Visualising Duals of Surfaces

Richard J. Morris

Department of Applied Mathematics and Theoretical Physics,
University of Liverpool, P.O.Box 147, Liverpool, L69 7BR

May 15, 2002

Abstract

The dual of a surface is the set of all tangent planes to a surface. Duals are a useful tool for gaining geometric information about a surface. Whilst duals theoretically lie in $\mathbf{RP}^3$ they can be projected into $\mathbf{R}^3$ where they can be visualised on a computer.

We present a general scheme for visualising duals, and discuss the advantages and disadvantages of different types of projections. Parts of the visualising process involves rotations in $\mathbf{R}^4$ and we describe one method for specifying such a rotation. Any projection from $\mathbf{RP}^3$ to $\mathbf{R}^3$ will have some degenerate points and we discuss how such points can be easily removed.

Categories and Subject Descriptors: I.3.5 [**Computer Graphics**]: Computational Geometry and Object Modeling - *curves, surface and object representation, geometric algorithms.*

General Terms: Algorithms, Experementation.

Additional Key Words and Phrases: Duals, Hough transforms, scientific visualisation, projective space.

1   Introduction

The dual of a surface is the set of all tangent planes to the surface. We can think of the dual of a surface as the set of points $(a; b; c; d)$ in projectivised 3-space, $\mathbf{RP}^3$, such that the plane $ax+by+cz = d$ is tangent to the surface. Here the semi-colon indicates that we are using homogeneous coordinates, we have $(a; b; c; d) = (\alpha a; \alpha b; \alpha c; \alpha d)$ for all non-zero α. Typically the dual of a surface S will form an other surface in $\mathbf{RP}^3$ called the *dual surface*, S^*. The dual has a number of interesting properties which reveal much of the geometry of the original surface. These include:

1 If x lies on a parabolic line on S then the corresponding point on the dual, x^*, lies on a cuspidal edge on S^*.

2 If there is a plane tangent to S at x and y then S^* has a self intersection at $x^* = y^*$.

3 If S has a cusp of Gauss at x, then the dual has a swallowtail point at x^*. (At a cusp of Gauss the parabolic line on S is tangent to the principal direction with zero principal curvature, and the surface has a high order contact with the tangent plane. For a detailed investigation of cusps of Gauss see [?].)

If we think of our surface as lying in $\mathbf{RP}^3$ then the dual of the dual gives us back our original surface. This fact gives us the following *dual* results:

1^* If x lies on a cuspidal edge on S then x^* lies on a parabolic line.

2^* If S has a self-intersection then S^* has a bi-tangent plane.

3^* If S has a swallowtail point at x then x^* is a cusp of Gauss.

These properties make the dual a useful tool for investigating problems in computer vision and differential geometry. For instance in [?] the dual plays an important part in the study of families of surfaces. The dual is closely linked to the Hough transform [?] which is widely used in computer vision.

The study of dual surfaces is greatly enhanced by being able to construct a visual representation of the dual on a computer. Once we have such a representation it can be examined using a 3D viewing package such as *Geomview* [?], and questions about the pattern of cuspidal edges, swallowtail points, and self intersections can be quickly answered. In this paper we describe a method for constructing such a representation. This consist of four main steps:

1. Calculate the dual in $\mathbf{RP}^3$ (§2).

2. Project into $\mathbf{R}^4$ and apply a rotation in $\mathbf{R}^4$ (§4).

3. Trim the surface in $\mathbf{R}^4$ to remove problem facets (§5).

4. Project from $\mathbf{R}^4$ into $\mathbf{R}^3$. (§3, §6, §7).

In fact we can use this algorithm to visualise any object in $\mathbf{RP}^3$, simply by omitting the first step. We have implemented this process as a computer program forming part of the *Liverpool Surface Modelling Package* [?, ?], other programs in the package are used to generate the surfaces to be dualised. In §8 we look at how this visualisation has helped with investigations into flat umbilics [?] and cross-caps [?].

2 Calculating the Dual

Throughout this paper we will approximate all surfaces by a set of (not necessarily flat) polygons. If we have a surface S represented by a set of polygons in $\mathbf{R}^3$ with vertices (x_i, y_i, z_i), and corresponding normals (l_i, m_i, n_i); we can approximate the dual as the set of polygons in $\mathbf{RP}^3$ with vertices $(l_i; m_i; n_i; l_i x_i + m_i y_i + n_i z_i)$. Each point of this new set of polygons corresponds the the tangent plane $l_i x + m_i y + n_i z = l_i x_i + m_i y_i + n_i z_i$, the first three coordinates give the normal and the fourth is the distance from the origin to the tangent plane. In practice we represent the dual by the points $(l_i, m_i, n_i, l_i x_i + m_i y_i + n_i z_i)$ in $\mathbf{R}^4$, this simplifies the following steps.

The real problem with visualising the dual comes with finding projections from $\mathbf{RP}^3$ into $\mathbf{R}^3$. It is impossible to find a one to one map between these spaces; typically we can either expect a) the image of some points lie at infinity (projective map), b) some points in $\mathbf{RP}^3$ have two images in $\mathbf{R}^3$ (conformal map), or c) many points in $\mathbf{RP}^3$ are mapped onto the the same point in $\mathbf{R}^3$ (pedal curve). The following sections discuss the various projections and how we can deal with the degenerate points.

We should note at this point that many 3D graphics systems use homogeneous coordinates internally, and perform a mapping from $\mathbf{RP}^3$ to $\mathbf{R}^3$ as part of the graphics pipeline (this mapping is normally equivalent to the projective map discussed below). Typically these system have no special treatment of points at infinity so do not help us with the special problems of visualising duals. Furthermore the transformations used by these systems are not the most appropriate for visualising objects in $\mathbf{RP}^3$.

3 The Projective Map

The most frequently encountered projection is the *projective* or *Klein* map [**?**, **?**]. This is obtained from the map $(x; y; z; w) \rightarrow (x/w, y/w, z/w)$ often called the *dehomogenizing map*. We note that under this map all the points $(x; y; z; 0)$ are mapped to infinity. For many surfaces such as the double twisted Möbius band shown in figure **??**, the points of interest will be mapped to infinity and cannot be examined. In the case of the band which is defined by $x = \cos(\theta) + tu_x$, $y = \sin(\theta) + tu_y$, $z = tu_z$ where θ runs between $-\pi$ and π, t runs from -0.2 to 0.2, and (u_x, u_y, u_z) is the tangent vector $\cos(\theta)(\cos(\theta), \sin(\theta), 0) + \sin(\theta)(0, 0, 1)$, the points $(1, 0, 0)$ and $(-1, 0, 0)$ both have the same normal, $(0, 0, 1)$, hence the dual has a self intersection at the point $(0; 0; 1; 0)$, which we will not be able to see under this projection.

This problem can be partially solved by dividing by one of the other coordinates, which still gives a projective map. Here $(x; y; z; w)$ would be mapped to one of the points $(y/x, z/x, w/x)$, $(x/y, z/y, w/y)$ or $(x/z, y/z, w/z)$. Two projections of the twisted band are shown in figure **??**. In the first we divide by w and in the second
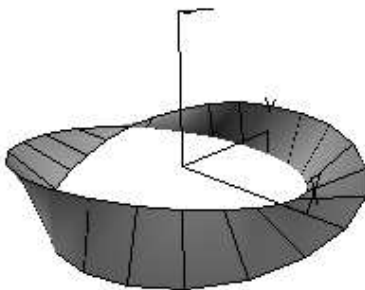
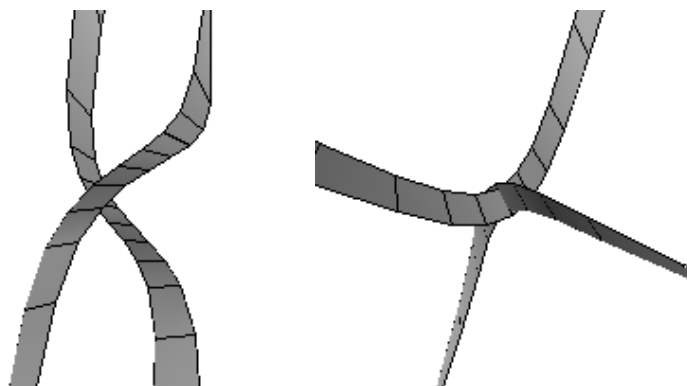Figure 1: A double twisted Möbius band

Figure 2: Two different projections of the dual of a double twisted Möbius band, using the projective model

we divide by z. Note that in the first picture there is no indication of the self-intersection, which is clearly shown in the second. Even in this family there may be surfaces where at least one of the points of interest on the dual will be mapped to infinity.

A far more general family of projections is obtained by mapping the dual into $\mathbf{R}^4$ using the map $(x; y; z; w) \rightarrow (x, y, z, w)$, applying a rotation in $\mathbf{R}^4$ and then projecting into $\mathbf{R}^3$ using the dehomogenizing map. Using this scheme we can find a rotation such that any chosen point (a, b, c, d), is rotated onto $(0, 0, 0, \sqrt{a^2 + b^2 + c^2 + d^2})$. The image of the point after the projection will be the origin, which is typically the center of our view.

4   Specifying Rotations in $\mathbf{R}^4$

Rotations in $\mathbf{R}^4$ are represented by 4 x 4 orthonormal matrices. To aid the viewing process we need to find a way the user can choose a particular matrix.

4

One way to specify a rotation matrix, is to define four numbers a, b, c, d with $a^2 + b^2 + c^2 + d^2 = 1$, and uses these as the elements of the fourth row. The rest of the components of the matrix can be computed automatically to ensure the orthonormal properties of the matrix. This fourth row is the most important as its entries give the coordinates of a point which will be rotated onto $(0, 0, 0, 1)$ and hence projected to the origin. One method for specifying the four points, used by C. Gunn in his 4-Axis graphical widget, is to choose a point (a, b, c) in the unit ball and set $d = \sqrt{1 - a^2 - b^2 - c^2}$.

We have used a different method, which is perhaps more intuitive. This involves a generalisation of the "Glass Sphere" model for rotations in three dimensions. If we are looking at a sphere in $\mathbf{R}^3$ along the z-axis, we can bring any point on the sphere into view by a repeated applications of small rotations R_x, R_y about the x and y axis. Rotations about the z axis are less important as they do not alter the parts of the sphere which are visible.

In four dimensions we define three rotations, R_{xw}^+, R_{yw}^+, R_{zw}^+ given by the matrices

$$
\begin{pmatrix}
\alpha & 0 & 0 & \beta \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
-\beta & 0 & 0 & \alpha
\end{pmatrix},
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & \alpha & 0 & \beta \\
0 & 0 & 1 & 0 \\
0 & -\beta & 0 & \alpha
\end{pmatrix},
\begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & \alpha & \beta \\
0 & 0 & -\beta & \alpha
\end{pmatrix},
$$

where $\alpha = \cos(\theta)$, $\beta = \sin(\theta)$ and θ is some small angle (10 degrees in our implementation). Rotations involving the first three coordinates are again of little interest and are better handled by a viewing program after the projection. By repeated applications of the rotations R_{xw}^+, R_{yw}^+, R_{zw}^+ and their inverses R_{xw}^-, R_{yw}^-, R_{zw}^- we can bring any point on the four sphere into the vicinity of $(0, 0, 0, 1)$. As each of the rotations is only by a small amount the resulting image does not change significantly with each application. This makes it easy to see the effect of the changes and navigate through the family of rotations. In fact when the R_{xw} rotation is applied the image will appear to move along the x-axis, with part of the image disappearing off one side of the screen and reappearing on the other side. Hence we have a fairly intuitive method of navigation. These rotations can be controlled by a set of six buttons: each time a button is pressed the corresponding rotation is applied. In our implementation we actually use twelve buttons, six with $\theta = 10°$ and six with $\theta = 1°$. The user interface for the program is shown in figure ??. At the far right are four buttons for selecting the default rotations which map $(0, 0, 0, 1)$, $(1, 0, 0, 0)$, $(0, 1, 0, 0)$ or $(0, 0, 1, 0)$ onto $(0, 0, 0, 1)$. To the left of these are the twelve buttons for applying the incremental rotations. One of the menu items displays the elements of the matrix which helps keep track of the rotations.
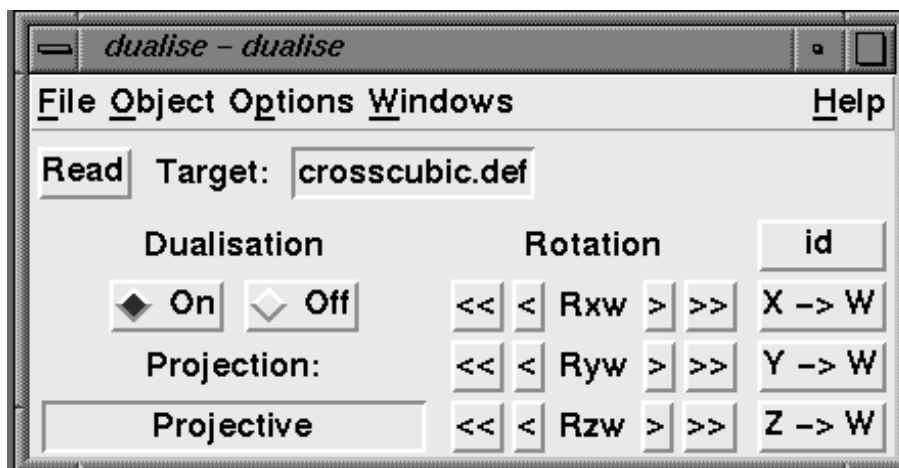
Figure 3: The user interface for the dualisation program

5   Problems with faces crossing the plane $w = 0$

If we consider the line in $\mathbf{R}^4$ joining the points $(1, 0, 0, 0.5)$ and $(1, 0, 0, -0.5)$ the endpoints will be projected onto $(2, 0, 0)$ and $(-2, 0, 0)$ using the dehomogenizing map. Connecting these points by a straight line would give a line passing through the origin, however the actual image should really be a curve which passes through infinity. When calculating the projection we really want to find an approximation to the curve. A related problem is that points with small values of w will be projected a great distance from the origin. Visually this is distracting, and improved results are obtained by only showing the part of the surface lying inside a sphere of radius R.

These problems can be overcome by first cutting the surface into two halves with $w > 0$ and $w < 0$ and then finding the points of the surface where $f(x, y, z, w) = R^2 w^2 - x^2 - y^2 - z^2 \geq 0$. This last equation defines a generalisation of a double cone in $\mathbf{R}^4$, the image of this cone will give a sphere of radius R. Any line which crosses this cone will be cut into two parts and we discard the part with $f < 0$. Figure ?? shows a line AB with $A = (1, 0, 0, 0.5)$, $B = (1, 0, 0, -0.5)$ and the cone corresponding to a sphere of radius 4. The line will first be cut at the point $C = (1, 0, 0, 0)$, and then at the points $D = (1, 0, 0, 0.25)$, $E = (1, 0, 0, -0.25)$. We will keep the segments AD and EB which project to the lines $(2, 0, 0)..(4, 0, 0)$ and $(-2, 0, 0)..(-4, 0, 0)$ respectively. Straight lines offer a fair representation of the actual image in the region inside the sphere. The segments DC and CE are discarded as they project to curves outside the sphere, and also have one point at infinity. The pictures in figure ?? show surfaces which have been sliced in this way which can be compared with figure ?? where this dual is shown without trimming, note the spurious facets.

The situation is further complicated if we consider the duals of surfaces where
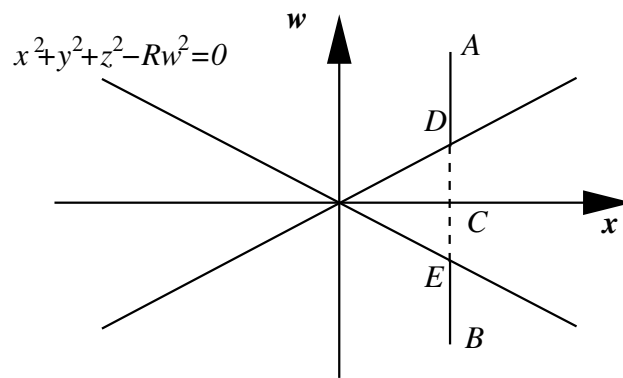
Figure 4: Cutting a line by a cone

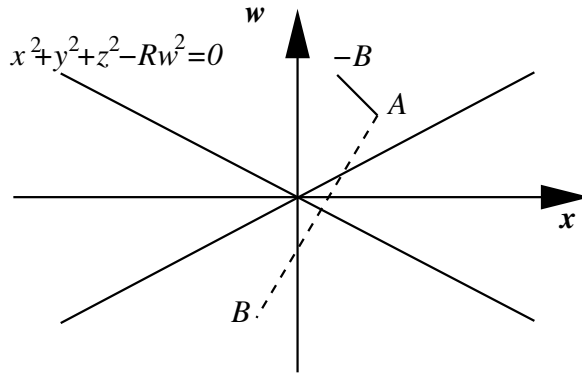Figure 5: The dual of the Möbius band without the trimming step.

Figure 6: Choosing local orientations.

the normals have not been given a consistent orientation. One example of this is the cuspidal edge $X = -s^2 - 0.3t^2$, $Y = s^3 + 0.5t^2$, $Z = t$ shown in figure **??**a. Here the normals change sign through $180°$ as we cross the edge, so points on either side of the edge will correspond to diametrically opposite points in $\mathbf{R}^4$, say (x, y, z, w) and $(-x', -y', -z', -w')$ with $x \approx x'$ etc. Using the above trimming algorithm would give several spurious points as show in figure **??**b. This problem can be overcome by choosing local orientations. Say we have two points A, B in $\mathbf{R}^4$ such that $A \cdot B < 0$, rather than considering the line connecting A and B we instead work with the line joining A and $-B$. This line will give the correct intersection points with the cone (Fig. **??**). For trimming a face A, B, C we actually work with the face A, B', C' where $B' = B$ if $A \cdot B > 0$ and $B' = -B$ if $A \cdot B < 0$ and similarly for C'. This convention only works locally, i.e. considering each face separately, where it is a reasonable assumption that the points are close together in $\mathbf{RP}^3$. Figure **??**c shows the correct dual of a cuspidal edge obtained in this way.

6 The conformal map

We can think of $\mathbf{RP}^3$ as the set of lines in $\mathbf{R}^4$ which pass through the origin. If we take the intersection of these lines with the lower hemisphere of the 3-sphere in $\mathbf{R}^4$, we can represent any object in $\mathbf{RP}^3$ as a set of points in the three-sphere. Explicitly we use the map $(x; y; z; w) \rightarrow sign(w)/l * (x, y, z, w)$ where $l = \sqrt{x^2 + y^2 + z^2 + w^2}$. We can then use any of the well-known projections of the 3-sphere in to $\mathbf{R}^3$ (see for example [**?**, **?**]). The most useful of these is the *conformal* or *stereographic* projection where we project from the point $(0, 0, 0, 1)$ onto the hyperplane $w = -1$ i.e. $(x, y, z, w) \rightarrow 2/(1 - w)(x, y, z)$. We can see that the lower hemisphere will be projected into the interior of the ball of radius 2. This gives us an important property for visualisation: *the whole of of a dual surface can be seen at the same time*, with the cautionary note that antipodal points on the boundary of the ball

8

(a)

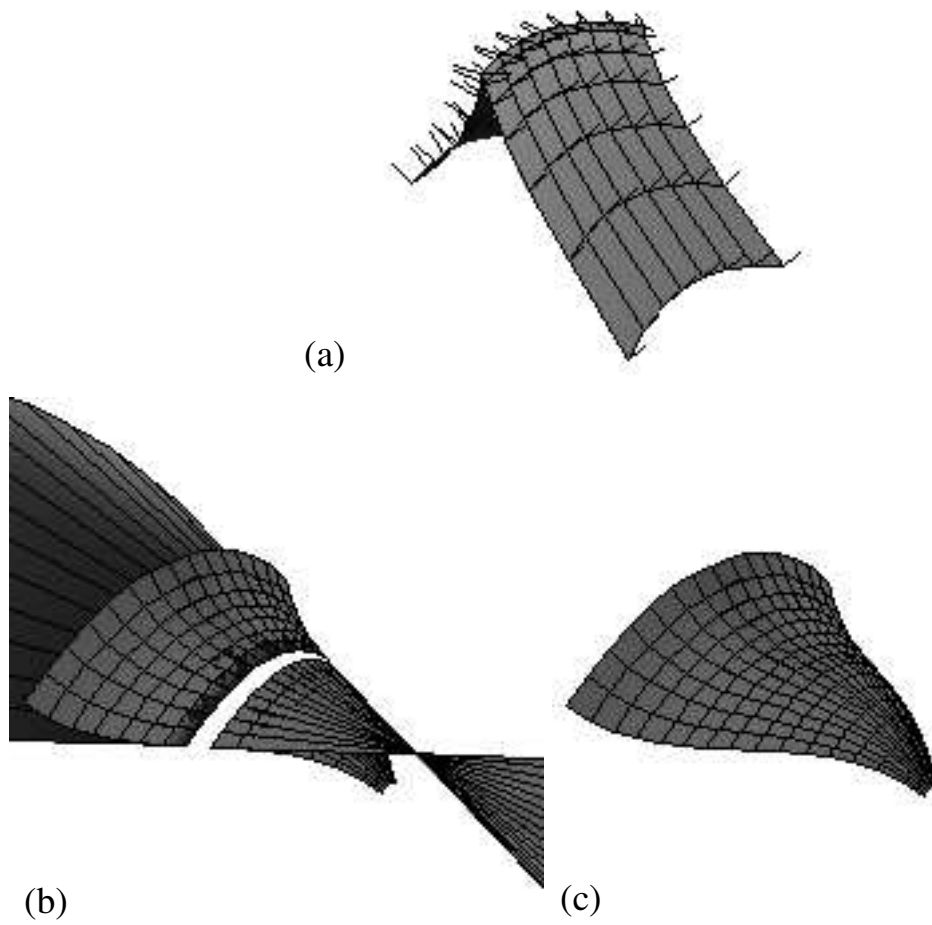(b)   (c)

Figure 7: A cuspidal edge (a) and its dual which has been trimmed using a naive method (b) and taking local orientations into account (c).
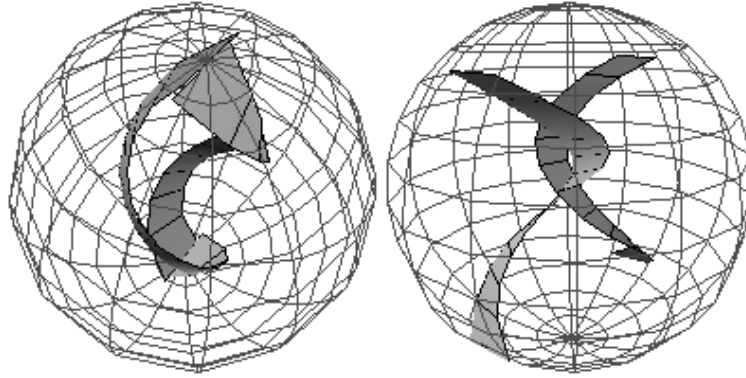
Figure 8: Two different projections of the double twisted Möbius band using the conformal model.

should be identified. We can use the same rotations and trimming algorithms from the previous sections to give a whole family of projections. In this case intersecting with a cone insures the image lies strictly in the interior of the ball and any problems with the antipodal points are avoided. Figure **??** shows the image of the twisted band using the conformal projection and the same two rotations as in figure **??**. Even though though the self intersection lies at infinity in the first of these pictures, there are still strong clues as to its existence which were not indicated using the projective map.

Using the conformal map most of the properties of the dual are preserved. However conformal maps typically map straight lines to circles, so we need to re-write 2^*:

2^* If S has a self-intersection then the dual S^* has a bi-tangent sphere.

We also lose the property that the dual of a dual gives us our original surface. In many application the ability to see the whole of the dual outweighs the loss of this information. As straight lines map to circles we should really use curves for the boundaries of our polygons. In practice this step is not necessary provided we have a fine enough mesh of polygons approximating our original surface.

7   Other Projections

Once the mechanism for calculating the dual, performing rotations, and trimming the surface have been set up it is an easy matter to add other projections. Let (x, y, z, w) be a point on the dual in $\mathbf{R}^4$. Some other examples of projections include:

The Gauss Map: $(x, y, z, w) \to 1/\sqrt{x^2 + y^2 + z^2}(x, y, z)$. This maps normals at each point on S onto its spherical image on the unit sphere. This map has been extensively studied [**?**] and makes apparent many features of the original

surface. Parabolic lines become folds in the image and *Cusps of Gauss* become cusps on the fold line. However bi-tangent lines can not be identified.

The Pedal Curve: $(x, y, z, w) \rightarrow (xw, yw, zw)$. This map picks out the point on the tangent plane to S closest to the origin. Most of the features of the dual are represented here, and for a compact surface the image will be compact, however all the tangent planes through the origin will be represented by a single point, the origin.

The Geomview program allows hyperbolic, and spherical spaces to be viewed as if the viewer lived inside the space [**?**]. Instead of using the normal Euclidean matrix to determine the straight lines for the light paths, the appropriate metric for hyperbolic or spherical space determine the geodesics in that space and hence what parts of the surfaces can be seen. Motions through this space are also determined using these metrics. This offers a very different method for studying the surface, rather than the dual being viewed as an object we could pick up and examine, the dual becomes an object which we fly around to get to the points of interest. For questions like "How many cusps does the dual of this surface have?" the previous methods may be more appropriate. In conjunction with this *virtual* method we can represent a dual surface in a number of ways:

Positive Hemisphere: $(x, y, z, w) \rightarrow sign(w)/l * (x, y, z, w)$;

Negative Hemisphere: $(x, y, z, w) \rightarrow -sign(w)/l * (x, y, z, w)$;

Oriented version: $(x, y, z, w) \rightarrow 1/l * (x, y, z, w)$;

where $l = \sqrt{x^2 + y^2 + z^2 + w^2}$. The first two of these give images which do not depend on the orientation, i.e. if we reverse the directions of the normals on the original surface we will still get the same points on the four sphere. The third projection respects orientation, reversing the direction of the normal will give an antipodal point. One final method is obtained by taking the unions of the images in the positive and negative hemispheres, this creates a double version of the dual which appears to be continuous as we fly around it.

So far we have always constructed the dual with respect to the origin: when constructing the dual we take the components of the normal vector and the distance from the origin as our point in $\mathbf{RP}^3$. It may well be the case that the origin is not the best place to dualise about, for example when the surface actually passes through the origin. It is a simple matter to change the definition of the dual, and use $(l; m; n; l(x - x_0) + m(y - y_0) + n(z - z_0))$ as the dual of the point (x, y, z) with normal (l, m, n). The values of x_0, y_0, z_0 can then be specified by the user. Changing the center in this way has no effect on the structure of cuspidal edges and self intersections of the dual, but may help to avoid problems with the $w = 0$ plane. This is equivalent to translating the surface and then calculating the dual.
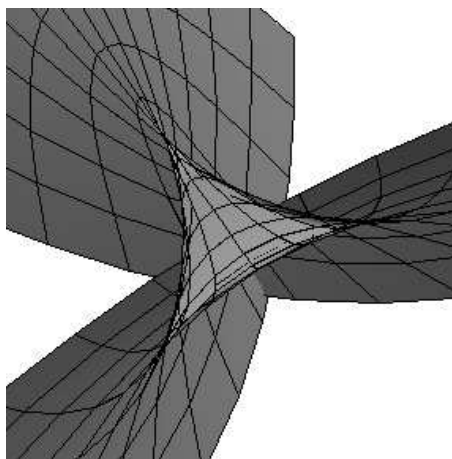
11

Figure 9: The dual of a slight deformation of a flat elliptical umbilic.

8 Conclusion

So far the computer program described here has proved useful in a number of investigations, in singularity theory. In the study of flat umbilics [?] Tari and myself quickly produced pictures of duals of the two cases, (elliptical and hyperbolic), and also of slight deformations of these surfaces which show a more generic geometry. Figure ?? show the dual of a slight deformation of an flat elliptical umbilic, note how the dual has three swallowtail points linked by a cuspidal edge. Furthermore the dual also contains a triple point (hidden from view). This allowed us to quickly deduce the geometry of the original surface, which must have three cusps of Gauss, and one tri-tangent plane. Whilst the structure of parabolic lines on a surface is easy to calculate, finding bi-tangent planes typically involves solving three equations in four unknowns which can take a considerable time, such curves are very apparent when we look at the dual.

A trickier investigation was the duals of cross-caps [?], parameterised by $(u, uv, au^2 + buv + v^2 + O(3))$. Here the dual of the actual cross-cap point should give a line in $\mathbf{RP}^3$ (taking the limiting tangent planes as we approach the cross-cap point from different direction give a one parameter family of planes). Whilst the above program can not deal with such singular points, we can remove a small neighbourhood of the origin and calculate the dual of the resulting surface. If we take a small enough neighbourhood the dual will have an edge corresponding to this exceptional line. Here the conformal projection is particularly useful as it enables us to see the whole of this edge, which will always cut the plane $w = 0$ no matter what rotation we use, (Fig. ??).
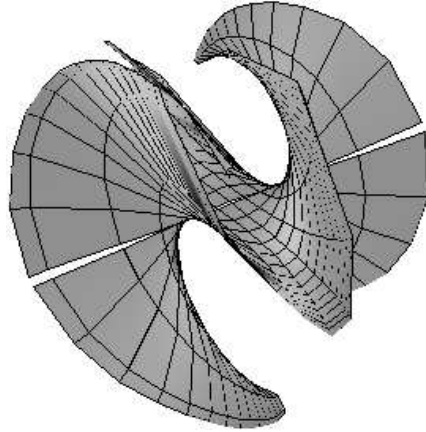
12

Figure 10: The dual of a cross-cap.

References

[1] T.Banchoff, T.J.Gaffney, C.McCrory, *Cusps of Gauss Mappings*, Pitman, New York, 1982.

[2] J.W.Bruce, P.J.Giblin, F.Tari, *Families of Surfaces: Height Functions, Gauss Maps and Duals*, To appear in Proc. 2nd International Conf. on Singularities, Sao Carlos, Brazil.

[3] C.Gunn, *Visualising Hyperbolic Geometry*, in *Computer Graphics and Mathematics*, pp 299-313, Eurographics, Springer Verlag, 1992. Geometry Center research report GCG 34.

[4] D.Hilbert, S.Cohn-Vossen, *Geometry and the Imagination*, Chelsea Publishing Company, 1952.

[5] J.Illingworth, J.Kittler, *A survey of the Hough transform*, Computer Vision, Graphics and Image Processing, Academic Press, New York, Vol. 44 pp87-116, 1988.

[6] R.J.Morris, *The Liverpool Surface Modelling Package reference manual*, preprint University of Liverpool, 1995.

[7] R.J.Morris, *Tools and Techniques for the visual analysis of sing theory*

[8] M.Philips, *Geomview reference manual.* Geometry Center research reports GCG 62, Geometry Center, Minneapolis.

[9] J.M.West, *The Differential Geometry of the Crosscap* Thesis, University of Liverpool, 1995.