0$ and $Q'' = 0$ have common real roots, these roots are entered into the lists.

It may occur that a curve does not extend clear across from the lower to the upper limit of the varying coordinate. In this case two branches must terminate by joining instead of by running out of bounds. Unless the point of meeting appeared

in the list for each branch, a gap would occur at the junction. This point of meeting is calculated by using one of the methods 3, 4 or 5, for determining natural bounds as explained in the section on minima and maxima.

The subroutine for calculating intersections contains many subsections for special cases. It was divided in this way to avoid unnecessarily lengthy calculations.

D. *Edges* [1]. In general any quadric surface will have visible curves (herein known as edges) which are not intersections with other surfaces. Every surface is tested for second degree terms to find these extremal points which are not a boundary. These extremal points are points at which the line of view is tangent to the quadric surface. The following theorem is used to obtain these points.

THEOREM. *If the point (α, β, γ) lies on the quadric surface $Q(x, y, z)$, the line $x = \alpha + \lambda s, y = \beta + \mu s, z = \gamma + \nu s$ will be the tangent to the surface at the point (α, β, γ) if and only if its direction cosines, λ, μ and ν satisfy the equation*

$$\lambda Q_1(\alpha, \beta, \gamma) + \mu Q_2(\alpha, \beta, \gamma) + \nu Q_3(\alpha, \beta, \gamma) = 0,$$

where Q_1, Q_2 and Q_3 are the partial derivatives of Q with respect to x, y and z respectively.

If the line is the line of view and its direction cosines are λ_0, μ_0, ν_0 , and if $Q(x, y, z)$ is the surface for which an edge is wanted and (x, y, z) is the point we are looking for, the equation to be solved is

$$\begin{aligned} \lambda_0(2a_1x + b_3y + b_2z + c_1) + \mu_0(b_3x + 2a_2y + b_1z + c_2) \\ + \nu_0(b_2x + b_1y + 2a_3z + c_3) = 0, \\ x(2a_1\lambda_0 + b_3\mu_0 + b_2\nu_0) + y(b_3\lambda_0 + 2a_2\mu_0 + b_1\nu_0) \\ + z(b_2\lambda_0 + b_1\mu_0 + 2a_3\nu_0) + (c_1\lambda_0 + c_2\mu_0 + c_3\nu_0) = 0. \end{aligned}$$

This gives us a linear equation in x, y and z . We already have a second degree equation (the equation of the surface) and we solve these numerically for edges just as we did for intersections.

E. *Visibility* [1]. When considering the feasibility of writing BE VISION, one of the first problems to be encountered was that of deciding whether a line is visible or not. For a picture with a fixed number of known surfaces and a predetermined view, large sections of curves are known to be visible or invisible; but the whole point of writing BE VISION was to have a general program for drawing *any* reasonable number of quadric surfaces from *any* view. In order to decide on the visibility of a curve (say of the intersection of surfaces Q' and Q'') or any arbitrarily small part of that curve, all the surfaces including Q' and Q'' have to be examined to see if any of them hides it from view. (Two good examples of this difficulty can be seen in Figure 9, the ellipsoidal coordinates, and Figure 6c, an oblique view of Mickey Mouse.) Changing the angle of the picture plane even slightly may make a large difference in the amount of visibility of a curve (as much as from completely visible to completely invisible or vice versa as in Figure 2). To assure an accurate and general way of deciding on the visibility of a curve, the author chose the following method; namely, calculate it point by point and test each point against all the surfaces to see if any of them is hiding it from view. Unfortunately this is a time-consuming process.

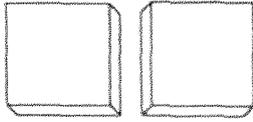


FIG. 2

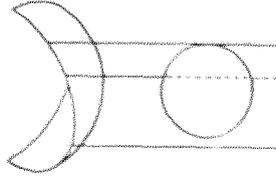


FIG. 3

Suppose all the points of a curve have been calculated; that is, the x, y, z coordinates for each point are in storage in order. It is now necessary to decide which points are visible. BE VISION uses orthographic projection⁸ only, which means that all the lines of view are parallel. The direction cosines, λ, μ and ν of the lines of view are obtained from the input parameters of the picture plane (see Section IIF) and these are constant for one picture. By solving a quadratic equation it can be determined if a line through the point (x, y, z) with direction cosines λ, μ and ν (a) pierces a quadric surface in two places, (b) is tangent to the surface, or (c) does not touch the surface. However, this depends on whether there are (a) two real roots, (b) a double root, or (c) no real roots (Figure 3).

If there are real roots, the piercing points are calculated and tested to see if they actually exist; for they may have been cut off by a boundary. If they are real points, a simple inequality is used to determine whether either one lies nearer the picture plane than the point (x, y, z) . This test is made against all the surfaces in the picture for each point (x, y, z) on the line. When two adjacent points in the list are visible they are connected. If a point P_n is visible and its adjacent point P_{n+1} is not visible, one twentieth of the difference between them is added to P_n until the first invisible point is found. The last visible point is then connected to P_n .

F. Point of View. In its canonical form the picture plane is assumed to be parallel to the $y-z$ plane and the \bar{x} and \bar{y} axes of the picture plane, parallel to the y and z axes respectively.

Then if α is the azimuth of the vantage point measured counterclockwise from the x -axis about the vertical z -axis in the horizontal $x-y$ plane, β is the elevation of the vantage point above the $x-y$ plane, and γ is the counterclockwise rotational position of the picture about the vantage direction, the direction cosines of the line of view, λ, μ and ν and the coordinates, \bar{x}, \bar{y} in the picture plane corresponding to the point x, y, z in three-space are derived from the three input angles α, β, γ in the following way:

$$\lambda = \cos \alpha \cos \beta, \quad \mu = \sin \alpha \cos \beta, \quad \nu = \sin \beta,$$

$$\bar{x} = x(-\sin \alpha \cos \gamma - \cos \alpha \sin \beta \sin \gamma) + y(\cos \alpha \cos \gamma - \sin \alpha \sin \beta \sin \gamma) + z(\cos \beta \sin \gamma),$$

$$\bar{y} = x(\sin \alpha \sin \gamma - \cos \alpha \sin \beta \cos \gamma) + y(-\cos \alpha \sin \gamma - \sin \alpha \sin \beta \cos \gamma) + z(\cos \beta \cos \gamma).$$

Figure 4 shows the orientation of the axes and the angles of the picture plane.

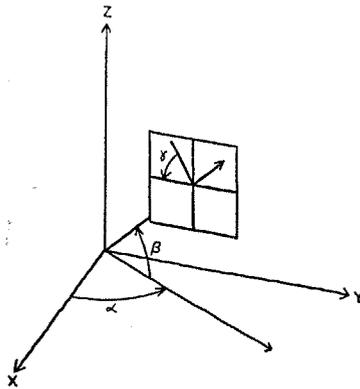


FIG. 4. Coordinate system

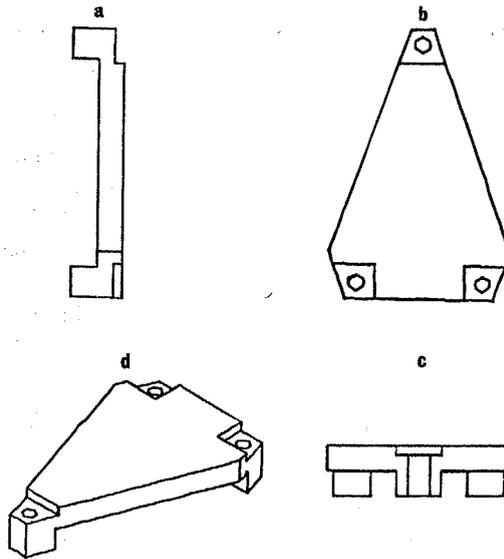


FIG. 5a-d. Four views of a mounting plate

III. Discussion

1. At the time BE VISION was written literature on the subject of drawing quadric surfaces was rather sketchy. The most important treatises on the subject of the analytic geometry of 3 dimensions did not cover such practical problems (for BE VISION) as which coordinate to vary for a curve, how to get the general numerical solution for the intersection of two quadric surfaces, or how to find the exact point at which the solution of the two equations becomes imaginary. These rules had to be derived and this fact undoubtedly made the subroutines which calculate the intersection of curves the most difficult to write.

More recently several papers have been written suggesting solutions to some of these problems. Luh and Krolak [6] use a pair of inequalities to obtain the intersection of a pair of quadric surfaces, but they do not explain how to get the exact end points of this intersection. Roberts [7, 8] has done extensive work in the field

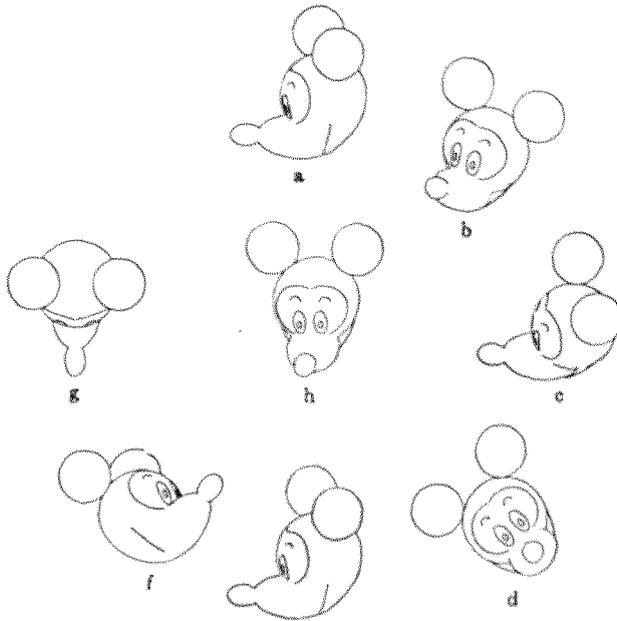


FIG. 6a-h (© Walt Disney)

of drawing 3-dimensional solids, but in his May 1965 paper he does not tackle the problem of the intersection of two quadric surfaces.

2. The roundoff error had to be bypassed if good points were not to be discarded (for instance, if the error made a point appear to be slightly behind a surface instead of in front of it). Every coordinate was given a tolerance limit when tested for its minimum or maximum or when tested for visibility. It would not be too intolerable to drop one point perhaps, but even this would leave a gap in the curve; there are cases, however, where a whole curve might be missed if, for example, one of the coordinates is a constant.

3. Right from the beginning it was obvious that machine time would be a problem. It takes time to calculate every single point of every curve, but it was necessary to do this because of the visibility problem. The time for calculating the points (a maximum of 1 millisecond per point and very much less than this where Q' and Q'' both are linear in all three variables) turned out to be surprisingly low. Even this time can be appreciable in a picture with several thousand points, however. The time for all the other operations necessary to get a picture takes anywhere from 3 to 15 times that for getting the points alone, depending on the number of surfaces involved. Most of this time is spent in testing for visibility.

4. The program is large. It occupies almost all of the approximately 60,000 octal locations available to a program running under the Bell Laboratories monitor, BESYS7. It includes the microfilm and mathematical functions subroutines but not the input-output which is included in BESYS7. The program can draw a picture with as many as 50 principal surfaces and a total of 180 principal and bounding surfaces. Experience has shown that this is a good ratio.

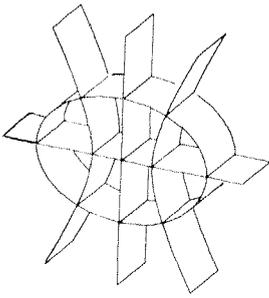


FIG. 7

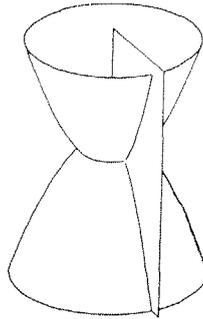


FIG. 8

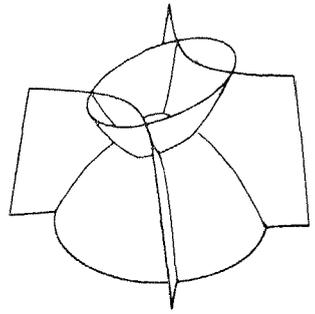


FIG. 9

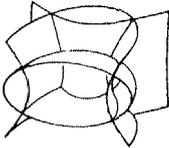


FIG. 10

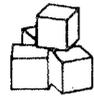
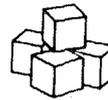
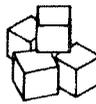


FIG. 11a-c. Three views of a pile of 4 boxes

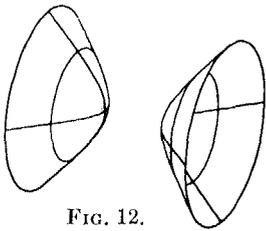


FIG. 12.
Hyperboloid of two sheets (1 surface)

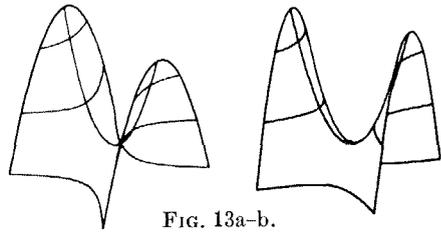


FIG. 13a-b.
Two views of hyperbolic paraboloid

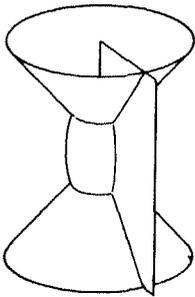


FIG. 14



FIG. 15

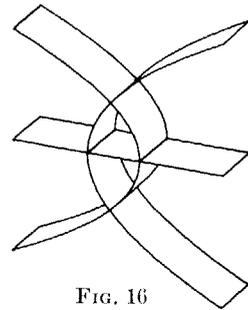


FIG. 16

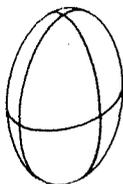


FIG. 17

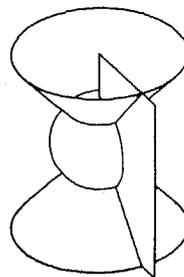


FIG. 18

The subprogram which gets the curves of intersection is by far the longest and most complex, and it does not include the use of Sylvester's Method which is in a separate subprogram.

5. Originally BE VISION was rather difficult to use as this involved writing out all the equations of the surfaces in rectangular form. Now there is a picture drawing language written by M. D. McIlroy of Bell Telephone Laboratories, which greatly facilitates the task of specifying the surfaces. The two programs are available as a single system on a single run.

IV. Pictures Produced by BE VISION

This section contains a few examples of pictures produced by BE VISION (see Figures 5-18).

In the case of Mickey Mouse all the views were produced with the same equations of type (1) and type (2); only the picture plane parameters were changed for different views. Mickey's head and ears are spheres, his snout and nose are ellipsoids, and the jaw which connects the head and snout is a piece of an ellipsoid. All the intersections (as defined in the Introduction) were omitted to give the artistic effect (see Figures 6a-g).

The general equations for the various coordinate systems were taken from *Field Theory Handbook* [4]. They were chosen as illustrations of the many ways in which quadric surfaces can be combined.

All the bounds of the mounting plate were omitted. If they were included superfluous lines would be drawn between sections of the same surface. The subroutine for getting intersections ignores parallel planes. In general, if a picture represents a completely enclosed solid body whose surfaces consist of planes, it is not necessary and indeed undesirable to draw the bounds.

Acknowledgments. The author wishes to thank M. D. McIlroy for suggesting such a fascinating project and to thank M. D. McIlroy and H. O. Pollak for their encouragement and interest throughout its implementation.

RECEIVED APRIL, 1965; REVISED SEPTEMBER, 1965

REFERENCES

1. DRESDEN, A. *Solid Analytical Geometry and Determinants*. John Wiley & Sons, New York City, 1930.
2. ROEVER, W. H. *Fundamental theorems of orthographic axonometry and their value in picturization*. Washington U. Studies—New Series, St. Louis, Mo., 1941.
3. WEISSNER, L. *Introduction to the Theory of Equations*. The MacMillan Co., New York, 1938.
4. MOON, P., AND SPENCER, D. E. *Field Theory Handbook; Including Coordinate Systems, Differential Equations and Their Solutions*. Springer-Verlag, Berlin, Gottinger, Heidelberg, 1961.
5. GRAY, M. C. QUARTC, a 7090 subroutine to get the roots of a quartic equation with real coefficients. Bell Telephone Laboratories Program Library, Murray Hill, N. J.
6. LUH, J. Y. S., AND KROLAK, R. J. A mathematical model for mechanical part description. *Comm. ACM* 8 (Feb. 1965), 125-129.
7. ROBERTS, L. G. Machine perception of three-dimensional solids. Tech. Rep. No. 315, Lincoln Laboratory, MIT, May 1963.
8. ROBERTS, L. G. Homogeneous matrix representation and manipulation of N -dimensional constructs. MS-1405, Lincoln Laboratory, MIT, May 1965.

APPENDIX

Theorem Used to Determine Visibility

THEOREM. (1) $Q(x, y, z) = 0$ is the equation of a quadric surface; (2) Q_1, Q_2, Q_3 are the partial derivatives of Q with respect to x, y, z ; and (3) $q(x, y, z)$ is the part of Q which is homogeneous in the second degree. Then the parameter values of the points in which the line $x = \alpha + \lambda s, y = \beta + \mu s, z = \gamma + \nu s$ meets the quadric surface $Q(x, y, z) = 0$, are the roots of the equation

$$L_0 s^2 + 2L_1 s + L_2 = 0,$$

where

$$L_2 = Q(\alpha, \beta, \gamma),$$

$$L_1 = \frac{1}{2}[\lambda Q_1(\alpha, \beta, \gamma) + \mu Q_2(\alpha, \beta, \gamma) + \nu Q_3(\alpha, \beta, \gamma)],$$

$$L_0 = q(\lambda, \mu, \nu).$$

COROLLARY. The line $x = \alpha + \lambda s, y = \beta + \mu s, z = \gamma + \nu s$ (a) will meet the quadric surface Q in two distinct real points iff $L_1^2 - L_0 L_2 > 0$; (b) will be tangent to Q iff $L_1^2 = L_0 L_2$; and (c) will not have any real points in common with the surface iff $L_1^2 - L_0 L_2 < 0$.

If (α, β, γ) is a point being tested for visibility, (x, y, z) is one of the points of (a) above and λ_0, μ_0, ν_0 are the direction cosines of the line of view; then (α, β, γ) is hidden by the surface $Q(x, y, z) = 0$ and not drawn if

$$[(\lambda_0 \alpha + \mu_0 \beta + \nu_0 \gamma) - (\lambda_0 x + \mu_0 y + \nu_0 z)] < 0.$$